Predicting Implementation of the PAX Good Behavior Game + MyTeachingPartner Interventions

Summer S. Braun[1*], Catherine P. Bradshaw[2], Lydia A. Beahm[2], Alexa C. Budavari[2], Jason Downer[2], Nicholas S. Ialongo[3], Patrick, H. Tolan[2]

[1]Department of Psychology and Center for Youth Development and Intervention, University of Alabama

[2]School of Education and Human Development, University of Virginia

[3]Bloomberg School of Public Health, Johns Hopkins University

*** Correspondence:**
Summer S. Braun
ssbraun@ua.edu

Braun, S. S., Bradshaw, C. P., Beahm, L. A., Budavari, A. C., Downer, J., Ialongo, N. S., & Tolan, P. H. (2023). Predicting implementation of the PAX Good Behavior Game + MyTeachingPartner interventions. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1059138

Published in *Frontiers in Psychology*

Funding: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305A190162 and R305A130107 (to the University of Virginia). The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education.

Keywords: Implementation; Good Behavior Game; Feasibility; Teacher Well-Being; Occupational Health

Abstract

Introduction. Effective classroom management is critical to creating a classroom environment in which social, emotional, and academic learning can take place. The present study investigated the association between early career, early elementary teachers' occupational health (job stress, burnout, and perceived teaching ability) and perceptions of program feasibility in relation to their implementation dosage and quality of two evidence-based classroom management programs implemented together: the PAX Good Behavior Game (GBG) and MyTeachingPartner (MTP) intervention.

Methods. Teachers provided information on their occupational health at the start of the school year and were then randomized to the PAX GBG + MTP condition or control condition. Teachers' perceptions of the feasibility of the program, implementation dosage, and implementation quality of the intervention were measured at the end of the school year for the 94 intervention teachers.

Results. Teachers participated in more MTP coaching cycles when they reported that the combined PAX GBG + MTP program was feasible. Although there were no main effects of occupational health on implementation, the associations between job stress and implementation quality were moderated by perceptions of feasibility.

Discussion. Findings highlight the complexity of factors influencing the implementation of evidence-based programs in school settings.

Contribution to the Field

The present study is the first, to our knowledge, to investigate factors influencing the implementation of the combination of two empirically supported classroom management interventions when implemented together: PAX Good Behavior Game (GBG) and MyTeacherPartner (MTP). Results highlight the importance of teachers' perceptions of the program's perceived ease of use in predicting implementation dosage. Further, results indicate that although teachers' occupational health did not directly relate to their implementation of the program, teachers' stress and perceived feasibility of the program may work together to influence implementation. Taken together, the results illustrate the complex nature of factors that influence the implementation of evidence-based programming in school settings.

1 Introduction

Social and emotional learning programs and classroom management interventions are known to improve children's social and emotional skills, attitudes, behavior, and academic performance (Durlak et al., 2011; Korpershoek et al., 2016). Yet, effective implementation of such evidence-based intervention programs is critical to producing these targeted outcomes (Domitrovich et al., 2010; Durlak & DuPre, 2008; Kam et al., 2003). Schools are inherently complex systems, and there is noted variation in the implementation of such programs within this context (Hicks et al., 2014; Ringwalt et al., 2009; Sanetti & Collier-Meek, 2019; Sanetti & Kratochwill, 2009; Witt et al., 1997). To advance the science of implementation, it is essential to understand the factors influencing the adoption of evidence-based programs. The present study aimed to contribute to the practical understanding of implementation in the school context by exploring the effects of teachers' occupational health and their perceptions of feasibility of two evidenced-based interventions employed together: PAX Good Behavior Game (GBG) and MyTeachingPartner (MTP).

1.1 Importance of Evidence-Based Programs for Youth

Evidence-based programs are those which have undergone rigorous scientific testing and are found to have beneficial effects for the target population (What Works Clearinghouse, 2017). Evidence-based social and emotional learning programs are known to have widespread positive effects on children: improving social and emotional skills, attitudes towards self and others, positive social behavior, mental health, and academic performance, as well as preventing conduct problems and emotional distress (Durlak et al., 2011). Yet, effective classroom management is critical to creating a classroom environment where such learning can occur (Evertson & Weinstein, 2011). In fact, classroom management interventions themselves have been shown to benefit not only classroom behavior but also students' academic outcomes and social and emotional development (Korpershoek et al., 2016). Further, evidence-based social and emotional learning and classroom management interventions can have effects long after program participation, even in areas not directly targeted by the intervention (e.g., graduation; Taylor et al., 2017). Based on evidence that early-career teachers struggle specifically with classroom management (Kukla-Acevedo, 2009; Wei et al., 2010), which is foundational to student learning, this project combined two interventions with a strong evidence-base for improving classroom management: a classroom management program, the PAX GBG, and a teacher coaching intervention, MTP.

1.1.1 PAX GBG

The GBG is an evidence-based program originally developed by Barrish et al. (1969) that aims to promote teachers' classroom behavior management, increase on-task behaviors, and decrease disruptive behaviors (Bowman-Perrott et al., 2016; Smith et al., 2021). The GBG incorporates several behavior management strategies such as positive behavior praise, explicit instruction, feedback, and positive reinforcement, thus making it a high-quality practice for teachers to implement in their classrooms. The GBG is an interdependent group contingency that requires all students in the group to meet the requirements of the contingency as a group and individually (Embry, 2002). As such, students typically encourage their peers to meet expectations, thus reducing some of the demands on teachers (Hopman et al., 2018). GBG includes identifying the target behavior (e.g., completing the worksheet silently), posting the GBG expectations, dividing the class into equal teams, and awarding points to teams meeting expectations or removing points for infractions (Barrish et al., 1969; Embry, 2002). The team with the most points receives a non-tangible group reinforcement. The PAX GBG augmented the original version of GBG by integrating additional activities and components to improve compliance and classroom management (e.g., soliciting student input on classroom

expectations; Embry, 2002). Previous research has demonstrated positive effects of the GBG and the PAX GBG for both students (e.g., reduced aggressive/disruptive behaviors and improved academic outcomes; (Ialongo et al., 2019; Johansson et al., 2020; Leidig et al., 2022; Smith et al., 2021) and teachers (e.g., decreased emotional exhaustion; (Hopman et al., 2018).

1.1.2 MTP

MTP is an evidence-based coaching intervention that targets effective classroom management through the quality of teachers' interactions with students (Allen et al., 2015). Throughout the program, coaches provide video-based, individualized feedback to teachers as they develop classroom management skills and the capacity to provide emotional and instructional support to students. Previous studies of MTP have demonstrated positive effects on the quality of student-teacher interactions, peer interactions, social functioning, behavioral engagement, and academic outcomes (Allen et al., 2011, 2015; Gregory et al., 2014; Mikami et al., 2011; Pianta et al., 2008).

Given the evidence behind the PAX GBG and MTP, recent research has combined both programs in an effort to support the development of early-career teachers' classroom management skills and capacity for high-quality interactions with students (Tolan et al., 2020). In this combined approach, the PAX GBG and MTP work in tandem, incorporating unique and overlapping classroom management strategies that aim to improve teachers' interactions with students, their classroom management practices, and subsequent student outcomes. Early-career teachers are a particularly suitable population for these interventions because, in comparison to more experienced teachers, they are actively developing new habits, may be more open to feedback, and may need new classroom management skills, given that many pre-service preparation programs are known to provide insufficient training in this area (Freeman et al., 2014).

1.2 Importance of Implementation

Importantly, effective implementation of evidence-based programs such as the PAX GBG and MTP is critical to their success. A meta-analysis of 542 studies of interventions for youth concluded that implementation had profound effects on outcomes; programs implemented well resulted in effect sizes several times higher than those with poorer implementation (Durlak & DuPre, 2008). Despite their potential convenience for improving public health, school-based interventions are often at risk for poor implementation (Domitrovich et al., 2008; Sanetti et al., 2014). The emergence of the field of implementation science has brought an explicit focus on understanding and in turn, addressing the barriers that jeopardize the effective implementation of such programs (Durlak, 2015; Eccles & Mittman, 2006).

1.3 Conceptualization of Implementation

Implementation refers to the content of the program and how it is delivered in a specific setting (Durlak & DuPre, 2008). Durlak and Dupre (2008) describe eight dimensions of implementation: 1) Fidelity, also known as adherence or compliance, is the extent to which a program aligns with the originally intended curriculum; 2) Dosage is the amount of the original program that was provided, often measured by the number of program sessions delivered; 3) Quality is how well, clearly, and correctly the program was delivered; 4) Participant responsiveness is the extent to which the program stimulates interest and garners the attention of the participants; 5) Program differentiation refers to the uniqueness of the program from other interventions; 6) Monitoring of comparison conditions is the documentation of the services received by those outside of the intervention group; with researchers primarily focused on the intervention condition, the control group often goes

unmonitored, yet knowing the activities of both groups is important when drawing conclusions about the comparative effect of a program; 7) Program reach refers to the proportion of involvement of individuals in a population and the representativeness of program participants, which is particularly important when considering program scale-up; 8) Adaptation refers to the changes made to the program that result in differences between that implementation and the original intervention. Most school-based implementation research has focused on the first two dimensions – fidelity and dosage – little is known about the effects of these other dimensions on program outcomes (Gould et al., 2016). The present study expands the field of implementation science by examining whether teachers' occupational health and perceptions of program feasibility influence both the dosage and quality of implementation of the PAX GBG and MTP.

1.4 Predictors of Implementation

Domitrovich et al. (2008) offered a multi-level conceptual framework outlining factors influencing the implementation of school-based interventions to guide implementation research in this context. Informed by ecological systems models (e.g., Bronfenbrenner, 1994), the Domitrovich et al. conceptual model posits that the implementation of school-based programs is impacted by influences specific to the context in which the program is being implemented. These influences are described in three main categories: individual-level factors relating to the program implementer (e.g., occupational health, perceptions of the program), school-level factors (e.g., school culture, resources), and macro-level factors (e.g., federal, state, and district policies). As most school-based intervention programs are implemented by classroom teachers instead of an external provider such as a clinician or school counselor (Forman et al., 2009), it is important to understand the role that teachers play in impacting implementation. As such, the present study focused on how the characteristics of the teachers implementing the PAX GBG + MTP may impact their dosage and quality of implementation of the programs. We focus on teachers' occupational health and their perceptions of the feasibility of the intervention due to their theoretical and empirical relevance to teachers' capacity to implement intervention programming, described in detail below.

