

Effects of Teachers' Praise-to-Reprimand Ratios on Elementary Students' On-task Behaviour

Paul Caldarella, PhD, paul_caldarella@byu.edu

Ross A. A. Larsen, PhD, ross.larsen@byu.edu

Leslie Williams, EdS, leslie_williams@byu.edu

Kade R. Downs, EdS, kadedowns@gmail.com

David O. McKay School of Education

Brigham Young University

Provo, UT 84602

Howard P. Wills, Ph.D., hpwills@ku.edu

Juniper Gardens Children's Project

University of Kansas

444 Minnesota Ave, Suite 300

Kansas City, KS 66101

Joseph H. Wehby, Ph.D., joseph.wehby@vanderbilt.edu

Department of Special Education

Peabody College

Vanderbilt University

Nashville, TN 37203

Author Notes

Address all correspondence to Paul Caldarella, PhD, Center for the Improvement of Teacher Education and Schooling, Brigham Young University, 149 MCKB, Provo, UT 84602.

Telephone: 801-422-5081. Fax: 801-422-0199. Email: paul_caldarella@byu.edu

Acknowledgement

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.

The citation for this paper in its' final published version is as follows:

Caldarella, P., Larsen, R. A. A., Williams, L., *Downs, K. R.*, Wills, H. P., & Wehby, J. H. (2020). Effects of teachers' praise-to-reprimand ratios on elementary students' on-task behaviour. *Educational Psychology*. Advance online publication. <https://doi.org/10.1080/01443410.2020.1711872>

Abstract

Though many educators are familiar with a suggested 3:1 or 4:1 praise-to-reprimand ratio (PRR), relatively little research has been conducted on this subject. Three years of data collected across three states in the United States, from 19 elementary schools and 151 classrooms, during a multi-site efficacy trial were used to analyse the effect of teachers' PRRs on their students' on-task behaviour. Though no PRR threshold (e.g., 3:1, 4:1) was found where behaviour dramatically improved, a positive linear relationship was evident, showing that the higher the teachers' PRR, the higher the students' on-task behaviour percentage. Limitations and implications are discussed.

Keywords: praise, reprimands, classroom management, elementary schools

Effects of Teachers' Praise-to-Reprimand Ratios on Elementary Students' On-task Behaviour

Teachers' patterns of praise and reprimands have been studied in schools to help develop a clearer understanding of how teacher and student behaviours interact. Student participants in these studies have varied in age (Haydon & Musti-Rao, 2011; Merrett & Wheldall, 1992; M. A. White, 1975), culture (Naylor, Kamps, & Wills, 2018; Stormont, Smith, & Lewis, 2007), and disability status (Caldarella, Williams, Hansen, & Wills, 2015; Ferrari, 1980).

There is no universal definition for praise, which may make it difficult to compare outcomes of praise research across studies. However, despite some differences, common roots of praise found in seminal works (Brophy, 1981) have influenced many current definitions and appear to yield positive outcomes (Floress, Beschta, Meyer, & Reinke, 2017). During data collection for this project, praise was defined as a verbal indication of approval following student behaviour more than acknowledging a correct response (see Method section), similar to definitions used in past studies (Brophy, 1981; Cameron & Pierce, 1994; Chalk & Bizo, 2004; Maag, 2001). Praise has been acknowledged as the simplest classroom management strategy to implement (Gable, Hester, Rock, & Hughes, 2009). It has also been noted as a tool to reinforce student behaviour (Conroy, Sutherland, Snyder, Al-Hendawi, & Vo, 2009; Weeden, Wills, Kottwitz, & Kamps, 2016) and an opportunity to recognise student engagement (Embry & Biglan, 2008). Its use has been correlated with positive academic and social outcomes (Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008) as well as teacher self-efficacy (Reinke, Herman, & Stormont, 2013). Behaviour-specific praise, in which a precise behaviour and how it met a teacher expectation is stated, is considered a potentially evidence-based practice, according to the guidelines of the Council for Exceptional Children (Royer, Lane, Dunlap, & Ennis, 2019).

However, relatively low rates of teacher praise (general praise or behaviour-specific praise) have been observed in general education settings during natural whole-class observations (Floress, Jenkins, Reinke, & McKown, 2018) as well as during observations of specific students displaying disruptive behaviours (Reinke, Lewis-Palmer, & Martin, 2007). Research suggests that praise is greatly under-utilised in schools, and its use tends to decrease as students age (Floress et al., 2018; Reddy, Fabiano, Dudek, & Hsu, 2013; M. A. White, 1975).

Similar to praise, there is no universal definition for reprimands. Reprimands have been referred to as *negative communication* (Rathel, Drasgow, & Christle, 2008) or *contingent punishment* (Merrell, Ervin, & Gimpel Peacock, 2012). During data collection for this project, a reprimand was defined as verbal disapproval (including a threat or scolding) in response to inappropriate behaviour or instruction that the behaviour must stop (see Method section) similar to definitions used in some past studies (Allday et al., 2012; Downs et al., 2019). A relatively small amount of research has been devoted to teacher reprimands, with most of this research showing negative associations with student behaviour. Teacher reprimands have been found to predict noncompliance and negative student behaviour (Van Acker, Grant, & Henry, 1996). Additionally, reprimands have been correlated positively with students' problem behaviour (Downs et al., 2019; Kodak, Northup, & Kelley, 2007) and teachers' emotional exhaustion (Reinke et al., 2013). Reprimand rates in elementary classrooms have been observed to be higher than praise rates (Reinke et al., 2013; Van Acker et al., 1996). While praise tends to decrease as grade level increases, there appears to be an increase in reprimands (Reddy et al., 2013; M. A. White, 1975).

