Developmental Psychology

A Randomized Controlled Trial of the Combination of Two School-Based Universal Preventive Interventions

Nicholas S. Ialongo, Celene Domitrovich, Dennis Embry, Mark Greenberg, April Lawson, Kimberly D. Becker, and Catherine Bradshaw

Online First Publication, March 21, 2019. http://dx.doi.org/10.1037/dev0000715

American Psychological Association

Developmental Psychology

0012-1649/19/$12.00

http://dx.doi.org/10.1037/dev0000715

A Randomized Controlled Trial of the Combination of Two School-Based Universal Preventive Interventions

Nicholas S. Ialongo
Johns Hopkins Bloomberg School of Public Health

Celene Domitrovich
Georgetown University School of Medicine

Dennis Embry
Paxis Institute, Tucson, Arizona

Mark Greenberg
Pennsylvania State University

April Lawson
Johns Hopkins Bloomberg School of Public Health

Kimberly D. Becker
University of South Carolina

Catherine Bradshaw
University of Virginia

The Good Behavior Game (GBG, Barrish, Saunders, & Wolf, 1969) and the PATHS Curriculum (Promoting Alternative Thinking Strategies; Greenberg, Kusche, Cook, & Quamma, 1995) represent 2 universal, elementary school, preventive interventions which have been shown in large-scale, randomized controlled trials to have an immediate and beneficial impact (GBG, Dolan et al., 1993; PATHS, Conduct Problems Prevention Research Group [CPPRG], 1999) on aggressive/disruptive and off-task behavior. Importantly, both risk behaviors are strong predictors of antisocial behavior, drug abuse, and low educational and occupational attainment in adolescence and young adulthood (Kellam et al., 2008). What has yet to be explored within a randomized controlled trial is whether the combination of these interventions would yield significantly greater impact on aggressive/disruptive and off-task behavior than the GBG alone. One reason for expecting additive if not synergistic effects as a result of combining the two interventions is that the GBG, by increasing attention to task and reducing disruptive behavior in the classroom, may facilitate the acquisition of the emotion regulation, social problem-solving, and conflict resolution skills taught in PATHS. To that end, a group randomized, effectiveness trial was carried out, wherein 27 schools were randomly assigned to one of 3 conditions, (a) the PAX GBG Alone (Embry, Staatemeier, Richardson, Lauger, & Mitich, 2003), (b) PATHS to PAX (that is, the PAX GBG + PATHS), or (c) a standard setting (control) condition. Classroom observations and teacher ratings of student behavior were carried out at pretest and 6 months later at posttest. Limited evidence of the superiority of the combined approach was found and potential reasons why and future directions are discussed.

Keywords: prevention, randomized trial, school-based, aggression

Durlak and colleagues (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Taylor, Oberle, Durlak, & Weissberg, 2017) have carried out what are likely the most extensive meta-analyses to date of the published findings on the proximal and distal effects of K-12 socioemotional interventions. As with their earlier review of the proximal effects of these interventions (Durlak et al., 2011),

Nicholas S. Ialongo, Department of Mental Health, Johns Hopkins Bloomberg School of Public Health; Celene Domitrovich, Department of Psychiatry, Georgetown University School of Medicine; Dennis Embry, Paxis Institute, Tucson, Arizona; Mark Greenberg, College of Human Development & Family Studies, Pennsylvania State University; April Lawson, Department of Mental Health, Johns Hopkins Bloomberg School of Public Health; Kimberly D. Becker, Department of Psychology, University of South Carolina; Catherine Bradshaw, Curry School of Education, University of Virginia.

This work was supported by a grant from the Institute of Education Sciences (R305A080326). We would like to acknowledge the significant contributions of Joseph Wehby, Vanderbilt Peabody College, to the development of the classroom observation protocol used and the training of the observers in its use. Celene Domitrovich is an author on the PATHS Curriculum and has a royalty agreement with Channing-Bete, Inc. This has been reviewed and managed by Penn State's Individual Conflict of Interest Committee. Dennis Embry is an author of the PAX Good Behavior Game and the president and senior scientist of PAXIS Institute, which owns the intellectual property of the PAX Good Behavior Game. PAXIS Institute supplies the materials, training and technical supports for the PAX Good Behavior Game. He did not participate in data collection or data analysis. Mark Greenberg is an author of the PATHS Curriculum and has a royalty agreement with Channing-Bete, Inc. He did not participate in data collection or data analysis.

Correspondence concerning this article should be addressed to Nicholas S. Ialongo, Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, 624 North Broadway, Baltimore, MD 21205. E-mail: nialong1@jhu.edu

their most recent meta-analysis (Taylor et al., 2017) of over 80 socioemotional interventions revealed a wide breadth of beneficial preventive intervention effects in the behavioral, academic, and mental health domains over follow-up periods as long as 15 years. The breadth and range of beneficial effects did not appear to vary as a function of students' race, socioeconomic background, or school location. Durlak and colleagues (2011) also provide estimates of the substantial economic benefits associated with these distal preventive effects.