1.4.1 Teachers' Occupational Health

Teachers' occupational health refers to their evaluations of various aspects of their job (van Horn et al., 2004). The multifaceted construct incorporates affective, cognitive, professional, social, and psychosomatic dimensions (van Horn et al., 2004). Notably, Kazdin (1993) posited that the absence of dysfunction does not reflect the presence of optimal functioning. Thus, it is important to consider both negative experiences of distress (e.g., stress) and positive experiences of well-being (e.g., perceived ability) in assessing occupational health. Indeed, researchers have used a variety of measures to assess this construct, including job stress, burnout, self-efficacy, and others (Bakker & Rodríguez-Muñoz, 2010; van Horn et al., 2004). Although research often considers teachers' occupational health as an outcome of an intervention (e.g., Ross et al., 2012) teachers' occupational health may also influence their classroom practice. Specifically, the conceptual model of the Prosocial Classroom (Jennings & Greenberg, 2009) posits that teachers' occupational health and well-being likely support effective implementation of intervention programs, while feelings of stress and associated experiences jeopardize implementation. We focus on three salient experiences of occupational health in the present study: teachers' experiences of job stress, burnout, and perceptions of teaching ability.

1.4.1.1 Job Stress

Teachers report one of the highest levels of stress of any profession (Johnson et al., 2005). The Jobs Demands-Resource model of occupational stress posits that when the demands of a job exceed the resources of the individual and organization, stress can occur (Bakker & Demerouti, 2007). Theoretically, when teachers are under stress their emotional resources, attention, and cognitive energy are devoted to coping, leaving fewer resources for maintaining healthy relationships with students, supporting student development, and effectively implementing programs (Boekaerts, 1993; Roeser et al., 2012, 2021). Although empirical studies investigating these associations are still emerging, initial evidence provides support for this theory. Stressed teachers report more barriers to implementing evidence-based programs, such as lack of time to implement the program than their less-stressed counterparts (McGoey et al., 2014). Further, another study found a well-being intervention for teachers reduced stress and also improved their implementation of an evidence-based program for their students (Larson et al., 2018). Finally, one recent study found high levels of teacher stress to be associated with poor implementation quality of a mindfulness curriculum for students (Braun et al., under review). However, this association was attenuated when teachers were provided with in-depth training, suggesting that highly stressed teachers may need more hands-on support than is typical to implement interventions well.

1.4.1.2 Burnout

Prolonged exposure to stressors (e.g., job demands exceeding resources) are often associated with experiences of burnout (Maslach & Jackson, 1981). Burnout is often characterized by three dimensions: emotional exhaustion, depersonalization (e.g., feeling disconnected), and personal accomplishment (e.g., feelings of competence in the classroom; (Maslach & Jackson, 1981). Burnout is an unfortunately common experience for teachers (García-Carmona et al., 2019; Salmela-Aro et al., 2019). Teachers' feelings of burnout have been known to be associated with other salient experiences for themselves (e.g., depression, job dissatisfaction, disengagement; Bakker & Schaufeli, 2000; Leiter & Durup, 1994; Pines & Keinan, 2005), their classroom practices (e.g., poor student-teacher interactions, classroom management characterized by harsh discipline; Reinke et al., 2013) and student outcomes (e.g., impaired academic achievement; Breeman et al., 2015; Chang & Davis, 2009; Herman et al., 2020).

One process through which burnout may have such effects on students is by reducing their capacity to effectively implement evidence-based intervention programs, as posited by the Prosocial Classroom model (Jennings & Greenberg, 2009). Indeed, research from the field of social and emotional learning has supported this conceptual model; teachers' feelings of burnout have been related to lower dosage implementation of several different school-based programs, including the PAX GBG (Domitrovich et al., 2015; Ransford et al., 2009; Swift et al., 2017). Although one study found no main effect of burnout, the effect of burnout on implementation dosage was moderated by teacher-coach alliance: burnout was associated with a lower dosage of the PAX GBG, specifically when teacher-coach alliance was low (Wehby et al., 2012). Notably, the negative effect of burnout was found in the PAX GBG dosage but not the quality of implementation (Domitrovich et al., 2015). The present study is the first to assess the role of burnout in the implementation of the combined PAX GBG + MTP interventions.

1.4.1.3 Perceived Ability

Self-efficacy refers to teachers' belief in their capability to successfully accomplish a specific teaching task (Tschannen-Moran et al., 1998). Two primary domains comprise self-efficacy: 1) self-perception of teaching competence (i.e., a teacher's assessment of their own skills and knowledge),

and 2) beliefs about the demands of a specific teaching task (e.g., a teacher's context-specific assessment of external resources and barriers). Teachers' sense of self-efficacy is associated with other indices of occupational health (e.g., burnout, job satisfaction; Brouwers & Tomic, 2000; Skaalvik & Skaalvik, 2010, 2014). In addition, teachers' self-efficacy is associated with the use of more supportive classroom management practices, higher quality interactions with students, and student achievement (Swars et al., 2006; Zee & Koomen, 2016). Of focus in the present study is teachers' perceived teaching ability. Perceived teaching ability captures the teachers' views toward their own abilities as a teacher, which may be likened to the self-perception of teaching competence domain of self-efficacy.

Although no research has investigated the association between teachers' perceived ability and intervention implementation specifically, previous studies have explored the association between overall self-efficacy and implementation of school-based interventions. Teachers' self-efficacy has been associated with both the quality (Kallestad & Olweus, 2003; Rohrbach et al., 1993) and dosage (Clayback et al., 2022; Ransford et al., 2009) of school-based interventions. Some studies have focused specifically on the association between self-efficacy for classroom management and implementation, with mixed results. One study of early childhood teachers' implementation of a SEL program found self-efficacy for classroom management predictive of dosage, but not the quality of implementation (Thierry et al., 2022). Perhaps most relevant to the present investigation is a study of the PAX GBG, which found self-efficacy for classroom management was unrelated to implementation dosage and quality (Domitrovich et al., 2015). Further research is needed to clarify whether distinct aspects of self-efficacy, such as perceived teaching ability (vs. efficacy for classroom management, etc.), are associated with implementation dosage and quality.

1.4.2 Perceptions of Program Feasibility

Perceptions of the feasibility of a program are a core component of the social validity of an intervention. Social validity refers to the extent to which an intervention is useable, valuable, and favorably viewed by interested parties (Horner et al., 2005; Kazdin, 1977; Wolf, 1978). Although teachers' ratings of the feasibility of a specific program may be averaged to reflect the feasibility of the program as a whole, teachers themselves may vary in the extent to which they personally find the program to be feasible to implement (Han & Weiss, 2005). This teacher-level variation in perceptions of feasibility has important implications for implementation: Conceptual understandings and empirical evidence indicate that teachers who have positive perceptions of an intervention attend more training sessions (dosage) and implement the program with higher fidelity (Clayback et al., 2022; Han & Weiss, 2005). With regard to the social validity of PAX GBG specifically, previous research has found that teachers who perceive the program more favorably implement the program with greater fidelity and quality (Wehby et al., 2012). The predictive utility of social validity in the context of the combined PAX GBG + MTP has not yet been assessed and is important to consider as the combined high-quality implementation of these programs could be profound. Further, as social validity is in response to the program itself, these perceptions are potentially malleable, and results could inform amendments to the program to maximize social validity if found to be an important predictor of implementation. To this end, the present study focused on teachers' perceptions of the feasibility of the combined PAX GBG + MTP program (i.e., how easy it was to use).

1.4.3 The Potential Moderating Effect of Perceptions of Feasibility on the Association Between Stress and Implementation

Although stress is theorized to be a barrier to implementation, it is possible that the effects of these predictors of implementation are more complex; a selection of factors may work together to impact

implementation (Jennings & Greenberg, 2009; Ransford et al., 2009). For example, Dreer et al. (2017) found teachers' perceptions of the program to moderate the association between teachers' readiness to engage in the program and their commitment to utilizing new skills, where teachers experienced the greatest commitment to utilizing new skills when they were both ready to engage with the program and had positive perceptions of the program. In the context of the present study, teachers' positive perceptions of the feasibility of the programs could serve to buffer against the negative effect of stress on implementation. Conceptually, teachers' perceptions of feasibility may motivate teachers' engagement with the program (Wehby et al., 2012), despite their stress and function as a protective factor to lessen the impact of stress on implementation. In contrast, high levels of stress and perceptions that the program is difficult to implement may indicate compounding risks for poor implementation. Yet, most research examining the predictors of teachers' implementation has focused on the main effects (e.g., Domitrovich et al., 2015). Although few studies have investigated these more complex associations, one such study did find that teachers with high levels of burnout and negative perceptions of a social and emotional learning program exhibited the lowest implementation dosage and quality (Ransford et al., 2009), suggesting that these associations may be more complex than initially proposed (Jennings & Greenberg, 2009). Thus, additional research probing these more complex effects will contribute to our understanding of how combinations of factors may work together to impact implementation. Although we may hypothesize that perceptions of feasibility may buffer against the negative effects of burnout in the same way as it might for stress, as burnout emerges from experiences of chronic stress (Maslach & Jackson, 1981), stress, rather than burnout, is likely a more salient experience for early career teachers. Thus, of interest in this study was whether early career teachers' perceptions of feasibility may attenuate the negative effects of stress on implementation.