Given the potential for negative outcomes associated with teacher reprimands and potential for positive outcomes associated with teacher praise, it is helpful to understand how

they interact and relate to student behaviour in school settings. There may be a need for occasional specific teacher reprimands, as an all-praise approach may not be realistic or most effective in managing student behaviour (Pfiffner, Rosén, & O'Leary, 1985; Rosén, O'Leary, Joyce, Conway, & Pfiffner, 1984). Pfiffner and colleagues (1985) conducted a study where they examined on-task behaviour changes for eight elementary students with behaviour challenges during three conditions: (a) *regular positives* (e.g., praise, posting work) *and negatives* (e.g., reprimands, withdrawal of privileges), (b) *regular positives* (with no negatives), and (c) *enhanced positives* (e.g., praise plus special recess activities, songtime, drawing; without any negatives). On-task behaviour was highest during the *regular positives and negatives* condition and lowest during the *regular positives* condition, indicating that positive consequences alone were not effective in maintaining the on-task behaviour in the classroom. The *enhanced positives* more closely approximated the on-task behaviour levels maintained in the condition with *regular positives and negatives*. The *regular positives and negatives* condition in this study did have higher praise with fewer reprimands (approximately 3.5 praises per 1 reprimand, though this varied per student) which may not be a 'regular condition' for some teachers (Reinke et al., 2007). As teacher reprimands may sometimes be needed, it is recommended that teachers use the evidence-based practice of teacher praise to mitigate potential negative effects of teacher reprimands (Shores, Gunter, & Jack, 1993).

Praise-to-Reprimand Ratios

Of increasing interest for promoting improved student behaviour is the study of teachers' praise-to-reprimand ratios (PRRs); this is a simple strategy or intervention where one considers the number of praise statements given compared to the number of reprimands given. There is no universal definition or method of calculation for PRRs (Sabey, Charlton, & Charlton, 2019).

Various ideal PRRs, calculated by frequency counts, have been discussed theoretically in the literature ranging from 3:1 (Shores et al., 1993), to 4:1 (Trussell, 2008), and 5:1 (Flora, 2000). However, a ratio calculated with frequency counts is difficult to use in statistical analyses requiring scaled variables. Accordingly, PRRs were defined in the current study as total teacher praise statements divided by the sum of total teacher praise statements plus total teacher reprimands (creating a scaled coefficient between 0 and 1; see Results section). As educators search for feasible evidence-based interventions to fit their students' needs, interventions that result in positive outcomes and require little or no cost, such as PRRs, are particularly encouraging. However, the empirical research supporting a particular PRR threshold (or an optimal ratio at which students' behaviour dramatically improves) is still developing (Sabey et al., 2019).

A PRR threshold would have potential to influence a number of other variables. For example, researchers have found PRRs to correlate negatively with students' disruptive behaviour and positively with their classroom enthusiasm and on-task behaviour (Leff et al., 2011). When PRRs increase, students' on-task behaviour has been shown to increase (R. L. White, 2010), indicating how changing these ratios could be a tool to potentially influence student behaviour. Increasing these ratios has also shown to improve class-wide behaviour management interventions (Caldarella, Williams, Jolstead, & Wills, 2017). Such patterns have motivated researchers to continue searching for an optimal PRR threshold.

Sutherland and Wehby (2001) implemented a self-evaluation intervention with a group of 20 Kindergarten to 8th grade teachers and 216 students with emotional and behavioural disorders (EBD) attending self-contained classrooms. Average PRRs increased for the treatment group between pretest (2.02:1) and treatment (6.03:1) phases and remained stable during maintenance

(6.01:1). In the treatment group, students generated a greater number of correct academic responses during the treatment phase, though rates did not maintain in the maintenance phase. The authors noted how reprimands seemed to decrease, although they were not a focus of the study, possibly due to improved academic instruction leading to less disruptive behaviour and consequently requiring fewer teacher reprimands.

Use of PRRs has increased in school-based intervention studies (Utley & Obiakor, 2015; Wills, Iwaszuk, Kamps, & Shumate, 2014), as well as in classroom climate research (Leff et al., 2011). For example, after utilising Class-Wide Function-related Intervention Teams (CW-FIT; Wills et al., 2010), a proactive classroom management intervention that addresses common functions of problem behaviour by teaching social expectations and praising and rewarding positive behaviour, increases in PRR were shown in preschool classrooms (2.64:1 to 9.95:1; Jolstead et al., 2017), general education kindergarten through second-grade classrooms (1.04:1 to 6.77:1; Caldarella et al., 2015), third- and fifth-grade art classrooms (0.91:1 to 2.65:1; Nelson et al., 2018), and a sixth-grade music classroom (1.65:1 to 4.50:1; Caldarella et al., 2017). Increased PRR was associated with increased on-task behaviours by students in all of these studies, but more research is necessary to understand specific PRRs as they affect student behaviour.

However, not all research supports the search for or use of a PRR threshold. In a descriptive study including 24 teachers serving Native American students (McComas et al., 2017), only opportunities to respond and reprimands significantly predicted student on-task behaviour (praise did not), though only 18% of observations yielded PRRs higher than 3:1. The average PRR across observations was 1.25:1; accordingly, the authors hypothesised correlations including praise might have been stronger had more PRRs met their optimal 3:1 standard. In

another descriptive study including 268 elementary students with or at risk for EBD in general education and self-contained classrooms (Maggin, Wehby, Moore Partin, Robertson, & Oliver, 2011), researchers observed how students across settings remained highly engaged despite a lower PRR than expected (near 1:1). In a third study, researchers measured the effects of a social skills intervention called The Cool Tool on teacher and student behaviour in an urban elementary school (Utley & Obiakor, 2015). There was no significant change in PRR between pretest (0.9:1) and posttest (1.2:1), but student on-task behaviour increased from 71% to 87%.

The field seems to lack consensus regarding the importance of high PRRs. Variable PRRs ranging from 1:1 by Maggin et al. (2011) to as high as 6:1 by Sutherland and Wehby (2001) have been reported to be effective for improving students' behaviour. Additionally, outcomes following interventions have been inconsistent as well. For instance, Jolstead et al. (2017) found increased PRR and increased student on-task behaviour, while Utley and Obiakor (2015) found increased student on-task behaviour with no change in PRR. Such differences could be due to a number of factors including differing operational definitions. Rather than discouraging future PRR research, divergent findings suggest a need for clear operational definitions. For example, it is possible that differences in the definition of praise and how it is delivered offer an explanation for the divergent findings in prior studies. It is also possible that PRR thresholds might not be the same across all populations (e.g., culture, age, disability status). For example, students with or at risk for EBD have been observed to be more sensitive to teacher praise and reprimands than students without EBD in general education settings (Downs et al., 2019).

