Rates of Common Classroom Behavior Management Strategies and Their Associations With Challenging Student Behavior in Elementary School

Julie Sarno Owens, PhD[1], Alex S. Holdaway, MS[1], Jessica Smith[1], Steven W. Evans, PhD[1], Lina K. Himawan, MS[1], Erika K. Coles, PhD[2], Erin Girio-Herrera, PhD[3], Clifton S. Mixon, MA[1], Theresa E. Egan, MA[1], and Anne E. Dawson, MS[1]

Abstract

In this study, we (a) describe patterns of challenging student behaviors (classwide and for a target student with attention deficit hyperactivity disorder [ADHD]) and teacher behaviors (i.e., praise, commands, and responses to challenging behavior) in kindergarten through Grade 5 classrooms, (b) examine the relations between these behaviors, and (c) describe a threshold of teacher behaviors most associated with low levels of challenging student behavior. Participants were 55 teachers observed using a modified version of the Student Behavior Teacher Response (SBTR) system. Across grades, there was variability in rates of classwide challenging behavior per hour ($M = 35.81$ to 102.62) and rates of praise per hour ($M = 10.90$ to 37.70). The percentage of challenging behaviors to which teachers responded appropriately was generally low ($M = 27\%$ to 47%) and stable across grades. For classwide challenging behavior, higher percentages of appropriate teacher response were significantly associated with lower rates of challenging behavior ($b = -.43$; $p < .01$), but effective commands and labeled praise were not. Classwide challenging behaviors dropped to 30 per hour once teachers reach a threshold of 51% appropriate response, with little incremental benefit at higher levels. Implications for professional development and future study of behavior management practices are discussed.

Keywords

classroom management, behavior management, rule violations, praise, commands

In recent years, there has been increased attention on teacher evaluation and the relations between instruction, classroom management, and student outcomes (e.g., Every Student Succeeds Act, 2015; Kane, Taylor, Tyler, & Wooten, 2011). Indeed, many state-level educator performance evaluation systems (e.g., Ohio Department of Education, 2015) have attempted to formalize evaluation of teachers' instructional behaviors and classroom management practices. However, the science behind these evaluations is lacking, leaving school administrators without data-derived benchmarks for evaluating and promoting the teacher behaviors most associated with positive student outcomes.

There is a strong history of empirical study of effective classroom management, most of which is grounded in behavioral theory (Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). Classic studies (e.g., Acker & O'Leary, 1987; Witt & Elliot, 1982) and modern classroom management resources (e.g., Gable, Hester, Rock, & Hughes, 2009; Marzano, Marzano, & Pickering, 2003; Wong & Wong, 2001) highlight the effectiveness of contingent positive reinforcement of appropriate behavior, use of clear and explicit commands, use of classroom rules, and strategic responses to violations of these rules in contributing to important student outcomes including math and reading achievement, social development, and behavioral outcomes. In contrast, poor management of challenging student behaviors is associated with greater time off-task and lower student achievement. As such, current multitiered frameworks for positive behavioral supports recommend that schools identify a set of behavioral expectations, and teachers reinforce those expectations and

[1]Ohio University, Athens, USA
[2]Florida International University, Miami, USA
[3]Towson University, MD, USA

Corresponding Author:
Julie Sarno Owens, Ohio University, Porter Hall 200, Athens, OH 45701, USA.
Email: owensj@ohio.edu

provide corrective feedback to help children meet these expectations (e.g., Reinke, Herman, & Stormont, 2013).

However, there are limitations to the scientific literature that preclude the development of benchmarks for teacher effectiveness. First, in many studies, researchers aggregated data about teacher and student behaviors across grade levels (Burnett & Mandel, 2010; Reinke et al., 2013), only provide data for one grade level (Floress & Jenkins, 2015), or use single-subject designs (Sutherland, Wehby, & Copeland, 2000). Second, only a few studies examine relations between teacher and student behaviors (Reddy, Fabiano, Dudek, & Hsu, 2013; Reinke et al., 2013), which is necessary to identify benchmarks for success. Given the inclusiveness of the 21st century classroom (60% of students with disabilities spend 80% of each day in the general education classroom; U.S. Department of Education, National Center for Education Statistics, 2015), and the high prevalence of challenging student behaviors in the general population (i.e., elevated rates of inattention, impulsivity, and noncompliance are present in 12% to 20% of students; (Fabiano, Pelham, et al., 2013; Kamphaus, Huberty, DiStefano, & Petoskey, 1997), it is important to examine the impact of teacher behavior on the class as a whole, as well as on target students. This distinction is important because students with elevated rates of challenging behaviors, like students with attention deficit hyperactivity disorder (ADHD), are stressful for teachers (Greene, Beszterczey, Katzenstein, Park, & Goring, 2002), have a negative impact on student–teacher relations (Birch & Ladd, 1998), detract time from instruction (Robb et al., 2011), and significantly contribute to teacher job dissatisfaction and attrition from the profession (Bibou-nakou, Stogiannidou, & Kiosseoglou, 1999). Furthermore, the academic outcomes for such students can be severe, particularly in the absence of effective classroom management skills or behavioral intervention (Fabiano et al., 2010; Owens, Murphy, Richerson, Girio, & Himawan, 2008). Given that students with these behaviors place a unique burden on teachers, impact the overall classroom, and are likely uniquely impacted by poor quality classroom management, research on the relations between teacher classroom management strategies and student behavior at both the classwide level and with students with or at risk of ADHD is needed.

Thus, the goal of this study was to contribute to this effort by (a) examining student and teacher behaviors by grade level, (b) examining the relative strength of the relations between teacher behaviors and student behaviors (classwide and target students), and (c) exploring thresholds of key teacher behaviors most associated with low rates of challenging student behavior. Below, we critique the literature related to three common classroom management strategies (praise, effective commands, and appropriate response to challenging behavior) as they pertain to the entire class and students with or at risk of ADHD. Although there are many strategies that could be examined, these three strategies were selected for study because of the evidence supporting their effectiveness in influencing student behavior and because they could all be captured by the observation system selected. We are unaware of other observation systems that capture the frequency of these strategies and the explicit temporal connection between student and teacher behaviors (e.g., student rule violations and teacher's response to them).

Evidence-Based Classroom Management Strategies

Praise

Classwide level. Praise is generally characterized as favorable verbal or nonverbal recognition directed toward a student following desirable behavior (see Jenkins, Floress, & Reinke, 2015, for review). Consistent with operant conditioning, praise theoretically increases desired behavior by providing recognition as reinforcement and subsequently reduces behaviors that are incompatible with the desired behavior. Teachers' use of praise toward shaping behavioral expectations in the classroom has been studied since the 1970s (Brophy, 1981; White, 1975) and interest in this topic continues today (Floress & Jenkins, 2015; Reddy et al., 2013). Published data document wide variability in teachers' rates of praise both across and within studies. Average rates of praise have been reported to range from less than five praise statements per hour (Brophy, 1981) to more than 40 per hour (Floress & Jenkins, 2015; White, 1975). Data across studies suggest that praise is more common in primary grades (K to Grade 3; Floress & Jenkins, 2015) than in intermediate elementary grades (Grades 4 and 5; see Brophy, 1981), and that unlabeled praise ("good job") is more common than labeled praise ("I like the way you are sitting quietly"; Floress & Jenkins, 2015).