1.5 Present Study

The present study aimed to expand our understanding of the role that teachers play in the implementation of evidence-based programs for youth by investigating predictors of implementation. Guided by the Domitrovich et al. (2015) conceptual model of implementation of school-based interventions and the Prosocial Classroom model (Jennings & Greenberg, 2009), the present study explored the association between teachers' occupational health and perceptions of feasibility in relation to their dosage and quality of implementation of the PAX GBG + MTP program. Specifically, we addressed the following two specific research questions: RQ_1) Do teachers' own occupational health and perceptions of the feasibility (i.e., ease of use) of the program impact their implementation of the PAX GBG + MTP? RQ_2) Is the effect of teachers' stress on their implementation of the PAX GBG + MTP moderated by their perceptions of the feasibility of the program? We hypothesized that low levels of job stress, low levels of burnout, high levels of perceived ability, and positive perceptions of feasibility would be associated with greater dosage and implementation quality. We also hypothesized that the negative association between stress and implementation would be weaker for teachers who had positive perceptions of feasibility.

2 Method

2.1 Study Design and Recruitment

The present study draws from a longitudinal, teacher-level randomized controlled trial of the PAX GBG + MTP program. Early career teachers ($\leq$ 3 years of teaching experience) hired by three participating public school districts in Kindergarten-3rd grade were identified by the districts. These teachers were recruited into the project by project staff during district-wide professional development events for early-career teachers held prior to the start of the 2013 school year. To limit heterogeneity

in teaching demands, eligible teachers included those in early grades (Kindergarten-3rd Grade) and excluded Teach for America engaged teachers, given the variation in their educational backgrounds from typical teachers. Project staff conducted all recruitment sessions, either through attendance at new teacher training and orientation events or through individual or small group sessions. Participation was voluntary, and teachers provided written informed consent consistent with IRB procedures approved at the investigators' universities and school divisions. Participating teachers received an honorarium (e.g., gift cards) for their participation and completion of data collection activities. This recruitment and randomization procedure was repeated the following two years (i.e., the fall of 2014 and 2015, respectively), for a total of three cohorts.

Recruitment efforts resulted in 272 interested teachers, of which 236 teachers consented to participate. Of those, eight withdrew before randomization, 15 were ineligible due to being assigned to an ineligible classroom (i.e., not Kindergarten-3rd Grade, special education classroom, resource class), not being permitted to attend training (based on the principal's decision), having already been trained in the PAX GBG or MTP, leaving the participating districts, or leaving the teaching profession altogether. Of those eligible, 25 left the project prior to baseline data collection. The final sample in the RCT intent-to-treat analyses included 188 teachers (69% of initially interested teachers; 80% of those who consented) recruited from 72 schools (Median number of teachers per school = 2, Range = 1-13 teachers). Cohort 1 consisted of 56 teachers (30 control condition, 26 intervention condition) from 34 schools, Cohort 2 consisted of 51 (25 control condition, 26 intervention condition) teachers from 30 schools, and Cohort 3 consisted of 81 teachers (39 control condition, 42 intervention condition) from 36 schools. Note that the same school could be represented in multiple cohorts if they had new teachers in subsequent years, as was the case in several instances. See Tolan et al. (2020) for the full consort diagram. Attrition during the first study year was low, with 11% (10 from control, 11 from intervention) discontinuing participation before the Year 1 post-intervention timepoint.

2.2 Participants

Due to the present study's focus on implementation outcomes, only the 94 teachers randomized to the intervention condition were included in the analytic sample in this study. Randomization was effective as there were no significant differences in baseline demographics, occupational health, nor implementation outcomes between teachers in the intervention and control conditions (see Downer et al., under review for full sample descriptives). The majority of teachers in the intervention condition were female (93%) and White (80%), with 1-3 years of teaching experience, and most teachers were in their first year of teaching (60%). Teachers were approximately evenly distributed across the Kindergarten-3rd Grade classes.

2.3 Procedure

Baseline (Time 1) data collection occurred in the fall of the school year as close to the beginning of school as possible, in October. Post-intervention (Time 2) data collection occurred seven months later, in May, shortly before the end of the school year. At each timepoint, teachers completed an online survey, and trained observers conducted classroom observations. Following baseline data collection, teachers were randomly assigned (blocking on school and district) to the intervention or control conditions. School-level demographic data were obtained from the state department of education.

2.3.1 Classroom observation procedures

Observations were conducted in accordance with the protocol for the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008). At each timepoint, certified CLASS observers conducted 4-6, 15-minute observation cycles. Observations were conducted over two separate days balancing observations in the morning and the afternoon. Immediately following each observation cycle, observers stepped out of the classroom and completed the CLASS ratings. There were no calculations of inter-rater agreement following the CLASS training and certification process, but previous research has demonstrated high inter-rater agreement for the CLASS (79%-94% within 1 point; Mantzicopoulos et al., 2018; Pianta et al., 2008).

2.3.2 PAX GBG

The GBG allows teachers to utilize social learning principles within a team-based, game-like context to reduce aggressive, disruptive, and off-task behavior and facilitate academic instruction. The current project used the PAX GBG, an augmented version of GBG which integrates ancillary components known to improve compliance and classroom management (Embry, 2002).

Prior to the implementation of the PAX GBG at the beginning of the school year, the teachers and students collaborated to define their vision of a "PAX" (Latin, meaning peaceful or ideal) classroom. Toward that end, they identified the behaviors that were necessary for creating a focused, productive, and peaceful classroom. During this collaboration, the teacher explained to the students that the positive behaviors they listed were referred to as "PAX" behaviors, and the negative behaviors were referred to as "spleems." After jointly defining PAX and spleems, teachers assigned students to one of three or four teams. The teams worked cooperatively to maintain PAX behavior in the classroom. Teachers gave points to the team when a member displayed a spleem. Teachers were trained to respond unemotionally to rule-breaking and when marking points against a child's team. At the end of the game period, all teams with three or fewer spleems won the game. The students were rewarded for displaying self-control, emotion regulation, and group regulation while not attending to or reinforcing the misbehavior of others. The team-based nature of the game allowed teachers to take advantage of positive peer pressure to improve academic and pro-social student behavior at the individual as well as at the classroom level.

Training in the PAX GBG. Professional development was provided to the intervention teachers over the course of one weekend day, during which teachers received intensive training and practice using the PAX GBG approach aligned with the MTP framework. Seventy-seven percent of the intervention teachers completed this training, with the remainder attending a small group or one-on-one make-up trainings. Teachers were asked to play approximately three PAX GBG games each school day with increasing length and in increasingly varied settings over the course of the year.

2.3.3 MTP

MTP is grounded in an evidence-based framework for thinking about teacher-student interactions that contribute to student behavior and achievement, called Teaching through Interactions (Hamre et al., 2013). This framework emphasizes that interactions should be emotionally supportive, well-organized, and cognitively enriching. The Teaching through Interactions framework is based on these three core domains of classroom interactions as captured by an observational approach called the Classroom Assessment Scoring System (CLASS): emotional support, instructional support, and classroom organization (Pianta et al., 2008). During the intervention, the CLASS is used as a lens for viewing and providing feedback on a teacher's practice in the classroom.

Training in MTP. Teachers in the intervention condition also participated in one day of training in MTP. Following the training, teachers participated in biweekly MTP coaching cycles throughout the training year, with initial contact in-person that then shifted to web-mediated training. See Tolan et al. (2020) for a detailed description of coaching steps. Over the course of the school year, these coaching cycles focused on all three CLASS domains as well as elements of the PAX GBG that would help teachers optimize their implementation of the PAX GBG by attending to their interactions with students. The coaching cycles were intended to be collaborative, supportive, constructive, and to help teachers develop CLASS and PAX GBG knowledge, improve observation skills, develop analysis skills, feel supported in these endeavors, and increase their sense of agency and efficacy in the classroom. Teachers were asked to incorporate new strategies related to both the CLASS domains and PAX GBG elements into their teaching practice to improve both their implementation of the PAX GBG and their overall teaching practice.

2.4 Measures

2.4.1 Implementation Outcomes

The present study reports implementation dosage and quality data for the PAX GBG and MTP collected at post-intervention.

2.4.1.1 Implementation Dosage

The number of *coaching cycles completed* is an indicator of the dosage of MTP teachers received. These data were collected on each participating teacher through the online MTP coaching platform. The number of *PAX GBG games played* is an indicator of the dosage of the PAX GBG the students received. For the school year following the training, teachers self-reported on the number of PAX GBG games played each week, which was averaged across the school year.

2.4.1.2 Implementation Quality

Following the CLASS procedure above, after each classroom observation period, observers provided a rating from 1 to 7 (1 = Low, 7 = High) for each of the 11 CLASS dimensions. In accordance with contemporary uses of the CLASS (e.g., Mashburn et al., 2008), three scales were created to reflect the core CLASS domains: classroom management, emotional support, and instructional support. Teachers' scores within each timepoint (baseline, post-intervention) were averaged across cycles of observation. Cronbach's α for all domains and timepoints were acceptable (ranging from .84-.92). As MTP aims to improve the quality of teachers' interactions with students using the CLASS as a guide to anchor coaches' feedback to teachers, the CLASS was used as an indicator of the quality of teachers' implementation of MTP.

2.4.1.2.1 Emotional Support

Emotional Support was calculated as the average of the positive climate, negative climate, teacher sensitivity, and regard for student perspectives dimensions.

2.4.1.2.2 Instructional Support

Instructional support reflects teachers' facilitation of academic learning, measured as the average of the quality of feedback, concept development, and language modeling.

2.4.1.2.3 Classroom Organization

Classroom organization, also referred to as classroom management, assesses the quality of teachers' interactions with students while they are managing the students in the room. It is the average of the behavior management, productivity, and instructional learning formats dimensions.

2.4.2 Occupational Health and Perceptions of Feasibility

Main predictors of interest included indices of teachers' occupational health collected at baseline, namely job stress, burnout, and perceived ability, along with teachers' perceptions of intervention feasibility collected post-intervention.

2.4.2.1 Job Stress

Teachers' experiences of job stress were assessed using items from the National Institute for Occupational Safety and Health survey of work-related stress (National Institute for Occupational Safety and Health., 1999). Teachers rated five items about their current feelings of stress (e.g., "In my job, I feel like I am under great stress") on a 1 to 4 scale (1 = Strongly Disagree; 4 = Strongly Agree; $\alpha = .82$).