Study Purpose

Considering the current state of knowledge about PRR, there is a need for research conducted with large samples and clear operational definitions. The current study analyses data

from a large sample of students across three states, from multiple classrooms, during various classroom activities, across multiple grade levels, using clear operational definitions, while simultaneously measuring students' on-task behaviour. Given the different PRR recommendations in the literature, this study examined whether PRRs can be used to predict students' on-task behaviour, and if so whether there is a PRR threshold at which behaviour dramatically improves. Specific research questions were as follows: (a) Do teacher PRRs predict students' on-task behaviour and (b) Is there a particular PRR threshold for improving students' on-task behaviour in the classroom?

Method

Settings and Participants

The data utilised in the current study were gathered over three years as part of a multi-site randomised control efficacy trial of CW-FIT (Wills, Kamps, Caldarella, Wehby, & Swinburne-Romine, 2018). During the efficacy trial, teachers' PRRs and students' on-task behaviour percentages were measured in 19 U.S. elementary schools across urban and suburban areas of Missouri, Tennessee, and Utah in treatment ($n = 79$) and control ($n = 72$) classrooms.

Table 1 shows demographic information for each school. The total number of participating students was 2,536 (53% male, 47% female) across 151 elementary classrooms (approximate ages 5-12), broken down by grade level as follows: kindergarten = 19.20%, 1st = 19.21%, 2nd = 13.91%, 3rd = 19.87%, 4th = 14.57%, 5th = 9.27%, 6th = 3.97%. The average class size was 19.69 ($SD = 6.01$), with an average of 2.09 ($SD = 2.14$) students with individualised education programmes and 2.74 ($SD = 3.81$) students identified as English language learners. Students were observed during subjects including language arts (56%), math

(38%), social studies (3%), science (1%), and other (2%). Approximately 7% of classrooms were special education settings. Participating teachers ($n = 151$) were predominantly female (95%) and White/Caucasian (83%). Teacher age ranged from 21 to 65 years ($M = 38.51$, $SD = 12.51$). Teachers' highest educational level was often a master's (46%) or a bachelor's degree (43%). Total years teaching ranged from 0 to 44 ($M = 9.28$; $SD = 9.13$).

INSERT TABLE 1 ABOUT HERE

Procedures and Measures

Participating schools were referred by their respective districts. School administrators provided opportunities for researchers to request voluntary participation from teachers and parents, who then completed appropriate informed consent procedures per institutional review board requirements at their school districts and participating universities. Researchers assigned teachers along with their classrooms to treatment or control conditions, stratifying them by grade level (K–2, 3–6) and classroom type (general or special education) and then randomising with a selection function in Microsoft Excel.

Teachers in the treatment condition were asked to use the CW-FIT intervention as their primary classroom management tool during the time of day they experienced the most challenging student behaviour. CW-FIT is implemented by (a) instructing students about expected social skills, (b) organising a group contingency with a timer, (c) providing differential reinforcement (e.g., publicly awarded points, tangible rewards) and praise for use of expected social skills, and (d) offering a secondary tier of more immediate support via self-management and help cards for students not responding to differential reinforcement. See Caldarella et al., 2018 and Wills et al., 2018 for more information regarding how the specific CW-FIT

intervention procedures were implemented. Teachers in the control condition used their typical classroom management practices (e.g., praise, token economies, reprimands, clip charts, positive behaviour tickets, daily behaviour report cards, class rewards), creating an opportunity for naturalistic observations during the times of day teachers reportedly experienced the most challenging student behaviour. Data were collected in both conditions across four to six months via direct observation and then entered and stored electronically on researchers' password-protected computers.

Direct observations. Researchers and trained observers completed 20-minute observations in both treatment and control classrooms. In each classroom, observations were completed at the same time of day, the time chosen by the teacher as having the most challenging behaviour, as CW-FIT aimed to improve challenging behaviour in the classroom by praising and rewarding positive student behaviours. Teachers' PRR and students' on-task behaviour were measured an average of 16 times across treatment and control classrooms during both baseline and intervention phases.

Teacher praise and reprimand. Teachers' praise statements and reprimands to individuals and groups were measured using a frequency count concurrently during the 20-minute observations. Praise was defined as verbal indication of approval following student behaviour more than acknowledging a correct response (see Caldarella et al., 2017; M. A. White, 1975; Wills et al., 2018 for additional details). Examples included, 'Well done class, you all followed directions and got in line quietly!' and 'Way to go, Robyn!' Vague statements were not included (e.g., 'Thanks, Rodrigo.'). Reprimand was defined as verbal disapproval (including a threat or scolding) in response to inappropriate behaviour or instruction that the behaviour must

stop (see Caldarella et al., 2017; M. A. White, 1975; Wills et al., 2018). Examples included, 'Everyone needs to keep their hands and feet to self,' or 'Kevin, I told you to stop throwing paper.' Vague disapproval without instruction (e.g., 'No, that's not it,') or the action of staring at students and silently waiting for them to stop talking were not included.

Student on-task and off-task behaviour. Student on-task and off-task behaviour was defined and measured as in other studies (see Caldarella et al., 2015; Wills et al., 2018). Student behaviour was measured using momentary time sampling at 30-second intervals during the 20-minute observations. At the end of each interval, observers recorded student groups (e.g., rows, teams, or clusters of students that could be easily identified by the observer as a group) as engaging in either on-task (+) or off-task (-) behaviour. If one or more students in the group were off-task at that moment of observation, the entire group was considered off-task for that interval. On-task was defined as students working appropriately on a class activity approved or assigned by the teacher: (a) attending to the teacher or task (e.g., following along in a book or keeping eyes on teacher), (b) asking for help appropriately (e.g., raising hand), or (c) waiting quietly. Off-task was defined as students not participating appropriately: disrupting class, disengaging, or working on an unapproved/unassigned activity. Following each observation, observers calculated an on-task behaviour percentage for the entire class by adding together the number of intervals each group was on-task and then dividing by the total number of intervals possible.

