SELF-MANAGEMENT FOR CLASSROOM TEACHERS

by

LINDSEY E. POWELL

(Under the Direction of Scott Ardoin)

ABSTRACT

The first study of this dissertation (Chapter 2) contains a systematic review of single-case design literature evaluating the effects of self-management (SM) interventions on teachers' rates of opportunities to respond (OTRs) and/or specific/contingent praise. The quality and rigor of single-case design studies utilizing SM to increase rates of OTRs and/or specific/contingent praise are discussed using the Single Case Analysis and Review Framework (SCARF; Ledford et al., 2018). We examined the effects of a video-recording self-evaluation intervention on teachers' rates of behavior-specific praise (BSP) in the second study of this dissertation (Chapter 3). Results suggested that the self-evaluation intervention increased rates of teachers' BSP. Implications for future research and study limitations are discussed.

INDEX WORDS: Self-management, Self-monitoring, Self-evaluation, Opportunities to respond, Behavior-specific praise

SELF-MANAGEMENT FOR CLASSROOM TEACHERS

by

LINDSEY E. POWELL

BS, University of Georgia, 2018

MS, University of Georgia, 2019

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2024

SELF-MANAGEMENT FOR CLASSROOM TEACHERS

by

LINDSEY E. POWELL

Major Professor:	Scott Ardoin
Committee:	Amy Reschly
	Kristin Sayeski
	Chitra Pidaparti

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2024

ACKNOWLEDGEMENTS

I would like to thank the following people, without whom I could not have finished my dissertation and made it thought my doctoral degree. First, thank you to my advisor, Scott, who has helped me develop into a behavior analyst and psychologist over the past 7 years. Thank you to my patient and flexible committee, who has gone above and beyond in helping me meet my deadlines. Thank you to my wonderful participants for taking on extra work in your classrooms in order to help me graduate. It was such a pleasure getting to see you teach your students. To Paulina, thank you for being my friend and always making sure I am fed when I am working into the night. To Greg, thank you for sticking with me throughout 5 years of graduate school. It has meant sacrificing some fun and surviving many sleepless nights. Most importantly, thank you to my parents for supporting me in my goals, helping me financially, and loving me unconditionally. I simply could not have done this without you.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

A REVIEW OF THE CLASSROOM AND BEHAVIOR MANAGEMENT LITERATURE

Teachers document classroom and behavior management as significant challenges (Reinke et al., 2013). Teacher preparation programs rarely provide future educators with adequate behavior management training or teach them how to utilize evidence-based classroom and behavior management strategies (Meister & Melnick, 2003; Levine, 2006; Stevenson et al., 2020). As a result of this lack of training, Reinke et al. (2011) found that early childhood and elementary school teachers reported disruptive behaviors were the most significant challenge that students experienced. Teachers simultaneously report low levels of training in classroom and behavior management with approximately 20% of the teachers rating their training and experience using behavioral interventions as none or minimal (Reinke et al., 2011). Both teachers and administrators perceive classroom and behavior management and the lack of training in teacher preparation programs as significant concerns (Ladd, 2000). In absence of sufficient training and support, teachers are likely to perceive that their professional demands exceed their ability to cope, thus leading to unnecessary stress, job dissatisfaction, and emotional exhaustion (McCarthy et al., 2015), eventually leading to teacher burnout (Otero-López et al., 2010). Student discipline problems linked to poor management practices are one of the largest barriers to teacher-professional success (Fideler & Haskelhorn, 1999) and one of the most common reasons that teachers leave the profession (Ingersoll & Smith, 2003; Otero-López et al., 2010).

In addition to the negative consequences that a lack of classroom and behavior management skills has on teachers, it also results in both negative academic and nonacademic outcomes for students of these teachers (Owens et al., 2017). Ineffective classroom and behavior management impacts academic instruction, as having to manage students' disruptive behaviors draws away from the time that teachers could otherwise be teaching (Freiberg et al., 2009; Flower et al., 2017). Simultaneously, ineffective classroom and behavior management can create an unstructured and unpredictable environment, reducing the occurrence of student on-task behavior (Simonsen et al., 2008) and the likelihood of teachers providing engaging instruction (Reinke et al., 2022).

In contrast, effective classroom and behavior management can decrease disruptive behaviors and increase student engagement (Reinke et al., 2008; Collier-Meek et al., 2019). Teachers who effectively implement classroom and behavior management maximize structure, reinforce expectations, engage students, acknowledge appropriate student behavior (Simonsen et al., 2008), and maximize instructional time (Flower et al., 2014). Furthermore, these teachers report higher levels of job satisfaction and are less likely to experience burnout (Canrinus et al., 2012; Caprara et al., 2006). Ultimately, poor classroom and behavior management and disruptive behaviors negatively impact students, teachers, classrooms, and schools (Freiberg et al., 2009). Evidence-based classroom and behavior management strategies must address these issues (Flower et al., 2014).

As teacher preparation programs generally fail to provide teachers with the skills to utilize evidence-based classroom and behavior management strategies (Levine, 2006), teachers ultimately develop behavior management skills once they are in the classroom. Low levels of training in classroom and behavior management significantly reduce the effectiveness of new

teachers (Oliver & Reschly, 2010). Teachers who feel underprepared at the outset of their careers are more likely to experience burnout (O'Brennan et al., 2017), and studies show that burnout is related to low teacher self-efficacy (O'Brennan et al., 2017; Skaalvik & Skaalvik, 2010). Data from O'Brennan et al. (2017) suggest burnout is more likely to occur when teachers feel they cannot manage student behavior. Thus, teachers are at higher risk for burnout when their self-efficacy and confidence is low (O'Brennan et al., 2017), and these individuals may be less likely to utilize effective strategies (Reinke et al., 2013).

Behavioral Consultation

To ensure teachers are using evidence-based classroom and behavior management strategies, administrators, school psychologists, and other school personnel may be required to take on the responsibility of disseminating knowledge about these interventions to teachers. Unfortunately, researchers suggest teachers have difficulty demonstrating adequate treatment integrity after intervention training (Noell et al., 2005; Collier-Meek et al., 2019). Given teachers' inconsistent intervention implementation, school personnel also monitor teachers' use of classroom and behavior management strategies and provide support in implementing interventions (Fallon et al., 2015). Not only are teachers expected to utilize evidence-based classroom and behavior management strategies, but school personnel should use evidence-based strategies to support teachers in their implementation of interventions (Fallon et al., 2015). Supporting teachers' implementation of evidence-based classroom and behavior management strategies can boost confidence and self-efficacy (O'Brennan et al., 2017) and positively affect student outcomes (Fallon et al., 2015).

School personnel, including school psychologists, might employ a consultation model to support teachers' implementation of evidence-based classroom and behavior management

strategies (Kratochwill & Bergan, 1990). Kratochwill and Bergan (1990) identified the three models of consultation that are used most often: behavioral consultation, mental health consultation, and organizational development consultation. Behavioral consultation is based on behavior modification (Kazdin, 1989) and utilizes indirect service delivery to create behavior change (Kratochwill & Bergan, 1990). Behavioral consultation involves a consultant (i.e., school personnel), a consultee (i.e., teacher or paraprofessional), and a client (i.e., student or students; Kratochwill & Bergan, 1990). Mental health consultation and organizational development consultation involve direct service delivery to treat an individual's problem and promote organizational productivity, respectively (Kratochwill & Bergan, 1990).

Behavioral consultation can be implemented in various settings, including schools (Kratochwill & Bergan, 1990). Studies document the effectiveness of behavioral consultation to treat academic underachievement as well as emotional and behavioral problems (Sheridan et al., 1996; Kratochwill et al., 2002). During school-based behavioral consultation, the school personnel's responsibilities include guiding the teacher through the steps of behavioral consultation and providing important information and resources from the applied literature in areas such as psychology, education, and behavior therapy (Kratochwill & Bergan, 1990). Although the main goal in behavioral consultation is to produce change in the student's behavior, behavioral consultation can be used as a targeted intervention for teachers to create positive change in the teacher's behavior (Kratochwill & Bergan, 1990). In order to alter a teacher's behavior in the school setting, school personnel may introduce novel interventions, modify existing strategies, and increase the teacher's confidence in their behavior management skills (Kratochwill & Bergan, 1990).

Performance Feedback

Performance feedback is a critical component of the behavioral consultation model that has been extensively researched as a means of changing adult behavior (Alvero et al., 2001) and maintaining treatment integrity during consultation (McKenney et al., 2013). Performance feedback involves monitoring a behavior of concern and informing the teacher of whether they correctly engaged in the target behavior (Noell et al., 2005). During behavioral consultation with performance feedback, the school consultant and teacher discuss intervention implementation, and the school consultant provides performance feedback, reviewing the integrity with which the intervention was implemented and outcome data (Fallon et al., 2015). Performance feedback may include positive reinforcement of intervention implementation or corrective feedback on intervention steps that were implemented incorrectly (Fallon et al., 2015). Often, behavioral consultation is not effective without performance feedback (Jones et al., 1997) with several studies suggesting that behavioral consultation alone results in low levels of treatment integrity (Jones et al., 1997; Wickstrom, 1995; Robbins & Gutkin, 1994). Despite its effectiveness, performance feedback has its drawbacks; the approach is time intensive and requires substantial assistance from school personnel, resulting in it often not being feasible for schools to implement due to limitations in time and staff (Rispoli et al., 2017). It is therefore crucial to explore effective and efficient strategies to provide performance feedback to teachers regarding classroom and behavior management (Rispoli et al., 2017).

Self-management (SM)

Self-management (SM) is a promising strategy to circumvent the time and intensive nature of performance feedback within the behavioral consultation model (Rispoli et al., 2017). SM involves the individual application of operant conditioning techniques (i.e., reinforcement and punishment) that produce a desired change in one's behavior (Malott, 1989). SM is a broad

strategy that encompasses groups of interventions such as (a) self-monitoring; (b) self-evaluation; (c) self-instruction; and (d) self-reinforcement (Eva & Regehr, 2011; Callahan et al., 1998; Nelson et al., 1991). Using self-monitoring, individuals self-assess their behavior and record the results (Moxley, 1998; Rock, 2005). Self-evaluation requires individuals to compare their behavior to a predetermined performance criterion (Spates & Kanfer, 1977). Both self-monitoring and self-evaluation have emerged as strategies to change teacher behavior, despite being used more frequently with children (Mouzakitis et al., 2015). Self-instruction involves individuals talking themselves through a task or activity and using self-induced statements to change behavior (IRIS, n.d.). In addition, when individuals select a reinforcer and reward themselves for reaching a predetermined criterion, it is considered self-reinforcement (Smith & Rivera, 1993). The benefits of self-monitoring and self-evaluation specifically include increasing the individuals' awareness of their behavior, improving target behaviors for teachers and students, and providing immediate feedback (Moxley, 1998; Rock, 2005). Self-monitoring and self-evaluation procedures are effective (Rispoli et al., 2017), minimize reliance on school personnel (Bruhn et al., 2015), and allow teachers to provide performance feedback to themselves.

Evidence-Based Classroom Management Techniques

SM strategies including self-monitoring and self-evaluation are effective and evidence-based (Sutherland & Wehby, 2001; Keller et al., 2005; Hager, 2012), but school personnel acting as consultants ultimately decide what evidence-based classroom and behavior management strategy they will implement with teachers. In a review of the classroom and behavior management literature, Simonsen et al. (2008) described 20 classroom and behavior management strategies with sufficient evidence that teachers can use. The strategies were organized into five

empirically-supported groups, including strategies to (a) maximize structure; (b) post, teach, review, monitor, and reinforce expectations; (c) actively engage students in observable ways; (d) use a continuum of strategies for responding to appropriate behaviors; and (e) use a continuum of strategies to respond to inappropriate behaviors (Simonsen et al., 2008). Specific strategies included opportunities to respond (OTRs), direct instruction, specific and contingent praise, class-wide group contingencies, behavior contracts, token economies, error correction, and performance feedback (Simonsen et al., 2008). Data from a review by Rispoli et al. (2017) suggested that the most common classroom and behavior management strategy targeted using SM strategies was teacher praise, as 10 of 17 studies aimed to increase rates of general and/or specific teacher praise. Teacher implementation of token economies, the Good Behavior Game, and use of augmentative and alternative communication to replace problem behaviors were also targeted through SM (Rispoli et al., 2017). Additionally, Haydon et al. (2012) provided specific information about how teachers can increase rates of OTR's using SM procedures.

Purpose

The first study of this dissertation (Chapter 2) extends the SM literature by systematically reviewing single-case design literature evaluating effects of SM interventions on teachers' rates of OTRs and/or specific/contingent praise. Chapter 2 assesses the quality and rigor of single-case design studies utilizing SM to increase rates of OTRs and/or specific/contingent praise using the Single Case Analysis and Review Framework (SCARF; Ledford et al., 2018). The second study of this dissertation (Chapter 3) evaluates the effects of a video-recording self-evaluation intervention on teachers' rates of behavior-specific praise (BSP) in the classroom. Chapter 3 proposes a simple, low-cost method for video self-evaluation in the school setting and extends the literature evaluating the effects of self-evaluation on teacher use of praise.

References

Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings. *Journal of Organizational Behavior Management, 21,* 3-29.

Bruhn, A., McDaniel, S., & Kreigh, C. (2015). Self-monitoring interventions for students with behavior problems: A systematic review of current research. *Behavioral Disorders*, *40*(2), 102-121.

Callahan, K., Rademacher, J. A., & Hildreth, B. L. (1998). The effect of parent participation in strategies to improve the homework performance of students who are at risk. *Remedial and Special Education*, *19*(3), 131-141.

Canrinus, E. T., Helms-Lorenz, M., Beijaard, D., Buitink, J., & Hofman, A. (2012). Self-efficacy, job satisfaction, motivation and commitment: Exploring the relationships between indicators of teachers' professional identity. *European Journal of Psychology of Education*, *27*, 115-132.

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*(6), 473-490.

Collier-Meek, M. A., Johnson, A. H., Sanetti, L. H., & Minami, T. (2019). Identifying critical components of classroom management implementation. *School Psychology Review*, *48*(4), 348-361.

Eva, K. W., & Regehr, G. (2011). Exploring the divergence between self-assessment and self-monitoring. *Advances in Health Sciences Education*, *16*(3), 311-329.

Fallon, L. M., Collier-Meek, M. A., Maggin, D. M., Sanetti, L. M., & Johnson, A. H. (2015). Is performance feedback for educators an evidence-based practice? A systematic review and evaluation based on single-case research. *Exceptional Children*, *81*(2), 227-246.

Fideler, E. F., & Haselkorn, D. (1999). *Learning the ropes: Urban teacher induction programs and practices in the united states and executive summary.* Recruiting New Teachers, Inc.

Flower, A., McKenna, J. W., Bunuan, R. L., Muething, C. S., & Vega Jr., R. (2014). Effects of the good behavior game on challenging behaviors in school settings. *Review of Educational Research*, *84*(4), 546-571.

Flower, A., McKenna, J. W., & Haring, C. D. (2017). Behavior and classroom management: Are teacher preparation programs really preparing our teachers? *Preventing School Failure: Alternative Education for Children and Youth*, *61*(2), 163-169.

Freiberg, H. J., Huzinec, C. A., & Templeton, S. M. (2009). Classroom management—a pathway to student achievement: A study of fourteen inner-city elementary schools. *The Elementary School Journal*, *110*(1), 63-80.

Hager, K. D. (2012). Self-monitoring as a strategy to increase student teachers' use of effective teaching practices. *Rural Special Education Quarterly*, *31*(4), 9-17.

Haydon, T., Macsuga-Gage, A. S., Simonsen, B., & Hawkins, R. (2012). Opportunities to respond: A key component of effective instruction. *Beyond Behavior*, *22*(1), 23-31.

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational Leadership*, *60*(8), 30-33.

IRIS. (n.d.). *Self-Instruction. https://iris.peabody.vanderbilt.edu/module/sr/cresource/q2/p05/*

Jones, K. M., Wickstrom, K. F., & Friman, P. C. (1997). The effects of observational feedback on treatment integrity in school-based behavioral consultation. *School Psychology Quarterly*, *12*(4), 316.

Kazdin, A. E. (1989). *Behavior modification in applied settings* (4th ed.). Thomson Brooks/Cole Publishing Co.

Keller, C. L., Brady, M. P., & Taylor, R. L. (2005). Using self-evaluation to improve student teacher interns' use of specific praise. *Education and Training in Developmental Disabilities*, 368-376.

Kratochwill, T. R., & Bergan, J. R. (1990). *Behavioral consultation in applied settings: An individual guide*. Springer Science & Business Media.

Kratochwill, T. R., Elliott, S. N., & Callan-Stoiber, K. (2002). Best Practices in School-Based Problem-Solving Consultation. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology IV* (pp. 583–608). National Association of School Psychologists.

Ladd, K. L. (2000). *A comparison of teacher education programs and graduates' perceptions of experiences*. University of Missouri-Columbia.

Ledford, J. R., Lane, J. D., & Tate, R. (2018). Evaluating quality and rigor in single case research. *Single Case Research Methodology* (pp. 365-392). Routledge.

Levine, A. (2006). Educating school teachers. *Education Schools Project*.

Malott, R. W. (1989). The achievement of evasive goals. *Rule-governed behavior: Cognition, contingencies, and instructional control.* (pp. 269-322). Springer.

McCarthy, C. J., Lineback, S., & Reiser, J. (2015). Teacher stress, emotion, and classroom management. *Handbook of Classroom Management*, *2*, 301-321.

McKenney, E. L., Waldron, N., & Conroy, M. (2013). The effects of training and performance feedback during behavioral consultation on general education middle school teachers' integrity to functional analysis procedures. *Journal of Educational and Psychological Consultation*, *23*(1), 63-85.