2.4.2.2 Burnout

Teachers' experiences of burnout were assessed using four items from the emotional exhaustion subscale of the Maslach Burnout Inventory (Maslach et al., 1996). Teachers rated four items about their feelings of emotional exhaustion ("I feel burned out from my work," "I feel like I am at the end of my rope," I feel emotionally drained from my work," and "I feel used up at the end of the work day") on a 1 to 4 scale (1 = Strongly Disagree; 4 = Strongly Agree; $\alpha = .85$). This subscale was abbreviated for use in this study due to practical considerations to reduce participant burden.

2.4.2.3 Perceived Ability

Teachers' perceptions of their ability as a teacher were assessed using the Perceived Ability subscale of the Factors Influencing Teaching Choice (FIT-Choice) measure (Watt & Richardson, 2007). Teachers rated three items about their perceived ability ("I have the qualities of a good teacher," "I have good teaching skills," and "Teaching is a career suited to my abilities") on a 1 to 4 scale (1 = Strongly Disagree; 4 = Strongly Agree; $\alpha = .69$).

2.4.2.4 Perceptions of Feasibility

Perceptions of the feasibility of the combined PAX GBG + MTP intervention were assessed using items from the Teacher Perceptions of the Intervention Attributes scale (Domitrovich et al., 2015) with adapted wording to be relevant to the PAX GBG and MTP programs. Teachers rated five items assessing their perceptions of how feasible the combined program was to implement (e.g., "The GBG + MTP coaching process was easy to participate in") on a 1 to 4 scale (1 = Strongly Disagree; 4 = Strongly Agree; $\alpha = .72$).

2.4.3 Demographics

At the teacher-level, teachers self-reported the grade that they taught and their years of teaching experience. School-level demographic data regarding the enrollment of the school and percent of students eligible for free and reduced-priced meals (FARMS) were obtained from the state department of education.

2.5 Analytic Plan

2.5.1 Missing Data

Full information maximum likelihood (FIML) was used to incorporate all participants with baseline data (including those who did not provide data at post-intervention) into the intent-to-treat analyses. This approach accounts for the missing data while obtaining minimally biased estimates (e.g., Little et al., 2014; Witkiewitz et al., 2014).

2.5.2 Preliminary Analyses

All analyses were run in R studio. Preliminary analyses included descriptive statistics and bivariate correlations among study measures.

2.5.3 RQ1: The Association Between Occupational Health, Perceptions of Feasibility, and Implementation.

Multiple linear regression models with cluster robust standard errors were employed to test the association between teachers' occupational health (i.e., stress, burnout, perceived ability) and perceptions of feasibility and their implementation of the PAX GBG + MTP. Because FIML was invoked to account for missing data, models were run in the latent framework using the lavaan package (Rosseel, 2012). A separate model was run for each outcome. Cluster robust standard errors were used to account for the nesting of teachers in schools in these analyses, where multilevel modeling is not appropriate given the average cluster size was so small (M = 2.44 teachers/school). Models predicting the number of coaching cycles completed used estimator = "MLR" to account for the non-normal distribution of this count outcome. Measures of occupational health and perceptions of feasibility were grand mean centered. Models controlled for grade level (continuous, where 0 = Kindergarten) and years teaching (continuous, where 0 = 1st year), as the participants in this study ranged from Kindergarten-3rd Grade teachers and were in their 1st-3rd years of teaching. Models predicting implementation quality (CLASS outcomes) controlled for teachers' CLASS scores at baseline, which were grand mean centered. School-level covariates included the school enrollment and the percent of students eligible for FARMs, which were standardized, and grand mean centered, respectively. Centering in this way results in an intercept that can be interpreted as the predicted level of implementation for a teacher who is experiencing an average level of job stress, burnout, perceptions of their teaching ability, and perceptions of feasibility, and is a Kindergarten teacher in their 1st year in the classroom, who is in a school of average enrollment and eligibility for FARMs.

2.5.4 RQ2: The Moderating Effect of Perceptions of Feasibility on the Association Between Stress and Implementation.

To test the potentially moderating role of perceptions of feasibility, the interaction of stress and perceptions of feasibility was added to each of the models above.

3 Results

3.1 Preliminary Analyses

Descriptive statistics of measures are provided in Table 1. Teachers demonstrated a relatively high dosage of the MTP elements of the combined program. On average, teachers completed 8.24 coaching cycles, which exceeded the target number of cycles of 8, which previous research has shown to impact teacher practice and student outcomes (Allen et al., 2011, 2015). Implementation of the PAX GBG elements of the combined program was also relatively high, with teachers playing an average of 9 games per week (i.e., an indicator of dosage).

Bivariate correlations among study measures are provided in Table 2. Notably, burnout was significantly negatively correlated with the number of MTP coaching cycles completed (r = -.21, p = .048), whereas perceptions of feasibility were significantly positively correlated with the number of MTP coaching cycles completed (r = .36, p = .001). These correlations indicated that teachers who reported higher levels of burnout at baseline completed fewer coaching cycles than their peers who were more burned out, and that teachers who reported the program was more feasible to implement (i.e., easy to use) completed more coaching cycles than their peers who reported lower levels of program feasibility.

3.2 RQ1: The Association Between Occupational Health, Perceptions of Feasibility, and Implementation

3.2.1 Coaching Cycles Completed

Teachers' perceptions of the feasibility of the program were associated with attending more coaching cycles (B = 1.66, SE = 0.53, p = .002). Grade was negatively associated with the number of coaching cycles completed (B = -0.43, SE = 0.19, p = .03; Table 3), such that teachers in lower grades attended more coaching sessions. No other effects were significant.

3.2.2 Number of Games Played

Occupational health and perceptions of feasibility were unrelated to the number of PAX GBG games played. At the school level, the percent of students eligible for FARMS was associated with playing more games (B = 5.39, SE = 1.47, p < .001). No other effects were significant.

3.2.3 Emotional Support

Occupational health and perceptions of feasibility were unrelated to observations of teachers' emotional support. Teachers' emotional support at baseline was strongly associated with their emotional support at post-intervention (B = 0.40, SE = 0.12, p < .001). At the school level, the percent of students eligible for FARMS was marginally negatively associated with emotional support (B = -0.66, SE = 0.39, p = .09). No other effects were significant.

3.2.4 Instructional Support

Teachers' instructional support at baseline was strongly associated with their emotional support at post-intervention (B = 0.40, SE = 0.12, p < .001). No other effects were significant.

3.2.5 Classroom Organization

Occupational health and perceptions of feasibility were unrelated to observations of teachers' classroom organization. Grade level was marginally associated with classroom organization (B = 0.34, SE = 0.09, p < .001) such that teachers in higher grades were observed to have higher levels of classroom organization. Teachers' classroom organization at baseline was strongly associated with their classroom organization at post-intervention (B = 0.34, SE = 0.09, p < .001). At the school level, school enrollment (B = -0.11, SE = 0.05, p = .03) and the percent of students eligible for FARMS (B = -0.74, SE = 0.34, p = .03) were negatively associated with classroom organization such that the larger the school and the more students eligible for FARMS, the lower the observed classroom organization. No other effects were significant.

3.3 RQ2: The Moderating Effect of Perceptions of Feasibility on the Association Between Stress and Implementation.

3.3.1 Coaching Cycles Completed

The interaction between stress and feasibility was not significant in predicting the number of coaching cycles completed (B = 0.17, SE = 1.04, p = .87; Table 4).

3.3.2 Number of Games Played

The interaction between stress and feasibility was not significant in predicting the average number of PAX GBG games played each week (B = 0.51, SE = 0.86, p = .55).

3.3.3 Emotional Support

The interaction between stress and feasibility was not significant in predicting observations of teachers' emotional support (B = -0.33, SE = 0.27, p = .21).

3.3.4 Instructional Support

The interaction between stress and feasibility was significant in predicting observations of teachers' instructional support (B = -0.77, SE = 0.26, p = .003). This effect, visualized in Figure 1a, indicates that teachers who reported high levels of stress and lower levels of program feasibility (i.e., perceptions that the program was harder to use) implemented the program with higher quality than those who were highly stressed and reported the program was more feasible to implement (i.e., easy to use).

3.3.5 Classroom Organization

The interaction between stress and feasibility was marginally significant in predicting observations of teachers' classroom organization (B = -0.41, SE = 0.24, p = .09). This effect, visualized in Figure 1b, shows that although job stress was relatively unrelated to implementation quality for those teachers who reported lower levels of program feasibility (i.e., perceptions that the program was harder to use), the opposite was true for teachers who reported that the interventions were feasible to implement. That is, teachers who had greater perceptions of program feasibility (i.e., perceptions that the program was easy to use) and experienced higher stress had poorer quality implementation than those who had greater perceptions of program feasibility and experienced low stress.

4 Discussion

The present study investigated the associations among teachers' occupational health and perceptions of the feasibility of the combined PAX GBG + MTP program and two dimensions of implementation: dosage and quality. Contrary to our hypotheses, there were no main effects of any occupational health indicator on implementation dosage or quality. However, teachers who reported that the program elements were easier to use did complete more coaching cycles. In addition, the effect of teachers' job stress on two dimensions of implementation quality, instructional support and classroom organization, was moderated by teachers' perceptions of how feasible the program was, suggesting that the effect of job stress on implementation may be more nuanced than initially proposed (Jennings & Greenberg, 2009). The present study expands the field of implementation science in education research by: 1) investigating several potential predictors of implementation with a particular focus on teachers' occupational health and perceptions of program feasibility; 2) considering two aspects of implementation: dosage and quality; and 3) testing these effects within the context of the combination of two evidence-based programs (the PAX GBG and MTP) with strong potential for scaling up and which have seldom been combined or tested together. Given that early-career teachers may have fewer resources (e.g., training, on the job experience) compared to more

experienced teachers, evidence-based programs for early-career teachers may be particularly useful resources for meeting the job demands they face.