The conceptual basis for expecting preventive effects from school-based socioemotional interventions is consistent with 2 major tenets of life-course developmental theory. The first is that success at an earlier stage of development forms the foundation for success at a later stage of development. The second tenet is that success or failure in any one developmental domain is likely influenced by success or failure in other domains. For example, social development is unlikely to proceed normally in the presence of a lag in language development (Rutter, 1988). Durlak and colleagues (Taylor et al., 2017) offer a framework for understanding the mechanisms by which socioemotional interventions result in preventive effects that is in keeping with these developmental tenets. More specifically, they propose that socioemotional interventions should serve to facilitate the development of students' socioemotional "assets" in the form of social and emotional skills and improved attitudes about themselves, others, and school. These assets are then theorized to translate into positive behavioral, academic, and mental health outcomes in later stages of development (Taylor et al., 2017).

Durlak and colleagues (Taylor et al., 2017)'s findings with regard to the distal benefits of school-based, socioemotional interventions are also in keeping with the considerable evidence from long-term longitudinal studies that aggressive/disruptive and off-task behaviors as early as entrance into elementary school predict later antisocial behavior, violence, substance abuse and low educational and occupational attainment in adolescence and young adulthood (Kellam et al., 2008; Petras et al., 2008). This evidence prompted a series of randomized trials of promising school-based universal preventive interventions targeting aggressive/disruptive behavior in elementary school that were carried out in a large, mid-Atlantic, urban school district. One of the interventions featured in these trials—the Good Behavior Game (Barrish et al., 1969)–was included in Durlak and colleagues' most recent meta-analysis (Taylor et al., 2017).

Besides the Good Behavior Game, which was aimed at aggressive-disruptive behavior, the first of these trials (Dolan et al., 1993) included the evaluation of an intervention that targeted poor school achievement. In terms of the immediate impacts of the interventions, the GBG resulted in significant—albeit modest—reductions relative to controls in aggressive/disruptive and off-task behavior based on independent observations by the end of first grade (Brown, 1993), whereas the academic intervention resulted in significant but modest improvement in standardized reading achievement (Dolan et al., 1993). Dolan et al. (1993) also reported significant beneficial impact of the GBG in terms of teacher ratings and peer nominations of aggressive/disruptive behavior in 1st grade. The GBG appeared to have its greatest proximal impact among children manifesting mild to moderate elevations in aggressive/disruptive behavior at pretest in the fall of first grade (Dolan et al., 1993).

With respect to longer term impact, Kellam and colleagues reported significant and beneficial impact of the GBG on antisocial behavior, violent crime, substance abuse/dependence, and high school graduation at ages 19–20 (Kellam et al., 2008; Petras et al., 2008). Consistent with the variation seen in the immediate impact of the GBG (Dolan et al., 1993), the greatest effects were found for participants with mild to moderate elevations in pretest levels of aggressive/disruptive behavior in 1st grade. The fact that students who manifested elevated pretest levels of the targeted risk behaviors benefited the most from the GBG is consistent with the universal nature of the intervention. That is, given universal interventions target the entire population regardless of initial level of risk behaviors, the expectation is that only those who manifest at least a modest to moderate level of risk will be in a position to benefit from the intervention (Greenberg & Abenavoli, 2016).

Subsequently, a second randomized trial was fielded to test whether the magnitude of the effects found in the first trial could be improved upon (Ialongo et al., 1999). One of the interventions evaluated represented the combination of the GBG with an academic intervention. The rationale for this was that whereas the GBG was associated with increased high school graduation rates relative to controls in the 1st trial (Kellam et al., 2008), the GBG only had minimal proximal impact on standardized achievement scores in first and second grades. On the other hand, the academic intervention had an impact on early achievement, but had only a modest to moderate significant crossover, or indirect, effect on aggressive/disruptive behavior. Consequently, it was reasoned that both academic achievement and off-task and aggressive/disruptive behaviors needed to be targeted in a single intervention.

In addition to combining the GBG with an academic intervention in this second trial, a universal, family school partnership (FSP) intervention was developed and fielded to contrast with the combined GBG/academic intervention. Like the GBG/academic intervention, the proximal targets of the FSP intervention were poor achievement and aggressive/disruptive and off-task behavior in 1st grade. Significant proximal (elementary school; Ialongo et al., 1999; Petras, Masyn, & Ialongo, 2011), intermediate (middle school; Ialongo, Poduska, Werthamer, & Kellam, 2001; Petras et al., 2011), and longer term (high school/early adulthood) intervention impacts were found on aggressive/disruptive behavior/conduct problems (Petras et al., 2011) and academic outcomes (Bradshaw, Zmuda, Kellam, & Ialongo, 2009). As in the 1st trial, the benefits of the interventions were more apparent for those students with elevated pretest levels of the targeted risk behaviors. Importantly, however, at least half of the students in upper end of the distribution of the targeted risk behaviors failed to demonstrate intervention impact. Also of note, Petras et al. (2011) provided some—albeit limited—evidence that students at the low end of the risk distribution benefit from the GBG universal programs by not progressing to elevated levels of conduct problems in middle and high school. But the preponderance of the findings to date with respect to randomized trials of the GBG have shown the greatest effects for the students at the upper end of the distribution of pretest risk factors.