Meister, D. G., & Melnick, S. A. (2003). National new teacher study: Beginning teachers' concerns. *Action in Teacher Education*, *24*(4), 87-94.

Mouzakitis, A., Codding, R. S., & Tryon, G. (2015). The effects of self-monitoring and performance feedback on the treatment integrity of behavior intervention plan implementation and generalization. *Journal of Positive Behavior Interventions*, *17*(4), 223-234.

Moxley, R. A. (1998). Treatment-only designs and student self-recording as strategies for public school teachers. *Education and Treatment of Children*, 37-61.

Nelson, J. R., Smith, D. J., Young, R. K., & Dodd, J. M. (1991). A review of self-management outcome research conducted with students who exhibit behavioral disorders. *Behavioral Disorders*, *16*(3), 169-179.

Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., ... & Duhon, G. J. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review*, *34*(1), 87-106.

O'Brennan, L., Pas, E., & Bradshaw, C. (2017). Multilevel examination of burnout among high school staff: Importance of staff and school factors. *School Psychology Review*, *46*(2), 165-176.

Oliver, R. M., & Reschly, D. J. (2010). Special education teacher preparation in classroom management: Implications for students with emotional and behavioral disorders. *Behavioral Disorders*, *35*(3), 188-199.

Otero-López, J. M., Bolaño, C. C., Mariño, M. J. S., & Pol, E. V. (2010). Exploring stress, burnout, and job dissatisfaction in secondary school teachers. *International Journal of Psychology and Psychological Therapy*, *10*(1), 107-123.

Owens, J. S., Coles, E. K., Evans, S. W., Himawan, L. K., Girio-Herrera, E., Holdaway, A. S., ... & Schulte, A. C. (2017). Using multi-component consultation to increase the integrity with which teachers implement behavioral classroom interventions: A pilot study. *School Mental Health*, *9*(3), 218-234.

Reinke, W. M., Lewis-Palmer, T., & Merrell, K. (2008). The classroom check-up: A class-wide teacher consultation model for increasing praise and decreasing disruptive behavior. *School Psychology Review*, *37*(3), 315-332.

Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly*, *26*(1), 1.

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*(1), 39-50.

Reinke, W. M., Herman, K. C., & Copeland, C. B. (2022). Student engagement: The importance of the classroom context. *Handbook of Research on Student Engagement*, 529-544.

Rispoli, M., Zaini, S., Mason, R., Brodhead, M., Burke, M. D., & Gregori, E. (2017). A systematic review of teacher self-monitoring on implementation of behavioral practices. *Teaching and Teacher Education*, *63*, 58-72.

Robbins, J. R., & Gutkin, T. B. (1994). Consultee and client remedial and preventive outcomes following consultation: Some mixed empirical results and directions for future researchers. *Journal of Educational and Psychological Consultation*, *5*(2), 149-167.

Rock, M. L. (2005). Use of strategic self-monitoring to enhance academic engagement, productivity, and accuracy of students with and without exceptionalities. *Journal of Positive Behavior Interventions*, *7*(1), 3-17.

Sheridan, S. M., Kratochwill, T. R., & Bergan, J. R. (1996). *Conjoint behavioral consultation: A procedural manual*. New York: Plenum Press.

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*, 351-380.

Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, *26*(4), 1059-1069.

Smith, D. D., & Rivera, D. P. (1993). *Effective discipline* (2nd ed.). Austin, TX: PRO-ED.

Spates, C. R., & Kanfer, F. H. (1977). Self-monitoring, self-evaluation, and self-reinforcement in children's learning: A test of a multistage self-regulation model. *Behavior Therapy*, *8*(1), 9-16.

Stevenson, N. A., VanLone, J., & Barber, B. R. (2020). A commentary on the misalignment of teacher education and the need for classroom behavior management skills. *Education and Treatment of Children*, *43*(4), 393-404.

Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *The Journal of Special Education*, *35*(3), 161-171.

Wickstrom, K. F. (1995). *A study of the relationship among teacher, process, and outcome variables within school-based consultation*. Louisiana State University and Agricultural & Mechanical College.

CHAPTER 2

INCREASING CONTINGENT PRAISE AND OPPORTUNITIES TO RESPOND WITH SELF-MANAGEMENT: A SYSTEMATIC REVIEW OF THE LITERATURE

Introduction

Self-management (SM) is a procedure that has the potential for behavior change (Koegel & Koegel, 1990). It is defined as the individual application of operant techniques that produce a desired change in one's behavior (Malott, 1989). In the early 1970s, research in the area of behavioral self-control began to grow (McDougall, 1990). A behavioral self-control study by Glynn et al. (1973) suggested a self-control model consisting of (a) self-assessment; (b) self-recording; (c) self-determination of reinforcement; and (d) self-administration of reinforcement. The first two components (i.e., self-assessment and self-recording) would later combine to accomplish what researchers now refer to as self-monitoring. By the late 1980s, researchers primarily used the term "self-management" instead of self-control (McDougall, 1990).

Self-management (SM)

SM procedures promote an individual's awareness of their behavior and ability to function independently (Nelson et al., 1991). SM encompasses groups of interventions, including self-monitoring and self-evaluation (Callahan et al., 1998; Nelson et al., 1991). An individual can accomplish behavior change through self-monitoring by systematically observing their behavior and recording occurrences of a target behavior (Dunlap et al., 1991). Researchers have effectively used self-monitoring to both increase desirable and decrease undesirable behaviors among children and adults (Hager, 2012). For example, researchers have successfully increased

student on-task behavior (Gulchak, 2008) and decreased disruptive classroom behavior (Webber, 1993) using self-monitoring. In addition, teachers have effectively used self-monitoring procedures in the classroom to change their own behavior (Hager, 2012). To illustrate, Simonsen et al. (2013) explored the effectiveness of three self-monitoring strategies on five teachers' use of behavior-specific praise (BSP). Teachers' rates of BSP were higher during all self-monitoring conditions compared to baseline. Additional studies utilizing self-monitoring interventions also resulted in increased rates of BSP by teachers (Kalis et al., 2007; Markelz et al., 2019).

Another frequently studied and empirically validated SM strategy is self-evaluation. Self-evaluation involves an individual comparing their own performance to a predetermined performance criterion (Spates & Kanfer, 1977). Spates and Kanfer (1977) suggested that self-evaluation was potentially the most important but least researched component of SM. Self-evaluation reliably improves students' classroom and academic behavior (Nelson et al., 1995; Sainato et al., 1992); however, research examining the use of self-evaluation interventions to change teacher behavior is lacking. One exception is a study by Keller et al. (2005) that investigated the effects of an audiotape self-evaluation intervention on three pre-service teachers' use of social BSP. The audiotape self-evaluation intervention increased all three teachers' use of praise. Another study by Sutherland and Wehby (2001) examined the effect of an audiotape self-evaluation intervention on 20 teachers' rates of general praise. A positive, short-term effect was observed after intervention implementation. In addition, the rates of teacher reprimands decreased during intervention and follow-up, and the number of correct academic responses from students increased. Although results of studies utilizing self-evaluation interventions to change teacher behavior are promising (Keller et al., 2005; Sutherland & Wehby, 2001), Keller et al. (2005) is the only study to date known to utilize single-case design methodology to evaluate the

effect of self-evaluation on teachers' use of praise; thus, further research is warranted to demonstrate the efficacy of self-evaluation as a means to change teacher behavior (Sutherland & Wehby, 2001).

Effective Classroom and Behavior Management

Classroom and behavior management is a critical teacher preparation skill (Emmer & Stough, 2001), as poor classroom and behavior management can lead to teacher job dissatisfaction, teacher burnout (Otero-López et al., 2010), off-task student behavior, and lower academic achievement for students in the classroom (Owens et al., 2017). Despite its importance, teachers often cite classroom and behavior management as a significant challenge (Reinke et al., 2013). In a review of the evidence-based classroom and behavior management literature, Simonsen et al. (2008) grouped practices into five critical features of effective classroom and behavior management, including (a) maximize structure; (b) post, teach, review, monitor, and reinforce expectations; (c) actively engage students in observable ways; (d) use a continuum of strategies for responding to appropriate behaviors; and (e) use a continuum of strategies to respond to inappropriate behaviors. Two practices included in Simonsen et al.'s review were opportunities to respond (OTRs) and specific/contingent praise.

Researchers define an OTR as a teacher behavior that prompts a response from a student or groups of students (Simonsen et al., 2008; Sayeski et al., 2019). Giving students frequent opportunities to actively respond to academic requests (i.e., questions, tasks, demands) is an effective teaching practice and can decrease disruptive behavior and increase academic achievement (Macsuga-Gage & Simonsen, 2015; Sutherland & Wehby, 2001). Simonsen et al. (2008) categorized rates of OTRs as a classroom and behavior management feature that actively engages students in an observable way. When students are actively engaged in the classroom, it

is difficult to engage in incompatible problem behaviors, such as off-task or disruptive behavior (Greenwood et al., 2002).

Equally important to OTRs is specific/contingent praise provided by the teacher, which informs the student about what they did well (specific) and occurs directly after the target behavior (contingent; Simonsen et al., 2008). Providing specific/contingent praise is a simple strategy for teachers and has a strong evidence base (Brophy, 1981; Markelz & Taylor, 2016; Simonsen et al., 2008). According to Simonsen et al. (2008), specific/contingent praise is a way for teachers to respond to appropriate behavior. It is part of a continuum of strategies that focuses on identifying and recognizing appropriate classroom behavior (Simonsen et al., 2008). High rates of both OTRs and specific/contingent teacher praise can result in more effective instruction and increased student achievement (Sutherland, 2000; Sutherland & Wehby, 2001). Students benefit from these two classroom and behavior management strategies (Cavanaugh, 2013), and teachers can increase both OTRs and specific/contingent praise using SM procedures like self-monitoring (Haydon, 2012; Hager, 2012) and self-evaluation (Haydon, 2012; Sutherland & Wehby, 2001).

SM for Classroom and Behavior Management

Researchers have incorporated SM into practices to support teachers' overall classroom and behavior management skills (Simonsen et al., 2020; Keller et al., 2005; Sutherland & Wehby, 2001). Rispoli et al. (2017) conducted a systematic review and summarized the research on teacher SM of behavioral interventions. Although Rispoli et al. utilized the term self-monitoring throughout the review to describe the studies' procedures, both self-monitoring and self-evaluation studies were included in the review. The most common teacher target behavior (*n* = 10) was contingent praise of student behavior. Positive teacher outcomes were reported in 10

studies, and no studies reported negative outcomes (7 studies reported mixed results; Rispoli et al., 2017). Positive student outcomes, including a reduction of challenging behaviors and an increase in task engagement and academic readiness, were reported in half of studies that collected student outcome data. Rispoli et al. concluded that teacher SM is an effective means of measuring and improving teachers' fidelity of classroom and behavior management strategies.

More recently, Layden et al. (2023) assessed the quality of the teacher self-monitoring literature using the Council for Exceptional Children's Quality Indicators. Although several studies failed to meet quality indicators in one or more areas, eight studies met all quality indicators; thus, Layden et al. concluded that teachers and administrators should consider self-monitoring as a complementary intervention to classroom and behavior management strategies. Another systematic review by Scheibel et al. (2023) examined the SM literature to investigate the conditions under which teacher-directed self-monitoring improves implementation fidelity. Scheibel et al. used the Single Case Analysis and Review Framework 2.0 (SCARF; Ledford et al., 2018) to assess the quality and rigor of single-case designs included in 10 studies. Results suggested that most designs (75%) included in the Scheibel et al. review were of sufficient quality and rigor and demonstrated that the use of teacher-directed self-monitoring improved implementation fidelity across participants.

Despite several comprehensive reviews of the teacher SM literature (Layden et al., 2023; Rispoli et al., 2017; Scheibel et al., 2023), the current study is the only one of its kind. Rispoli et al. (2017) and Layden et al. (2023) included both single-case and group designs in their reviews and did not assess for the quality and rigor of single-case designs using the SCARF. The dependent variable in Scheibel et al. (2023) was implementation fidelity, and the review did not focus on specific classroom and behavior management strategies. Researchers have not

conducted an updated review of the teacher SM literature focusing on single-case design methodology, which is important to evaluate further if SM is an effective procedure to increase teachers' use of classroom and behavior management strategies (Rispoli et al., 2017).

Purpose

OTRs and specific/contingent praise are effective classroom and behavior management strategies (Simonsen et al., 2008), and high rates of both lead to positive student outcomes (Sutherland & Wehby, 2001). However, it is less clear exactly how teachers can increase their use of classroom and behavior management strategies in a feasible and accurate manner. The purpose of the current study was to systematically review single-case design literature on the effects of SM interventions on teachers' rates of OTRs and/or specific/contingent praise and update Rispoli et al. (2017). There are no systematic reviews that evaluate the quality and rigor of single-case design studies using SM interventions to increase teachers' rates of OTRs and specific/contingent praise in the current literature. Thus, the current study fills this gap in the literature and provides information regarding the effectiveness of SM interventions on two effective classroom and behavior management strategies (i.e., OTRs and specific/contingent praise) and under which conditions SM interventions are effective for these two variables. The current study will also compare the effect of self-monitoring and self-evaluation independently on rates of OTRs and specific/contingent praise.

Method

Search Procedures

We utilized the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 Guidelines (Page et al., 2021) to screen records. A PRISMA flow diagram detailing search procedures is available in Figure 1. We conducted a search of electronic

databases in education and psychology, including APA PsycArticles, Education Research Complete, ERIC, JSTOR, and PubMed, to identify existing SM studies that met the criteria for inclusion in the current review. Search terms included "*self-manag* OR self-monitor* OR self-evaluat* AND contingent praise AND opportunities to respond*" and all articles published prior to July 2024 were screened for inclusion. The search was limited to peer-reviewed articles written in English. The database search yielded a total of 3,223 records. We conducted a forward and backward search of articles that met inclusion criteria by reviewing sources that were cited in and cited by the included articles (n = 7). Studies were also identified by reviewing citations in the reference list of Rispoli et al. (2017) in addition to the database search (n = 0). Three duplicate articles were removed, for a total of 3,227 records screened.

Inclusion and Exclusion Criteria

We reviewed journal articles identified within the search (n = 3,227) by title and abstract to determine fit with the inclusion criteria (Page et al., 2021). These efforts resulted in 30 studies for full-text screening. Studies were included if they (a) included human participants; (b) were written in English; (c) were peer-reviewed; (d) utilized experimental single-case design procedures; and (e) evaluated the effects of a SM intervention (i.e., self-monitoring or self-evaluation) on teachers' rates of OTRs and/or specific/contingent praise. Studies were excluded if they utilized group design procedures (n = 5) or did not include teachers (i.e., pre-service or certified) or paraprofessionals as participants (n = 2). Research syntheses (e.g., systematic reviews and meta-analyses), books, and non-peer-reviewed primary studies were excluded from the search (n = 3). The systematic search process identified 21 studies for inclusion.

Coding Procedures

We coded participant descriptions and instructional context separately for each participant. For each study, the intervention and design characteristics and dependent variable(s) were coded. We coded each single-case design using the Single Case Analysis and Review Framework 2.0 (SCARF; Ledford et al., 2018). That is, if two ABAB designs were included in a single study, each single-case design was coded independently and referred to as two separate studies in the SCARF. Similarly, multiple-baseline-across-settings designs with more than one participant were evaluated as separate studies, but multiple-baseline-across-participants designs were considered one study. See Table 1 for variations in criteria necessary for a code of "Yes" in the present SCARF.

Participant Descriptions

We coded the following information for each participant: age, gender, race/ethnicity, number of school years teaching, current classroom placement (i.e., grade, number of students, general/special education), and reason for inclusion in the study (e.g., first-year teacher). If this information was not reported, it was coded as "No." The total number of participants in each study was also recorded.

Instructional Context

We coded where sessions took place (i.e., classroom, another setting) and who collected data (i.e., teacher, researcher, another person). For materials, we coded whether the SM intervention utilized tally (i.e., paper and pencil), a hand-held counter, or video-recording technology to collect data on teachers' rates of OTRs and specific/contingent praise.

Intervention Characteristics

For each study, we coded what kind of SM intervention was used (i.e., self-monitoring or self-evaluation). SM interventions were differentiated by the presence of a comparison of teacher

performance to a predetermined performance criterion. If the participant systematically observed their behavior and recorded the occurrence or nonoccurrence of a target behavior, it was considered self-monitoring. However, if the participant systematically observed their behavior and recorded occurrence or nonoccurrence, in addition to comparing their performance to a predetermined performance criterion, it was considered self-evaluation. The inclusion of additional intervention components, such as goal-setting and performance feedback, was also recorded.

Design Characteristics

We coded the single-case design used (e.g., withdrawal, multiple-baseline-across-participants) in each study and determined if the study met the basic What Works Clearinghouse (WWC) standards for single-case designs (i.e., made publicly available, released within 20 years of our review, used eligible populations, examined eligible interventions, showed eligible outcomes). Each single-case design was evaluated against general and design-specific WWC standards (What Works Clearinghouse, 2022). Standards included the presence of outcome measures (i.e., validity, reliability), data provided in graphical or tabular format, systematic manipulation of the independent variable, no residual treatment effects, design-specific assessment, and limited risk of bias. Using these standards, we determined whether each single-case design (a) met without reservations; (b) met with reservations; or (c) did not meet WWC standards (What Works Clearinghouse, 2022). One study (Szykula & Hector, 1978) was not evaluated with WWC standards because it is dated more than 20 years prior to the current review.