4.1 Teachers' Occupational Health Did Not Predict Implementation

Regarding RQ_1, although the bivariate correlations indicated that burnout at baseline was negatively associated with the number of coaching cycles completed, this association did not hold in the more complex models. It is worth noting that neither stress, burnout, nor perceived ability at baseline predicted implementation dosage or quality as assessed at post-program. Thus, the hypothesis that greater occupational health at baseline would be associated with greater implementation dosage and quality of the interventions was not supported. The lack of significant associations between occupational health and implementation found in the current study is mostly inconsistent with previous findings of school-based intervention programs, including the PAX GBG, which suggested that higher levels of occupational health were associated with greater implementation dosage and quality (Domitrovich et al., 2015; Ransford et al., 2009; Wehby et al., 2012). These previous studies have included qualitatively different populations than the current study, including teachers of up to 5th grade, and those beyond their first three years in the classroom. Yet, results are consistent with existing studies of the PAX GBG which found no association between self-efficacy for behavior management and implementation (Domitrovich et al., 2015), and a recent study of early childhood educators which found inconsistent associations between stress and implementation (Clayback et al., 2022).

The non-significant associations between stress, burnout, perceived ability, and implementation outcomes in this study should be interpreted in the context of several key considerations. First, occupational health may yet be related to implementation, just the present study's assessment of stress, burnout, and perceived ability may not be the relevant occupational health indicators that are important for implementation. For example, perceived ability, a specific aspect of self-efficacy measured in this study, may be too nuanced; with previous research indicating that general self-efficacy is associated with the quality and dosage of school-based interventions (Kallestad & Olweus, 2003; Ransford et al., 2009), it may be that general self-efficacy or specific self-efficacy around implementing new programs, rather than perceived teaching ability, may be related to implementation. Specifically, positive dimensions of occupational health such as job satisfaction and feelings of personal accomplishment are known to be salient experiences for teachers (e.g., Hakanen et al., 2006; Maslach et al., 2001). These indices of occupational health were not measured in the present study but may influence teachers' implementation of intervention programs. Future research assessing positive indices of occupational health (e.g., general self-efficacy, job satisfaction, etc.) will shed light on this possibility.

Moreover, the current project was focused solely on early career teachers. Experiences of burnout may be less salient than other measures of occupational health in this population teachers given it results from experiences of chronic stress, which these teachers may not have had time to experience yet. In addition, there may be other factors besides occupational health that exert a greater influence on implementation quality for early career teachers, such as administrative support or openness to interventions. Future research should explore other such teacher-specific factors that may influence implementation quality among early career teachers (Domitrovich et al., 2008). Further, because these teachers are still actively developing their teaching practices, efforts to optimize implementation during this period may particularly impactful. For example, interventions addressing barriers to implementation could have a positive effect on these teachers' implementation of evidence-based programs across their career. It is also worth noting that the teachers in this study

tended to report relatively high levels of occupational health. For example, teachers reported high levels of ability with limited variation (M = 3.34; SD = 0.44), potentially precluding the opportunity to detect significant differences across the spectrum of ability. Levels of burnout in this sample (M = 2.71; SD = 0.69) were also lower than those in other studies (e.g., Roeser et al., 2022), perhaps due to their early career status. Regardless, the limited variability in these measures of occupational health may have also limited their predictive utility; it may be that higher levels of stress and burnout are necessary in order to impair implementation. It is also possible that the indicators of occupational health measured in this study may have effects on other domains of implementation described by Durlak & DuPre (2008) that were not assessed in this study (e.g., participant responsiveness). Finally, the moderation of the effect of stress on implementation quality by perceptions of feasibility suggests that the association between occupational health and implementation may be more nuanced than direct effects, a finding which we explore further in the subsequent sections.

4.2 Teachers' Perceptions of the Feasibility of the PAX GBG + MTP Program Predicted MTP Dosage

The hypothesis that greater perceptions of program feasibility would be associated with greater implementation was partially supported. Teachers' perceptions of how feasible the program was, an indicator of the social validity of the intervention, were not predictive of implementation quality, but were predictive of implementation dosage, assessed here as the number of coaching cycles completed. These results are consistent with existing evidence that positive perceptions of the program are associated with greater implementation (e.g., Clayback et al., 2022; Wehby et al., 2012). Implementation dosage is an important outcome to consider as existing literature has found dosage of MTP to be related to program outcomes (e.g., Pianta et al., 2014, 2022). This finding is informative for interventionists as it indicates that designing programs in ways that are simple to implement could be an effective strategy to increase sustained engagement in the program and subsequent targeted outcomes. Findings from successfully implemented school-based interventions have highlighted that school administration can be important champions for interventions (Forman et al., 2009). In this vein, it may be advantageous for school leadership to frame these programs as easy to use and easy to integrate into teaching, which could set the program up for success from the start (Forman et al., 2009).

Importantly, we measured perceptions of program feasibility alongside post-intervention assessments of implementation. This decision was made to not overburden participants with another survey in the middle of the school year. However, this design decision precludes definitive conclusions about the directionality of this association; yet, it is anticipated that teachers' perceptions of program feasibility preceded their implementation of the games and their attendance in coaching cycles. Finally, although we conceptualized teachers' perceptions of how easy the program elements were to use as an assessment of feasibility and social validity (Kazdin, 1977; Wolf, 1978), it could also be conceptualized as a component of implementation, namely participant responsiveness, a little studied dimension of implementation described in Durlak and DuPre (2008). The conceptualization of the construct underlies important design and analytic decisions, such as situating it as a predictor of implementation or an implementation outcome.

4.3 The Association between Stress and Implementation Quality was Moderated by Perceptions of Feasibility

Regarding RQ_2, a significant interaction between teacher-reported stress at baseline and perceptions of feasibility at post-intervention emerged in models predicting both instructional support and classroom organization, indicating that the effect of stress on implementation quality differed

according to perceptions of the feasibility of the program. Although our hypothesis regarding feasibility moderating the effect of stress on implementation was supported, the direction of these effects was contrary to our hypotheses. We hypothesized that the negative association between stress and implementation would be weaker for teachers who had positive perceptions of feasibility. Yet, results indicated that highly stressed teachers demonstrated greater instructional support and classroom organization when they found the program was *harder* to use compared to teachers who found it easy to use. For instructional support, we found that teachers reporting low levels of stress had higher implementation quality when they found the program easy to use compared to low-stress teachers who found it hard to use. The findings among low-stress teachers are consistent with previous literature, which has found that positive perceptions of social validity are associated with increased implementation (McNeill, 2019; Wehby et al., 2012). Contrary to previous research, the findings among highly stressed teachers may be capturing a particular subset of highly conscientious teachers who devoted more time and effort to learning and implementing the program, thus making it more difficult to use due to the high resource burden. It may also be the case that highly stressed teachers may have perceived the program as more valuable or useful due to its perceived complexity and difficulty, thus leading these teachers to implement increased instructional support and classroom management techniques. The complexity of these findings is aligned with recent evidence that the association between teachers' stress and the implementation of a mindfulness-based program for students was moderated by the amount of training they received (Braun et al., under review). Together, these results suggest that there is more nuance to these associations than suggested in conceptual models, such as the Prosocial Classroom Model, in that the effect of stress on implementation may differ according to other teacher-, school-, and program-specific factors (Dreer et al., 2017; Jennings & Greenberg, 2009). Results should also be interpreted in light of the timing of these measures, which is elaborated more in the Limitations and Future Research Directions section.

4.4 Teacher- and School-Level Demographics and Implementation

Although not a main focus of these analyses, the effects of teacher- and school-level demographics included as covariates yielded findings also worth discussing. Teacher-level demographics of grade level and years of teaching experience were primarily unrelated to teachers' implementation dosage and quality. The exception was that teachers of lower grade levels completed more MTP coaching cycles. Teaching in the lower grade levels, particularly, is highly relational and high-quality interactions with students are as important as didactic instruction (Burchinal et al., 2008; Pianta & Stuhlman, 2004). Teachers of younger students could have been more motivated to attended MTP coaching because the relational content was particularly salient given the age of their students.

At the school level, teachers working in schools where more (vs. fewer) students were eligible for FARMS played a higher number of games, indicating increased implementation dosage in these schools. At the same time, implementation quality across all CLASS domains was lower for teachers in schools with more (vs. fewer) students eligible for FARMS. These results are consistent with existing research demonstrating that students experiencing the greatest socioeconomic need have teachers with lower-quality interactions (e.g., St Clair & Stone, 2016). Similarly, teachers in larger schools were observed to have lower levels of classroom organization. These findings indicate that although teachers in schools with high levels of FARMS may recognize the need for such interventions and employ more PAX GBG games than their peers from other schools, the quality of their implementation of MTP may be lower. These findings highlight that predictors of implementation dosage are not necessarily redundant with predictors of implementation quality, suggesting that researchers should continue to investigate dimensions of implementation as related yet separate outcomes. These teacher- and school-level findings could be useful in identifying

teachers who may be at risk for a lower dosage of implementation and lower quality implementation of interventions.

4.5 Limitations and Future Research Directions

There were certain limitations of the perceived ability measure used in the current analysis, evidenced by the relatively low internal consistency of the measure (i.e., $\alpha = .69$). This may be due to the fact that the items were drawn from a scale intended to measure individual's motivations for becoming a teacher, such that the items only capture the teaching competence domain of self-efficacy. Based on these findings, future research should incorporate measures that assess both domains of self-efficacy in order to capture both internal and external influences on teachers' perceived self-efficacy. Despite this limitation, the current findings demonstrate the importance of incorporating task- and context-specific measures of self-efficacy when examining factors that influence implementation quality.

An assumption underlying the interpretations of the feasibility findings is that perceptions of feasibility were stable across the course of the intervention. As feasibility was only assessed at post-intervention in this study, we were unable to test the variability nor directionality of these effects. Perceptions of the feasibility, or more generally, the social validity, of interventions could shift over the course of the program (Clayback et al., 2022). Future research should administer measures of social validity throughout the intervention in order to understand the potentially bidirectional influence between social validity and implementation dosage and quality, and what might predict more favorable changes in social validity over the course of the intervention.

In addition, the current findings should be interpreted within the context of early career, elementary school teachers since the identified associations with implementation quality and dosage may be specific to this population. Furthermore, these findings should be contextualized within the sociodemographic makeup of the sample, given that the sample was predominantly white (80%) and female (93%). Future research should build upon these findings in order to clarify whether similar factors influence implementation among teachers who teach middle and high school, have a greater number of years of experience, and are from more sociodemographically diverse backgrounds. Future research could employ enriched samples to improve racial/ethnic and gender identity diversity in order to capture the broader experience of all teachers. Further, despite the early career status of teachers in this study, participants reported slightly higher averages of burnout than stress. Future research could continue to explore whether other indices of occupational health (e.g., burnout) may also interact with perceptions of feasibility to impact teachers' implementation of evidence-based programs.