Interobserver agreement (IOA). Observers were trained to use the on-task observation form by studying definitions, passing a quiz, and watching videos of classrooms previously scored by the researchers. When observers achieved 90% accuracy with training videos, they practiced observing in classrooms with a research coordinator until they achieved 90% accuracy,

exceeding the minimum recommendation of at least 80% accuracy (Cooper, Heron, & Heward, 2007). Once observers were proficient, they began data collection. During data collection a reliability observer was present during 24% of observations across classrooms, exceeding the minimum recommendation of collecting IOA for at least 20% of observations (Cooper et al., 2007). IOA was calculated following reliability observations by dividing the number of observer agreements, by interval, by the total number of agreements and disagreements. IOA averaged 97% ($SD = 3.77$) across data collection, ranging from 75% to 100%.

Data Analyses

We began with descriptive statistics and correlations of the variables of interest used in the analyses. We then created a multilevel linear regression model (MLM) with time nested within teacher to answer the research questions regarding teachers' PRR and students' on-task behaviour. MLM has been previously used to analyse data from educational settings (Thum & Bhattacharya, 2001) and was found sufficient to study relations among the variables of interest (e.g., PRR and students' on-task behaviour percentages). Studies to date have created a scaled PRR coefficient (R. L. White, 2010) or used MLM (Gage, Scott, Hirn, & MacSuga-Gage, 2018), but the current study is the first to utilise both of them together.

Nested classroom data were addressed with the use of `CLUSTER=studentID` and `TYPE=twolevel` in the ANALYSIS section of Mplus 8.1 (Muthén & Muthén, 2017). At the within-teacher level we controlled for (a) observation beginning, (b) CW-FIT exposure, (c) daily praise statements, (d) daily reprimands, (e) daily PRR, and (f) grade level. At the teacher level we controlled for (a) average number of praise statements, (b) average number of reprimands, (c) school level free/reduced price lunch, (d) observation site (Missouri, Tennessee, Utah), (e)

receipt of CW-FIT intervention, (f) average PRR, and (g) grade level. A Bayesian estimator was used, as it has been shown to produce trustworthy results, especially with smaller sample size (Muthén, 2010). The Bayesian estimator in this study used uninformed priors to make its parameter estimates. This was done to mimic the maximum likelihood results as much as possible but with the advantages that the Bayesian methodology (e.g. no normality assumptions) provides. Assumptions of linearity, variance equality, and multi-collinearity were tested using scatter plots, residual plots, and histograms. Missing data were addressed using the Gibbs algorithm, which is similar to the full information maximum likelihood (FIML) method in Mplus.

Results

Descriptive Statistics and Correlations

Table 2 shows the descriptive statistics and correlations of the variables of interest used in the analysis. As expected, the percentage of students' on-task behaviour was negatively associated with total teacher reprimands ($r = -.32$) and positively associated with both total teacher praise statements ($r = .37$) and PRR ($r = .52$). The CW-FIT intervention was positively associated with on-task behaviour ($r = .40$). As expected, PRR was also positively correlated with total teacher praise statements ($r = .62$) and negatively correlated with total teacher reprimands ($r = -.42$). The p values for all statistically significant correlations were less than .001. PRR ranged from .05 (1:20) to .87 (7:1), and on-task behaviour ranged from 31.66% to 86.42%, indicating a broad range of classroom environments relevant to the number of praise statements, reprimands, and instances of students' on-task behaviour.

INSERT TABLE 2 ABOUT HERE

MLM results

To answer the question of whether there is a PRR threshold, we included a squared (nonlinear) term in the MLM. If a PRR threshold existed, the relation or line between PRR and on-task behaviour would be flat until this threshold was met and then a steep incline would be apparent. The nonlinear term was found to be non-significant ($p > .05$), revealing only a linear relation between PRR and on-task behaviour, with no steep incline (see Figure 1). This answers both Research Questions 1 and 2, as it shows that there was a relation between PRR and on-task behaviour but not an observed threshold or tipping point. In order to understand Figure 1, we need to unpack the PRR, which is a ratio (or percentage) of the total number of praise statements a teacher delivers to the group (as a whole, to individuals, or small groups) over the total number of interactions she/he has with the group (praise statements and reprimands). PRR is defined as total teacher praise statements divided by the sum of total teacher praise statements plus total teacher reprimands.

INSERT FIGURE 1 ABOUT HERE

A value of 1.00 on the PRR would mean all the interactions with the teacher were praise, while a 0 would mean that none of the interactions with the teacher were praise. A PRR value of 0.50 would mean half the interactions were praise statements and half were reprimands (e.g., 1:1 PRR). A PRR value of 0.75 would mean 75% of the interactions were praise statements (e.g., 3:1 PRR). In Figure 1, floor and ceiling effects of PRR are evident, as it is not possible to have more than 100% praise or less than 0% praise. This is a slight modification (but is mathematically equivalent) to the classic PRR, as this new metric allows for the cases of no praises or no reprimands.

As can be seen in Table 3, the effect of the CW-FIT intervention was significant ($b = 8.55$), meaning intervention classrooms scored an average of 8.55 percentage points higher than control classrooms on the students' on-task behaviour metric. The effect of overall praise was also significant but negative ($b = -0.34$). This counterintuitive result can be understood as this relationship is estimated in the presence of the PRR which already contains the teacher praises and reprimands, therefore what is leftover is not what would be captured in a bivariate correlation between teacher praise and student on-task behaviour. Reprimands were not associated with on-task behaviour, while the PRR was positively associated with on-task behaviour ($b = 29.37$), meaning as classroom PPR changed from near 0 (no praises) to near 1 (all praises) student on-task behaviour percentage was predicted to increase by almost 30 percentage points. An analysis revealed that the CW-FIT intervention did not interact with the effects of teacher praise, reprimands, or PRR. Several control variables (see Data Analyses section) were included in the final analysis but not included in Table 3 because of space limitations and because between classroom results were the focus of the study.