Dependent Variables

The current study provides information on the effectiveness of SM interventions on two dependent variables, OTRs and specific/contingent praise. OTRs were defined as any teacher behavior that prompts a response from a student or groups of students (Simonsen et al., 2008; Sayeski et al., 2019). Specific/contingent praise was defined as praise that addresses positive behaviors demonstrated by the student and occurs directly after the target behavior (Simonsen et al., 2008). Teacher-provided OTRs and specific/contingent praise were primarily measured using frequency (count) or rate (total number of occurrences divided by total observation time in minutes) and were coded accordingly. Additionally, we coded whether studies measured student behavior change (e.g., on-task behavior, academic achievement) as a secondary variable in addition to teacher OTRs and/or specific/contingent praise.

Quality and Rigor

We utilized SCARF to assess the quality and rigor of each single-case design by evaluating (a) rigor and (b) quality and breadth of measurement (QBM; Ledford et al., 2018). Using SCARF allowed for us to evaluate rigor by examining areas including dependent variable reliability (i.e., interobserver agreement; IOA), the procedures utilized to ensure and report implementation fidelity, and the sufficiency of the data. We evaluated dependent variable reliability by assessing for the presence of IOA data, as well as how often IOA was collected, how it was calculated, if agreement between data collectors was above 80% for each condition, and if the researchers reported that data collectors were naïve to the study's purpose. The area of implementation fidelity contained questions regarding (a) whether authors reported fidelity for intervention implementation; (b) how fidelity data were collected (e.g., observation) and how often it was collected (i.e., at least 20% of sessions); and (c) if data collectors calculated agreement on fidelity assessments. The last area included in the assessment of rigor was the

sufficiency of the data for each study. We coded each design on the following sufficiency criteria: at least three data points in each condition, at least three potential demonstrations of effect, type of single-case design (i.e., multiple-baseline-across-participants, ABAB), and absence of threats to internal validity. These three components (i.e., dependent variable reliability, fidelity, sufficiency of data) are the most highly weighted SCARF components.

We assessed the QBM of each design by evaluating participant descriptions, dependent variable descriptions, condition descriptions, and the presence of social and ecological validity indicators (Ledford et al., 2018). Participant descriptions were coded "Yes" if studies reported (a) participant age and gender; (b) years of teaching experience; (c) grade, classroom type (i.e., general/special education), and number of students; and (d) knowledge of self-management skills. Adequate condition descriptions meant other researchers could easily replicate procedures; the setting, a description of intervention implementers, and the frequency/duration of sessions were reported. Lastly, we answered questions about the presence of feasibility/acceptability ratings in each study, how social validity was measured (e.g., questionnaires, interviews), and the psychometric properties of social validity measures.

We answered questions regarding rigor and QBM components listed above by coding "Yes" or "No" for each single-case design. Researcher codes from rigor and QBM components resulted in a populated score ranging from 0-4, with 0 indicating very low quality/rigor and 4 indicating very high quality/rigor. Once we entered all of the necessary codes into the SCARF tool, we obtained populated scores using the following equation: [2x (average rigor score) + (average QBM score)]/3. Scores of 2 and above generally indicate high quality and rigor (Ledford et al., 2018).

Study Outcomes

We coded primary outcomes from each single-case design with a score ranging from 0-4 using the SCARF (Ledford et al., 2018) and used visual analysis to evaluate each design for evidence of a therapeutic, contratherapeutic, or no effect. The following scores were assigned based upon the effect:

- Score of 0: non-effects or contratherapeutic effects
- Score of 1: two demonstrations of effect and one non-effect
- Score of 2: three or more demonstrations of effect, with no more than one non-effect
- Score of 3: three or more demonstrations of effect, with no non-effects and no more than one weak effect
- Score of 4: three or more demonstrations of effect, with no non-effects or weak effects

Scores of 3 and 4 are consistent with the potential for a functional relation (Ledford et al., 2018). See Table 2 for design-specific definitions with corresponding scores. Graphs were automatically produced for primary outcomes, generalization, and maintenance according to scores assigned by researchers. The SCARF scatterplot included each single-case design with the researcher assigned primary outcome score on the y-axis and the overall quality and rigor score on the x-axis. The graph is divided into four quadrants that pertain to (a) high-quality evidence of positive effects (top right); (b) high-quality evidence of negative or minimal effects (bottom right); (c) low-quality evidence of positive effects (top left); and (d) low-quality evidence of negative or minimal effects (bottom left).

Generalization and maintenance outcomes were also coded, each using a 0-4 scale (Ledford et al., 2018). For generalization outcomes, a score of 0 indicated no measurement, 1 indicated consistent non-effects or contratherapeutic effects, 2 indicated mixed effects, 3

indicated positive effects that may include weak effects, and 4 indicated all strong, positive effects. Maintenance outcomes were evaluated on a scale from 0-4 in an identical fashion to generalization outcomes and pertained to the latency of maintenance measurement (Ledford et al., 2018).

Results

The review included 21 studies resulting in 30 single-case designs. Articles spanned the years 1978 to 2023. A total of 70 participants conducted SM interventions in classroom settings, including 52 certified teachers, 6 paraprofessionals, and 12 student teachers. The participants in these studies ranged in age from 21 to 50 years of age (M = 29.8 years old). Participants were 90% female and 10% male. Of 12 studies that included demographic information regarding race and/or ethnicity, 90.4% of participants identified as White or Caucasian, 2.4% identified as Black or African American, 2.4% identified as Hispanic, 2.4% identified as Asian American, and 2.4% identified as Native American.

Participants had an average of 5.2 years of teaching experience, though the range varied from 1 to 25 years across teachers. Specific classroom setting was not reported for 11 participants; when reported, 52.5% of participants taught general education, 38.9% of participants taught special education, 5% taught inclusion, and 3.6% taught non-academic instruction (i.e., chorus). Participants taught a range of grade levels from Pre-Kindergarten to 12th grade, with the most commonly reported grade levels being elementary school (K-5). The number of students in participants' classrooms ranged from 2-31 students (M = 18 students per classroom). Reasons for teacher participation in the studies included volunteering (n = 33), being recommended by school administration (n = 5), being a new teacher or having no prior training

in classroom and behavior management (n = 8), having difficulty with classroom and behavior management (n = 5), or requesting support from school administration (n = 8).

All participants implemented SM interventions in a classroom setting. The majority of teachers (88%) not only self-monitored their behavior but participated in data collection and recording as well. Teachers used various methods for SM including tally (n = 8), a hand-held counter (n = 8), a phone application (n = 3), or audio/video-recording technology (n = 2) to collect data on their rates of OTRs and specific/contingent praise. Most studies reported using self-monitoring (85.7%), while 4.7% used self-evaluation, and 9.6% used a combination of the two SM interventions. Goal-setting (52.4%) and performance feedback (42.9%) were often included in SM intervention packages. All studies measured teachers' use of specific/contingent praise. Only two studies (Hager, 2012; Sallese & Vannest, 2020) included OTRs as a dependent variable, though these studies also evaluated teachers' use of praise. Teachers' rates of OTRs and specific/contingent praise were primarily measured using frequency (count) or rate, with one study (Rivera et al., 2015) utilizing partial-interval recording. Additionally, 12 studies measured changes in student behavior as a result of teacher-implemented SM interventions. Student behaviors included engagement (n = 4), on-task behavior (n = 6), off-task behavior (n = 1), and disruptive behavior (n = 4). Two studies (Knochel et al., 2021; Myers et al., 2011) measured multiple student behaviors.

The studies included in this review utilized various single-case designs including multiple-baseline-across-participants (n = 11), multiple-baseline-across-behaviors (n = 1), multiple-baseline-across-settings (n = 2), multiple-probe-across-participants (n = 3), withdrawal (n = 2), changing criterion (n = 1), and alternating treatments (n = 1). According to WWC

standards, 37 single-case designs met without reservations, 16 met with reservations, and 6 did not meet (What Works Clearinghouse, 2022).

Analysis of Quality and Rigor

The average SCARF quality/rigor score across 30 designs in the 21 included studies was 2.81. A total of 27 designs were of adequate quality/rigor regardless of effect. Figure 2 displays the assessment of quality/rigor as generated by SCARF. The SCARF graphic display is separated into four quadrants. The top left quadrant represents studies with low quality/rigor and positive effects. The top right quadrant represents studies with high quality/rigor and positive effects. The lower quadrants represent studies with low or minimal effects with low quality/rigor (bottom left) and high quality/rigor (bottom right; Ledford et al., 2018). The majority of the data points (56.67%) are in the top right quadrant, representing studies with high quality evidence of positive effects. Additional data points are in the lower left and right quadrants, which represent studies with both low- and high-quality evidence of minimal effects. More specifically, 17 single-case designs received a primary outcome score of 3 (n = 6) or 4 (n = 11), as the potential for a functional relationship between the independent and primary dependent variables was identified. The remaining designs received a score of 1 (n = 11) and 0 (n = 2), representing designs with at least one non-effect or contratherapeutic effect. No design received a rating of 2 for primary outcomes.

The studies included in this review were rigorous in terms of reliability, as 100% of studies collected IOA. The majority of designs (86.6%) included data related to the fidelity of intervention implementation. All but one study (Keller et al., 2005) reported social validity data in order to describe the feasibility and acceptability of SM interventions from participants' perspectives. Few studies (n = 4) reported generalization outcomes (Hager, 2012; Justus et al.,

2023; Keller et al., 2005; Markelz et al., 2021). Of the four studies (see Figure 3), Keller et al. (2005) collected generalization data pre- and post-intervention, Hager (2012) and Justus et al. (2023) collected data intermittently throughout the intervention, and Markelz et al. (2021) measured generalization experimentally. Maintenance outcomes are displayed in Figure 4. Eleven studies (i.e., 14 single-case designs) reported collecting maintenance data post-intervention. Maintenance data were collected immediately following intervention (n = 4), at least one week post-intervention (n = 5), and one or more months post-intervention (n = 5). Overall, the evaluation of studies included in the SCARF suggests that the majority of studies assessing the effect of SM interventions on teachers' rates of OTRs and specific/contingent praise demonstrate positive effects and are of high quality.

Discussion

OTRs and specific/contingent praise are effective classroom and behavior management strategies that can be increased through SM interventions (Cavanaugh, 2013; Simonsen et al., 2008). This review evaluated the SM literature to determine for whom, in what settings, with what method, and for what reason SM interventions have been used. Further, this review provided information about the effect of SM interventions (i.e., self-monitoring and self-evaluation) on teachers' implementation of OTRs and specific/contingent praise. The single-case SM studies included in this review were of mostly sufficient quality/rigor, as the majority of single-case designs (n = 17) fall in the top right quadrant of the SCARF scatterplot. Similar to conclusions reported by Rispoli et al. (2017), the SCARF primary outcome data support teacher SM as an effective intervention to improve teachers' use of classroom and behavior management strategies.

Strengths of SM Research

This review of the literature identified a variety of strengths in the available research. Most notably, a variety of data collection methods have been utilized when implementing SM interventions to improve teachers' use of OTRs and specific/contingent praise. This finding suggests that self-monitoring and self-evaluation can successfully be implemented through a variety of means, from pencil and paper to video-recording technology. Teacher SM interventions have also been implemented across a variety of settings successfully, including across grade levels (Pre-Kindergarten-12th) and across general and special education classrooms with class sizes ranging from 2-31 students.

Furthermore, the characteristics of teachers for whom the SM interventions are seemingly effective are also widespread. Although being a new teacher was one reason for participant inclusion, there is evidence to suggest that SM interventions effectively improve teachers' classroom and behavior management even when teachers have an abundance of teaching experience. Snyder (2017) discusses potential reasons that veteran teachers (i.e., teachers with 20 or more years of experience) may be resistant to change. Reasons for resistance may include changes that negatively affect relationships and decrease teacher autonomy, as well as a long reinforcement history engaging in the same classroom and behavior management strategies and observing their potential effect (Snyder, 2017). The effectiveness of SM strategies with veteran teachers is likely related to the individualized nature of SM, as each participant's change in behavior is compared to their own baseline levels in order to determine the effect. Even if teachers frequently utilize OTRs or specific/contingent praise, their rate can always be improved relative to their current performance.

In addition to generalizability of teacher SM interventions, almost half (46.67%) of single-case designs evaluated in this review reported maintenance outcomes. For the purposes of

the SCARF, maintenance was defined as "evidence of continued behavior change post-intervention" (Ledford et al., 2018). Thirteen of the designs reported adequately positive effects, suggesting potential lasting effects of an SM intervention on teachers' rates of OTRs and specific/contingent praise. However, currently reviewed studies have not included maintenance data revealing if teachers continue to utilize SM strategies after intervention has concluded.

Identified Weaknesses

The current review also identified some weaknesses in the current literature. First, the SCARF resulted in the identification of six studies (Justus et al., 2023; Markelz et al., 2021; Myers et al., 2021; Niwayama et al., 2020; Simonsen et al., 2013; Wills et al., 2019) producing 10 single-case designs with low or minimal effects on teachers' rates of OTRs or specific/contingent praise despite high quality/rigor (i.e., lower right quadrant). When individual SCARF items were reviewed for similarities between these six studies, there were no apparent similarities in participant descriptions, instructional context, intervention characteristics, or study design. One potential explanation is that, although all six studies were considered rigorous and measured implementation fidelity, a majority (n = 5) did not report collecting fidelity data in at least 20% of sessions; thus, there is a possibility that the effects of the intervention were influenced by inconsistent implementation across conditions. OTRs and specific/contingent praise are considered evidence-based practices for classroom and behavior management; however, six rigorous studies produced low or minimal effects on teachers' behavior. Thus, more research may be needed displaying positive effects from high quality/rigorous studies to allow for a more convincing conclusion that SM can be effectively utilized to increase OTRs and specific/contingent praise. Second, studies included in the current review reported a fairly homogenous group of participants in regard to age, gender, and race/ethnicity. Participants

tended to identify as young adults, White, and female. Researchers should continue to evaluate the effects of SM interventions including a more diverse population of teacher participants. In addition, only four studies collected generalization data. Planning for generalization is an important future direction as it is difficult to predict the effectiveness of SM interventions outside of the classroom context, using different materials, and on different individuals' behavior without intentionally measuring it.

Limitations

Limitations of the current review should be considered when interpreting the findings. Interobserver agreement (IOA) was not collected on (a) the decision to include a study in this review or (b) the accuracy of coding of the SCARF. Furthermore, studies written in languages other than English were excluded, as well as dissertations, which has the potential to limit the diversity of studies reviewed as well as more recent, unpublished research. Finally, this review did not include group design studies. There were several group design studies identified during initial search procedures that examined the effect of SM interventions on teachers' rates of OTRs and specific/contingent praise that were removed.

Future Directions

A majority of studies (85.7%) in this review used self-monitoring, while only one study (4.7%; Keller et al., 2005) utilized self-evaluation alone to increase teachers' rates of specific/contingent praise. Future research should evaluate the effectiveness of self-evaluation to improve teachers' classroom and behavior management skills. Goal-setting and performance feedback were common additions to SM intervention packages utilizing self-monitoring and/or self-evaluation, which is consistent with the current SM literature (Moore et al., 2001). However, researchers should consider evaluating the effectiveness of SM interventions (i.e., self-

monitoring and self-evaluation) without goal-setting and performance feedback components in order to determine the SM interventions' effect on the dependent variable alone. Without experimental studies that evaluate the effect of SM interventions without additional package components (e.g., goal-setting and performance feedback), it is difficult to determine the effectiveness of SM to change teacher behavior as an independent intervention. Furthermore, teachers have increased access to technology in the classroom which can be used to effectively implement SM interventions and collect data (Hager, 2012). Although video-recording is a promising way to collect SM data, only one study (Hager, 2012) was identified in this review of the literature that employed video-recording technology to collect data on participants' rates of OTRs and specific/contingent praise. Additional research is warranted to determine the effect of video-recording SM interventions on teachers' rates of OTRs and specific/contingent praise.

Additional research is also needed to evaluate the effect of SM interventions on teachers' use of OTRs, as research in this area is lacking. Only two studies in the current review (Hager, 2012; Sallese & Vannest, 2020) included OTRs as a dependent variable, and these studies also evaluated the effect of SM on teachers' use of praise. No single-case studies to date have evaluated the effect of SM interventions on teachers' rates of OTRs alone. Further, a majority of the studies included in this review reported fidelity of implementation (86.6%). Although this is an important factor to measure to ensure that the intervention is implemented in the way it is intended, there is likely value in determining if teachers must implement SM strategies at a high-level of fidelity (i.e., 80% or above) in order to see improvements in their rates of OTRs and specific/contingent praise, as well as positive student outcomes. There is evidence to suggest that SM interventions bring awareness to individuals about their behavior (Nelson et al., 1991). Future research may evaluate whether a teacher's awareness of his or her behavior is enough to

create behavior change, even if their implementation fidelity is not at or above research standards.

Although maintenance data was collected by approximately half of the studies in this review, maintenance measurement focused on teachers' utilization of the SM strategies after intervention has concluded has not been evaluated. This is a potentially important endeavor as it may reveal information about the continued feasibility and acceptability of the SM intervention once the intervention phase is complete. Additionally, all but one study (95.23%) reported social validity data describing the feasibility of SM interventions for teachers to implement in their classrooms. In addition to teacher-rated feasibility and acceptability data, it would be interesting for future studies to examine the amount of teacher training that is required for them to implement SM interventions in the school setting. If teachers can quickly learn SM interventions and implement them with sufficient fidelity, this information would add to the feasibility of introducing SM interventions in school settings.

Finally, researchers should consider including student outcomes in future studies evaluating the effectiveness of SM interventions on teacher behavior. Twelve of the studies in this review included student outcomes as secondary variables; however, there is continued room for improvement for future studies to include this important outcome data. Though teachers are the target audience for intervention in these studies, research suggests SM interventions reliably improve students' classroom and academic behavior (Hager, 2012). Ultimately, improving student behavior and outcomes is the goal of targeting teachers' classroom and behavior management skills.