To summarize, both the immediate (Ialongo et al., 1999) and longer term results of the second trial (e.g., Bradshaw et al., 2009; Petras et al., 2011) supported the hypothesis that the combination of the GBG with an academic intervention would yield greater impact in terms of effect size on academic achievement than that

seen in the 1st trial, wherein the GBG and the academic intervention were examined, separately (Dolan et al., 1993). However, greater impact was not seen in terms of off-task and aggressive/disruptive behaviors as a result of combining the GBG with an academic intervention. The family school partnership intervention did yield significant immediate benefits in terms of aggressive/disruptive behavior and academic achievement, but the breadth of the effects and their size were smaller than for the combination of the GBG and the academic intervention (Ialongo et al., 1999, 2001).

The Combination of the GBG and PATHS

Given the combination of the GBG with an academic intervention did not yield greater improvement in off-task and aggressive/disruptive behavior than that seen in the 1st trial of the GBG alone, particularly in terms of decreasing the proportion of nonresponders to the interventions among those with the highest level of the targeted risk behaviors, one logical next step would have been to combine the GBG + academic intervention condition with the FSP intervention, This would be in keeping with the finding that the latter did have an impact on aggressive/disruptive behavior—albeit modest. But the costs and logistics involved in mounting a universal family school partnership intervention proved daunting.

Consequently, a decision was made to pursue the combination of the GBG with a second and complementary, evidenced-based, universal preventive intervention: Promoting Alternative Thinking Strategies (PATHS, Greenberg et al., 1995). Like the Good Behavior Game, PATHS was included in Durlak and colleagues' meta-analyses (Durlak et al., 2011; Taylor et al., 2017) and both PATHS and the GBG have been found to yield a 60 + percent return on investment based on the economic benefits associated with their long term outcomes relative to their initial training and materials costs (Washington State Institute for Public Policy, 2017). Both interventions have also been found efficacious in economically disadvantaged, ethnic-minority populations (e.g., Ialongo et al., 1999; CPPRG, 1999).

With respect to the conceptual basis for the expected preventive benefits of the GBG and PATHS, Patterson and colleagues (Patterson, Reid, & Dishion, 1992) offer perhaps the most comprehensive life course, social–cognitive learning model of the development of antisocial behavior, substance abuse and academic and occupational failure. According to Patterson and colleagues' coercion theory, one major pathway to serious antisocial behavior, substance abuse, and academic and occupational failure begins in the toddler years, when parental success in teaching their child to interact within a normal range of compliance and aversive behavior is a prerequisite for the child's development of social survival skills. Alternatively, the parents' failure to consistently and effectively punish child coercive behavior during these formative years and to teach reasonable levels of compliance comprises the first step in a process that serves to "train" the child to become progressively more coercive and antisocial over time and development. The transition to elementary school also represents a crucial developmental milestone for the coercive child. If their elementary school teachers fail to consistently discipline such children in the face of their coercive behavior, the coercive cycle seen in the home will spill over to the classroom. Subsequently, the coercive and noncompliant child will not only be rejected by their parents, but by their teachers and well-adjusted peers over the course of elementary school. As a result, the coercive child will fail to develop the "survival skills"—or what Durlak and colleagues term "social assets"—necessary for academic, social and occupational success.

In keeping with Patterson et al.'s (1992) coercion theory and Greenberg and colleagues' ABC model of socioemotional development (Greenberg, Kusche, & Riggs, 2004), both the GBG and PATHS seek to reduce the early antecedent risk behaviors of aggressive/disruptive/coercive behavior and its distal correlates via changing teacher behavior. In the case of the GBG, the focus is on providing teachers with an efficient way of reinforcing the inhibition of aggressive/disruptive/coercive and off-task behavior in a "game" like context. Whereas with PATHS, teachers are trained to provide explicit instruction in the development of emotion regulation, self-control, social problem solving and conflict resolution skills in the context of weekly didactic lessons across the school year.