The authors of many classroom management studies (e.g., Reinke et al., 2013; Sutherland et al., 2000) assert that labeled praise is more effective than unlabeled praise for reinforcing the desired behavior. However, most of the evidence supporting this claim comes from single subject research (e.g., Sutherland et al., 2000), studies with small samples sizes (e.g., Chalk & Bizo, 2004), or intervention studies that attempt to modify teacher behavior (e.g., Anderson, Evertson, & Brophy, 1979) rather that naturalistic classroom observation. Thus, the conclusions drawn from these studies may not generalize to or be representative of typical classroom conditions. Furthermore, recent studies that have delineated between labeled and unlabeled praise have either not included measurement of child behavior (Reddy et al., 2013; Reinke et al., 2013) or have been limited by small sample size (e.g., four teachers in Floress & Jenkins, 2015). As such, studies that include larger samples, both

teacher and student behavior, and examination of praise under naturalistic classroom conditions may clarify the potential importance of labeled versus unlabeled praise.

With target students. Because students with or at risk of ADHD tend to demonstrate a higher frequency of negative behavior and lower frequency of positive behavior (Atkins, Pelham, & Licht, 1985), it has been argued that praise is a powerful tool for shaping individual student behavior toward developmentally appropriate expectations. Indeed, the topic of praise is embedded in most training programs for parents of youth with ADHD or oppositional behavior (Barkley, 2013; McMahon & Forehand, 2003). However, teachers' use of praise at the classwide level may not be representative of praise used toward target students with disruptive behavior. First, some experts encourage teachers to praise at a rate that represents two to four times the amount of correction the student receives (e.g., Kalis, Vannest, & Parker, 2007). If teachers follow this recommendation, rates of praise toward targets students would exceed that toward typical peers. Second, some theorists (e.g., Brophy, 1981) argue that the higher rate of praise needed to shape the behavior of students with behavior problems may be intrusive for typical students.

Effective Commands

Classwide level. Another important classroom management behavior is effective instructions or effective commands. Effective commands are typically defined as proactive (i.e., used prior to disruptive behavior) verbal or nonverbal communication that provides a clear expectation for behavioral change (initiation, cessation, or modification; Colvin, Sugai, Good, & Lee, 1997). Common recommendations include using commands that are (a) developmentally appropriate, (b) stated in the positive, (c) stated in one-to-two steps, (d) specific, and (e) allow time for compliance. Although these indicators of effectiveness have been widely disseminated, much of this work is based on the parent training literature (e.g., Forehand, 1977; Forehand & McMahon, 1981) with disruptive or oppositional children and less work has documented the utility of teachers' use of effective commands in the classroom.

Theoretically, the use of effective commands in the classroom ensures the efficient progression of activities and minimizes time off-task by providing explicit developmentally appropriate expectations and guidance. We are aware of only two studies that have examined rates of effective commands with sample sizes greater than 20 classrooms (Reddy et al., 2013; Vujnovic et al., 2014). Vujnovic et al. found that, on average, during a 30-min observation period, preschool teachers used effective commands ($M = 15.67$; $SD = 10.20$) at a higher rate than ineffective commands ($M = 0.77$; $SD = 1.44$). Reddy and colleagues found a similar pattern across two 30-min observations in elementary classrooms ($M = 17.09$; $SD = 8.77$ for effective commands; $M = 3.67$; $SD = 3.92$ for vague commands). Reddy et al. also found that the use of commands declined slightly across grades (19 per observation in kindergarten to 13 in Grade 5). Yet, replication is necessary to establish the generalizability of results. Furthermore, neither study examined the relations between each type of instruction and challenging student behavior.

With target students. Several studies document the relationship between the use of parent's effective commands and increased compliance in children with noncompliant behavior (Roberts, McMahon, Forehand, & Humphreys, 1978; Williams & Forehand, 1984). Several small-scale, multiple-baseline trials have demonstrated increased student compliance or academic engagement when effective commands are used in the classroom or during transitions, as compared with baseline periods (Colvin et al., 1997; Ford, 1998; Neef, Shafer, Egel, Cataldo, & Parrish, 1983). Matheson and Shriver (2005) isolated the effects of commands (i.e., in the absence of reinforcement) in a sample of three children selected for noncompliant behavior. The data showed that teachers' use of effective commands produced improved compliance, with effective commands plus praise resulting in additional improvements.

Teacher Response to Challenging Student Behaviors

Classwide level. As with praise and effective commands, response to challenging student behaviors has been a focus of classroom management research since the advent of the field (Anderson et al., 1979; Madsen, Becker, & Thomas, 1968). Although teachers have autonomy in establishing expectations for appropriate classroom behavior and there is some variability across student developmental level, common classroom expectations include be respectful, obey adults, work quietly, use materials appropriately, remain in seat, raise hand to speak, and stay on task. Because of the common focus on these expectations, many classroom observation systems have developed operational definitions for these student behaviors (e.g., Abikoff & Gittelman, 1985; Saudargas, 1997; Vujnovic et al., 2014). Similarly, most best-practice recommendations for appropriately responding to disruptive behavior include (a) gaining the attention of the offending student, (b) providing a brief description of the alternative desired behavior, (c) using a neutral tone of voice, and (d) allowing the student time to exhibit the desired behavior (Lane, Gresham, & O'Shaughnessy, 2002; Reinke et al., 2013; Rhode, Jensen, & Reavis, 1992).

However, questions remain as to the typical rate at which teachers appropriately respond to violations of expectations and the magnitude of the relation between response to violations and overall rates of violations. Early studies provide evidence that classroom expectations are ineffective when insufficiently enforced and when students perceive teachers as lax enforcers of the expectations, suggesting that consistent enforcement of the expectations may be key to affecting challenging behavior (e.g., Madsen et al., 1968). However, several early studies (e.g., Madsen et al., 1968; Neef et al., 1983) were based on small sample sizes (i.e., less than 10 teachers) and thus offer limited information about the relations between teacher responses and student behavior. More recent studies shed light on teacher behavior in the modern classroom, but are not without similar limitations. For example, in a sample of 33 classrooms (K–Grade 3), Reinke and colleagues (2013) reported the rates at which teachers responded to violations (overall average = 0.67 per min; or approximately 20 per 30-min observation); however, these data were not linked to student behavior. Similarly, Reddy and colleagues (2013), using data from two 30-min observations per teacher, reported the frequency of response to violations at the classroom level among a sample of 317 teachers. In this sample, response to violations (termed behavioral corrective feedback) occurred an average of 8.86 (SD = 6.98) times per 30-min observation. A general decrease in frequency was observed between younger (kindergarten: M = 11.69) and older grades (Grade 5: M = 5.93). However, the relation between teacher behavior and child behavior was not assessed.

Vujnovic and colleagues (2014) examined the percentage of violations responded to appropriately by preschool teachers (N = 88), reporting that teachers responded appropriately to 59% of violations per observation. However, the authors did not provide data on the relation between this teacher behavior and student violations. Furthermore, the generalizability of these data to elementary school is limited given that the sample was preschool teachers and Reddy et al. (2013) reported that teacher response to violations decreases across elementary grades.

With target students. In addition to classwide response to violations, some researchers have examined teacher response to violations for students with disruptive behavior. Pfiffner and O'Leary (1987) showed that to achieve acceptable rates (80%) of on-task behavior among students (N = 8) with academic and behavior problems, teachers needed to use negative consequences and appropriate response to violations in addition to positive encouragement (i.e., use of positive encouragement alone was insufficient). Not surprisingly, these students require greater consistency and potency of teacher behavior management than most other students.