INSERT TABLE 3 ABOUT HERE

Discussion

Prior research has addressed the positive effects of teacher praise (Conroy et al., 2009; Reinke et al., 2013; Royer et al., 2019) as well as the negative effects of teacher reprimands (Kodak et al., 2007; Van Acker et al., 1996) on student outcomes. Researchers have made promising discoveries by combining increased praise and decreased reprimands into a single construct (higher PRRs) and observing how it relates to student outcomes, such as on-task behaviour (Caldarella et al., 2017; Leff et al., 2011). However, PRR research in educational

settings is not currently sufficient to produce a widely accepted empirically based threshold ratio similar to the 5:1 ratio of positive to negative statements noted in marital research (Gottman & Levenson, 1999). The current study endeavoured to analyse how teachers' PRRs affected students' behaviour in a large, diverse sample of elementary schools in the United States as well as to identify a PRR threshold.

Predicting On-Task Behaviour

We found a linear relationship between students' on-task behaviour and teachers' PRR, which suggests that as PRR increases, students' on-task behaviour increases in a predictable linear manner. Additionally, PRR was more highly associated with praise, which makes sense as CW-FIT, the intervention forming the backdrop for this study, set out to increase teacher praise, which would increase the PRR, while the intervention did not seek to directly affect teacher reprimands. Both the CW-FIT and control group showed 20% to 30% improvements in student on-task behaviour as PRR varied from near 0 to near 1. Further, the higher mean percentage of students' on-task behaviour observed in the CW-FIT group provides additional support for proactive classroom interventions intended to improve student outcomes. Figure 1 suggests that if teachers can simply reach a 1:1 PRR ratio, students' on-task behaviour could reasonably be expected to reach approximately 60% in the absence of other interventions. Results from these analyses are similar to previously published research showing that increased PRR is associated with increased student on-task behaviour (Jolstead et al., 2017; Leff et al., 2011; R. L. White, 2010).

A PRR Threshold

In this study, the quadratic term was not significant, which suggests that there is not a PRR threshold, or 'tipping point' (e.g., 3:1 or 4:1), at which on-task behaviour drastically improves. Instead, study results suggest that as teachers' PRRs increase, students' on-task behaviour linearly increases. Results suggest that if teachers can increase their PRR to even a 1:1 ratio, students' on-task behaviours will improve, similar to the findings of Pisacreta, Tincani, Connell, and Axelrod (2011). This knowledge can be helpful for teachers, as some find it difficult to continually maintain high rates of praise (Dufrene, Lestermau, & Zoder-Martell, 2014). If teachers can simply praise more and reprimand less, student behaviour is likely to improve. This approach is simpler than trying to keep track of a specific target PRR.

Study results add a great deal more specificity to past research. Though a general rule of thumb in education is to strive for a PRR threshold of approximately 3:1 (Shores et al., 1993) or even 6:1 (Sutherland & Wehby, 2001), our analysis did not confirm this or any ratio as a tipping point. We found instead of a specific PRR; the higher the PRR (more praise, fewer reprimands) the better the students' on-task behaviour. The CW-FIT intervention was necessary to get to PRR ratios at or above 3:1.

In positive behaviour intervention and support theory, a realistic expectation is that about 80% of students will respond to a Tier 1 or universal level of support (Sugai & Horner, 2002). Thus teachers should expect that 80% of their class should respond to the use of a universal intervention such as increased PRR. In our research, we found that as PRR increased, students did improve to a desired level of on-task behaviour. However, PRR alone is unlikely to achieve consistently high levels of on-task behaviour for all students: Sound instructional techniques and

other evidence-based classroom management strategies must also be used (Simonsen et al., 2008).

Limitations and Directions for Future Research

Various limitations became apparent during this study. Generalisation of the study was limited because data were collected from an efficacy trial involving only one behavioural intervention programme in elementary schools in the United States. Though we accounted for the treatment condition, future research would be stronger utilising data from other behaviour intervention programmes and including secondary school settings in other geographical areas to determine whether the results are similar. Further, data were not collected during a uniform subject (e.g., math, science) or instructional activity (e.g., large group instruction, individual seat work). As a result, there may have been more or less opportunities to praise or reprimand based on the subject or instructional activity occurring in a classroom. Research conducted during a uniform period may yield additional results on the effects of teacher PRR and students' on-task behaviours.

Another limitation was that we did not assess the effects of PRR on individual students versus the whole class, students' disruptive behaviour, or the particular effects on students with or at risk for EBD, though these applications are worthy of future research as recently explored by Caldarella, Larsen, Williams, Wills, and Wehby (2019). Additional research could examine whether there is a difference for students who receive individual praise/reprimands compared to those only receiving whole class praise/reprimands, as well as behaviour specific praise compared to different types of praise (e.g., praise for ability or effort). Also, we did not examine the effects of PRR on classroom climate, other than the effects on student on-task behaviour, though this may be an area for future research.

In this study, though we noted higher PRR is more effective on improving student behaviour than lower PRR, it is unclear as to whether students need reprimands. In the study by Pfiffner et al. (1985), they found that an all praise approach was not as effective as one with both praise and reprimands. The effects of reprimands on student on-task behaviour is an area worthy of future research.

While we did not find any particular effects of grade level as a control variable on PRR, we did not test for an interaction effect between grade level and PRR as it was not a focus of this study. It would be worthwhile to further explore the effects of grade level on PRR and on-task behaviour. Though teachers often praise less and reprimand more (lower PRR) as grade level increases (Reddy et al., 2013; M. A. White, 1975), it is unknown what the effects are on student on-task behaviour across grade levels.

Finally, while the efficacy trial utilised a randomised control design to strengthen the causal results of increasing PRR on students' on-task behaviour, verifying the findings with other samples would strengthen results overall. As a particular PRR threshold of effectiveness was not detectable does not necessarily mean one does not exist--with a more diverse sample and more statistical power such a threshold could become evident.

Conclusion

Results of the present study are promising and suggest that as elementary teachers' PRR increases, students' on-task behaviour increases as well. Teachers may utilise study results to improve classroom environments by increasing their PRR thus improving the behaviour of elementary students, though additional research is needed. Future PRR studies might focus on understanding the relationship between teacher and student variables across a number of settings, which could be accomplished by involving additional diverse samples, by using consistent

operational definitions across multiple settings (e.g., classroom activities, grades), and by measuring both teacher and student behaviour. However, we agree with Sabey et al. (2019) that ratio research should focus less on identifying a single rule and more on responding appropriately to specific classroom and individual student needs. While increasing teachers' PRRs can improve students' on-task behaviour, other classroom management strategies are also recommended such as teaching and reminding students of classroom expectations, increasing opportunities to respond, using a continuum of strategies to acknowledge appropriate and inappropriate behaviour, and positively supporting students while managing classroom behaviour (Simonsen et al., 2008). The current study results are hopeful, as they indicate that any increase of teachers' PRR appears to lead to increased students' on-task behaviour in elementary school classrooms.