The present study provides some support for the Prosocial Classroom Model and model of factors impacting the implementation of school-based interventions (Domitrovich et al., 2008; Jennings & Greenberg, 2009). However, the interactions between occupational health and perceptions of feasibility found in this study also highlight that those models may be too simplified for the complexity of school-based research. Based on these findings, future research should continue to explore the multitude of program-, teacher- and school-level factors that influence the quality of intervention implementation among teachers with a range of experience and across varying intervention programs.

4.6 Implications for Practice

Higher dosages of coaching cycles frequently lead to improved implementation fidelity and, ultimately, better student outcomes (Becker et al., 2013; Pas et al., 2022). Therefore, it is important for educators to be motivated to participate in coaching cycles. Results of this study indicate that teachers who perceived PAX GBG + MTP as feasible also participated in more coaching cycles. As such, efforts to increase perceptions of program feasibility may result in greater program dosage. One way to increase perceptions of feasibility is to ensure the program aligns with the school's core values (Forman et al., 2009). If teachers feel as though the program is a good "fit" to their own goals and philosophies, they are more likely to view the program in a positive way (Forman et al., 2009). Additionally, researchers may consider sharing findings regarding the positive perceptions of PAX GBG + MTP with teachers interested in implementing the program, as teachers respond well to learning new information from other teachers (Beahm et al., 2021; Forman et al., 2009).

Although the results provided no evidence that teachers' occupational health predicted their dosage and quality of implementation of the PAX GBG + MTP, we are cautious in our interpretation of these findings given that these associations have been found in previous research (e.g., Domitrovich et al., 2015; Ransford et al., 2009). Regardless of its predictive utility for teachers' implementation of evidence-based programs, experiences of occupational health are salient and meaningful experiences for teachers. When indicating that teachers are suffering from poor occupational health, schools should be motivated to intervene not just because poor occupational health could impact teaching practices and implementation of evidence-based programs but also from a compassionate perspective to alleviate suffering.

The teacher and school demographics included as covariates in this study shed light on who and in what contexts implementation is notably high. Given that teachers of higher grades completed fewer coaching cycles, these teachers may be in need of greater support from coaching staff in order to increase engagement in the program. Although teachers in schools where a higher percentage of students were eligible for FARMS had greater implementation dosage, they simultaneously had lower implementation quality. These findings indicate that these teachers may be in need of additional support, potentially beyond the existing scope of the PAX GBG + MTP program, in order to reach high-quality implementation of PAX GBG + MTP. Taken together, future research should continue to explore teacher and school characteristics that influence both implementation quality and dosage in order to improve student and teacher outcomes.

4.7 Conclusion

The present study advances the field of implementation science in school-based research by investigating the association between teachers' occupational health and perceptions of program feasibility in relation to the dosage and quality of implementation of two evidence-based programs implemented together. Results provided some support for conceptual models of factors that influence the implementation of school-based interventions (Domitrovich et al., 2015), and highlight the complexity of optimizing implementation in this context. With the growing emphasis on the implementation of evidence-based programs in schools, efforts to scale-up such programs with fidelity should continue to attend to teacher- and school-level contextual factors. This study provides additional empirical evidence of particular characteristics that that may hinder implementation, while identifying potential factors such as program feasibility that may be important for facilitating the implementation of evidence-based programs.

5 References

Allen, J. P., Hafen, C. A., Gregory, A. C., Mikami, A. Y., & Pianta, R. (2015). Enhancing secondary school instruction and student achievement: Replication and extension of the My Teaching Partner-Secondary intervention. *Journal of Research on Educational Effectiveness, 8*(4), 475–489. https://doi.org/10.1080/19345747.2015.1017680

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science, 333*(6045), 1034–1037. https://doi.org/10.1126/science.1207998.An

Bakker, A. B., & Schaufeli, W. B. (2000). Burnout contagion processes among teachers. *Journal of Applied Social Psychology, 30*(11), 2289–2308.

Bakker, A B, & Rodríguez-Muñoz, A. (2010). Positive occupational health psychology. In *Occupational Health Psychology* (pp. 194–224). http://www.scopus.com/inward/record.url?eid=2-s2.0-84857025810&partnerID=40&md5=4f4c3480384cdcd09bc88b724e4ad08d

Bakker, Arnold B, & Demerouti, E. (2007). The Job Demands-Resources model: State of the art. *Journal of Managerial Psychology, 22*(3), 309–328. https://doi.org/10.1108/02683940710733115

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good Behavior Game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*(2), 119–124.

Beahm, L. A., Yan, X., & Cook, B. G. (2021). Where do teachers go for behavior management strategies? *Education and Treatment of Children, 44*, 201–213.

Becker, K. D., Bradshaw, C. P., Domitrovich, C., & Ialongo, N. S. (2013). Coaching teachers to improve implementation of the Good Behavior Game. *Administration and Policy in Mental Health, 40*, 482–493. https://doi.org/10.1007/s10488-013-0482-8

Boekaerts, M. (1993). Being concerned with well-being and with learning. *Educational Psychologist, 28*(2), 149–167.

Bowman-Perrott, L., Burke, M. D., Zaini, S., Zhang, N., & Vannest, K. (2016). Promoting positive behavior using the Good Behavior Game: A meta-analysis of single-case research. *Journal of Positive Behavior Interventions, 18*(3), 180–190. https://doi.org/10.1177/1098300715592355

Breeman, L. D., Wubbels, T., Lier, P. A. C. Van, Verhulst, F. C., Ende, J. Van Der, Maras, A., Hopman, J. A. B., & Tick, N. T. (2015). Teacher characteristics, social classroom relationships, and children's social, emotional, and behavioral classroom adjustment in special education. *Journal of School Psychology, 53*(1), 87–103. https://doi.org/10.1016/j.jsp.2014.11.005

Bronfenbrenner, U. (1994). Ecological models of human development. In International Encyclopedia of Education (2nd ed., Vol. 3, pp. 37–43). Elsevier.

Brouwers, H., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education, 16*, 239–253.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science, 12*(3), 140–153. https://doi.org/10.1080/10888690802199418

Chang, M. L., & Davis, H. A. (2009). Understanding the role of teacher appraisals in shaping the dynamics of their relationships with students: Deconstructing teachers' judgments of disruptive behavior/students. In Advances in Teacher Emotion Research: The Impact on Teachers' Lives. https://doi.org/10.1007/978-1-4419-0564-2_6

Clayback, K. A., Williford, A. P., & Vitiello, V. E. (2022). Identifying teacher beliefs and experiences associated with curriculum implementation fidelity in early childhood education. *Prevention Science*, 1–12. https://doi.org/10.1007/s11121-022-01414-z

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Buckley, J. A., Olin, S., Romanelli, L. H., Leaf, P. J., Greenberg, M. T., & Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion, 1*(3), 6–28.

Domitrovich, C. E., Gest, S. D., Jones, D., Gill, S., & Derousie, R. M. S. (2010). Implementation quality: Lessons learned in the context of the Head Start REDI trial. *Early Childhood Research Quarterly, 25*, 284–298. https://doi.org/10.1016/j.ecresq.2010.04.001

Domitrovich, C. E., Pas, E. T., Bradshaw, C. P., Becker, K. D., Keperling, J. P., Embry, D. D., & Ialongo, N. (2015). Individual and school organizational factors that influence implementation of the PAX good behavior game intervention. *Prevention Science, 16*(8), 1064–1074. https://doi.org/10.1007/s11121-015-0557-8

Dreer, B., Dietrich, J., & Kracke, B. (2017). From in-service teacher development to school improvement: Factors of learning transfer in teacher education. *Teacher Development, 4530*, 1–17. https://doi.org/10.1080/13664530.2016.1224774

Durlak, J. A. (2015). What everyone should know about program implementation. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), Handbook of Social and Emotional Learning (pp. 395–405). Guilford.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3–4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Eccles, M. P., & Mittman, B. S. (2006). Welcome to Implementation Science. *Implementation Science, 3*, 1–3. https://doi.org/10.1186/1748-5908-1-1

Embry, D. D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review, 5*(4), 273–297.

Evertson, C. M., & Weinstein, C. S. (Eds.). (2011). Handbook of classroom management: Research, Practice, and Contemporary Issues. Routledge.

Forman, S. G., Olin, Æ. S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health, 1*, 26–36. https://doi.org/10.1007/s12310-008-9002-5

Freeman, J., Simonsen, B., Briere, D. E., & MacSuga-Gage, A. S. (2014). Preservice teacher training in classroom management: A review of state accreditation policy and teacher preparation programs. *Teacher Education and Special Education, 37*(2), 106–120. https://doi.org/10.1177/0888406413507002

García-Carmona, M., Marín, M. D., & Aguayo, R. (2019). Burnout syndrome in secondary school teachers: A systematic review and meta-analysis. *Social Psychology of Education, 22*(1), 189–208. https://doi.org/10.1007/s11218-018-9471-9

Gould, L. F., Dariotis, J. K., Greenberg, M. T., & Mendelson, T. (2016). Assessing fidelity of implementation (FOI) for school-based mindfulness and yoga interventions: A systematic review. *Mindfulness, 7*, 5–33. https://doi.org/10.1007/s12671-015-0395-6

Gregory, A., Allen, J. P., Kimami, A., Hafen, C. A., & Pianta, R. C. (2014). Effects of a professional development program on behavioral engagement of students in middle and high school. *Psychology in the Schools, 51*(2), 143–163. https://doi.org/10.1002/pits

Hakanen, J. J., Bakker, A. B., & Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *Journal of School Psychology, 43*, 495–513. https://doi.org/10.1016/j.jsp.2005.11.001

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., Brown, J. L., Cappella, E., Atkins, M., Rivers, S. E., Brackett, M. A., & Hamagami, A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms.