Consideration of Developmental Level

As mentioned above, students at different grade levels may have different needs and teachers at different grade levels likely have different expectations for academic performance and behavioral control. For example, kindergarten students may need more supervision and assistance to follow classroom rules than Grade 5 students. Thus, teacher praise may be common when shaping early socialization patterns of younger students, but may be used less frequently with Grade 5 students who are very familiar with the school setting. Similarly, teachers of older students may respond to a higher percentage of violations than teachers of younger students, because there may be fewer overall violations, and/or violations of older students may be more egregious (e.g., disrespect versus leaving seat). Finally, for some children (e.g., those demonstrating frequent deviance behaviors), the increasing influence of peers as children age may decrease the relative impact of the teacher's appropriate response to rule violations, if peer attention is more reinforcing than the consequence given by the teacher. However, it also possible that grade level can be collapsed into broader, yet meaningful categories, such as primary and intermediate grades, as is evident in many teacher education programs and teacher licensure categories. Primary classrooms (kindergarten–Grade 3) differ from intermediate classrooms (Grades 4 and 5) in that they are often self-contained and have lower demands for independence and autonomy from students; these characteristics may affect teachers' use of praise and commands and response to violations. Thus, examining behavior by grade level and/or in meaningful developmental segments may reveal important patterns.

Although rates of student and teacher behaviors may vary by developmental level, the key components of effective classroom management practices (e.g., praise, effective commands, response to violations of expectations) are similar across all developmental levels, and the expectations of students (e.g., respect, obey adults) are similar enough across kindergarten to Grade 5 that one set of operational definitions can be used for observing and coding student and teacher behaviors in elementary school settings (e.g., Abikoff & Gittelman, 1985; Vujnovic et al., 2014).

Summary and This Study

Collectively, the literature on teacher classroom management practices has several limitations. Namely, in some studies, researchers only studied one grade level, limiting our understanding of the potentially differential impact of teacher behavior on student behavior across development. Several studies used single-subject designs, rather than portraying natural rates of behaviors in typical classrooms. Many studies failed to examine the relation between teacher

behavior and student behavior, limiting our understanding of the utility of, or relative importance of, each commonly recommended classroom management skill.

Thus, the goals of this study were to advance the literature by examining the rates of student behavior and teacher behavior by grade level (Aim 1), examining the strength of the relations between key teacher behaviors and student rule violations at a classwide level and for a student with or at risk of ADHD (Aim 2), and exploring thresholds of teacher behaviors most associated with low rates of student rule violations (Aim 3). We expected rates of teacher and student behaviors to vary by grade (e.g., a decline in the frequency of praise and violations by grade) and we expected appropriate response to rule violations, use of labeled (but not unlabeled) praise, and use of effective (but not ineffective) commands to be associated with lower student rule violations.

Method

Participants

Teacher participants were 55 elementary school teachers (25 from Ohio, 30 from Florida) who were participating in a grant-funded multisite consultation study designed to facilitate teachers' implementation of effective classroom management strategies (i.e., those previously described) and a daily report card (DRC) intervention. (The DRC is a well-established school intervention for children with or at risk of ADHD; Evans, Owens, & Bunford, 2014; Owens et al., 2012; Vannest, Davis, Davis, Mason, & Burke, 2010). Most teachers were women (94.5%) and identified as Non-Hispanic White (49.1%) or Hispanic (any race) (47.3%). They had an average of 14.43 years ($SD = 8.58$) of teaching experience and 8.90 years ($SD = 7.54$) teaching at their current school. Most (62%) had obtained a master's degree or higher.

Target student participants were 55 elementary school students (76.5% male, 56.4% Hispanic, any race). Most target students (92.7%) met criteria for ADHD (69.1% combined presentation, 21.8% inattentive presentation, 1.8% hyperactive/impulsive presentation) and the remaining 7.3% were at risk of ADHD (elevated symptoms plus impairment). The sample had an average IQ estimate of 98.04 ($SD = 12.46$), as assessed by the *Wechsler Abbreviated Scales of Intelligence–Second Edition* (WASI-II; Wechsler, 2011). The socioeconomic status of their families was low to middle class (20% had a household income of less than US$15,000, 60% had an income between US$15,000 and US$49,999, and 20% were above US$50,000). Per parent report at intake, 9.1% had been diagnosed with a learning disability and 23.6% had a medication prescription.

Teachers and students were recruited from eight participating schools across the two sites. The Ohio schools had an average of 377 students and 16 general education teachers per school, with 12% to 29% of students receiving special education services and 35% to 75% receiving free or reduced lunch services. The Florida schools had an average of 1,024 students and 50 general education teachers, with 4% to 11% receiving special education services and 76 to 95% receiving free or reduced lunch services. The racial makeup of schools was predominantly Caucasian (range: 90%–98%) in Ohio and primarily Latino in Florida (range: 94%–98%).

Procedures

Procedures were approved by the Institutional Review Boards at the two universities and within all participating school districts. To recruit teachers, the primary investigator at the site met with the administrator and staff at each school to describe the project. To enroll, teachers had to refer a student with inattentive and/or disruptive behavior and academic impairment and the student had to meet eligibility requirements. Namely, investigators conducted an assessment to determine that the student had an IQ estimate score of 80 or above within the 95% confidence interval, was in the general education classroom at least 50% of the day, was demonstrating classroom impairment and symptoms of inattention and/or hyperactivity/impulsivity, and had no prior diagnosis of an Autism Spectrum Disorder, Bipolar Disorder, or Cognitive or Developmental Disability per parent report.

Once the student (henceforth referred to as target student) was enrolled, baseline observations of the teacher and the student began (see description of observation system below). Most (approximately 90%) of baseline observations occurred in the fall (i.e., prior to the end of November); however, a few cases were referred later in the year and their baseline observations occurred at the time of referral (i.e., December through March). Observations lasted between 15 and 45 min; on average, the observations lasted for 27.68 min for kindergarten, 33.51 min for Grade 1, 30.08 min for Grade 2, 29.52 min for Grade 3, 32.39 min for Grade 4, and 24.83 min for Grade 5. During the weeks in which baseline observations were occurring, teachers also participated in three meetings with the consultant to provide information on their classroom management practices and identify the student's target behaviors to be addressed by the school-based intervention. All data presented in this article are derived from observations that occurred *prior* to the teacher's receipt of consultation about classroom management practices.

Volpe, McConaughy, and Hintze (2009) found that reliable estimates of disruptive student behaviors can be achieved with 50 to 60 min of observation, and reliable estimates of inattention and off-task behaviors can be achieved with 110 to 140 min of observation. Hill, Charalambous, and Kraft (2012) found that reliable estimates of teacher behaviors for instruction was achieved within four 30-min

observations. To align with these findings, teacher–student dyads were included in the analyses if they had at least three baseline observations totaling at least 120 min prior to the start of the consultation intervention. (This criterion resulted in the exclusion of four cases from the total available sample in the larger grant-funded project, $N = 59$). Teachers retained in the sample had between three and nine baseline observations ($M = 4.60$, $SD = 1.23$).