References

Briere, D. E., Simonsen, B., Sugai, G., & Myers, D. (2015). Increasing new teachers' specific praise using a within-school consultation intervention. *Journal of Positive Behavior Interventions*, *17*(1), 50-60.

Brophy, J. E. (1981a). On praising effectively. *The Elementary School Journal*, *81*(5), 269-278.

Cavanaugh, B. (2013). Performance feedback and teachers' use of praise and opportunities to respond: A review of the literature. *Education and Treatment of Children*, *36*(1), 111-136.

Callahan, K., Rademacher, J. A., & Hildreth, B. L. (1998). The effect of parent participation in strategies to improve the homework performance of students who are at risk. *Remedial and Special Education*, *19*(3), 131-141.

Dunlap, L. K., Dunlap, G., Koegel, L. K., & Koegel, R. L. (1991). Using self-monitoring to increase independence. *Teaching Exceptional Children*, *23*(3), 17-22.

Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist, 36*(2), 102-112.

Glynn, E. L., Thomas, J. D., & Shee, S. M. (1973). Behavioral self-control of on-task behavior in an elementary classroom. *Journal of Applied Behavior Analysis*, *6*(1), 105-113.

Greenwood, C. R., Horton, B. T., Utley, C. A. (2002). Academic engagement: Current perspectives in research and practice. *School Psychology Review, 31,* 328-349.

Gulchak, D. J. (2008). Using a mobile handheld computer to teach a student with an emotional and behavioral disorder to self-monitor attention. *Education and Treatment of Children*, *31*(4), 567-581.

Hager, K. D. (2012). Self-monitoring as a strategy to increase student teachers' use of effective teaching practices. *Rural Special Education Quarterly*, *31*(4), 9-17.

Haydon, T., Macsuga-Gage, A. S., Simonsen, B., & Hawkins, R. (2012). Opportunities to respond: A key component of effective instruction. *Beyond Behavior*, *22*(1), 23-31.

Justus, J., Hott, B. L., & Heiniger, S. (2023). Using self-monitoring to increase behavior specific praise in elementary classrooms. *Behavior Analysis in Practice*, *16*(3), 885-891.

Kalis, T. M., Vannest, K. J., & Parker, R. (2007). Praise counts: Using self-monitoring to increase effective teaching practices. *Preventing School Failure: Alternative Education for Children and Youth*, *51*(3), 20-27.

Keller, C. L., Brady, M. P., & Taylor, R. L. (2005). Using self-evaluation to improve student teacher interns' use of specific praise. *Education and Training in Developmental Disabilities*, 368-376.

Knochel, A. E., Blair, K. S. C., & Sofarelli, R. (2021). Culturally focused classroom staff training to increase praise for students with autism spectrum disorder in Ghana. *Journal of Positive Behavior Interventions*, *23*(2), 106-117.

Knochel, A. E., Blair, K. S. C., Kincaid, D., & Randazzo, A. (2022). Promoting equity in teachers' use of behavior-specific praise with self-monitoring and performance feedback. *Journal of Positive Behavior Interventions*, *24*(1), 17-31.

Koegel, R. L., & Koegel, L. K. (1990). Extended reductions in stereotypic behavior of students with autism through a self-management treatment package. *Journal of Applied Behavior Analysis*, *23*(1), 119-127.

Layden, S. J., Crowson, T. G., & Hayden, K. E. (2023). A 30-year systematic review of self-monitoring as a strategy to improve teacher performance. *The Journal of Special Education*, *57*(1), 47-58.

Ledford, J. R., Lane, J. D., & Tate, R. (2018). Evaluating quality and rigor in single case research. *Single Case Research Methodology* (pp. 365-392). Routledge.

MacSuga-Gage, A. S., & Simonsen, B. (2015). Examining the effects of teacher-directed opportunities to respond on student outcomes: A systematic review of the literature. *Education and Treatment of Children*, *38*(2), 211-239.

Malott, R. W. (1989). The achievement of evasive goals. *Rule-governed behavior: Cognition, contingencies, and instructional control.* (pp. 269-322). Boston, MA: Springer.

Markelz, A. M., & Taylor, J. C. (2016). Effects of teacher praise on attending behaviors and academic achievement of students with emotional and behavioral disabilities. *Journal of Special Education Apprenticeship*, *5*(1), n1.

Markelz, A. M., Taylor, J. C., Kitchen, T., Riccomini, P. J., Catherine Scheeler, M., & McNaughton, D. B. (2019). Effects of tactile prompting and self-monitoring on teachers' use of behavior-specific praise. *Exceptional Children*, *85*(4), 471-489.

McDougall, D. (1990). *Use of behavioral self-control to promote academic responding of exceptional students in regular education classrooms: Performance and generalization effects*. University of Houston.

Moore, D. W., Prebble, S., Robertson, J., Waetford, R., & Anderson, A. (2001). Self-recording with goal setting: A self-management programme for the classroom. *Educational Psychology*, *21*(3), 255-265.

Myers, D. M., Simonsen, B., & Sugai, G. (2011). Increasing teachers' use of praise with a response-to-intervention approach. *Education and treatment of children*, *34*(1), 35-59.

Nelson, J. R., Smith, D. J., Young, R. K., & Dodd, J. M. (1991). A review of self-management outcome research conducted with students who exhibit behavioral disorders. *Behavioral Disorders*, *16*(3), 169-179.

Nelson, J. R., Smith, D. J., & Colvin, G. (1995). The effects of a peer-mediated self-evaluation procedure on the recess behavior of students with behavior problems. *Remedial and Special Education*, *16*(2), 117-126.

Niwayama, K., Maeda, Y., Kaneyama, Y., & Sato, H. (2020). Increasing teachers' behavior-specific praise using self-monitoring and a peer teacher's feedback: The effect on children's academic engagement. *Preventing School Failure: Alternative Education for Children and Youth*, *64*(4), 271-280.

Owens, J. S., Coles, E. K., Evans, S. W., Himawan, L. K., Girio-Herrera, E., Holdaway, A. S., ... & Schulte, A. C. (2017). Using multi-component consultation to increase the integrity with which teachers implement behavioral classroom interventions: A pilot study. *School Mental Health*, *9*(3), 218-234.

Otero-López, J. M., Bolaño, C. C., Mariño, M. J. S., & Pol, E. V. (2010). Exploring stress, burnout, and job dissatisfaction in secondary school teachers. *International Journal of Psychology and Psychological Therapy*, *10*(1), 107-123.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*(1), 1-11.

Pinter, E. B., East, A., & Thrush, N. (2015). Effects of a video-feedback intervention on teachers' use of praise. *Education and Treatment of Children*, *38*(4), 451-472.

Reinke, W. M., Lewis-Palmer, T., & Merrell, K. (2008). The classroom check-up: A class wide teacher consultation model for increasing praise and decreasing disruptive behavior. *School Psychology Review*, *37*(3), 315-332.

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*(1), 39-50.

Rispoli, M., Zaini, S., Mason, R., Brodhead, M., Burke, M. D., & Gregori, E. (2017). A systematic review of teacher self-monitoring on implementation of behavioral practices. *Teaching and Teacher Education*, *63*, 58-72.

Rivera, C. J., Mason, L. L., Jabeen, I., & Johnson, J. (2015). Increasing teacher praise and on task behavior for students with autism using mobile technology. *Journal of Special Education Technology*, *30*(2), 101-111. https://doi.org/10.1177/0162643415617375

Sainato, D. M., Goldstein, H., & Strain, P. S. (1992). Effects of self-evaluation on preschool children's use of social interaction strategies with their classmates with autism. *Journal of Applied Behavior Analysis*, *25*(1), 127-141.

Sallese, M. R., & Vanneste, K. J. (2020). The effects of a multicomponent self-monitoring intervention on the rates of pre-service teacher behavior-specific praise in a masked

single-case experimental design. *Journal of Positive Behavior Interventions*, *22*(4), 207-219.

Sayeski, K. L., Hamilton-Jones, B., Cutler, G., Earle, G. A., & Husney, L. (2019). The role of practice and feedback for developing teacher candidate's opportunities to respond expertise. *Teacher Education and Special Education*, *42*(1), 18-35.

Scheibel, G., Chen, P. Y., Zaeske, L. M., Wills, H. P., & Zimmerman, K. N. (2023). Improving Implementation Fidelity with Teacher-Directed Self-Monitoring Interventions: A Systematic Review. *Journal of Positive Behavior Interventions*, *25*(4), 253-269.

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*, 351-380.

Simonsen, B., MacSuga, A. S., Fallon, L. M., & Sugai, G. (2013). The effects of self-monitoring on teachers' use of specific praise. *Journal of Positive Behavior Interventions*, *15*(1), 5-15.

Simonsen, B., Freeman, J., Dooley, K., Maddock, E., Kern, L., & Myers, D. (2017). Effects of targeted professional development on teachers' specific praise rates. *Journal of Positive Behavior Interventions*, *19*(1), 37-47.

Simonsen, B., Freeman, J., Myers, D., Dooley, K., Maddock, E., Kern, L., & Byun, S. (2020). The effects of targeted professional development on teachers' use of empirically supported classroom management practices. *Journal of Positive Behavior Interventions*, *22*(1), 3-14.

Snyder, R. R. 2017. Resistance to change among veteran teachers: Providing voice for more effective engagement. *International Journal of Educational Leadership Preparation*, *12*(1), n1.

Spates, C. R., & Kanfer, F. H. (1977). Self-monitoring, self-evaluation, and self-reinforcement in children's learning: A test of a multistage self-regulation model. *Behavior Therapy*, *8*(1), 9-16.

Sutherland, K. S. (2000). *Effects of self-evaluation on rates of teacher behaviors in classrooms for students with emotional and behavioral disorders*. Vanderbilt University.

Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *The Journal of Special Education*, *35*(3), 161-171.

Szykula, S. A., & Hector, M. A. (1978). Teacher instructional behavior change through self-control. *Psychology in the Schools*, *15*(1), 87-94.

Thompson, M. T., Marchant, M., Anderson, D., Prater, M. A., & Gibb, G. (2012). Effects of tiered training on general educators' use of specific praise. *Education and Treatment of Children*, *35*(4), 521-546.

VanLone, J., Freeman, J., Simonsen, B., Everett, S., Sugai, G., & Whitcomb, S. (2022). The effects of a video self-analysis package on pre-service teachers' use of behavior-specific praise. *Journal of Special Education Apprenticeship*, *11*(1), n1.

Webber, J., Scheuermann, B., McCall, C., & Coleman, M. (1993). Research on self-monitoring as a behavior management technique in special education classrooms: A descriptive review. *Remedial and Special Education*, *14*(2), 38-56.

What Works Clearinghouse. (2022). *What Works Clearinghouse Procedures and Standards Handbook* (Version 5.0). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Wills, H. P., Mason, R., Gregori, E., & Veatch, M. (2019). Effects of self-monitoring on the praise rates of paraprofessionals for students with emotional and behavioral disorders. *The Elementary School Journal*, *119*(4), 562-579.

Table 1

Variations in SCARF Criteria

Section	Items	Interpretation
Participant Descriptions	Demographic Information (P1)	Studies that included information about participant age and gender were coded as YES.
	Formal Test Results (P2)	Studies that included information about number of schoolyears participants have been teaching were coded as YES.
	General Information (P3)	Studies that included information about participant's classroom placement, including grade, classroom type, and # of students were coded as YES.
	Inclusion Criteria (P4)	Studies that included any information about prerequisite skills necessary for self-management were coded as YES.
Condition Descriptions	Descriptions of Comparison Conditions (C1)	Studies that included sufficient information for researchers to replicate baseline and intervention condition(s) were coded as YES.
	Dosage (C2)	Studies that included information about the number of sessions and how long they lasted were coded as YES.
Stimulus Generalization	Different Context (SG1)	Studies that included information about the use of multiple classrooms or different times of day were coded as YES.
	Different Materials (SG2)	N/A
	Social Partners (SG3)	Studies that included information about the use of multiple teachers or school staff were coded as YES.

Table 2

Specific Score Definitions for Single-Case Designs

Design	Score	Definition
Alternating Treatments Design	0	Data paths undifferentiated; approximately half or more of data paths are overlapping (i.e., approximately the same values or with some higher values in one condition and some higher values in another condition).
	1	Approximately half or more data are overlapping as described for a score of 0 but overlap decreases over time.
	2	Less than half of data points are overlapping but there is a decreasing or variable differentiation between conditions (i.e., difference in values between conditions decreases over time or is not consistent).
	3	Less than half of data points are overlapping and there is increasing differentiation over time (i.e., difference in values between conditions increases over time).
	4	Minimal/no overlap occurs; consistent differentiation between conditions.
Multiple Baseline/Probe Design	0	>1 non-effect or any contratherapeutic effect; if vertical analysis suggests change in data in one tier is associated with condition change in another tier.
	1	<3 demonstrations of effect, 1 non-effect.
	2	>=3 demonstrations, >=1 non-effect.
	3	>=3 demonstrations, >=1 weak effects, 0 non-effects.
	4	>=3 demonstrations, 0 weak effects/non-effects.
Other Designs	0	>1 non-effect or any contratherapeutic effect.
	1	<3 demonstrations of effect, 1 non-effect.
	2	>=3 demonstrations, >=1 non-effect.
	3	>=3 demonstrations, >=1 weak effects, 0 non-effects.
	4	>=3 demonstrations, 0 weak effects/non-effects.

Table 3

Participant Descriptions by Study

Study	Age	Gender	Race/Ethnicity	Experience	Type	Grade	Students
Briere et al. (2015)	-	Female	-	1	-	1st	20
	-	Female	-	2	-	Pre-K	20
	-	Female	-	1	-	3rd	20
Hager (2012)	23	Female	Caucasian	1	SE	K-5th	-
Justus et al. (2023)	47	Female	White	8	GE	4th	-
	26	Female	White	4	GE	3rd	-
	24	Female	White	2	GE	1st	-
Kalis et al. (2007)	24	Female	-	1	SE	9th-11th	5
Keller et al. (2005)	24-36	Female	European Am.	1	SE	K-5th	-
	24-36	Female	European Am.	1	SE	K-5th	-
	24-36	Female	European Am.	1	SE	K-5th	-
Knochel et al. (2021)	22	Female	-	3	-	-	30
	28	Male	-	2	-	-	30
	31	Male	-	4	-	-	30
	24	Male	-	3	-	-	30
Knochel et al. (2022)	-	Female	White	20	GE	5th	20
	-	Female	White	1	GE	4th	22
	-	Female	White	16	GE	4th	23
	-	Female	White	14	GE	1st	21
Markelz et al. (2019)	36	Female	-	11	SE	2nd-3rd	8
	28	Female	-	3	SE	4th-5th	6
	26	Female	-	3	SE	3rd-5th	13
Markelz et al. (2021)	29	Female	White	2	-	Pre-K	-
	23	Female	White	2	-	Pre-K	-
	26	Female	African Am.	7	-	Pre-K	-
Myers et al. (2011)	-	Female	White	4	SE	7th	-
	-	Female	White	11	Incl.	5th	-
	-	Female	White	8	GE	6th	-
	-	Female	White	1	GE	7th	-
Niwayama et al. (2020)	-	Male	-	5	GE	2nd	27
	-	Male	-	4	GE	1st	30
Pinter et al. (2015)	-	Female	Caucasian	13	SE	9th-12th	12
	-	Female	Caucasian	4	SE	9th-12th	10
	-	Female	Caucasian	2	SE	7th-8th	13
	-	Male	Caucasian	8	SE	9th-12th	7
Reinke et al. (2008)	-	Female	Caucasian	25	GE	1st	-
	-	Female	Caucasian	14	GE	2nd	-
	-	Female	Caucasian	13	GE	2nd	-
	-	Female	Caucasian	5	GE	5th	-
Rivera et al. (2015)	-	Female	-	10	SE	5th	-
	-	Female	-	10	SE	5th	-

Table 3 (Continued)

	-	Female	-	10	SE	5th	-
	-	Female	-	10	SE	K	2
Sallese & Vannest (2020)	22-24	Female	Caucasian	1	GE	3rd-5th	23-27
	22-24	Female	Caucasian	1	GE	3rd-6th	23-27
	22-24	Female	Caucasian	1	GE	3rd-7th	23-27
	22-24	Female	Caucasian	1	GE	3rd-8th	23-27
Simonsen et al. (2013)	-	Female	-	-	GE	8th	-
	-	Female	-	-	GE	5th	-
	-	Female	-	-	SE	5th-8th	-
	-	Female	-	-	GE	7th	-
	-	Female	-	-	GE	5th	-
Simonsen et al. (2017)	-	Female	-	15	GE	K	19
	-	Female	-	2	GE	3rd	17
	-	Female	-	13	Incl.	3rd	17
	-	Female	-	7	GE	3rd	20
	-	Female	-	5	GE	K	20
	-	Female	-	11	GE	3rd	20
Szykula & Hector (1978)	-	Female	-	-	-	1st	-
Thompson et al. (2012)	40-50	Female	White	11	GE	4th	31
	40-50	Female	White	13	GE	4th	26
	40-50	Female	White	1	GE	3rd	26
VanLone et al. (2022)	21	Female	Asian Am.	1	Incl.	2nd	-
	21	Female	Caucasian	1	NA	6th-8th	-
	22	Male	Caucasian	1	NA	9th-12th	-
	22	Female	Caucasian	1	SE	4th	-
Wills et al. (2019)	-	Female	Native Am.	18	SE	K-5th	4-6
	-	Female	Hispanic	1	SE	K-5th	4-6
	-	Female	White	1	SE	K-5th	4-6
	-	Female	White	1	SE	K-5th	4-6

Note. Experience = number of years teaching; Type = classroom designation (GE = general education, SE = special education, Incl. = Inclusion, NA = non-academic); Students = number of students in the participant's class