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology, 33*(6), 665–679. https://doi.org/10.1007/s10802-005-7646-2

Herman, K. C., Prewett, S. L., Eddy, C. L., Savala, A., & Reinke, W. M. (2020). Profiles of middle school teacher stress and coping: Concurrent and prospective correlates. *Journal of School Psychology, 78*, 54–68. https://doi.org/10.1016/j.jsp.2019.11.003

Hicks, T. B., Shahidullah, J. D., Carlson, J. S., & Palejwala, M. H. (2014). Nationally certified school psychologists' use and reported barriers to using evidence-based interventions in sc ... *School Psychology Quarterly*.

Hopman, J. A. B., Lier, P. A. C. Van, Ende, J. Van Der, Struiksma, C., Wubbels, T., Verhulst, F. C., Maras, A., Breeman, L. D., & Tick, N. T. (2018). Impact of the Good Behavior Game on special education teachers. *Teachers and Teaching: Theory and Practice, 24*(4), 350–368. https://doi.org/10.1080/13540602.2017.1379389

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165–179.

Ialongo, N. S., Domitrovich, C., Embry, D., Greenberg, M. T., Lawson, A., Becker, K. D., & Bradshaw, C. P. (2019). A randomized controlled trial of the combination of two school-based universal prevention programs. *Developmental Psychology*, 1–13. https://doi.org/10.1037/dev0000715

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research, 79*(1), 491–525. https://doi.org/10.3102/0034654308325693

Johansson, M., Biglan, A., & Embry, D. (2020). The PAX Good Behavior Game: One model for evolving a more nurturing society. *Clinical Child and Family Psychology Review, 23*(4), 462–482. https://doi.org/10.1007/s10567-020-00323-3

Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P. J., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology, 20*(2), 178–187. https://doi.org/10.1108/02683940510579803

Kallestad, J. H., & Olweus, D. (2003). Predicting teachers' and schools' implementation of the Olweus bullying prevention program: A multi-level study. *Prevention & Treatment, 6*(1).

Kam, C.-M., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science, 4*(1), 55–63. papers3://publication/uuid/4E1D3845-11A2-4146-BA55-768FB6ABAFBD

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification, 1*(4), 427–452.

Kazdin, A. E. (1993). Adolescent mental health. *American Psychologist, 48*(2), 127–141. http://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research, 86*(3), 643–680. https://doi.org/10.3102/0034654315626799

Kukla-Acevedo, S. (2009). Leavers, movers, and stayers: The role of workplace conditions in teacher mobility decisions. *The Journal of Educational Research, 102*(6), 443–452. https://doi.org/10.3200/JOER.102.6.443-452

Larson, M., Cook, C. R., Fiat, A., & Lyon, A. R. (2018). Stressed teachers don't make good implementers: Examining the interplay between stress reduction and intervention fidelity. *School Mental Health, 10*(1), 61–76. https://doi.org/10.1007/s12310-018-9250-y

Leidig, T., Casale, G., Wilbert, J., Hennemann, T., Volpe, R. J., Briesch, A., Grosche, M., & Grosche, M. (2022). Individual, generalized, and moderated effects of the good behavior game on at-

risk primary school students: A multi-level multiple baseline study using behavioral progress monitoring. *Frontiers in Education*, July, 1–15. https://doi.org/10.3389/feduc.2022.917138

Leiter, M. P., & Durup, J. (1994). The discriminant validity of burnout and depression: A confirmatory factor analytic study. *Anxiety, Stress, and Coping*, *7*(4), 357–373.

Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, E. W. G. (2014). On the joys of missing data. *Journal of Pediatric Psychology*, *39*(2), 151–162.

Mantzicopoulos, P., French, B. F., Patrick, H., Watson, J. S., & Ahn, I. (2018). The stability of kindergarten teachers' effectiveness: A generalizability study comparing the Framework For Teaching and the Classroom Assessment Scoring System. *Educational Assessment*, *23*(1), 24–46. https://doi.org/10.1080/10627197.2017.1408407

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). Maslach burnout inventory manual.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*, 397–422.

Maslach, C., & Jackson, S. (1981). The measurement of experienced Burnout. *Journal of Occupational Behavior*, *2*(2), 99–113. https://doi.org/10.1002/job.4030020205

McGoey, K. E., Rispoli, K. M., Venesky, L. G., Schaffner, K. F., McGuirk, L., & Marshall, S. (2014). A preliminary investigation into teacher perceptions of the barriers to behavior intervention implementation. *Journal of Applied School Psychology*, *30*(4), 375–390. https://doi.org/10.1080/15377903.2014.950441

McNeill, J. (2019). Social validity and teachers' use of evidence-based practices for autism. *Journal of Autism and Developmental Disorders*, *49*(11), 4585–4594. https://doi.org/10.1007/s10803-019-04190-y

Mikami, A. Y., Gregory, A., Allen, J. P., Pianta, R. C., & Lun, J. (2011). Effects of a teacher professional development intervention on peer relationships in secondary classrooms. *School Psychology Review*, *40*(3), 367–385. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3379816&tool=pmcentrez&rendertype=abstract

National Institute for Occupational Safety and Health. (1999). Stress at work.

Pas, E. T., Borden, L., Debnam, K. J., Lucia, D. De, & Bradshaw, C. P. (2022). Exploring profiles of coaches' fidelity to Double Check's motivational interviewing-embedded coaching: Outcomes associated with fidelity. *Journal of School Psychology*, *92*(January), 285–298. https://doi.org/10.1016/j.jsp.2022.04.003

Pianta, R. C., La Paro, K. M., & Hamre, B. (2008a). Classroom Assessment Scoring System: Manual K-3. Paul H Brookes Publishing.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008b). Classroom assessment scoring system (CLASS): Manual pre-K. Paul H. Brookes Publishing Company.

Pianta, R. C., & Stuhlman, M. W. (2004). Teacher-child relationships and children's success in the first years of school. *School Psychology Review, 33*(3), 444–458.

Pianta, Robert C, Decoster, J., Cabell, S., Burchinal, M., Hamre, B. K., Downer, J., Locasale-crouch, J., Williford, A., & Howes, C. (2014). Dose-response relations between preschool teachers' exposure to components of professional development and increases in quality of their interactions with children. *Early Childhood Research Quarterly, 29*(4), 499–508. https://doi.org/10.1016/j.ecresq.2014.06.001

Pianta, Robert C, Lipscomb, D., & Ruzek, E. (2022). Indirect effects of coaching on pre-K students' engagement and literacy skill as a function of improved teacher-student interaction. *Journal of School Psychology, 91*(December 2021), 65–80. https://doi.org/10.1016/j.jsp.2021.12.003

Pianta, Robert C, Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly, 23*, 431–451. https://doi.org/10.1016/j.ecresq.2008.02.001

Pines, A. M., & Keinan, G. (2005). Stress and burnout: The significant difference. *Personality and Individual Difference, 39*, 625–635. https://doi.org/10.1016/j.paid.2005.02.009

Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., & Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on the implementation of a social and emotional learning curriculum. *School Psychology Review, 38*(4), 510–532.

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. https://doi.org/10.1177/1098300712459079

Ringwalt, C., Vincus, A. A., Hanley, S., Ennett, S. T., Bowling, J. M., & Rohrbach, L. A. (2009). The prevalence of evidence-based drug use prevention curricula in U.S. middle schools in 2005. *Prevention Science, 10*(1), 33–40. https://doi.org/10.1007/s11121-008-0112-y

Roeser, R. W., Mashburn, A. J., Skinner, E. A., Taylor, C., Choles, J. R., Rickert, N. P., & Sorenson, J. (2021). Mindfulness training improves middle school teachers' occupational health, well-being, and interactions with students in their most stressful classrooms. *Journal of Educational Psychology, 114*(2), 408–425.

Roeser, R. W., Skinner, E., Beers, J., & Jennings, P. A. (2012). Mindfulness training and teachers' professional development: An emerging area of research and practice. *Child Development Perspectives, 6*(2), 167–173. https://doi.org/10.1111/j.1750-8606.2012.00238.x

Rohrbach, L. A., Graham, J. W., & Hansen, W. B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Preventive Medicine, 22*(2), 237–260.

Ross, S. W., Romer, N., & Horner, R. H. (2012). Teacher well-being and the implementation of school-wide Positive Behavior Interventions and Supports. *Journal of Positive Behavior Interventions, 14*(2), 118–128. https://doi.org/10.1177/1098300711413820

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36.

Salmela-Aro, K., Hietajärvi, L., & Lonka, K. (2019). Work burnout and engagement profiles among teachers. *Frontiers in Psychology, 10*(October), 1–8. https://doi.org/10.3389/fpsyg.2019.02254

Sanetti, L., & Collier-Meek, M. (2019). Increasing implementation science literacy to address the research-to-practice gap in school psychology. *School Psychology Research, 76*(July), 33–47. https://doi.org/10.1016/j.jsp.2019.07.008

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Treatment integrity assessment in the schools: An evaluation of the Treatment Integrity Planning Protocol. *School Psychology Quarterly, 24*(1), 24–35.

Sanetti, Lisa M Hagermoser, Dobey, L. M., & Gallucci, J. (2014). Treatment integrity of interventions with children in School Psychology International from 1995 – 2010. *School Psychology International, 35*(4), 370–383. https://doi.org/10.1177/0143034313476399

Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education, 26,* 1059–1069. https://doi.org/10.1016/j.tate.2009.11.001

Skaalvik, E. M., & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological Reports: Employment Psychology and Marketing, 114*(1), 68–77. https://doi.org/10.2466/14.02.PR0.114k14w0

Smith, S., Barajas, K., Ellis, B., Moore, C., McCauley, S., & Reichow, B. (2021). A meta-analytic review of randomized controlled trials of the Good Behavior Game. *Behavior Modification, 45*(4), 641–666. https://doi.org/10.1177/0145445519878670

St Clair, L., & Stone, T. (2016). Who gets the better educators in afterschool? An analysis of teaching and learning interactions and student economic status. *School Community Journal, 26*(2), 71–81.