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
- Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research, 51*, 5–32. doi:10.2307/1170249
- Caldarella, P., Larsen, R. A. A., Williams, L., Wills, H., Kamps, D., & Wehby, J. H. (2018). Effects of CW-FIT on teachers' ratings of elementary school students at risk for emotional and behavioral disorders. *Journal of Positive Behavior Interventions, 20*, 78–89. doi:10.1177/1098300717723353
- Caldarella, P., Larsen, R. A. A., Williams, L., Wills, H. P., & Wehby, J. H. (2019). Teacher praise-to-reprimand ratios: Behavioral response of students at risk for EBD compared with typically developing peers. *Education and Treatment of Children, 42*, 447-468.
- 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
- Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64*, 363–423. doi:10.2307/1170677

- Chalk, K., & Bizo, L. A. (2004). Specific praise improves on-task behavior and numeracy enjoyment: A study of year four pupils engaged in the numeracy hour. *Educational Psychology in Practice, 20*, 335–351. doi:10.1080/0266736042000314277
- Conroy, M. A., Sutherland, K. S., Snyder, A., Al-Hendawi, M., & Vo, A. (2009). Creating a positive classroom atmosphere: Teachers' use of effective praise and feedback. *Beyond Behavior, 18*(2), 18–26.
- Cooper, J. O. Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson.
- Downs, K. R., Caldarella, P., Larsen, R. A. A., Charlton, C. T., Wills, H. P., Kamps, D. M., & Wehby, J. H. (2019). Teacher praise and reprimands: The differential response of students at risk for emotional and behavioral disorders. *Journal of Positive Behavior Interventions, 21*, 135–147. doi:10.1177/1098300718800824
- Dufrene, B. A., Lestremau, L., & Zoder-Martell, L. (2014). Direct behavioral consultation: Effects on teachers' praise and student disruptive behavior. *Psychology in the Schools, 51*, 567–580. doi:10.1002/pits.21768
- Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review, 11*, 75–113. doi:10.1007/s10567-008-0036-x
- Ferrari, M. (1980). Effects of praise and reprimand on the topography and probability of occurrence of stereotypies in autistic children. *Psychological Reports, 46*, 517–522. doi:10.2466/pr0.1980.46.2.519
- 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., Beschta, S. L., Meyer, K. L., & Reinke, W. M. (2017). Praise research trends and future directions: Characteristics and teacher training. *Behavioral Disorders, 43*(1), 227–243. doi:10.1177/0198742917704648
- 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
- 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. doi:10.1177/1053451208328831
- Gage, N. A., Scott, T., Hirn, R., & MacSuga-Gage, A. S. (2018). The relationship between teachers' implementation of classroom management practices and student behavior in elementary school. *Behavioral Disorders, 43*, 302–315. doi:10.1177/0198742917714809
- Gottman, J. M., & Levenson, R. W. (1999). What predicts change in marital interaction over time? A study of alternative medicine. *Family Process, 38*, 143–158. doi:10.1111/j.1545-5300.1999.00143.x
- Haydon, T., & Musti-Rao, S. (2011). Effective use of behavior-specific praise: A middle school case study. *Beyond Behavior, 20*(2), 31–39.
- 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

- Kodak, T., Northup, J., & Kelley, M. E. (2007). An evaluation of the types of attention that maintain problem behavior. *Journal of Applied Behavior Analysis, 40*, 167–171.
doi:10.1901/jaba.2007.43-06
- 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
- Maag, J. W. (2001). Rewarded by punishment. Reflections on the disuse of positive reinforcement in schools. *Exceptional Children, 67*, 173–186.
doi:10.1177/001440290106700203
- Maggin, D. M., Wehby, J. H., Moore Partin, T. C., Robertson, R., & Oliver, R. M. (2011). Comparison of the instructional context for students with behavioral issues enrolled in self-contained and general education classrooms. *Behavioral Disorders, 36*, 84–99.
doi:10.1177/019874291103600201
- McComas, J. J., Downwind, I., Klingbeil, D. A., Petersen-Brown, S., Davidson, K. M., & Parker, D. C. (2017). Relations between instructional practices and on-task behavior in classrooms serving American Indian students. *Journal of Applied School Psychology, 33*, 89–108. doi:10.1080/15377903.2016.1236308
- Merrell, K. W., Ervin, R. A., & Gimpel Peacock, G. (2012). *School psychology for the 21st century: Foundations and practices* (2nd ed.). New York, NY: The Guilford Press.
- Merrett, F., & Wheldall, K. (1992). Teachers' use of praise and reprimands to boys and girls. *Educational Review, 44*, 73–79. doi:10.1080/0013191920440106

Muthén, B. (2010). Bayesian analysis in Mplus: A brief introduction. *Unpublished manuscript*.

Retrieved from www.statmodel.com/download/IntroBayesVersion.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (7th ed.). Los Angeles, CA:

Muthén & Muthén.