Figure 1

PRISMA Flow Diagram

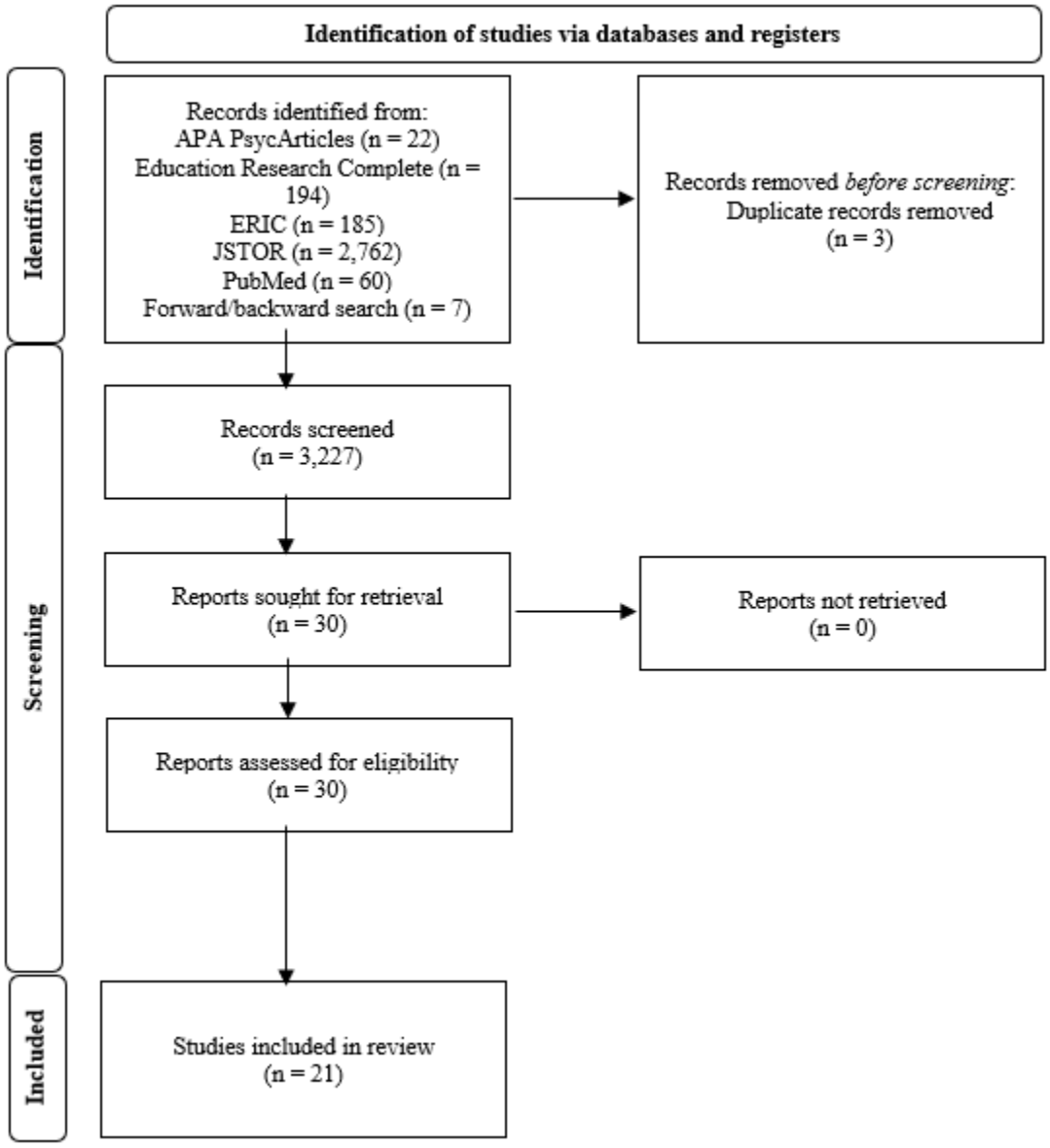


Figure 2

SCARF Scatterplot

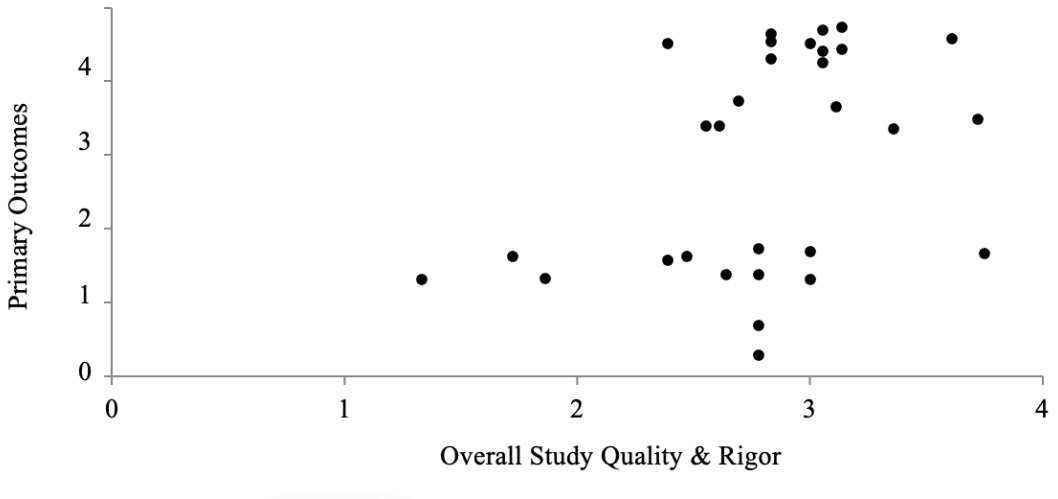


Figure 3

SCARF Generalization Scatterplot

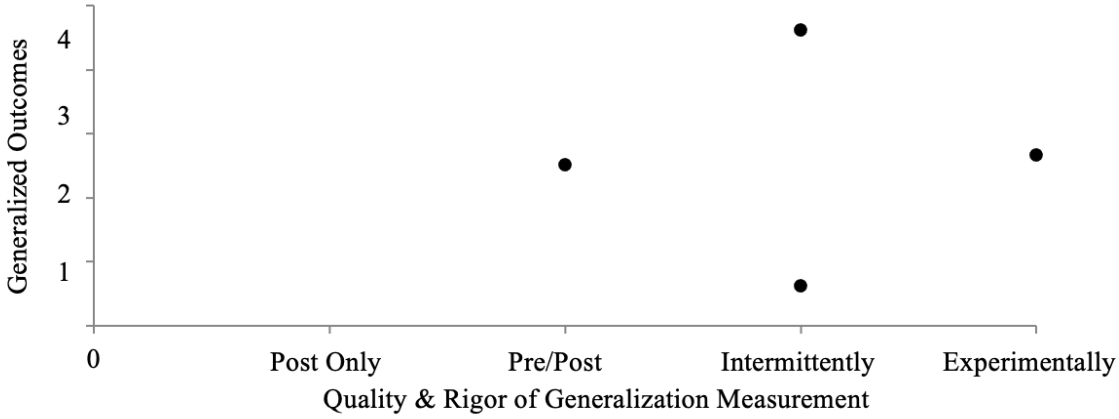


Figure 4

SCARF Maintenance Scatterplot

CHAPTER 3

EFFECTS OF A VIDEO SELF-EVALUATION INTERVENTION ON PRAISE STATEMENTS

Introduction

Teachers consider classroom and behavior management to be one of the most challenging aspects of their job (Reinke et al., 2011). Despite the challenge behavior management poses for teachers, classroom and behavior management is an area in which teachers report receiving little to no training (Reinke et al., 2011). Disruptive behavior resulting from poorly managed classrooms can lead to teacher stress (McCarthy et al., 2015), low self-efficacy (O'Brennan et al., 2017), and burnout (Otero-López et al., 2010), and are a primary reason teachers choose to leave the profession (Ingersoll & Smith, 2003). Furthermore, poor classroom and behavior management is associated with negative outcomes for students, such as exposure to less academic instruction (Weinstein, 2007), greater time off-task (Owens et al., 2017), lower academic achievement (Owens et al., 2017), and negative behavioral and social outcomes (National Research Council, 2002). Although professional development is often employed as a means of improving teachers' classroom and behavior management skills, it is generally ineffective in changing teachers' behavior in the absence of hands-on practice or performance feedback and goal-setting (Bruce, 2010; Chappuis et al., 2020). Thus, the current study examined a video self-evaluation intervention designed to improve teachers' classroom and behavior management skills and, more specifically, teachers' use of behavior-specific praise (BSP).

Teacher Praise

Teacher praise is a valuable and effective proactive classroom and behavior management strategy (Floress et al., 2017a; Reinke et al., 2022; Sutherland et al., 2000). Research as early as the 1960s indicates that teacher praise effectively increases appropriate student classroom behavior, such as task engagement (Hall et al., 1968), and decreases disruptive behavior (Madsen et al., 1968). Teacher praise also improves student outcomes (Floress et al., 2017a) and is associated with positive student-teacher interactions and a safe and engaging classroom atmosphere (Conroy et al., 2009; Reinke et al., 2022). Not only is teacher praise an effective classroom and behavior management strategy, but it is easy to implement, nonintrusive, and is not associated with any cost (Floress et al., 2017a; Jenkins et al., 2015). Unfortunately, despite the many benefits of teacher praise, research generally indicates teachers' use of praise in classrooms to be low (Floress et al., 2017b).

White (1975) was the first to examine natural rates of teacher praise in the classroom setting with their results indicating that rates were highest in first- and second-grade classrooms (M = 0.7 statements per min) and dropped in subsequent grades (3^{rd}-5^{th}, M = 0.4; middle school, M = 0.3; high school, M = 0.1 statements per min). It was suggested that rates of praise were highest during early elementary school because praising younger students was more likely to be associated with reinforcement of teacher behavior. Overall, White concluded that teachers, particularly in later grades, were not utilizing praise at an optimal rate for increasing positive student outcomes. Brophy (1981b) reported teachers' rates of praise in the classroom may be even lower than White observed. Brophy reviewed six studies using the Brophy-Good dyadic interaction coding system (Brophy & Good, 1970) and reported data on teachers' responses (i.e., praise and criticism) to students' behavior and academic performance. Teachers' rates of praise of appropriate classroom behavior and academic performance were low across studies (Brophy,

1981b). Consistent with results from White, Brophy reported praise of appropriate classroom behavior dropped to near-zero levels (M = 0.01 statements per hour) in later grades. These early studies examining teacher praise provide clear evidence that rates of praise are lower than might be desired, suggesting intervention is needed to increase teachers' use of praise (Brophy, 1981b; White, 1975).

Behavior-specific Praise (BSP)

Following early research on teacher praise (Brophy, 1981b; White, 1975), researchers began to categorize praise statements as general and BSP (Haydon & Musti-Rao, 2011). General praise definitions vary; however, they are typically in agreement with Brophy's (1981a) definition of "positive responses to students' good work or good conduct that goes beyond mere affirmation or positive feedback". General praise statements are broad and do not *specify* the praised behavior (Markelz et al., 2021). Praise is, however, believed to be most effective when it is (a) specific, meaning the praise statement explicitly defines the praised behavior (Brophy, 1981a; Moffat, 2011); (b) contingent, meaning teachers deliver praise after the target behavior (Haydon & Musti-Rao, 2011); and (c) sincere and varied to the situation (O'Leary & O'Leary, 1977). These conditions are believed to result in praise being more effective because they allow praise to function as contingent positive reinforcement for desired behavior. When BSP functions as a reinforcer, it increases the likelihood of the praised behavior occurring in the future and simultaneously reduces incompatible behavior (Owens et al., 2017).

Considering teachers fail to provide adequate rates of general praise (Brophy, 1981b; White, 1975), it is not surprising that they also provide inadequate rates of BSP. Burnett and Mandel (2010) reported teachers' rates of general praise (M = 0.48 statements per min) were significantly higher than rates of BSP (M = 0.03 statements per min) in first- through sixth-grade

general education classrooms. Reinke et al. (2013) also reported suboptimal rates of both general (M = 0.43 per min) and BSP (M = 0.13 per min) by 33 Kindergarten to third-grade teachers, especially when compared to their rates of reprimands (M = 0.67 per min). When comparing teachers' rates of general and BSP, a pattern is evident; teachers utilize general praise more often than BSP (Floress & Jenkins, 2015; Floress et al., 2017b). Teachers use general praise at low rates and BSP at even lower rates, with a downward trend in the frequency of praise as grade level increased (see Table 4; Floress et al., 2018).

Measurement Procedures

Researchers use two measurement procedures, rate (Markelz et al., 2021) and praise-to-reprimand ratio (Caldarella et al., 2020), to assess the frequency of praise. Rate is a ratio of count (frequency) to observation time, and researchers have calculated it by dividing the frequency of the target behavior by the number of standard units of time used in observation. Rate is the most common measurement procedure for collecting data on both general praise and BSP statements, likely because this is the simplest data collection procedure for teachers to use (Markelz et al., 2021). Praise-to-reprimand ratio consists of measuring the frequency of praise statements given compared to the number of reprimands (Caldarella et al., 2020). Although researchers have not experimentally tested an optimal rate of praise (Floress et al., 2017a), prior research recommends that teachers should provide, on average, 0.4 BSP statements per min (i.e., 3-5 times per 10 min; Floress et al., 2017a; Gage et al., 2017). If teachers utilized praise effectively at the recommended rate, disruptive behaviors would likely decrease and positive student outcomes would increase (Floress et al., 2017a).

To demonstrate the positive effect praise has on student outcomes, Sutherland et al. (2000) utilized an ABAB design to assess the relationship between teacher-delivered BSP and

student task engagement in a special-education classroom for students with emotional and behavioral disorders (EBD). High rates of BSP, which exceeded the recommended 0.4 BSP statements per min (Floress et al., 2017a), resulted in increased levels of task engagement for students in the classroom from baseline (*M* = 48.7% intervals on task) to intervention (*M* = 85.6% intervals on task). The results of Sutherland et al. (2000) are socially meaningful, as task engagement for participants with EBD fell within the typical range for task engagement in general education settings during intervention (*M* = 75% to 85% of intervals on task). Prior research (Brophy, 1981b; White, 1975) suggests intervention is needed to increase teachers' use of BSP because it is difficult to train teachers to implement even simple classroom and behavior management strategies (Dufrene et al., 2014). Floress et al. (2017a) suggested teachers self-manage their use of BSP as a means of increasing their BSP rates.

Self-management (SM)

Malott (1993) describes self-management (SM) as the individual application of operant techniques (i.e., reinforcement and punishment) that produce a desired modification in one's own behavior. Self-monitoring and self-evaluation are two SM strategies that are often combined. Self-monitoring consists of an individual systematically observing their behavior and recording the occurrence or nonoccurrence of a pre-specified target behavior (Dunlap et al., 1991). When adding self-evaluation to this process, the individual judges their performance adequacy and effectiveness relative to a predetermined performance criterion (Keller et al., 2005; Spates & Kanfer, 1977). Although there is substantial literature illustrating how teachers can support students in the use of self-monitoring and self-evaluation procedures (Harris et al., 1994; Otero-López et al., 2010), teachers are less likely to consider SM strategies to improve their own behavior (Hager, 2018). Research, however, documents SM as an effective strategy to increase

teachers' rates of praise (Sutherland & Wehby, 2001). For example, Kalis et al. (2007) taught a first-year teacher in a self-contained high-school classroom to self-monitor her use of praise using a hand-held counter during 10 min segments. Rates of general and BSP praise that occurred at a rate of near-zero per min during baseline increased during intervention to a mean of 1.68 per min and 0.44 per min, respectively (Kalis et al., 2007). Similarly, Haydon and Musti-Rao (2011) taught two first-year middle school teachers to self-monitor their use of general praise, BSP, and reprimands using a vibrating device. During baseline, rates of general and BSP were observed to be at 0 statements per min; however, the intervention resulted in increased rates of general (0.2-0.3 statements per min) and BSP (0.3-0.7 statements per min). Haydon and Musti-Rao concluded that SM effectively increased teachers' rates of praise and simultaneously decreased reprimands and disruptive behaviors in the classroom.

Simonsen et al. (2013) extended prior research (Haydon and Musti-Rao, 2011; Kalis et al., 2007) establishing the effectiveness of SM as a means of increasing teachers' rates of praise by examining the effects of three self-monitoring strategies (i.e., tally, hand-held counter, rating scale) on teachers' BSP rates. Results indicated all three strategies increased rates of BSP, with the tally and hand-held counter conditions resulting in greater treatment adherence and accuracy as well as the highest rates of BSP. Cumulatively, these studies suggest (Haydon and Musti-Rao, 2011; Kalis et al., 2007; Simonsen et al., 2013) that teachers can effectively use a variety of modalities to self-manage and ultimately change their behavior in a therapeutic direction.

In addition to evidence supporting self-monitoring as an effective strategy to change teachers' behavior (Haydon & Musti-Rao, 2011; Kalis et al., 2007), self-evaluation can also be used to increase teachers' rates of praise (Sutherland & Wehby, 2001). First, Sutherland and Wehby (2001) examined the effects of a self-evaluation intervention on the rates of teacher

praise in self-contained classrooms for students with emotional and behavioral disorders. Intervention requiring teachers to self-evaluate using an audiotape sample of their classroom instruction resulted in a mean increase in general praise (M = 1.47 statements per min) from baseline (M = 0.59 statements per min) and a mean decrease in reprimands (M = 0.43 statements per min to M = 0.30 statements per min). Similarly, Keller et al. (2005) implemented an audiotape self-evaluation intervention with pre-service teachers to assess the effects on BSP rates. All three pre-service teachers increased their use of BSP during the intervention (Keller et al., 2005). Although promising, a thorough search of the literature did not unveil any published studies other than Keller et al. and Sutherland and Wehby that evaluated the effects of self-evaluation on teachers' rates of praise.