Measures

Student Behavior–Teacher Response observation rating system (SBTR). The SBTR (Pelham, Greiner, & Gnagy, 2008) is a systematic classwide observation system that was developed using a behavior theory framework and designed to capture discrete student–teacher interactions in preschool and elementary classrooms (Fabiano et al., 2010; Fabiano, Vujnovic, et al., 2013; Vujnovic et al., 2014). In previous studies (Fabiano, Vujnovic, et al., 2013; Vujnovic et al., 2014) interrater reliability has been indexed by the correlations between observers; they were significant and of large magnitude (*r* for total challenging behaviors > .88, effective commands > .80, praise > .88), although appropriate response to rule violations was lower (*r* = .57). As evidence of convergent validity, the frequency of student rule violations and teachers' appropriate responses to these behaviors was correlated in the expected direction with other measures of classroom climate (*r*s range from .20 to .45; Massetti, Pelham, & Waschbusch, 2007; Vujnovic et al., 2014). Finally, scores on the SBTR have demonstrated sensitivity to change as a function of intervention (Fabiano et al., 2010). Observers obtained (a) frequency counts of specific rule violations by the target student, (b) frequency counts of total rule violations by all other students in the classroom collectively, (c) frequency counts of how the teacher responded to each of those types of violations, (d) frequency counts of the teacher's use of social reinforcement (i.e., labeled and unlabeled praise to the target student, other individual students, and groups of students), and (e) frequency counts of the teacher's use of commands (effective and ineffective to the target student, other individual students, and groups of students). The observation manual includes definitions and inclusion and exclusion examples for coding violations of seven expectations (i.e., be respectful, obey adults, work quietly, use materials appropriately, remain in seat, raise hand to speak, stay on task), the teacher's response to each violation (i.e., coded as appropriate, inappropriate, or no response), teacher labeled and unlabeled praise, and effective and ineffective commands. An appropriate response is defined as any verbal or nonverbal action that follows a rule violation to provide a response to the behavior. Appropriate responses contain appropriate content, are delivered with appropriate affect, with a neutral tone of voice of normal pitch and intensity, and without including any behavior included in the Inappropriate Response definition (i.e., verbal or nonverbal behavior that is antagonistic, accompanied by excessive or inappropriate gestures, or delivered with inappropriate affect or an inappropriate tone of voice.). All definitions are available upon request from the first author.

Observers were trained to reliability on the SBTR. They attended an initial training, were required to pass (100% accurate) a written definitions test, were required to pass (100% accuracy) coding of 2 to 5 min video clips, and were required to achieve at least 80% reliability across all coded behaviors in a classroom with a master observer. Maintenance of reliability was checked throughout the year. Interobserver assessments were conducted for 24% of all observations. To assess the interrater reliability, we computed intraclass correlations (ICC) of Type 1 for average of *k* raters, that is ICC (1, *k*) as outlined in Shrout and Fleiss (1979) or ICC (*k*) for Case 1 as outlined in McGraw and Wong (1996); henceforth, we use the notation of ICC (1, *k*) because we did not have a set of consistent raters across all ratees and we would like to observe the reliability across a set of raters rather than one rater only. Across all frequency count variables, the ICC (1, *k*) ranged from 0.78 to 0.98 with an average of 0.90. Specifically, the ICC (1, *k*) of the target child violations ranged from 0.79 to 0.98 with an average of 0.89. The ICC (1, *k*) for the teacher's appropriate response to the target child was 0.88 and to the other child was 0.94. The ICC (1, *k*) for all teacher's praise variables ranged from 0.88 to 0.95 with an average of 0.93. The ICC (1, *k*) all instruction variables ranged from 0.78 to 0.95 with an average of 0.89.

To prepare the data for analyses, the following calculations were conducted first for the target student, then for all other students in the classroom. First, each frequency count variable (e.g., total rule violations, total labeled praise, and total effective commands) was summed for a given observation and divided by the total duration of the observation in minutes. This number was multiplied by 60 to produce a rate per hour. These rates were averaged across all observations for the case. Second, the total number of appropriate teacher responses to target student rule violations for a given observation was divided by the total number of rule violations by the target student for that observation period. This produced the percentage of appropriate response to rule violations per observation. These percentages were averaged across all observations. This was repeated for appropriate response to rule violations by all other students.

Results

Aim 1: Teacher and Student Behaviors by Grade

In Table 1, we provide descriptive statistics of student behavior (i.e., classwide and target student violations) and

Table 1. Descriptive Data for Student and Teacher Behaviors by Grade.

Variable	Total sample N = 55 (M, SD)	K n = 11 (M, SD)	Grade 1 n = 8 (M, SD)	Grade 2 n = 12 (M, SD)	Grade 3 n = 14 (M, SD)	Grade 4 n = 6 (M, SD)	Grade 5 n = 4 (M, SD)
Student behavior							
Target student violations rate/h	14.91 (10.75)	17.94 (8.72)	14.35 (5.85)	15.21 (8.81)	11.45 (8.85)	21.26 (22.90)	9.33 (5.96)
Other child violations rate/h	64.37 (43.75)	102.62 (61.29)	57.27 (24.02)	45.12 (22.08)	66.65 (37.92)	55.95 (50.46)	35.81 (9.62)
Teacher behavior							
% appropriate response to target student	24% (17.2%)	31% (16.5%)	26% (22.3%)	29% (16.9%)	17% (14.5%)	11% (10.3%)	25% (17.1%)
% appropriate response to other children	37% (18.9%)	40% (18.5%)	32% (12.4%)	47% (21.7%)	31% (18.3%)	27% (15.6%)	42% (20.5%)
Praise (rate/h)							
Target student–labeled	1.35 (1.78)	2.36 (2.42)	1.09 (1.44)	1.54 (2.01)	1.14 (1.49)	0.28 (0.33)	0.81 (0.85)
Target student–unlabeled	1.23 (1.51)	1.79 (1.94)	0.80 (0.69)	1.21 (1.73)	1.13 (1.38)	1.05 (1.62)	1.12 91.46)
Total labeled praise[a]	12.12 (11.31)	22.64 (14.00)	9.64 (5.91)	11.78 (9.61)	8.93 (8.21)	3.49 (2.56)	13.27 (17.43)
Total unlabeled praise[b]	11.45 (6.97)	15.06 (6.78)	12.11 (4.02)	13.15 (8.04)	8.46 (6.69)	7.41 (7.46)	11.63 (5.10)
Total praise (labeled and unlabeled)	23.57 (15.72)	37.70 (18.54)	21.75 (9.06)	24.93 (15.27)	17.39 (11.78)	10.90 (9.00)	24.90 (16.51)
Total labeled praise/total praise	0.51	0.60	0.44	0.47	0.51	0.32	0.53
Commands (rate/h)							
Target student–effective	3.17 (2.25)	2.91 (1.63)	4.38 (3.23)	3.23 (2.18)	2.55 (1.98)	2.29 (1.09)	4.72 (3.27)
Target student–ineffective	0.55 (0.82)	.075 (1.25)	0.41 (0.31)	0.50 (0.56)	0.49 (0.66)	0.55 (1.13)	0.59 (1.18)
Total effective commands[c]	46.53 (20.56)	45.86 (18.13)	54.75 (21.93)	41.77 (21.96)	46.51 (21.47)	35.18 (11.09)	63.26 (22.82)
Total ineffective commands[d]	8.90 (8.05)	9.96 (6.57)	8.67 (4.54)	11.13 (14.62)	7.34 (2.83)	6.94 (6.49)	8.18 (6.55)
Total commands (effective and ineffective)	55.43 (22.09)	55.82 (18.35)	63.42 (23.74)	52.90 (26.40)	53.86 (22.02)	42.13 (8.59)	71.44 (25.74)
Total effective commands/total commands	0.84	0.82	0.86	0.79	0.86	0.84	0.89

Note. Class sizes ranged from 19 to 25 students per class.
[a]Sum of Target student–labeled, Other individual–labeled and Group–labeled. [b]Sum of Target student–unlabeled, Other individual–unlabeled and Group–unlabeled. [c]Sum of Target student–effective, Other individual–effective, and Group–effective. [d]Sum of target student–ineffective, Other individual–ineffective, and Group–ineffective.

Table 2. Correlations Between Student and Teacher Behaviors for the Total Sample.