Naylor, A. S., Kamps, D., & Wills, H. (2018). The effects of the CW-FIT group contingency on class-wide and individual behavior in an urban first grade classroom. *Education and Treatment of Children, 41*, 1–30. doi:10.1353/etc.2018.0000

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

Pfiffner, L. J., Rosén, L. A., & O'Leary, S. G. (1985). The efficacy of an all-positive approach to classroom management. *Journal of Applied Behavior Analysis, 18*, 257–261.

doi:10.1901/jaba.1985.18-257

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

Rathel, J. M., Drasgow, E., & Christle, C. C. (2008). Effects of supervisor performance feedback on increasing preservice teachers' positive communication behaviors with students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 16*, 67–

77. doi:10.1177/1063426607312537

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. doi:10.1016/j.jsp.2013.10.001

- 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
- Reinke, W. M., Lewis-Palmer, T., & Martin, E. (2007). The effect of visual performance feedback on teacher use of behavior-specific praise. *Behavior Modification, 31*, 247–263. doi:10.1177/0145445506288967
- Rosén, L. A., O'Leary, S. G., Joyce, S. A., Conway, G., & Pfiffner, L. J. (1984). The importance of prudent negative consequences for maintaining the appropriate behavior of hyperactive students. *Journal of Abnormal Child Psychology, 12*, 581–604. doi:10.1007/bf00916852
- Royer, D. J., Lane, K. L., Dunlap, K. D., & Ennis, R. P. (2019). A systematic review of teacher-delivered behavior-specific praise on K-12 student performance. *Remedial and Special Education, 40*, 112-128. doi:10.1177/0741932517751054
- Sabey, C. V., Charlton, C., & Charlton, S. R. (2019). The “magic” positive-to-negative interaction ratio: Benefits, applications, cautions, and recommendations. *Journal of Emotional and Behavioral Disorders, 27*, 154-164. doi:10.1177/1063426618763106
- 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
- Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31*, 351–380. doi:10.1353/etc.0.0007

- Stormont, M. A., Smith, S. C., & Lewis, T. J. (2007). Teacher implementation of precorrection and praise statements in Head Start classrooms as a component of a program-wide system of positive behavior support. *Journal of Behavioral Education, 16*, 280–290.
doi:10.1007/s10864-007-9040-3
- Sugai, G. & Horner, R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy, 24*(1-2), 23–50.
doi:10.1300/J019v24n01_03
- Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *Journal of Special Education, 35*, 161–171. doi:10.1177/002246690103500306
- Thum, Y. M., & Bhattacharya, S. K. (2001). Detecting a change in school performance: A Bayesian analysis for a multilevel join point problem. *Journal of Educational and Behavioral Statistics, 26*, 443–468. doi:10.3102/10769986026004443
- Trussell, R. P. (2008). Classroom universals to prevent problem behaviors. *Intervention in School and Clinic, 43*, 179–185. doi:10.1177/1053451207311678
- Utle, C. A., & Obiakor, F. E. (2015). Measuring the cool tool as a targeted intervention to minimize teacher reprimands and students' on-task behaviors in an urban elementary school. *Learning Disabilities: A Contemporary Journal, 13*, 43–57.
- Van Acker, R., Grant, S. H., & Henry, D. (1996). Teacher and student behavior as a function of risk for aggression. *Education and Treatment of Children, 19*, 316–334.
- Weeden, M., Wills, H. P., Kottwitz, E., & Kamps, D. (2016). The effects of a class-wide behavior intervention for students with emotional and behavioral disorders. *Behavioral Disorders, 42*, 285–293. doi:10.17988/bd-14-12.1

- White, M. A. (1975). Natural rates of teacher approval and disapproval in the classroom. *Journal of Applied Behavior Analysis*, 8, 367–372. doi:10.1901/jaba.1975.8-367
- White, R. L. (2010). *Demonstration of the effects of an increased praise ratio on student on-task behavior* (Doctoral dissertation [or Master's thesis]). Retrieved from ProQuest Digital Dissertations. (AAT 577644399)
- Wills, H. P., Iwaszuk, W. M., Kamps, D., & Shumate, E. (2014). CW-FIT: Group contingency effects across the day. *Education and Treatment of Children*, 37, 191–210. doi.org/10.1353/etc.2014.0016
- Wills, H. P., Kamps, D., Caldarella, P., Wehby, H. P., & Swinburne-Romine, R. S. (2018). Class-Wide Function-Related Intervention Teams “CW-FIT”: Outcomes from a multi-site randomized replication trial. *The Elementary Education Journal*, 119, 29–51. doi:10.1086/698818
- 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

Table 1

Descriptive Data for Schools (n = 19) by Sites and Years

School ID	School Size	School FRL	Ethnicity							Participating Teachers		
			Asian	Black	Caucasian	Hispanic	American Indian	Multi-Race	Pacific Islander	Exp.	Comp.	
Missouri												
Year 1												
1	220	71.90%	0.00%	55.50%	31.40%	9.10%	0.00%	4.00%	0.00%	5	3	
2	243	48.20%	0.00%	35.40%	42.80%	16.50%	0.00%	5.30%	0.00%	2	2	
Year 2												
3	319	95.80%	0.00%	83.40%	5.30%	0.00%	0.00%	11.30%	0.00%	4	4	
4	578	59.20%	0.00%	7.30%	67.30%	14.50%	0.00%	10.90%	0.00%	4	4	
5	290	81.20%	0.00%	65.50%	13.40%	13.40%	0.00%	7.00%	0.00%	3	2	
Year 3												
6	289	72.70%	0.00%	11.80%	65.70%	10.40%	0.00%	12.10%	0.00%	4	5	
7	515	65.00%	0.00%	16.70%	52.00%	16.90%	0.00%	14.40%	0.00%	8	7	
Tennessee												
Year 1												
8	519	52.00%	7.10%	23.70%	58.60%	10.20%	0.00%	0.00%	0.00%	4	4	
9	677	94.20%	1.50%	41.70%	24.50%	32.30%	0.00%	0.00%	0.00%	3	4	
Year 2												
10	475	98.10%	0.40%	91.40%	4.60%	3.60%	0.00%	0.00%	0.00%	5	4	
11	472	40.90%	3.80%	31.10%	59.50%	5.10%	0.20%	0.00%	0.00%	3	3	
Year 3												
12	317	91.50%	1.60%	69.10%	15.50%	13.90%	0.00%	0.00%	0.00%	4	3	
13	384	99.00%	1.00%	89.80%	7.80%	1.00%	0.00%	0.00%	0.00%	4	3	
Utah												
Year 1												
14	425	69.20%	0.00%	1.18%	80.24%	15.53%	0.71%	1.18%	1.18%	4	4	
Year 2												
15	490	81.00%	1.22%	0.61%	57.35%	36.12%	1.43%	2.04%	1.22%	5	4	
16	476	35.90%	1.05%	0.42%	81.51%	9.24%	0.84%	4.20%	2.73%	3	2	
17	409	55.30%	0.98%	1.96%	57.46%	35.21%	0.49%	2.44%	1.47%	3	3	
Year 3												
18	630	34.00%	2.06%	1.27%	79.84%	10.79%	0.48%	2.70%	2.86%	4	4	
19	504	82.70%	1.59%	1.19%	35.32%	54.76%	1.19%	3.77%	2.18%	7	7	

Note. FRL = free/reduced-price lunch; Exp. = number of teachers assigned to experimental condition; Comp. = number of teachers assigned to comparison condition.