Performance Feedback

SM intervention packages often incorporate performance feedback, which research describes as a promising method for creating teacher behavior change and increasing evidence-based intervention implementation (Noell et al., 2005; Reinke et al., 2007). Performance feedback involves monitoring a behavior and providing feedback to an individual about that behavior (Noell et al., 2005). Performance feedback provides teachers with increased awareness of their teaching behaviors, reinforces desired behaviors, and corrects undesired behaviors (Duchaine et al., 2011). A systematic review by Rispoli et al. (2017) reported that performance feedback was the second most common intervention component in SM studies targeting teacher behavior, and of 17 teacher SM studies reviewed, it was included in nine of the SM intervention packages (Szykula & Hector, 1978; Allen & Blackston, 2003; Keller et al., 2005; Bingham et al., 2007; Reinke et al., 2008; Briere et al., 2015; Mouzakitis et al., 2015; Oliver et al., 2015; Pinter et al., 2015). Although not reviewed by Rispoli et al., researchers have also successfully included

performance feedback as part of SM intervention packages targeting BSP (Grasley-Boy et al., 2021; Rila et al., 2024).

Despite the effectiveness of performance feedback on teacher behavior change, performance feedback requires assistance from school personnel (e.g., school psychologists, assistant principals) and is time consuming (Rispoli et al., 2017). School personnel employing performance feedback must schedule and conduct classroom observations, monitor teachers' behaviors, collect data on the behavior of concern, and schedule and attend meetings during which they provide feedback to the teacher on their behavior (Miller & Uphold, 2021; Noell, 2010; Fallon et al., 2015). Such complex and time-intensive intervention components may not be feasible for school personnel as resources, such as school staff and time, are often limited (Rispoli et al., 2017).

Video-Recording

Self-monitoring one's behavior while teaching is challenging due to the cognitive demand it places on teachers when having to simultaneously provide instruction to students and manage their behavior (Lan & Morgan, 2003). Self-evaluation using video recording can alleviate this issue (Lan & Morgan, 2003), as video recordings allow teachers to focus on instruction and self-evaluate at a later, more convenient, time (Hager, 2018). Additional benefits of video recording include allowing teachers the ability to evaluate their own performance as often and as many times as needed, and self-evaluating one's performance tends to be less threatening than evaluation by a supervisor (Struyk, 1993).

To date, Wright (2012) is the only study to empirically evaluate the effects of a video self-evaluation intervention on teachers' use of praise. Specifically, using a group comparison design, Wright evaluated the effect of an immediate or delayed video-recording self-evaluation

intervention package on 51 preschool teachers' use of general and BSP. Participants in the treatment group received a video-recording self-evaluation intervention package that consisted of (a) observation skills training, which involved giving teachers examples and nonexamples of praise and having teachers identify praise statements until they reached at least 80% agreement; (b) self-evaluation, where teachers observed 5-10 min video-recordings of themselves and recorded frequency (count) of their use of general and BSP; and (c) goal setting, where teachers created a personal goal to improve their rate of BSP based on their self-evaluation of video-recordings (Wright, 2012). Treatment groups either self-evaluated (a) immediately after recording or (b) in a delayed fashion the next day. Observation skills training and self-evaluation had a significant positive effect on teachers' rates of general and BSP. There were no significant differences in teachers' use of praise when comparing teachers in the immediate or delayed self-evaluation groups (Wright, 2012). Overall, the intervention resulted in a significant increase in the frequency of teacher praise statements and thus, is consistent with the existing literature (Wright, 2012).

With improvements in technology, teachers have increased access to reliable forms of technology in the classroom (Hager, 2018). Video-assessment platforms such as GoReact are used in teacher education programs and allow individuals to record and upload videos, rewatch videos, and receive targeted feedback from instructors (GoReact, 2022). One potential use for GoReact is for self-evaluation, as the platform allows teachers to upload, watch, and store videos. Despite a number of descriptive studies and claims that teachers can use GoReact to practice teaching methods, conduct observations, and provide personalized feedback, researchers have not empirically studied GoReact for the use of teacher self-evaluation (D. Rinn, personal communication, July 22, 2022).

Purpose

Teachers lack appropriate classroom and behavior management skills (Levine, 2006) and are in need of efficient, evidence-based strategies to use in their classrooms (Flower et al., 2017). Fortunately, there are effective, cheap, and easy-to-implement interventions (e.g., increasing rates of praise) that teachers can learn when provided with sufficient training (Floress et al., 2017a). Unfortunately, many schools lack the funding and resources to provide teachers with appropriate training in classroom and behavior management skills (Burns, 2019). Given evidence that SM interventions can increase teachers' rates of praise, together with preliminary findings suggesting video self-evaluation is an effective intervention strategy, further research is needed to evaluate the effects of video-assessment platforms to develop teachers' classroom and behavior management skills (Hager, 2018).

The purpose of the current study was to evaluate the effects of a video-recording self-evaluation intervention on teachers' rate of BSP in the classroom. The current study extends the literature by evaluating a simple and efficient method for self-evaluation in the classroom setting and contributes to the growing literature evaluating the effects of self-evaluation on teachers' rates of praise. To our knowledge, it is the only study to date that utilizes a video-recording self-evaluation procedure alone (i.e., without a performance feedback or goal-setting component) to evaluate teachers' rates of BSP. Conducting the study in this way is valuable as the effectiveness of video-recording self-evaluation procedures alone cannot be concluded from the current literature. The current study also offers an evaluation of a novel form of video-assessment technology (i.e., GoReact) that allows for self-evaluation and the collection of frequency data. The majority of studies included in the SM literature use paper and pencil or hand-held counters to collect frequency data; thus, evaluating data collection methods via GoReact technology adds

to the current literature. Finally, to train teachers in the basics and importance of BSP, this study utilizes a federally funded training website–the IRIS Center from Vanderbilt University–that is widely used by teachers for discovering and learning evidence-based academic and behavioral practices. The use of the IRIS Center website for the training component of this study further minimizes the need for district personnel to provide teachers with training.

Method

Participants and Setting

Participants were three certified teachers who provided instruction in two elementary schools in a southeastern state. Table 5 contains teacher demographic data. Erin did not report her age. Kellyn is an Early Intervention Program (EIP) teacher and typically works with small groups of students. She did not report the number of students in her class as she pushes in to several classes. Consent to participate using a university institutional review board (IRB) approved consent form was obtained from each participant. Inclusion criteria for participation included (a) providing instruction to preschool or elementary-aged children in a classroom setting; (b) owning a personal device with video-recording technology; and (c) utilizing a mean rate of BSP of less than 0.3 statements per min. Sessions occurred in each participant's classroom during small-group instruction. Amelia taught a math group with 5 students. Erin and Kellyn taught reading groups with 4 and 6 students, respectively. Amelia and Kellyn received a $100.00 gift card as compensation for participating in the study. Erin was unable to receive compensation for her participation in this research study due to rules in the county in which she taught.

Materials

Participants utilized GoReact–an online platform–to record and upload videos for data collection. Each participant was provided with a GoReact account and uploaded videos to which the experimenter (i.e., first author) had access. Participants used their own personal devices (i.e., phones) to record videos of themselves teaching during small-group instruction. Academic materials used by the participants during sessions varied based on their students' needs, the content area, and the academic context in which teaching occurred.

Target Behaviors

The primary dependent variable was the rate of BSP. BSP was defined as contingent verbal praise that addressed specific positive behaviors engaged in by students within the classroom (Simonsen et al., 2008; Sutherland et al., 2000). For example, statements such as "Good job sitting with your hands and feet to yourself!" and "Thank you for raising your hand!" were considered BSP statements. Statements such as "Good job." were not considered BSP. Secondary dependent variables included general praise and general and behavior-specific reprimands. General praise was defined as nonspecific verbal or nonverbal praise that did not address the specific positive behaviors engaged in by students within the classroom (Floress et al., 2018). For example, verbal statements such as "Great job." and "Nice work." were considered general praise statements as well as nonverbal behaviors such as a "thumbs up." General reprimands were defined as non-specific verbal or nonverbal behavior that expressed general disapproval of the behaviors engaged in by students within the classroom (Rathel et al., 2008). Non-specific verbal statements such as "No!", "Stop it!", and "Shh!" were considered general reprimands as were nonverbal behaviors like shaking their head or wagging a finger. Behavior-specific reprimands were defined as contingent verbal reprimands that addressed the specific problem behavior engaged in by students. Specific verbal statements such as "Stop

spinning in your chair." and "Don't interrupt your peers when they are speaking." were considered behavior-specific reprimands. Nonverbal behaviors were not considered behavior-specific reprimands as they cannot address the specific behavior engaged in by students.

Design and Procedure

The experimenter used a nonconcurrent multiple-baseline-across-participants design to evaluate the effect of a self-evaluation intervention on participants' rates of BSP. Slocum et al. (2022) defines multiple-baseline designs as "a single-case experimental design that evaluates causal relations through the use of multiple baseline-treatment comparisons with phase changes that are offset in (a) real time; (b) number of days in baseline; and (c) number of sessions in baseline". Experimental control is established when multiple-baseline designs demonstrate a functional relationship between the introduction of the independent variable and changes in the dependent variable (Slocum et al., 2022). In a nonconcurrent multiple-baseline design, the tiers are not organized in real time (Slocum et al., 2022). That is, the session numbers across tiers do not necessarily correspond to the same calendar date. When utilizing nonconcurrent multiple-baseline designs, threats to internal validity must be considered (Kazdin, 2021). Some researchers claim nonconcurrent multiple-baseline designs are weak because they do not reliably control for extraneous variables, such as coincidental events (Gast et al., 2018; Johnston et al., 2020). In order to reduce threats to internal validity, researchers should isolate participants across settings (i.e., tiers) and vary the lag between phase changes (i.e., days/sessions in baseline; Slocum et al., 2022). Importantly, multiple-baseline designs, including nonconcurrent multiple-baseline designs, do not involve a reversal to baseline which may be more acceptable to teachers and administrators (Harvey et al., 2004).

Phases included baseline, IRIS training, self-evaluation intervention, and maintenance. Data collection began in Spring 2023 and ended in Spring 2024. A predetermined minimum number of baseline data points were determined (i.e., 5, 7, 9) for each participant. When baseline data were determined to be stable by visual analysis, the IRIS training condition (i.e., an extended baseline) was implemented with Amelia (Ledford & Gast, 2018). When data collected during the IRIS training phase were observed to be stable, the self-evaluation intervention condition was implemented with Amelia. Amelia progressed to maintenance upon five stable self-evaluation intervention data points. Data collection for Erin and Kellyn occurred in the same fashion. Erin did not complete maintenance due to time constraints at the end of the school year.

Baseline

The experimenter met with each participant prior to data collection to informally train each participant how to record and upload videos to Go React. Participants conducted a predetermined number of 15 min video observations during small-group instruction during baseline. Participants were not informed of the behaviors the experimenter was collecting data on during baseline and no feedback was provided. The experimenter watched the pre-recorded videos and collected frequency data on primary (BSP statements used) and secondary (general praise statements and reprimands used) variables. Rate was calculated for each variable by dividing the number of occurrences by the total observation time in minutes.

IRIS Training

Participants were provided with educational resources through Vanderbilt University's IRIS Center (IRIS, 2010). These resources included a PDF containing a Fundamental Skill Sheet on BSP and links to two 2-min videos containing examples and non-examples of BSP (see Appendix A). After reviewing the materials, participants completed a researcher-developed five-

question quiz (see Appendix B) on the information contained in the packet and sent answers to the experimenter via email. Participants were required to obtain a score of 100% on the quiz to verify their understanding of BSP versus general praise. All participants scored 100% on the quiz. Once complete, participants continued recording videos in their classrooms and the experimenter continued collecting frequency data on primary and secondary variables. No feedback was provided to participants during IRIS training. IRIS training was considered an extended baseline condition because it was not expected that this phase would sufficiently alter participants' rates of praise, as research suggests that basic training is necessary, but alone rarely leads to sufficient behavior change (Lerman et al., 2004; Poduska & Kurki, 2014). A minimum of three data points were collected after IRIS training. Participants continued to the next phase if the data showed no change in level or trend or if the data displayed a contratherapeutic trend.

Self-evaluation Intervention

The experimenter first conducted behavioral skills training (DiGennaro Reed et al., 2018) with each participant on the use of self-evaluation procedures. During the one-to-one training, the experimenter provided teachers with instructions on how to self-evaluate and modeled how to collect data. Participants were required to rehearse the steps with the experimenter and the experimenter provided immediate feedback on participants' performance. The experimenter used a checklist containing the steps of the intervention to determine participants' competency with self-evaluation (see Appendix C). The number of steps implemented correctly was divided by the total number of steps and multiplied by 100. Participants were required to demonstrate competency at a level of 100%.

During the self-evaluation phase, participants recorded themselves in their classrooms using GoReact. Participants set up their devices (e.g., phones) in their classrooms so they could

be seen providing instruction and their statements could be heard. Participants began recording their teaching when they were in proximity to their students and at least half of their students were in place. When the lesson was finished or 15 min elapsed, participants stopped the recording and uploaded their video to GoReact. Participants logged on to GoReact during a free segment of the day (e.g., planning period, after school) to observe the video and collect frequency data on their use of BSP statements during minutes 1 through 6 of the video (for a total of 5 min). Participants were asked to collect self-evaluation data during minutes 1 through 6 to (a) allow children in the classroom to transition and settle before the beginning of the lesson; (b) allow the participant to begin the lesson; (c) minimize the time participants spent collecting data; and (d) minimize the effect that recording one's behavior could have on the rate of BSP at the beginning of instruction. Participants used video markers on GoReact to tag the precise moment they used BSP. Participants self-evaluated their behavior before the next scheduled video recording session.

Maintenance

Participants conducted maintenance probes by recording three additional 15 min videos in their classroom. Participants did not self-evaluate their use of BSP during maintenance. Maintenance probes for each participant were collected within (a) three days; (b) one week; and (c) two weeks following the completion of the self-evaluation intervention. The experimenter collected frequency data on primary and secondary variables.

Training Fidelity

Behavioral skills training was provided to participants at the onset of the self-evaluation phase. Training fidelity was collected to ensure the experimenter taught each participant the necessary steps to implement the self-evaluation intervention. Using a checklist with all of the

necessary steps of the intervention (see Appendix D), a second observer collected data on training fidelity by observing each step that the experimenter taught the participant via video. Training fidelity was calculated by dividing the number of steps taught to the participant by the total number of steps and multiplying by 100. Training fidelity was at 100% across participants.

Teacher Accuracy in Self-evaluation

Using participants' videos, the experimenter assessed teacher accuracy in self-evaluation by comparing the BSP markers on the participant's video to the experimenter's BSP markers on the same video. If the BSP markers were within 2 s of each other on the video, they were counted toward the same BSP statement and were considered an "agreement." BSP markers that differed by more than 2 s were considered "disagreements." Teacher accuracy in self-evaluation was calculated for every session across each participant during the self-evaluation phase. Teacher accuracy was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100.

Teacher Treatment Fidelity

During the self-evaluation intervention, the experimenter recorded whether the participant (a) recorded a video as scheduled and (b) self-evaluated their recorded video before the next scheduled video recording. Teacher treatment fidelity was calculated separately for each of the steps above. The percentage of sessions participants (a) recorded a video on the day they were scheduled and (b) self-evaluated before the next scheduled video recording was calculated by dividing the number of times each step was implemented correctly by the total number of opportunities to complete each step and multiplying by 100.

Social Validity

The participants provided evaluation of the perceived social validity of the self-evaluation intervention by completing the Usage Rating Profile – Intervention, Revised (URP-IR; Briesch et al., 2013) once data collection finished. The URP-IR is a 29-item self-report measure assessing six areas: acceptability, understanding, home-school collaboration, feasibility, system climate, and system support. Participants rated items on a six-point scale ranging from *strongly disagree* to *strongly agree*. Example items included, "I understand how to use this intervention," "The intervention procedures easily fit in with my current practices," and "I would be committed to carrying out this intervention."

Results

In the following sections, results are summarized by participant for primary (i.e., rates of BSP) and secondary variables (i.e., rates of general praise, general reprimands, and behavior-specific reprimands). Changes in primary and secondary variables are shown in Figure 5. In addition, mean and standard deviation (SD) were calculated for participants' BSP rates across conditions (Table 6). Social validity results are also summarized below.

Amelia

During baseline, Amelia demonstrated a low and stable BSP rate (M = 0.04 statements per min) and a mean rate of general praise of M = 0.267 statements per min. After Amelia watched the IRIS training video, increases in BSP (M = 0.134 statements per min) and general praise (M = 0.45 statements per min) were observed. Unfortunately, her mean rate of BSP was not sufficient when compared to the recommended rate of 0.4 BSP statements per min (Floress et al., 2017a), and thus the self-evaluation intervention was implemented after nine sessions. The level of Amelia's BSP statements (M = 0.52 statements per min) during self-evaluation both exceeded baseline levels, with no overlapping data points, and recommended rates of BSP for

practice. Amelia's mean rate of general praise (M = 0.893 statements per min) also remained high during the self-evaluation intervention. During maintenance, rates of BSP for Amelia remained high, as her mean rate of BSP increased to M = 0.867 statements per min; however, her mean rate of general praise (M = 0.533 statements per min) decreased after the conclusion of the self-evaluation intervention.

Amelia's rates of general reprimands were low and stable across phases. There was no clear change in level or trend. Mean rates of general reprimands for baseline, IRIS training, self-evaluation intervention, and maintenance were 0.067, 0.033, 0.027, and 0.089, respectively. Similarly, Amelia's rates of behavior-specific reprimands remained at near-zero levels across phases.