Swars, S. L., Daane, C. J., & Giesen, J. (2006). Mathematics anxiety and mathematics teacher efficacy: What is the relationship in elementary preservice teachers? *School Science and Mathematics, 106*(7), 306–315.

Swift, L. E., Hubbard, J. A., Bookhout, M. K., Grassetti, S. N., Smith, M. A., & Morrow, M. T. (2017). Teacher factors contributing to dosage of the KiVa anti-bullying program. *Journal of School Psychology, 65*(August 2016), 102–115. https://doi.org/10.1016/j.jsp.2017.07.005

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development, 88*(4), 1156–1171. https://doi.org/10.1111/cdev.12864

Thierry, K. L., Vincent, R. L., Norris, K. (2022). Teacher-level predictors of the fidelity of implementation of a social-emotional learning curriculum. *Early Education and Development, 33*(1), 92–106. https://doi.org/10.1080/10409289.2020.1849896

Tolan, P., Molloy, L. E., Bradshaw, C. P., Downer, J. T., & Ialongo, N. (2020). Randomized trial testing the integration of the Good Behavior Game and MyTeachingPartner: The moderating role of distress among new teachers on student outcomes. *Journal of School Psychology, 78*(2020), 75–95. https://doi.org/10.1016/j.jsp.2019.12.002

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202–248.

van Horn, J. E., Taris, T. W., Schaufeli, W. B., & Schreurs, P. J. G. (2004). The structure of occupational well-being: A study among Dutch teachers. *Journal of Occupational and Organizational Psychology, 77,* 365–375.

Watt, H. M. G., & Richardson, P. W. (2007). Motivational factors influencing teaching as a career choice: Development and validation of the FIT-Choice scale. *The Journal of Experimental Education, 75*(3), 167–202. https://doi.org/10.3200/JEXE.75.3.167-202

Wehby, J. H., Maggin, D. M., Partin, T. C. M., & Robertson, R. (2012). The impact of working alliance, social validity, and teacher burnout on implementation fidelity of the Good Behavior Game. *School Mental Health, 4*, 22–33. https://doi.org/10.1007/s12310-011-9067-4

Wei, R. C., Darling-Hammond, L., & Adamson, F. (2010). Professional development in the United States: Trends and challenges.

What Works Clearinghouse. (2017). What Works Clearinghouse Standards Handbook. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_standards_handbook_v4_draft.pdf

Witkiewitz, K., Falk, D. E., Kranzler, H. R., Litten, R. Z., & Hallgren, K. A. (2014). Methods to analyze treatment effects in the presence of missing data for a continuous heavy drinking outcome measure when participants drop out from treatment in alcohol clinical trials. *Alcoholism: Clinical and Experimental Research, 38*(11), 2826–2834. https://doi.org/10.1111/acer.12543

Witt, J. C., Noell, G. H., LaFleur, L. H., & Mortenson, B. P. (1997). Teacher use of interventions in general education settings: measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis, 30,* 693–696.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*(2), 203–214.

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research, 86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

SSB: Conceptualization of the Paper; Formal Analysis; Writing-Original Draft

CPB: Conceptualization of the Project; Funding Acquisition; Writing-Review and Editing

LAB: Writing-Original Draft

ACB: Writing-Original Draft

JTD: Conceptualization of the Project; Funding Acquisition; Writing-Review and Editing

NSI: Conceptualization of the Project; Funding Acquisition; Writing-Review and Editing

PHT: Conceptualization of the Project; Funding Acquisition; Writing-Review and Editing

8 Funding

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305A190162 and R305A130107 (to the University of Virginia). The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education.

9 Acknowledgments

The authors would like to thank the coaches who supported implementation of this project and April Lawson for her assistance with project management.

10 Tables

10.1 Table 1: Descriptive Statistics

	N	*Missing*	Mean	*SD*	Minimum	Maximum	Range
Implementation Outcomes							
Implementation Dosage							
Coaching Cycles Completed	86	8%	8.24	2.3	1	11	10
Number of Games Played	83	11%	9.16	2.81	2.11	15.43	13.32
Quality of Implementation							
Emotional Support	83	11%	4.88	0.68	2.39	6.25	3.86
Classroom Organization	83	11%	5.04	0.71	3.38	6.28	2.90
Instructional Support	83	11%	2.35	0.65	1.33	4.22	2.89
Occupational Health (Time 1)							
Job Stress	91	3%	2.63	0.61	1.00	3.80	2.80
Burnout	91	3%	2.71	0.69	1.00	4.00	3.00
Perceived Ability	91	3%	3.34	0.44	2.00	4.00	2.00
Feasibility (Time 2)							
Easy to Use	78	16%	3.16	0.56	1.60	4.00	2.40

10.2 Table 2: Bivariate Correlations Among Study Measures

		Program Implementation					Occupational Health			Feasibility	Demographics			
		Implementation Dosage		Quality of Implementation							Teacher-Level		School-Level	
		1	2	3	4	5	6	7	8	9	10	11	12	13
Implementation														
	Implementation Dosage													
1	Coaching Cycles Completed	-												
2	Games Played	**0.26**	-											
	Quality of Implementation													
3	Emotional Support	0.19	-0.09	-										
4	Classroom Organization	0.08	0.03	**0.52**	-									
5	Instructional Support	0.05	-0.11	**0.75**	**0.46**	-								
Occupational Health (Time 1)														
6	Job Stress	-0.17	-0.02	0.12	0.13	0.00	-							
7	Burnout	**-0.21**	0.00	0.06	0.17	-0.08	**0.84**	-						
8	Perceived Ability	0.10	0.03	-0.04	-0.06	-0.06	**-0.25**	**-0.29**	-					
Feasibility (Time 2)														
9	Feasibility	**0.36**	-0.02	-0.01	0.00	-0.06	0.05	-0.05	0.08	-				
Demographics														
	Teacher-Level													
10	Grade	-0.15	-0.05	0.09	0.16	0.14	0.17	0.19	-0.12	0.20	-			
11	Years Teaching	-0.16	0.05	0.13	0.01	0.03	0.19	0.15	-0.04	-0.12	-0.09	-		
	School-Level													
12	Enrollment	-0.02	-0.10	0.13	-0.07	-0.10	-0.01	0.05	0.08	-0.16	0.04	0.09	-	
13	Eligible for FARMS	-0.10	**0.35**	**-0.25**	-0.14	**-0.31**	0.08	0.05	0.18	-0.13	**-0.21**	0.12	-0.15	-

Note. Bold indicates significant at $p < .05$.

10.3 Table 3: Main Effects Models: Predicting Implementation Dosage and Quality of PAX GBG + MTP

	Implementation Dosage				Implementation Quality					
	Coaching Cycles Completed		Number of Games Played		Emotional Support		Instructional Support		Classroom Organization	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	9.10 *	0.41	9.13 *	0.50	4.74 *	0.12	2.31 *	0.15	4.89 *	0.11
Occupational Health										
Job Stress	-0.08	0.53	-0.38	0.86	0.13	0.21	0.07	0.28	0.18	0.22
Burnout	-0.35	0.55	0.16	0.87	0.00	0.17	0.16	0.22	-0.14	0.16
Perceived Ability	-0.01	0.61	-0.29	0.70	0.09	0.19	0.04	0.15	-0.03	0.12
Feasibility										
Feasibility	1.66 *	0.53	0.11	0.51	-0.07	0.13	-0.07	0.13	-0.20	0.15
Demographics & Covariates										
Teacher-Level										
Grade	-0.43 *	0.19	0.07	0.31	0.07	0.05	0.04	0.07	0.11 +	0.06
Years Teaching	-0.29	0.32	-0.08	0.34	0.09	0.08	0.00	0.09	0.00	0.09
CLASS (where appropriate)					0.40 *	0.12	0.40 *	0.12	0.34 *	0.09
School-Level										
Enrollment	0.12	0.20	-0.16	0.34	0.04	0.06	-0.09	0.07	-0.11 *	0.05
FARMS	-0.44	1.01	5.39 *	1.47	-0.66 +	0.39	-0.27	0.32	-0.74 *	0.34
R Squared	0.24		0.15		0.29		0.17		0.30	

Note. * indicates significant at $p < .05$, + indicates significant at $p < .10$. Cohort was omitted in final models because its inclusion did not substantively change the pattern of results.

10.4 Table 4: Moderation Models: Predicting Implementation Dosage and Quality of PAX GBG + MTP

	Implementation Dosage				Implementation Quality					
	Coaching Cycles Completed		Number of Games Played		Emotional Support		Instructional Support		Classroom Organization	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	9.12 *	0.40	9.16 *	0.50	4.72 *	0.12	2.28 *	0.15	4.86 *	0.12
Occupational Health										
Job Stress	-0.05	0.53	-0.32	0.89	0.09	0.21	-0.01	0.27	0.13	0.22
Burnout	-0.35	0.56	0.14	0.88	0.02	0.17	0.18	0.22	-0.13	0.16
Perceived Ability	-0.01	0.61	-0.28	0.70	0.08	0.19	0.03	0.16	-0.05	0.12
Feasibility										
Feasibility	1.64 *	0.51	0.06	0.53	-0.04	0.12	0.00	0.13	-0.16	0.15
Interactions										
Stress*Feasibility	0.17	1.04	0.51	0.86	-0.33	0.27	-0.77 *	0.26	-0.41 +	0.24
Demographics & Covariates										
Teacher-Level										
Grade	-0.44 *	0.18	0.05	0.31	0.08	0.06	0.05	0.07	0.12 *	0.06
Years Teaching	-0.30	0.31	-0.11	0.34	0.11	0.08	0.03	0.09	0.02	0.09
CLASS (where appropriate)					0.40 *	0.11	0.48 *	0.11	0.35 *	0.09
School-Level										
Enrollment	0.12	0.20	-0.16	0.35	0.03	0.06	-0.11 +	0.07	-0.11 *	0.05
FARMS	-0.47	1.02	5.32 *	1.43	-0.61	0.40	-0.10	0.34	-0.67 +	0.36
R Squared	0.24		0.15		0.30		0.26		0.33	

Note. * indicates significant at $p < .05$, + indicates significant at $p < .10$. Cohort was omitted in final models because its inclusion did not substantively change the pattern of results.