Table 2

Descriptive statistics and correlations of variables of interest used in multilevel linear regression

model at the teacher level (n = 151)

Variable	Students' On-task Behaviour Percentage	Total Teacher Praise	Total Teacher Reprimands	Treatment (CW-FIT) Effect	Average PRR
Students' On-task Behaviour Percentage	1				
Total Teacher Praise	.37**	1			
Total Teacher Reprimands	-.32**	.01	1		
Treatment (CW-FIT) Effect	.40**	.44**	-.06	1	
Average PRR	.52**	.62**	-.42**	.44**	1
Mean	65.12	9.31	5.76	.53	.52
SD	11.26	7.81	3.15	.50	.21
min	31.66	0.40	1.11	0.00	.05
max	86.42	47.50	17.29	1.00	.87

** $p < .001$; CW-FIT = Class-wide function-related intervention teams; PRR = praise-to-reprimand ratio.

Table 3

Multilevel linear regression model results predicting students' on-task behaviour percentage (n = 2,536 students, 151 classrooms)

Predictor Variable	Beta Estimate	Posterior <i>SD</i>	Standardised Beta
Total Teacher Praise	-0.34*	0.15	-0.23
Total Teacher Reprimands	-0.55	0.31	-0.14
Treatment (CW-FIT) effect	8.55*	1.80	0.39
Average PRR	29.37*	7.14	0.54

Note. Only between classroom results are shown as that was the focus of the study. The controls for daily praise, daily reprimands, daily PRR, when treatment started, observation site, school free/reduced priced lunch, and grade level are not shown for space limitations. *Parameter estimate of 0 was not in 95% Bayesian credibility interval, indicating statistical significance.

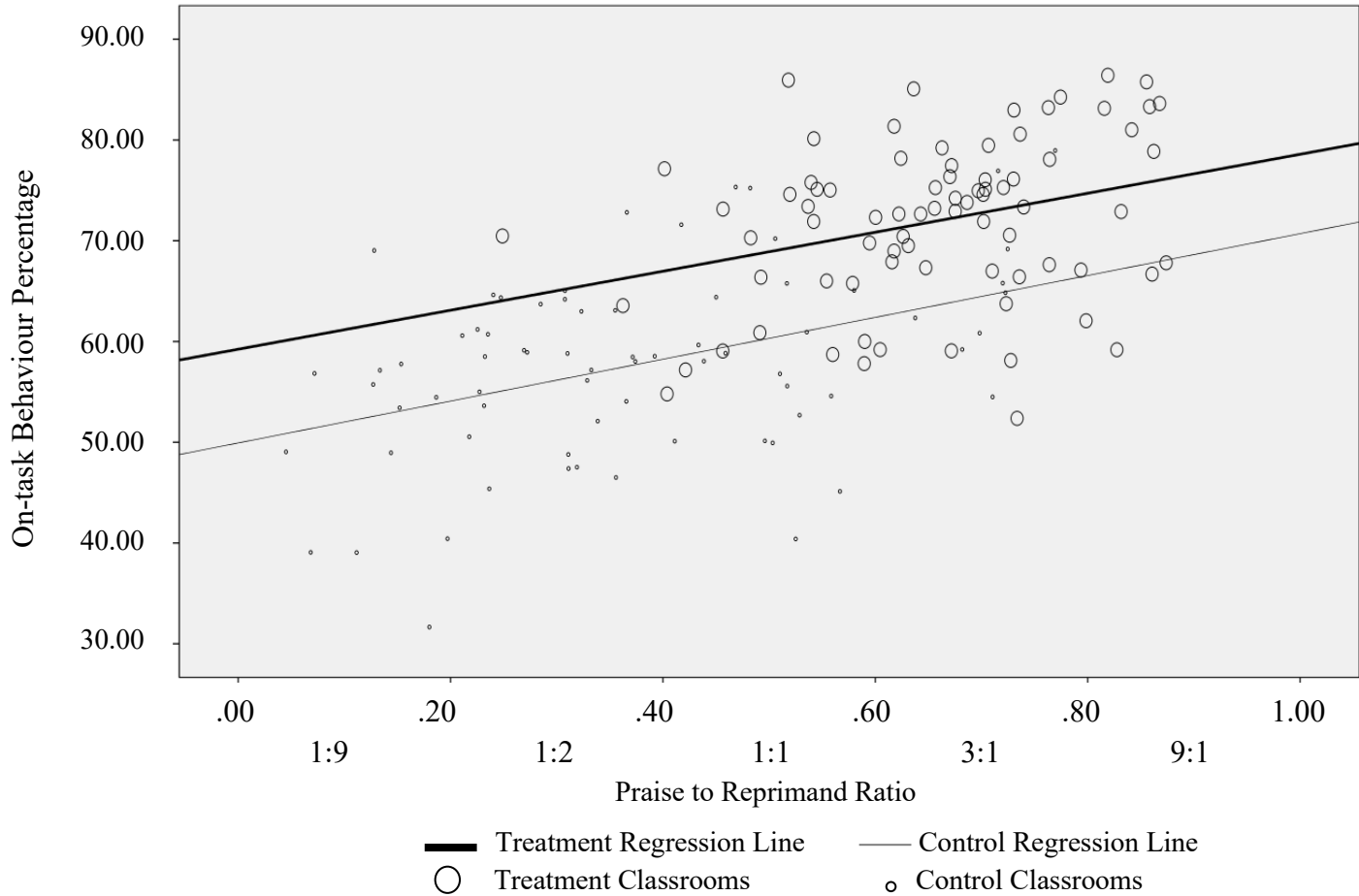


Figure 1. Simple linear relation between average praise-to-reprimand ratio (PRR) and average on-task behaviour percentage per classroom ($n = 151$).