Study variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Total violations–target student	—											
2. Total violations–other children	.07	—										
3. Appropriate response to target student violations	−.11	−.22	—									
4. Appropriate response to other children violation	−.20	−.40**	.53**	—								
5. Labeled praise–target student	.17	.19	.25	.05	—							
6. Unlabeled praise–target student	.13	.13	.01	−.18	.48**	—						
7. Total labeled praise	.10	.22	.29*	.15	.68**	.19	—					
8. Total unlabeled praise	−.04	.21	.13	−.09	.35**	.45**	.45**	—				
9. Effective commands–target student	.04	−.15	.31*	.08	.26†	.25†	.08	.08	—			
10. Ineffective commands–target student	.07	.03	.14	−.07	.14	.53**	.17	.28*	.30*	—		
11. Total effective commands	−.08	−.06	.22	.05	.08	−.07	.24	.03	.64**	−.04	—	
12. Total ineffective commands	.03	.12	.13	−.05	.18	.07	.32*	.39**	−.04	.34*	.00	—

Note. N = 55. With the exception of the two appropriate response-to-violations variables (which represent a percentage), all other variables represent rates per hour.
†p < .10. *p < .05. **p < .01.

teacher behavior (i.e., rate and type of praise, rate and type of commands, and percent response to violations) for the total sample and by grade level. (Although there were some differences across sites, the variability within schools at a given site was 2 to 10 times greater than the differences between schools when aggregated by site. Thus, data from the total sample, combined across sites, are presented and analyzed.) The data presented in Table 1 demonstrate the wide variability in both teacher and student behaviors across grades. For example, an average of approximately 102 classwide violations were observed per hour in kindergarten classrooms, and approximately 35 classwide violations were observed per hour at the Grade 5 level. With an average class size of 19 to 25, this translates to four to five violations per student per hour in a kindergarten classroom and approximately one to two violations per student per hour in a Grade 5 classroom. With the exception of four comparisons (i.e., those involving Grades 1 and 2 compared to Grades 3 and 4), the grade-by-grade comparisons indicated that classrooms with younger children had higher rates of classwide rule violations than classrooms with older children (Cohen's ds range from 0.37 to 1.56, M = 0.93). In comparison, target students demonstrated 9 to 17 violations per hour depending on the grade. The average percentage of violations to which teachers provided an appropriate response was rather low, never reaching 50% for target student or classwide violations at any grade.

The average rate of praise per hour for the total sample was 23.57 (SD = 15.72). With the exception of fifth grade, there was a general decline in rates of total praise across grades, with an average of 37 praise statements (labeled

and unlabeled) per hour in kindergarten classrooms, and approximately 10 praise statements per hour in the fourth grade. With the exception of four comparisons, the grade-by-grade comparisons indicated that teachers of lower grades had higher rates of praise than teachers of higher grades (ds range from 0.17 to 1.66, M = 0.88).

Interestingly, percentage of labeled praise relative to total praise showed no consistent trend across grades, ranging from 30% in fourth grade, to 40% in first and second grades, 50% in third and fifth grade, and 60% in kindergarten. Rates of commands across kindergarten through third grade, ranged from approximately 52 to 63 per hour; teachers in the fourth grade had the lowest rate per hour (42), and teachers in the fifth grade had the highest rate per hour (71). Notably, the percentage of commands that were effective was high across all grades (ranging from 79% to 89%).

Aim 2: Teacher Behaviors Predicting Student Violations

In Table 2, we provide correlations between all variables used in the regression models. Using linear regression models, we examined the extent to which teacher-appropriate response to violations, as well as teacher use of labeled and unlabeled praise, and effective and ineffective commands were associated with target student and classwide violations (see first column in Table 3). For classwide variables, the teacher behaviors explained 25% of the variation in total classroom violations. Consistent with expectations, standardized betas indicate that higher percentages of appropriate response to violations were associated with lower

Table 3. Regression Models of the Association Between Teacher Behaviors and Student Violations for Total Sample and Primary (K–3) Sample.

Model	Total sample[a]		Primary (K–3) sample[b]	
	Beta[c]	R^2	Beta[c]	R^2
Classwide model				
Rate effective commands—other children	−.11		−.05	
Rate ineffective commands—other children	−.01		−.01	
Rate labeled praise—other children	.29[†]		.22	
Rate unlabeled praise—other children	.05		.13	
% appropriate response to classwide RVs	−.43**		−.46**	
Model R^2		.25*		.30*
Target student model				
Rate effective commands–target student	.02		−.06	
Rate ineffective commands–target student	.06		.16	
Rate labeled praise–target student	.21		.30[†]	
Rate unlabeled praise–target student	.00		.08	
% appropriate response to target studentwide RVs	−.18		−.27[†]	
Model R^2		.06		.19

[a]$N = 55$. [b]$N = 45$. [c]Standardized betas.
[†]$p < .10$. *$p < .05$. **$p < .01$.

classroom violations. In contrast to expectations, neither effective commands nor labeled praise was significantly associated with classroom violations. For the target student variables, the teacher behaviors accounted for 6% of the variance in target student violations. The overall model was not significant, but the pattern of the betas for praise (labeled and unlabeled) and percentage of appropriate responses to violations mirrored the pattern found in the classwide model. (All two-way interactions between the predictor variables were tested, but none were significant.)

Because the rates of behaviors were variable across grades (see Table 1), we examined correlations between variables at each grade. A pattern was observed such that the direction of the relations between some variables (e.g., target student violations and other child violations, target student violations and labeled praise, and classwide violations and labeled praise) was positive for kindergarten and Grades 1, 2, and 3, but negative for Grades 4 and 5. Thus, in Table 4, we provide the correlations between all variables separately for primary elementary (kindergarten to Grade 3) and intermediate grades (Grades 4 and 5). In addition, we reexamined the linear regression models for the primary grades separately (see Table 3, right column). The pattern for the classwide variables was similar to that found for the total sample; the model accounted for 30% of the variance. In the model with the target student variables, the teacher behaviors accounted for 19% of the variance in target student violations. The pattern of the betas mirrored the pattern found in the classwide model, with one exception; for instruction, the direction of the relationship with violations was positive for ineffective commands and negative for

effective commands. Because the fourth and fifth grade subsample was small ($n = 10$), regression analyses could not be conducted on this subsample.

Aim 3: Possible Benchmark for Appropriate Response to Challenging Behavior

Given that response to violations was consistently associated with lower classwide violations for the total sample and for the primary elementary grades, we sought to determine a threshold where appropriate responses were most associated with low rates of classwide violations in primary elementary classrooms. We did this to establish a possible benchmark for percentage of appropriate responses for teachers to aim. From a behavioral perspective, the more likely a behavior will lead to a salient consequence, the less likely it is that the person will exhibit the behavior. Thus, we examined rates of classroom violations by decile of appropriate response to violations (see Table 5). Results in Table 5 show that rule violations drop from approximately 70 or 80 per hour when the percentage of appropriate response to violations is lower than 20%, to approximately 35 per hour once teachers reach a threshold of 51% appropriate response. The variance across classrooms also declines accordingly. Furthermore, rates of rule violations are not lower with thresholds greater than 51%.

Discussion

This study contributes to the literature by providing (a) an updated profile of student and teacher behavior across the

Table 4. Correlations Between Student and Teacher Behaviors by Primary (K–3) and Intermediate (Grades 4 and 5) Elementary Grades.