Amelia implemented the self-evaluation intervention on 100% of days scheduled. She self-evaluated before the next scheduled video recording for 80% of sessions. When self-evaluating her own videos, Amelia demonstrated a mean accuracy of 57.4% when BSP markers on her videos were compared to the experimenter's BSP markers on the same videos.

Erin

Erin demonstrated low rates of BSP during baseline (M = 0.093 statements per min) with a contratherapeutic trend and high rate of general praise (M = 1.59 statements per min). Variable rates of general (M = 0.16 statements per min) and behavior-specific reprimands (M = 0.45 statements per min) were observed during baseline. After IRIS training, Erin showed an increase in mean BSP (M = 0.229 statements per min) and a decrease in general praise (M = 1.08 statements per min). Erin's use of general reprimands decreased (M = 0.06 statements per min) and her use of behavior-specific reprimands increased (M = 0.60 statements per min) after watching the IRIS training videos. Erin's mean rate of BSP remained below the recommended

rate and a contratherapeutic trend was observed, thus the self-evaluation intervention was implemented after 12 sessions. During self-evaluation, Erin had a mean rate of BSP that was higher than baseline (M = 0.417 statements per min) and above the recommended rate, as well as a general praise rate of M = 0.95 statements per min. An increase in general reprimands was observed (M = 0.217 statements per min) during self-evaluation. Rates of behavior-specific reprimands during self-evaluation remained similar to the IRIS training condition (M = 0.633 statements per min).

Erin's implementation fidelity of the self-evaluation intervention was 100% for recording on days she was scheduled and 100% self-evaluating before the next scheduled video recording. Erin demonstrated a mean accuracy of 58.34% when self-evaluating her own use of BSP.

Kellyn

During baseline, Kellyn's rates of BSP and general praise were M = 0.215 statements per min and M = 0.75 statements per min, respectively. Her rates of general and behavior-specific reprimands were low during baseline and both occurred at a rate of M = 0.022 statements per min. After watching the IRIS training videos, Kellyn demonstrated an increase in mean BSP (M = 0.3 statements per min) and general praise (M = 1.13 statements per min); however, her mean BSP rate remained below the recommended rate and a contratherapeutic trend was present. Kellyn's rates of general reprimands increased slightly during the IRIS training condition (M = 0.083 statements per min), though her rates of behavior-specific reprimands remained near zero (M = 0.01 statements per min). Kellyn transitioned to the self-evaluation phase after 13 sessions because her mean BSP rate remained below 0.4 BSP statements per min (Floress et al., 2017a). Kellyn's mean rate of BSP during self-evaluation (M = 0.789 statements per min) represented a significant improvement compared to baseline, and her mean rate of general praise increased to

M = 1.23 statements per min. Rates of general reprimands increased (*M* = 0.133 statements per min) during self-evaluation, and rates of behavior-specific reprimands remained low (*M* = 0.01 statements per min). Kellyn demonstrated a slight decrease in mean BSP (*M* = 0.645 statements per min) and general praise (*M* = 0.911 statements per min) from self-evaluation to maintenance, though her mean rate of BSP remained above the recommended level. During maintenance, rates of general reprimands and behavior-specific reprimands were *M* = 0.11 statements per min and *M* = 0 statements per min, respectively.

Kellyn implemented the self-evaluation intervention on 60% of days originally scheduled; however, she self-evaluated before the next scheduled video-recording 100% of opportunities. Kellyn demonstrated a mean accuracy of 93.8% when self-evaluating her own use of BSP when watching her videos.

Social Validity

Participants found the self-evaluation intervention acceptable. All participants indicated that self-evaluation was a simple, effective, and fair intervention. Participants rated self-evaluation as manageable and minimally disruptive, with few materials needed to complete the intervention. Participants reported self-evaluation was consistent with job expectations and fit well with current classroom practices. All participants reported they would be enthusiastic about carrying out this intervention in their classrooms.

Discussion

Classroom and behavior management is a challenging aspect of teachers' jobs in which they receive little to no training or support (Reinke et al., 2011). Disruptive classroom behavior resulting from a lack of teacher training or support is associated with poor outcomes for teachers (Otero-López et al., 2010) and students (Owens et al., 2017). One proactive and effective

classroom and behavior management strategy is teacher praise (Floress et al., 2017a; Reinke et al., 2022). Research suggests that praise is most effective when it is specific and contingent (Brophy, 1981a; Haydon & Musti-Rao, 2011; Moffat, 2011). Although BSP is an effective classroom and behavior management strategy (Owens et al., 2017), teachers rarely use sufficient rates in practice (Floress et al., 2018). One strategy shown to increase teachers' rates of praise is self-management (Sutherland & Wehby, 2001), which includes self-monitoring and self-evaluation.

A critique from Moore et al. (2019) indicated there is not enough high-quality research to classify teacher praise as an evidence-based practice for K-12 students without severe disabilities. Out of 30 studies meeting inclusion criteria, 11 were rated as being methodologically sound using the Council for Exceptional Children and What Works Clearinghouse guidelines. Although Moore et al. concluded there is insufficient research to determine who may benefit from teacher praise and under what conditions, they identified teacher praise as a research-based practice (versus evidence-based practice) that has several benefits, which included requiring minimal teacher effort, time, and resources. One factor that impacts the methodological rigor of studies evaluating teacher praise is age. Moore et al. states that teacher praise has been researched for over 50 years; thus, many commonly cited studies do not meet contemporary methodological standards used to judge their quality and rigor. Additionally, although Moore et al. did not distinguish between studies evaluating general and BSP, they noted that no studies reported negative or contratherapeutic effects. Thus, we chose to measure the effect of a self-evaluation intervention on both general and BSP. We chose BSP as our primary dependent variable due to its efficacy cited in other studies (Brophy, 1981a; Haydon & Musti-Rao, 2011; Moffat, 2011) and the focus of BSP in the IRIS training module. We wanted to include the IRIS

training module as a condition in this study, and their module for addressing challenging classroom behavior explains BSP, not general praise.

This study examined the effects of a video-recording self-evaluation intervention on three elementary school teachers' rates of BSP in the classroom. Replication of effects across participants suggests a functional relation between the video-recording self-evaluation intervention and rates of BSP, such that teachers were observed to use a greater amount of BSP following the self-evaluation intervention. These findings are consistent with prior studies involving implementation of self-evaluation interventions to increase teachers' use of BSP (Sutherland & Wehby, 2001; Keller et al., 2005; Wright, 2012). This study extended the current literature as Sutherland and Wehby (2001) and Keller et al. (2005) used audiotapes to allow teachers to self-evaluate their behavior, whereas the current study used a novel video-assessment platform (i.e., GoReact). All three studies (Sutherland & Wehby, 2001; Keller et al., 2005; Wright, 2012) incorporated a goal-setting component, which prohibits an assessment of the effect of the self-evaluation intervention alone on teachers' rates of praise. This study adds to the growing literature evaluating the effectiveness of self-evaluation on teachers' rates of praise as it is, to our knowledge, the only study to date that utilizes a video-recording self-evaluation procedure without additional components (i.e., performance feedback, goal-setting).

All three participants demonstrated low rates of BSP and mean rates of general praise that exceeded BSP during baseline. After completing IRIS training, participants' BSP increased, at least initially, before a nontherapeutic trend was observed. IRIS training did not result in sufficient improvement in BSP rates across participants when compared to the recommended rate of 0.4 BSP statements per min (Floress et al., 2017a). These results are consistent with research suggesting that basic training on a skill is necessary, but rarely leads to sufficient behavior

change (Lerman et al., 2004; Poduska & Kukri, 2014). The self-evaluation intervention resulted in an increase in mean BSP for all participants. Therefore, self-evaluation may be a promising strategy for increasing teachers' use of BSP in the classroom.

Amelia's and Kellyn's rates of general praise increased in each condition from baseline to IRIS training to self-evaluation. Their rates of general praise dropped slightly during maintenance, though their rates of BSP remained well-above the recommended rate of 0.4 statements per min. Unfortunately, Erin's rate of general praise decreased throughout the course of the study, despite steady increases in her BSP rates. This could be due to the fact that Erin was primarily focused on delivering BSP, thus her rates of general praise decreased over time.

Amelia's rates of general and behavior-specific reprimands remained low across conditions. Her lowest rate of general reprimands was achieved during the self-evaluation intervention, and her rate increased slightly during maintenance. Kellyn's rates of behavior-specific reprimands decreased across conditions; however, her rate of general reprimands slowly increased throughout the study. Erin's rates of general reprimands and behavior-specific reprimands increased from baseline to the self-evaluation intervention. Although increasing teachers' rates of reprimands was not a goal of this study, it is plausible that when we reinforce one behavior (i.e., praise statements) we may also inadvertently increase another similar behavior (i.e., reprimand statements). We did not reinforce teachers' behaviors directly, as we did not provide them with any feedback on their behavior during the study; thus, the frequency of their praise and reprimand statements was reinforced by their students' behavior and their own perception of their behavior when watching their own videos. It appears that for Kellyn and Erin, their use of reprimand statements was also reinforced simultaneously with their use of BSP statements.

Maintenance data were collected for Amelia and Kellyn to examine whether effects of video-recording self-evaluation would maintain after the intervention was discontinued. Both participants' rates of BSP during maintenance were above baseline. Although the current study demonstrated maintenance of teachers' use of BSP following intervention, it is possible that participants' BSP rates may decrease over time (Keller et al., 2005). It would likely be beneficial for school personnel to provide teachers additional prompts to promote durable rates of BSP over time. Video-recording self-evaluation interventions could be implemented intermittently in the school setting in order to increase long-term efficacy and support positive classroom and behavior management strategies.

High social validity ratings by participants suggest that the self-evaluation intervention is appropriate, effective, and efficient for teachers. Minimal training was required to teach participants how to record and upload videos to GoReact as informal training sessions did not last longer than approximately 15 min. Engaging in behavioral skills training at the beginning of the self-evaluation intervention was not time intensive, as it took us approximately 5 min to explain the steps of the intervention to the teachers and approximately 5 min for the participants to rehearse the steps with us. Evidence of minimal training requirements add to the feasibility of SM interventions for teachers to implement in practice. Additionally, all participants did not accurately self-evaluate during the self-evaluation condition. Amelia's mean accuracy was 57.4%, Erin's was 58.34%, and Kellyn's was 93.8%. These results suggest that the self-evaluation intervention was effective despite teachers' inaccuracy in identifying their own use of BSP.

This study adds to the literature on video-recording self-evaluation as an effective way to increase teachers' use of BSP and other evidence-based classroom and behavior management

strategies. The results have important implications for teachers and students alike. Not only does video-recording self-evaluation have the potential to improve teachers' overall classroom and behavior management, but these interventions could indirectly increase teachers' self-efficacy and reduce burnout (Otero-López et al., 2010; O'Brennan et al., 2017). Students positively benefit from well-managed classrooms; thus, teachers' use of self-evaluation interventions have the potential to decrease disruptive behaviors and improve student engagement.

Limitations

Several limitations of the current study must be considered. First, teachers volunteered to participate in this study. Participants who volunteered to participate may be more interested in and open to trying new interventions, thus enhancing efficacy and social validity ratings (Myers et al., 2011). Second, Slocum et al. (2022) suggest that when using nonconcurrent multiple-baseline designs, researchers can strengthen the design by considering factors that could contribute to the isolation of tiers. This study utilized a nonconcurrent multiple-baseline-across participants design, such that there was a different participant in each tier. However, due to resources available, Amelia and Kellyn were not isolated across settings, as they taught the same grade at the same school. This may have introduced threats to internal validity as both Amelia and Kellyn could have been exposed to the same setting-level events (i.e., coincidental events).

Finally, although the study design allowed for flexibility in data collection, sessions did not occur concurrently across tiers. In order to reduce threats to internal validity when using nonconcurrent multiple-baseline designs, it is suggested that researchers predetermine the number of baseline data points for each participant (Slocum et al., 2022). Amelia and Kellyn were predetermined to have 5 and 9 baseline data points, respectively. We did not, however, follow through and complete 7 baseline data points with Erin as planned. Due to this, the results

need to be interpreted with extreme caution as not having a predetermined number of baseline data points when using nonconcurrent multiple baseline designs presents threats to internal validity. However, given Erin's levels of BSP were low and stable during baseline, and past research indicating that professional training alone does not permanently change behavior, we would not expect that additional time in baseline or IRIS training would increase rates of BSP.

Future Directions

The results of this study suggest that video-recording self-evaluation interventions are promising strategies to increase teachers' use of BSP, which can direct intervention implementation in the school setting and future studies. For instance, school administrators may consider including a video-recording self-evaluation intervention as initial training for all new teachers to provide adequate training in classroom and behavior management. In addition to initial training, including intermittent implementation of self-evaluation procedures may be helpful in maintaining teachers' rates of BSP above recommended rates.

Future research should continue to evaluate the effect of self-evaluation interventions to improve teachers' classroom and behavior management skills. Although this study contributed to the current literature by examining the effect of self-evaluation without additional components (i.e., performance feedback or goal-setting), researchers should continue to evaluate the effectiveness of SM procedures alone. This allows for further investigation regarding the effectiveness of SM as an independent intervention to change teachers' behavior.

Future studies may also benefit from assessing teachers' readiness to engage in behavior change prior to implementing video-recording self-evaluation, given individual factors may influence response to intervention. For instance, current literature has found that while self-management is a well-established, evidence-based intervention, it assumes the individual has an

intrinsic motivation to change (Orji et al., 2018). Given research that suggests that an individual must be prepared and motivated to change for intervention to be efficacious (Orji et al., 2018), future studies may benefit from consideration of individual differences and readiness for change when implementing self-evaluation interventions. Additionally, it would likely be beneficial for studies to include student outcomes in their findings to provide tangible feedback on the positive outcomes associated with increased BSP in the classroom.

Conclusion

Overall, the purpose of the present study was to examine the effects of a video-recording self-evaluation intervention on teachers' rates of BSP. Although there is robust literature that describes the efficacy of such SM interventions more broadly, few studies exist using a self-evaluation intervention. The results of this study found that a video-recording self-evaluation intervention is an effective and accessible intervention to increase teachers' rates of BSP in the classroom setting. Including video-based self-evaluation interventions in the school setting has the potential to not only enhance the teaching experience, with potential to reduce teacher burnout, but also improve student outcomes and experiences. These findings have important implications for both research and school-based practice, warranting additional examination of long-term implementation and feasibility in the school setting.

References

Allen, S. J., & Blackston, A. R. (2003). Training preservice teachers in collaborative problem solving: An investigation of the impact on teacher and student behavior change in real-world settings. *School Psychology Quarterly*, *18*(1), 22.

Bingham, M. A., Spooner, F., & Browder, D. (2007). Training paraeducators to promote the use of augmentative and alternative communication by students with significant disabilities. *Education and Training in Developmental Disabilities*, 339-352.

Briere, D. E., Simonsen, B., Sugai, G., & Myers, D. (2015). Increasing new teachers' specific praise using a within-school consultation intervention. *Journal of Positive Behavior Interventions*, *17*(1), 50-60.

Briesch, A. M., Chafouleas, S. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2013). Assessing influences on intervention use: Revision of the Usage Rating Profile-Intervention. *Journal of School Psychology, 51*, 81-96.

Brophy, J. E. (1981a). On praising effectively. *The Elementary School Journal*, *81*(5), 269-278.

Brophy, J. E. (1981b). Teacher praise: A functional analysis. *Review of Educational Research*, *51*(1), 5-32.

Brophy, J. E., & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, *61*(5), 365.

Bruce, J. R. (2010). *Teachers' perceptions of the inclusion of students with disabilities in the general education classroom*. Tennessee State University.

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.

Burns, M. K. (Ed.). (2019). *Introduction to school psychology: Controversies and current practice*. Oxford University Press.

Caldarella, P., Larsen, R. 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*, *40*(10), 1306-1322.

Chappuis, J., Stiggins, R. J., Chappuis, S., & Arter, J. (2020). *Classroom assessment for student learning: Doing it right--using it well* (p. 432). New York: Pearson.

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.

DiGennaro Reed, F. D., Blackman, A. L., Erath, T. G., Brand, D., & Novak, M. D. (2018). Guidelines for using behavioral skills training to provide teacher support. *Teaching Exceptional Children*, *50*(6), 373-380.

Duchaine, E. L., Jolivette, K., & Fredrick, L. D. (2011). The effect of teacher coaching with performance feedback on behavior-specific praise in inclusion classrooms. *Education and Treatment of Children*, *34*(2), 209-227.

Dufrene, B. A., Lestremau, L., & Zoder-Martell, K. (2014). Direct behavioral consultation: Effects on teachers' praise and student disruptive behavior. *Psychology in the Schools*, *51*(6), 567-580.

Dunlap, L. K., Dunlap, G., Koegel, L. K., & Koegel, R. L. (1991). Using self-monitoring to increase independence. *Teaching Exceptional Children*, *23*(3), 17-22.

Fallon, L. M., Collier-Meek, M. A., Maggin, D. M., Sanetti, L. M., & Johnson, A. H. (2015). Is performance feedback for educators an evidence-based practice? A systematic review and evaluation based on single-case research. *Exceptional Children*, *81*(2), 227-246.

Floress, M. T., & Jenkins, L. N. (2015). A preliminary investigation of kindergarten teachers' use of praise in general education classrooms. *Preventing School Failure: Alternative Education for Children and Youth*, *59*(4), 253-262.

Floress, M. T., Berlinghof, J. R., Rader, R. A., & Riedesel, E. K. (2017a). Preschool teachers' use of praise in general, at-risk, and special education classrooms. *Psychology in the Schools*, *54*(5), 519-531.

Floress, M. T., Beschta, S. L., Meyer, K. L., & Reinke, W. M. (2017b). Praise research trends and future directions: Characteristics and teacher training. *Behavioral Disorders*, *43*(1), 227-243.

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*(4), 411-422.

Flower, A., McKenna, J. W., & Haring, C. D. (2017). Behavior and classroom management: Are teacher preparation programs really preparing our teachers? *Preventing School Failure: Alternative Education for Children and Youth*, *61*(2), 163-169.

Gage, N. A., MacSuga-Gage, A. S., & Crews, E. (2017). Increasing teachers' use of behavior-specific praise using a multitiered system for professional development. *Journal of Positive Behavior Interventions*, *19*(4), 239-251.

Gast, D. L., Lloyd, B. P., & Ledford, J. R. (2018). Multiple baseline and multiple probe designs. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (pp. 288–335).

Grasley-Boy, N. M., Gage, N. A., Reichow, B., MacSuga-Gage, A. S., & Lane, H. (2021). A conceptual replication of targeted professional development to increase teachers' behavior-specific praise. *School Psychology Review*, *52*(1), 72-86.

Hager, K. D. (2018). Teachers' use of video self-monitoring to improve delivery of effective teaching practices. *Teaching Exceptional Children*, *50*(5), 283-290.

Hall, R. V., Lund, D., & Jackson, D. (1968). Effects of teacher attention on study behavior 1. *Journal of Applied Behavior Analysis*, *1*(1), 1-12.

Harris, K. R., Graham, S., Reid, R., McElroy, K., & Hamby, R. S. (1994). Self-monitoring of attention versus self-monitoring of performance: Replication and cross-task comparison studies. *Learning Disability Quarterly*, *17*(2), 121-139.

Harvey, M. T., May, M. E., & Kennedy, C. H. (2004). Nonconcurrent multiple baseline designs and the evaluation of educational systems. *Journal of Behavioral Education*, *13*, 267-276.

Haydon, T., & Musti-Rao, S. (2011). Effective use of behavior-specific praise: A middle school case study. *Beyond Behavior*, *20*(2).

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational Leadership*, *60*(8), 30-33.

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*(5), 463-476.

Johnston, J. M., Pennypacker, H. S., & Green, G. (2020). *Strategies and tactics of behavioral research and practice* (4th ed.). Routledge/Taylor & Francis Group.

Kalis, T. M., Vannest, K. J., & Parker, R. (2007). Praise counts: Using self-monitoring to increase effective teaching practices. *Preventing School Failure: Alternative Education for Children and Youth*, *51*(3), 20-27.

Kazdin, A. E. (2021). *Single-case research designs: Methods for clinical and applied settings* (3rd ed.). Oxford University Press.

Keller, C. L., Brady, M. P., & Taylor, R. L. (2005). Using self-evaluation to improve student teacher interns' use of specific praise. *Education and Training in Developmental Disabilities*, 368-376.

Lan, W. Y., & Morgan, J. (2003). Videotaping as a means of self-monitoring to improve theater students' performance. *The Journal of Experimental Education*, *71*(4), 371-381.

Ledford, J. R., & Gast, D. L. (Eds.). (2018). *Single case research methodology* (p. 377). New York, NY: Routledge.

Lerman, D. C., Vorndran, C. M., Addison, L., & Kuhn, S. C. (2004). Preparing teachers in evidence-based practices for young children with autism. *School Psychology Review*, *33*(4), 510-526.

Levine, A. (2006). Educating school teachers. *Education Schools Project*.

Madsen Jr., C. H., Becker, W. C., & Thomas, D. R. (1968). Rules, praise, and ignoring: Elements of elementary classroom control 1. *Journal of applied behavior analysis*, *1*(2), 139-150.

Malott, R. W. (1993). A theory of rule-governed behavior and organizational behavior management. *Journal of Organizational Behavior Management*, *12*(2), 45-65.

Markelz, A. M., Riden, B. S., Zoder-Martell, K. A., Miller, J. E., & Bolinger, S. J. (2021). Reliability assessment of an observation tool to measure praise characteristics. *Journal of Positive Behavior Interventions*, *23*(1), 17-29.

McCarthy, C. J., Lineback, S., & Reiser, J. (2015). Teacher stress, emotion, and classroom management. *Handbook of Classroom Management*, *2*, 301-321.

Miller, R. D., & Uphold, N. (2021). Using content acquisition podcasts to improve preservice teacher use of behavior-specific praise. *Teacher Education and Special Education*, *44*(4), 300-318.

Moffat, T. K. (2011). Increasing the teacher rate of behaviour specific praise and its effect on a child with aggressive behaviour problems. *Kairaranga*, *12*(1), 51-58.

Moore, T. C., Maggin, D. M., Thompson, K. M., Gordon, J. R., Daniels, S., & Lang, L. E. (2019). Evidence review for teacher praise to improve students' classroom behavior. *Journal of Positive Behavior Interventions*, *21*(1), 3-18.

Mouzakitis, A., Codding, R. S., & Tryon, G. (2015). The effects of self-monitoring and performance feedback on the treatment integrity of behavior intervention plan implementation and generalization. *Journal of Positive Behavior Interventions*, *17*(4), 223-234.

Myers, D. M., Simonsen, B., & Sugai, G. (2011). Increasing teachers' use of praise with a response-to-intervention approach. *Education and treatment of children*, *34*(1), 35-59.

Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., ... & Duhon, G. J. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review*, *34*(1), 87-106.

Noell, G. H. (2010). Empirical and pragmatic issues in assessing and supporting intervention implementation in school. *Practical handbook in school psychology*, 513-530.

O'Brennan, L., Pas, E., & Bradshaw, C. (2017). Multilevel examination of burnout among high school staff: Importance of staff and school factors. *School Psychology Review*, *46*(2), 165-176.

O'Leary, K. D., & O'Leary, S. G. (1977). *Classroom management: The successful use of behavior modification*. Pergamon.

Otero-López, J. M., Bolaño, C. C., Mariño, M. J. S., & Pol, E. V. (2010). Exploring stress, burnout, and job dissatisfaction in secondary school teachers. *International Journal of Psychology and Psychological Therapy*, *10*(1), 107-123.

Oliver, R. M., Wehby, J. H., & Nelson, J. R. (2015). Helping teachers maintain classroom management practices using a self-monitoring checklist. *Teaching and Teacher Education*, *51*, 113-120.

Orji, R., Lomotey, R., Oyibo, K., Orji, F., Blustein, J., & Shahid, S. (2018). Tracking feels oppressive and 'punishy': Exploring the costs and benefits of self-monitoring for health and wellness. *Digital Health*, *4*, 2055207618797554.

Owens, J. S., Holdaway, A. S., Smith, J., Evans, S. W., Himawan, L. K., Coles, E. K., ... & Dawson, A. E. (2017). Rates of common classroom behavior management strategies and their associations with challenging student behavior in elementary school. *Journal of Emotional and Behavioral Disorders*, *26*(3), 156-169.

Pinter, E. B., East, A., & Thrush, N. (2015). Effects of a video-feedback intervention on teachers' use of praise. *Education and Treatment of Children*, *38*(4), 451-472.

Poduska, J. M., & Kurki, A. (2014). Guided by theory, informed by practice: Training and support for the good behavior game, a classroom-based behavior management strategy. *Journal of Emotional and Behavioral Disorders*, *22*(2), 83-94.

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*(2), 67-77.

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*(3), 247-263.

Reinke, W. M., Lewis-Palmer, T., & Merrell, K. (2008). The classroom check-up: A class wide teacher consultation model for increasing praise and decreasing disruptive behavior. *School Psychology Review*, *37*(3), 315-332.

Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly*, *26*(1), 1.

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*(1), 39-50.

Reinke, W. M., Herman, K. C., & Copeland, C. B. (2022). Student engagement: The importance of the classroom context. *Handbook of Research on Student Engagement*, 529-544.

Rila, A., Bruhn, A. L., & Wang, L. (2024). The disaggregated effects of visual performance feedback on teachers' use of behavior-specific praise and reprimands. *Journal of Behavioral Education*, *33*(1), 53-80.

Rispoli, M., Zaini, S., Mason, R., Brodhead, M., Burke, M. D., & Gregori, E. (2017). A systematic review of teacher self-monitoring on implementation of behavioral practices. *Teaching and Teacher Education*, *63*, 58-72.

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*, 351-380.

Simonsen, B., MacSuga, A. S., Fallon, L. M., & Sugai, G. (2013). The effects of self-monitoring on teachers' use of specific praise. *Journal of Positive Behavior Interventions*, *15*(1), 5-15.

Slocum, T. A., Pinkelman, S. E., Joslyn, P. R., & Nichols, B. (2022). Threats to internal validity in multiple-baseline design variations. *Perspectives on Behavior Science*, *45*(3), 619-638.

Spates, C. R., & Kanfer, F. H. (1977). Self-monitoring, self-evaluation, and self-reinforcement in children's learning: A test of a multistage self-regulation model. *Behavior Therapy*, *8*(1), 9-16.

Struyk, L. R., & McCoy, L. H. (1993). Pre-service teachers' use of videotape for self-evaluation. *The Clearing House*, *67*(1), 31-34.

Sutherland, K. S. (2000). *Effects of self-evaluation on rates of teacher behaviors in classrooms for students with emotional and behavioral disorders*. Vanderbilt University.

Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *The Journal of Special Education*, *35*(3), 161-171.

Szykula, S. A., & Hector, M. A. (1978). Teacher instructional behavior change through self-control. *Psychology in the Schools*, *15*(1), 87-94.

Weinstein, C. S., Mignano, A. J., & Romano, M. E. (2007). Elementary classroom management: Lessons from research and practice.

White, M. A. (1975). Natural rates of teacher approval and disapproval in the classroom. *Journal of Applied Behavior Analysis*, *8*(4), 367-372.

Wright, M. R., Ellis, D. N., & Baxter, A. (2012). The effect of immediate or delayed video-based teacher self-evaluation on Head Start teachers' use of praise. *Journal of Research in Childhood Education*, *26*(2), 187-198.

Table 4

*Natural Mean Rates of Praise per Minute in the Classroom by Grade Level**

Study	General Praise				Behavior Specific Praise			
	Pre-K	K-2	3-5	6-9	Pre-K	K-2	3-5	6-9
White (1975)	-	0.7	0.4	0.3	-	-	-	-
Brophy (1981b)	-	-	-	-	-	-	-	-
Evertson et al. (1973)	-	0.04 (1)	-	-	-	-	-	-
Anderson et al. (1979)	-	0.03 (1)	-	-	-	-	-	-
Brophy et al. (1976)	-	0.03 (2-5)		-	-	-	-	-
Good & Grouws (1977)	-	-	0.04 (4)	-	-	-	-	-
Brophy et al. (1973)	-	-	0.03 (5)	-	-	-	-	-
Evertson et al. (1980)	-	-	-	0.02 (7-8)	-	-	-	-
Burnett and Mandel (2010)	-		0.48 (1-6)		-		0.03 (1-6)	
Reinke et al. (2013)	-	0.43 (K-3)		-	-	0.13 (K-3)		-
Floress and Jenkins (2015)	-	0.64	-	-	-	0.14	-	-
Floress et al. (2017b)	0.78	-	-	-	0.24	-	-	-

**Numbers in parentheses represent the grade or grade range that the value represents.*

Table 5

Participant Demographics

Participants	Age	Gender	Race	Experience	Type	Grade	Students
Amelia	24	Female	White	3	GE	2	21
Erin	-	Female	White	9	Incl.	1	19
Kellyn	49	Female	White	26	EIP	2	-

Note. Experience = number of years teaching; Type = classroom designation (GE = general education, Incl. = Inclusion, EIP = Early Intervention Program); Students = number of students in the participant's class

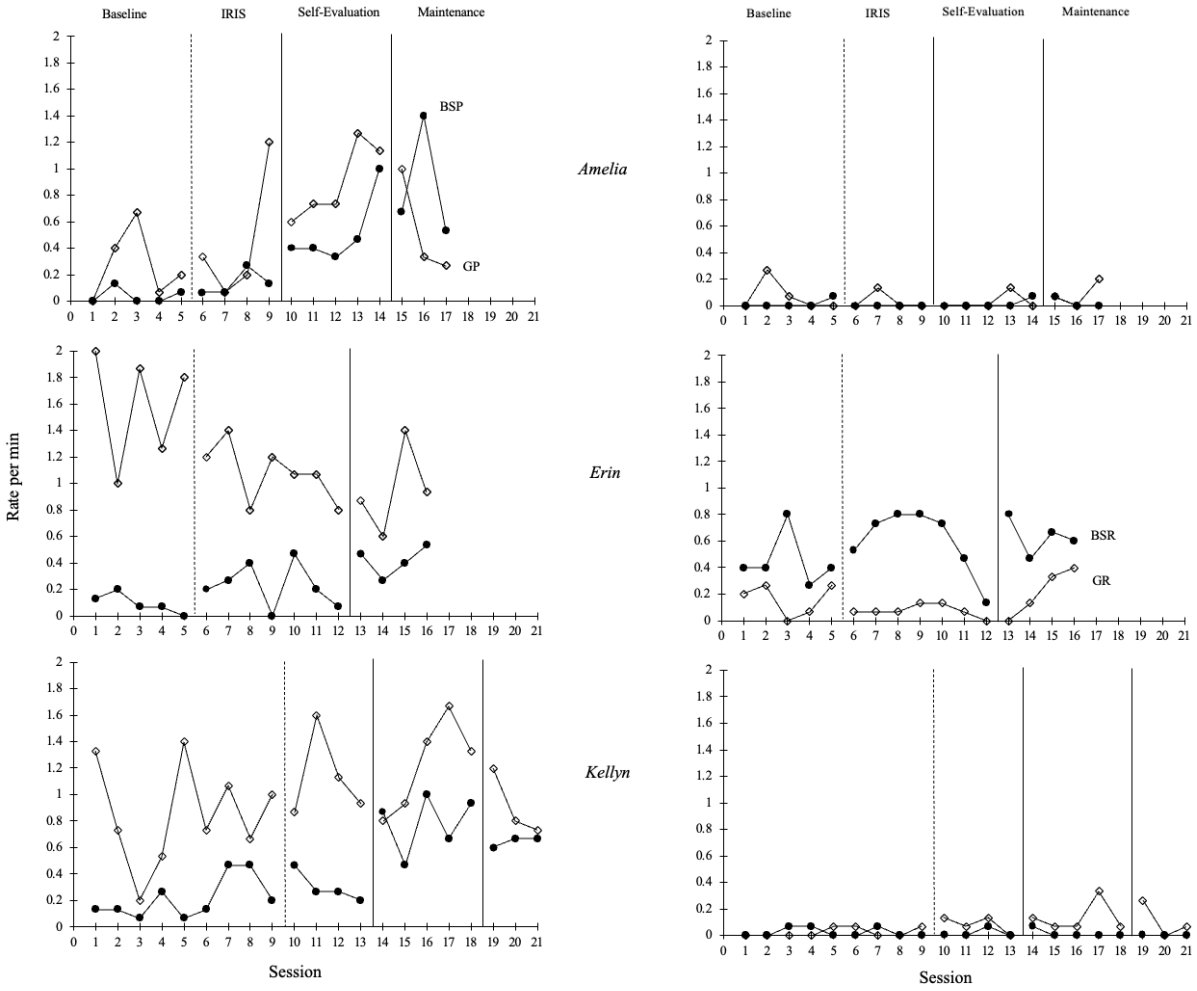
Table 6

Mean and Standard Deviations of BSP per min

Participants	Baseline		IRIS Training		Self-evaluation		Maintenance	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Amelia	0.040	0.060	0.134	0.094	0.520	0.272	0.867	0.467
Erin	0.093	0.076	0.229	0.167	0.417	0.114	-	-
Kellyn	0.215	0.156	0.300	0.116	0.789	0.218	0.645	0.039

Figure 5

Effects of Video Self-Evaluation Intervention on Teachers' Behavior



Note. BSP = behavior-specific praise; GP = general praise; BSR = behavior-specific reprimands; GR = general reprimands

Appendix A

Outline of Fundamental Skill Sheet and Video Examples

Behavior Specific Praise

- What is it?
- What do we know about this skill/practice?
- Procedures
- Tips for Implementation
- Things to Keep in Mind
- Implementation Examples
- Elementary Video Example
    - Example of BSP
    - Nonexample of BSP
- High School Video Example
    - Example of BSP
    - Nonexample of BSP
- Foundational Research and References

https://iris.peabody.vanderbilt.edu/wp-content/uploads/misc_media/fss/pdfs/2018/fss_behaviro_specific_praise.pdf

Appendix B

Quiz on IRIS Training Packet

1. BSP is delivered immediately following the behavior you would like to praise.	True	False
2. Using BSP is linked to positive student outcomes like staying on-task.	True	False
3. BSP can only be delivered to an individual student.	True	False
4. “Nice work, Anna!” is an example of BSP.	True	False
5. “Greg, I love how you are sitting quietly” is an example of BSP.	True	False

Appendix C

Behavioral Skills Training Checklist

Go to GoReact on web browser	
Log into GoReact platform	
Click dissertation folder	
Find correct activity (Day #)	
Click correct activity folder	
Start activity	
Add activity title	
Record video	
Conduct at least a 15 min recording	
Stop recording	
Upload video to GoReact	

Appendix D

Training Fidelity Checklist

Go to GoReact on web browser	
Log into GoReact platform	
Click dissertation folder	
Find correct activity (Day #)	
Click correct activity folder	
Start activity	
Add activity title	
Record video	
Conduct at least a 15 min recording	
Stop recording	
Upload video to GoReact	
Return to video later same day (or prior to next video recording)	
Fast forward to minute 1:00	
Watch own video from minute 1:00-6:00 (5 minutes total)	
Click BSP button when you hear yourself give BSP to students	