Study variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Total violations—target student	—	−.31	.13	−.11	−.16	−.19	−.11	−.39	.12	−.11	.12	−.35
2. Total violations—other children	.25	—	−.32	−.26	−.13	−.12	−.08	−.22	−.28	−.07	−.27	.06
3. Appropriate response to target student violations	−.20	−.26†	—	.33	.80*	.03	.76*	.30	.37	.04	.66*	.13
4. Appropriate response to other children violation	−.25	−.45**	.55**	—	.38	−.50	.46	−.43	−.02	−.42	.08	−.42
5. Labeled praise—target student	.30*	.18	.19	.00	—	.26	.89**	.33	.44	.28	.45	.18
6. Unlabeled praise—target student	.29†	.16	−.01	−.13	.52**	—	−.05	.84**	.50	.97**	−.05	.81**
7. Total labeled praise	.23	.24	.17	.07	.68**	.23	—	.09	.14	−.08	.40	−.18
8. Total unlabeled praise	.13	.26†	.07	−.04	.34*	.38**	.50**	—	.41	.74*	.233	.82**
9. Effective commands—target student	−.01	−.12	.31*	.10	.27†	.20	.08	.01	—	.56	.64*	.44
10. Ineffective commands—target student	.18	.06	.18	.03	.15	.42**	.26†	.16	.22	—	−.05	.81**
11. Total effective commands	−.17	−.03	.15	.04	.06	−.07	.21	−.01	.64**	−.04	—	.00
12. Total ineffective commands	.17	.12	.11	−.01	.17	−.04	.39**	.32*	−.11	.26	.00	—

Note. Correlations below the diagonal represent kindergarten to Grade 3; correlations above the diagonal represent Grades 4 and 5. With the exception of the two appropriate response-to-violations variables (which represent a percentage), all other variables represent rates per hour. $N = 45$ for primary grades; $N = 10$ for intermediate grades.
†$p < .10$. *$p < .05$. **$p < .01$.

Table 5. Rates of Classroom Violations per Hour as a Function of Teacher Appropriate Response to Violations.

Levels of appropriate response	Primary sample violations (M, SD)	Total sample violations (M, SD)
Less than 20% appropriate response	80.83 (38.84)[a]	70.96 (36.70)[g]
Less than 30% appropriate response	89.82 (38.30)[b]	83.51 (42.90)[h]
At least 40% appropriate response	53.82 (48.22)[c]	50.77 (44.32)[i]
At least 51% appropriate response	33.89 (13.29)[d]	33.86 (12.98)[j]
At least 60% appropriate response	39.13 (14.48)[e]	39.13 (14.48)[k]
At least 70% appropriate response	37.14 (17.96)[f]	37.14 (17.96)[l]

[a]$N = 8$. [b]$N = 19$. [c]$N = 19$. [d]$N = 11$. [e]$N = 6$. [f]$N = 3$. [g]$N = 11$. [h]$N = 24$. [i]$N = 23$. [j]$N = 13$. [k]$N = 6$. [l]$N = 3$.

elementary grades by which future observations may be compared, (b) rates of rule violations at the class level and as demonstrated by target students with or at risk of ADHD, (c) information about the classroom management strategies most associated with student rule violations and, (d) a possible minimum benchmark for percentage of appropriate response to violations. Our findings have implications for future research, teacher professional development in classroom management, and teacher evaluation.

As anticipated, results showed that rates of classwide rule violations and teacher praise varied by grade level. The finding that use of praise declines with increasing grade is consistent with previous research (Reddy et al., 2013; White, 1975) and with developmental socialization patterns. It may be that students in kindergarten and first grade are still learning social and behavioral control, and thus are violating many classroom expectations in the process. To counteract these violations, teachers of younger students may be using praise as a proactive tool to guide and shape socialization at a greater rate than in classrooms of older children. However, at older ages, typically developing students have greater behavioral control and less impulsivity, and following basic expectation has become routine; thus some types of violations (e.g., out of seat, interrupting) are less frequent. Furthermore, teachers of intermediate grades may use praise more sparingly, perhaps to enhance its potency and indicate that the standard of behavior necessary to receive reinforcement is higher than in grades of younger children. Regardless of the mechanisms producing these patterns, these data suggest that previous studies in which researchers have collapsed data across grade level have failed to account for meaningful differences in development and classroom structure and, thus, likely misrepresented actual rates of student and

teacher behaviors within each grade. The rates within the current study provide evidence that developmental considerations for child and teacher behavior should be highlighted rather than minimized.

Our examination of both classwide and target student violations also highlights an interesting contrast. For example, among target students with or at risk of ADHD, violations did not decline with increasing grade and their rate of violations (9 to 17 per hour) exceeded that of the average student (1 to 2 per hour) in the classroom by a large margin. This finding highlights the challenges these students pose for teachers, as well as the need for more individualized intervention within a multitiered system of support. More specifically, each target student was responsible for 15% to 28% of total violations in the classroom. Given that this is the unique contribution of only one child in the classroom, it is not surprising that many teachers report feeling stressed (Greene et al., 2002) and unprepared for managing students with or at risk of ADHD (Coalition for Psychology in Schools and Education, 2006). Because there are likely one to two students with ADHD or similar symptoms in every elementary classroom, these data underscore the need to adequately prepare teachers to implement evidence-based Tier 2 interventions, such as the daily report card (Evans et al., 2014; Owens et al., 2012; Vannest et al., 2010). Given that behavioral infractions detract from lesson focus and classroom instruction time, reducing violations can have a substantial impact on learning time for all students.

Rates of effective commands were fairly stable across grades. The stability of these rates is not surprising, given that commands are a natural part of leading a classroom in academic activities and that our SBTR definition of commands included those focused on academic behaviors (e.g., "open your book to page 12") and behavioral control (e.g., "sit in your seat"). To our knowledge, this is the first study to document rates of effective and ineffective commands in general classrooms; thus, replication is necessary, and differential coding of academic and behaviorally focused commands in future studies may offer data to support this interpretation.

Interestingly, despite varying rates of violations across grades, average rates of appropriate response to violations were low (never reaching 50%) and fairly stable across grades. However, there was striking variability across teachers within a grade (e.g., standard deviations of 15% to 20%). First, the results of the linear regression models showed that commonly recommended classroom management strategies (praise, commands, appropriate response to violations) in combination account for a substantial portion of the variance in classwide violations (see Table 3). Second, among these strategies, the percentage of appropriate teacher responses was most robustly associated with violations; a greater percentage of appropriate responses was associated with lower rates of violations. This pattern emerged in the classwide and target student models for the total sample and for the primary elementary sample. This consistency underscores the importance of focusing on this teacher behavior in future research, classroom management–related professional development, and teacher evaluation systems.

Unexpectedly, we found a positive relation between labeled praise and violations in both the classwide and target student models. One hypothesis is that teachers witnessing high rates of violations were attempting to use praise as a proactive strategy to address the inattentive and/or disruptive behavior. Indeed, there is evidence that teachers prefer positive strategies to perceived punitive strategies (Elliot, Witt, Galvin, & Peterson, 1984; Girio & Owens, 2009). Perhaps teachers were using praise to try to prevent misbehavior. The positive relation between praise and violations makes it difficult to compare our findings to the assertion that labeled praise is more effective than unlabeled praise for reinforcing the desired behavior. The regression models also suggest that, when considered with other teacher management practices, praise is less associated with students' violations than teachers' appropriate responses to a majority of those violations. Although praise may have relatively little impact on reducing violations, it may increase the rates of positive student outcomes (e.g., time on task; behavioral compliance) and enhance student–teacher relations. Future research is needed to evaluate these hypotheses.

Finally, our data help to identify a possible benchmark for appropriate response to violations. Teacher's appropriate response to violations was the teacher behavior most associated with student violations; however, the association did not appear to be linear. Data in Table 5 document that an appropriate teacher response to at least 51% of student violations was associated with the lowest rates of violations (about 1 per student per hour) and that higher rates of appropriate response were not associated with incremental benefit. In addition to the declines in the average rates of violations, we also observed declines in the average standard deviations (see Table 5). This suggests that not only was this associated with lower average student violations, it was associated with less extremes as well. This finding may challenge our assumptions that higher (than 51%) consistency is critical to achieving desired outcomes. On one hand, as long as teachers are appropriately responding to violations "more often than not," they are creating predictability in the classroom and following through on their expectations. Such a threshold is likely more feasible for teachers than one of perfection or near perfection (e.g., 80 to 100%). On the other hand, it is important to recognize that this may represent a minimum threshold; thus, training toward a higher threshold may be an important goal to allow for inconsistency in teachers' implementation of classroom management strategies and "backsliding" during stressful times. Furthermore, there may have been other teacher behaviors or contextual factors that were not measured that

may be contributing to this finding. Thus, this finding offers new insights for professional development and teacher evaluation systems, and generates new hypotheses for research. However, the finding warrants replication, as well as research that includes manipulation of this variable with a larger sample to determine causality.

Limitations

Given the important outcomes provided with this sample, these research questions should be examined with larger samples within each grade. Indeed, our smaller subsamples prevented the examination of Aims 2 and 3 by grade. Second, this study occurred with a sample of teachers willing to participate in a study that involved consultation. Thus, this subset of teachers may not represent the larger population. Third, there may be facets to the operational definitions used in the modified SBTR that limit some of our conclusions. For example, both academically and behaviorally focused commands and praise were included in our definitions of effective and ineffective commands and praise codes. Similarly, we coded a negative child behavior (i.e., violations), and did not include a positive child behavior (on-task). By further delineating academically focused and behaviorally focused actions and by including positive student behaviors, new patterns and hypotheses will likely emerge. Fourth, Aims 1 and 3 are descriptive; thus, these patterns and the conclusions drawn about them should be considered preliminary until replicated with a larger sample using an experimental design. Finally, these relations were examined in the context of baseline observations and do not represent how teachers' improvement in classroom management strategies may produce change in student behaviors.

Conclusions and Implications

Our findings offer several hypotheses for future study and have possible implications for teacher professional development and evaluation systems. First, additional study of teacher and student behaviors in typical classrooms is needed and such study should adopt a developmental lens, disaggregating data at either the grade level or in developmentally conceptual segments, given the differing needs and contextual influences on primary versus intermediate elementary school students. Second, should these findings be replicated, those involved in teacher training, professional development, consultation, and evaluation should prioritize developing the skill of appropriate response to violations by covering it first, in more depth, and/or more frequently. Recent studies (e.g., Coles, Owens, Serrano, Slavec, & Evans, 2015) suggest that, for some teachers, improving response to violations will require multicomponent consultation and training. Although there is research to support the use of labeled praise and effective commands especially for students with behavioral difficulties (Pfiffner & O'Leary, 1987), our results suggest these skills may be less closely associated with misbehavior than in responding appropriately to violations. Praise may be more important for establishing positive student–teacher relationships and/or on-task behavior, and this goal can be weighed alongside reduction in challenging behaviors. Future studies should consider examining the effects of praise and positive teacher–student interaction on the student–teacher relationship, on-task behavior, and student prosocial skills.

Finally, this is the first study to offer a possible empirically derived minimum benchmark for teachers' appropriate response to violations. This provides preliminary evidence that "more appropriate responding may not be empirically better" beyond a certain threshold. We hope that this finding stimulates additional work on empirically based benchmarking. For example, examination of student rule violations following experimental manipulation of different types of teacher responses would allow causal inferences to be drawn. Similarly replication with larger samples would provide greater confidence when recommending benchmarks. Such data would offer important guidelines for training preservice and in-service teachers and for creating consistency in policies related to teacher evaluation systems.

References

Abikoff, H., & Gittelman, R. (1985). Classroom observation code: A modification of the Stony Brook code. *Psychopharmacology Bulletin*, *21*, 909–981.

Acker, M. M., & O'Leary, S. G. (1987). Effects of reprimands and praise on appropriate behavior in the classroom. *Journal of Abnormal Child Psychology*, *15*, 549–557.

Anderson, L. M., Evertson, C. M., & Brophy, J. E. (1979). An experimental study of effective teaching in first-grade reading groups. *The Elementary School Journal*, *79*, 193–223.

Atkins, M. S., Pelham, W. E., & Licht, M. H. (1985). A comparison of objective classroom measures and teacher ratings of attention deficit disorder. *Journal of Abnormal Child Psychology*, *13*, 155–167.

Barkley, R. A. (2013). *Taking charge of ADHD: The complete authoritative guide for parents* (3rd ed.). New York, NY: Guilford Press.

Bibou-nakou, I., Stogiannidou, A., & Kiosseoglou, G. (1999). The relation between teacher burnout and teachers' attributions and practices regarding school behavior problems. *School Psychology International*, *20*, 209–217.

Birch, S. H., & Ladd, G. W. (1998). Children's interpersonal behaviors and the teacher-child relationship. *Developmental Psychology*, *34*, 934–946.

Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research*, *51*, 5–32.

Burnett, P. C., & Mandel, V. (2010). Praise and feedback in the primary classroom: Teachers' and students' perspectives. *Australian Journal of Educational & Developmental Psychology*, *10*, 145–154.

Chalk, K., & Bizo, L. A. (2004). Specific praise improves on-task behaviour and numeracy enjoyment: A study of year four pupils engaged in the numeracy hour. *Educational Psychology in Practice*, *20*, 335–351.

Coalition for Psychology in Schools and Education. (2006, August). *Report on the teacher needs survey*. Washington, DC: Center for Psychology in Schools and Education, American Psychological Association.

Coles, E. K., Owens, J. S., Serrano, V., Slavec, J., & Evans, S. W. (2015). From consultation to student outcomes: The role of teacher knowledge, skills, and beliefs in increasing integrity in classroom behavior management. *School Mental Health*, *7*, 34–48.

Colvin, G., Sugai, G., Good, I. I. I. R. H., & Lee, Y. Y. (1997). Using active supervision and precorrection to improve transition behaviors in an elementary school. *School Psychology Quarterly*, *12*, 344–363.

Elliot, S. N., Witt, J. C., Galvin, R., & Peterson, M. (1984). Acceptability of positive and reductive interventions: Factors that influence teachers' decisions. *Journal of School Psychology*, *22*, 353–360.

Evans, S. W., Owens, J. S., & Bunford, N. (2014). Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. *Journal of Clinical Child & Adolescent Psychology*, *43*, 527–551.

Every Student Succeeds Act (ESSA) of 2015, Pub. L. No. 114-95, S. 1177.

Fabiano, G. A., Pelham, W. E., Majumdar, A., Evans, S. W., Manos, M. J., Caserta, D., . . . Carter, R. L. (2013). Elementary and middle school teacher perceptions of attention-deficit/hyperactivity disorder incidence. *Child & Youth Care Forum*, *42*, 87–99.

Fabiano, G. A., Vujnovic, R. K., Pelham, W. E., Waschbusch, D. A., Massetti, G. M., Yu, J., . . . Volker, M. (2010). Enhancing the effectiveness of special education programming for children with ADHD using a daily report card. *School Psychology Review*, *39*, 219–239.

Fabiano, G. A., Vujnovic, R. K., Waschbusch, D. A., Yu, J., Mashtare, T., Pariseau, M. E., . . . Smalls, K. J. (2013). A comparison of workshop training versus intensive, experiential training for improving behavior support skills in early educators. *Early Childhood Research Quarterly*, *28*, 450–460.

Floress, M. F., & Jenkins, L. N. (2015). A preliminary investigation of kindergarten teacher's use of praise in general education. *Preventing School Failure*, *594*, 253–262.

Ford, A. D. (1998). *An evaluation of the contributions of effective instruction delivery to a compliance training package for children* (Doctoral dissertation). Retrieved from http://aquila.usm.edu/theses_dissertations/2012/

Forehand, R. (1977). Child noncompliance to parental requests: Behavioral analysis and treatment. *Progress in Behavior Modification*, *5*, 111–147.

Forehand, R. L., & McMahon, R. J. (1981). *Helping the noncompliant child: A clinician's guide to parent training*. New York, NY: Guilford press.

Gable, R. A., Hester, P. H., Rock, M. L., & Hughes, K. G. (2009). Back to basics: Rules, praise, ignoring and reprimands revisited. *Intervention in School and Clinic*, *44*, 195–205.

Girio, E. L., & Owens, J. S. (2009). Teacher acceptability of evidence-based and promising treatments for children with attention-deficit/hyperactivity disorder. *School Mental Health*, *1*, 16–25. doi:10.1007/s12310-008-9001-6

Greene, R., Beszterczey, S. K., Katzenstein, T., Park, K., & Goring, J. (2002). Are students with ADHD more stressful to teach? Patterns of teacher stress in an elementary school sample. *Journal of Emotional and Behavioral Disorders*, *10*, 79–89.

Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher*, *41*, 56–64.

Jenkins, L. N., Floress, M. T., & Reinke, W. (2015). Rates and types of teacher praise: A review and future directions. *Psychology in the Schools*, *52*, 463–476.

Kalis, T. M., Vannest, K. J., & Parker, R. (2007). Praise counts: Using self-monitoring to increase effective teaching practices. *Preventing School Failure*, *51*, 20–27.

Kamphaus, R. W., Huberty, C. J., DiStefano, C., & Petoskey, M. D. (1997). A typology of teacher-rated child behavior for a national US sample. *Journal of Abnormal Child Psychology*, *25*, 453–463.

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Identifying effective classroom practices using student achievement data. *Journal of Human Resources*, *46*, 587–613.

Lane, K. L., Gresham, F. M., & O'Shaughnessy, T. E. (2002). *Interventions for children with or at risk for emotional and behavioral disorders*. Boston, MA: Allyn & Bacon.

Madsen, C. H., Jr., Becker, W. C., & Thomas, D. R. (1968). Rules, praise, and ignoring: Elements of elementary classroom control. *Journal of Applied Behavior Analysis*, *1*, 139–150.

Marzano, R. J., Marzano, J. S., & Pickering, D. (2003). *Classroom management that works: Research-based strategies for every teacher*. Alexandria, VA: Association for Supervision and Curriculum Development.

Massetti, G. M., Pelham, W. E., & Waschbusch, D. A. (2007, June). *Teacher fidelity in use of behavior management strategies: Relationships among observations, self-report, and children's disruptive behavior*. Poster presented at the annual IES conference, *Washington, DC*.

Matheson, A. S., & Shriver, M. D. (2005). Training teachers to give effective commands: Effects on student compliance and academic behaviors. *School Psychology Review*, *34*, 202–219.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46.

McMahon, R. J., & Forehand, R. (2003). *Helping the noncompliant child* (2nd ed.). New York, NY: Guilford Press.

Neef, N. A., Shafer, M. S., Egel, A. L., Cataldo, M. F., & Parrish, J. M. (1983). The class specific effects of compliance training with "do" and "don't" requests: Analogue analysis and classroom application. *Journal of Applied Behavior Analysis*, *16*, 81–99.

Ohio Department of Education. (2015). *Ohio teacher evaluation system resource guide*. Retrieved from http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System

Owens, J. S., Holdaway, A. S., Zoromski, A. K., Evans, S. W., Himawan, L. K., Girio-Herrera, E., & Murphy, C. E. (2012). Incremental benefits of a daily report card intervention over time for youth with disruptive behavior. *Behavior Therapy*, *434*, 848–861.

Owens, J. S., Murphy, C. E., Richerson, L., Girio, E. L., & Himawan, L. K. (2008). Science to practice in underserved communities: The effectiveness of school mental health programming. *Journal of Clinical Child and Adolescent Psychology*, *37*, 434–447. doi:10.1080/15374410801955912

Pelham, W. E., Greiner, A. R., & Gnagy, E. M. (2008). *Student behavior teacher response observation code manual*. Unpublished manual.

Pfiffner, L. J., & O'Leary, S. G. (1987). The efficacy of all-positive management as a function of the prior use of negative consequences. *Journal of Applied Behavior Analysis*, *20*, 265–271.

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*, 683–700.

Reinke, W., Herman, C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS: Identifying areas for enhancement. *Journal of Positive Behavior Interventions*, *5*, 39–50.

Rhode, G., Jensen, W. R., & Reavis, H. K. (1992). *The tough kid book: Practical classroom management systems*. Longmont, CO: Sopris West.

Robb, J. A., Sibley, M. H., Pelham, Jr., W. E., Foster, E. M., Molina, B. S., Gnagy, E. M., . . . Kuriyan, A. B. (2011). The estimated annual cost of ADHD to the US education system. *School Mental Health*, *3*, 169–177.

Roberts, M. W., McMahon, R. J., Forehand, R., & Humphreys, L. (1978). The effect of parental instruction-giving on child compliance. *Behavior Therapy*, *21*, 257–271.

Saudargas, R. A. (1997). *State-Event Classroom Observation System (SECOS)* (Observation manual). Knoxville: University of Tennessee.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assesing rater reliability. *Psychological Bulletin*, *86*, 420–428.

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

Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders*, *8*(1), 2–8.

U.S. Department of Education, National Center for Education Statistics. (2015). *The digest of education statistics, 2013* (NCES 2015-011, Table 204.60). Retrieved from https://nces.ed.gov/fastfacts/display.asp?id=59

Vannest, K. J., Davis, J. L., Davis, C. R., Mason, B. A., & Burke, M. D. (2010). Effective intervention for behavior with a daily behavior report card: A meta-analysis. *School Psychology Review*, *39*, 654–672.

Volpe, R. J., McConaughy, S. H., & Hintze, J. M. (2009). Generalizability of classroom behavior problem and on-task scores from the Direct Observation Form. *School Psychology Review*, *38*, 382–401.

Vujnovic, R. K., Fabiano, G. A., Waschbusch, D. A., Pelham, W. E., Greiner, A., Gera, S., . . . Buck, M. (2014). Preliminary psychometric properties of an observation system to assess teachers' use of effective behavior support strategies in preschool classrooms. *Education and Treatment of Children*, *37*, 323–346.

Wechsler, D. (2011). *Wechsler abbreviated scale of intelligence* (2nd ed). Bloomington, MN: Pearson.

White, M. A. (1975). Natural rates of teacher approval and disapproval in the classroom. *Journal of Applied Behavior Analysis*, *8*, 367–372.

Williams, C. A., & Forehand, R. (1984). An examination of predictor variables for child compliance and noncompliance. *Journal of Abnormal Child Psychology*, *12*, 491–503.

Witt, J. C., & Elliot, S. N. (1982). The response cost lottery: A time efficient and effective classroom intervention. *Journal of School Psychology*, *20*, 155–161.

Wong, H. K., & Wong, R. T. (2001). *The effective teacher*. Mountain View, CA: Harry K. Wong Publishing.