There are also a number of reasons why additive, if not synergistic effects, were expected as a result of combining the two interventions. First, the efficacy of PATHS in terms of reducing off-task and aggressive/disruptive behavior and increasing prosocial behavior, social competence, inhibitory control and verbal fluency in the elementary school years had been demonstrated in a series of randomized controlled studies (CPPRG, 1999, 2010; Greenberg et al., 2004; Greenberg & Kusche, 2006; Riggs, Greenberg, Kusché, & Pentz, 2006). Second, PATHS seeks to accomplish reductions in aggressive/disruptive behavior via teacher led instruction aimed at facilitating emotion regulation, self-control, social problem-solving, and conflict resolution skills (Greenberg et al., 1995), whereas the GBG is based on social learning principles and provides teachers with an efficient means of managing student aggressive/disruptive and off-task behavior via reinforcement from teachers and peers for the inhibition of these behaviors and the reinforcement of prosocial behaviors. The GBG, by increasing attention to task and reducing disruptive behavior in the classroom, may facilitate the acquisition of the emotion regulation, self-control, social problem-solving, and conflict resolution skills taught in PATHS. Third, the social learning based GBG may increase the likelihood that students' newly acquired skills would be appropriately prompted and reinforced by teachers and peers. Consequently, the PATHS skills would be better learned and more frequently employed. Fourth, the increased teacher and student success, as a result of combining PATHS and the GBG, should minimize teacher and student discouragement and subsequent failure to participate or comply fully with the intervention regimens.

In this paper, the proximal results of a 27-school randomized controlled trial of the PAX GBG + PATHS are described—or as referred to from hereon, PATHS to PAX. The version of the GBG used in this study was the PAX GBG (Embry et al., 2003), which as described below, incorporated a number of enhancements into the GBG training and intervention protocols employed in Dolan et al. (1993) and Ialongo et al. (1999). The 27 schools (Grades K-5) were assigned to one or three conditions: (a) Comparison/Control; (b) PAX GBG Alone; or (c) PATHS to PAX (PAX GBG + PATHS). Consistent with the above, it was hypothesized that relative to the control condition both interventions would result in significantly lower levels of aggressive/disruptive and off-task behaviors at posttest. Moreover, the effects of the PATHS to PAX condition would be significantly greater than the PAX GBG con-

dition alone in terms of not only aggressive/disruptive and off-task behaviors, but with respect to social competence and emotion regulation, given the direct and explicit focus of the PATHS component of PATHS to PAX on these domains.

In addition to the test of the relative efficacy of PATHS to PAX versus the PAX GBG Alone, a number of potential moderators of intervention impact were examined. These included the pretest level of outcomes and age/grade levels. With regard to the former and as previously noted, due to the universal nature of the interventions, it was hypothesized that those students who manifested at least a modest to moderate level of the targeted risk behaviors would be in a position to benefit the most in terms of intervention impact (Greenberg & Abenavoli, 2016). In terms of the latter, there have been no studies to date of variation in GBG and PATHS intervention outcomes as a function of age and/or grade level. In the case of the GBG, this may be due to the fact that virtually all of the GBG large scale randomized trials have only included one or two grade levels. Nevertheless, in line with Patterson et al.'s (1992) coercion theory, it seems reasonable to expect that GBG and PATHS intervention outcomes may vary as function of grade level. More specifically, it might be expected that due to a shorter history of reinforcement of aggressive/coercive behavior among early versus late elementary school students, the early elementary students' aggressive/coercive behavior may prove more malleable. Poorer intervention response among older children would also be consistent with the concept of developmental cascades as reflected in Patterson and colleagues' (1992) model of the development of antisocial behavior over the life-course. More specifically, it may become increasing more difficult with age for an intervention limited to a single context—the classroom—and implemented over a single school year to overcome the "cumulative consequences for development of the many interactions and transactions occurring in developing systems that result in spreading effects across levels, among domains at the same level, and across different systems or generation" (Masten & Cicchetti, 2010, p. 491). Alternatively, it may be the case that late elementary school students may be more advanced in terms of their social–cognitive development than their early elementary school counterparts and, as such, may be more likely to comprehend and translate into action the PATHS socio-emotional learning concepts being taught to them.

Method

Design

The design consisted of three cohorts of nine elementary schools, with schools randomly assigned to one of the three intervention conditions within each cohort: the PAX GBG only, PATHS to PAX and a control, or standard setting condition where teachers conducted their usual practice. All elementary schools in one, large, urban public school district were eligible to participate in the trial with the exception of charter schools, schools exclusively serving special education students, and schools that participated in pilot efforts aimed at determining the feasibility of combining the PAX GBG and PATHS. Just under 100 schools were sent a recruitment letter and a rolling admission process was followed to identify 27 schools willing to participate. Only two of the schools contacted refused to participate. The administrators of these two schools reported that their teachers were satisfied with their current classroom behavior management programs. The schools that agreed to participate were ranked in terms of the proportion of student suspensions in the prior school year and triads were formed based on schools closest in suspension rank. Three triads were randomly selected for inclusion in the study each year for three consecutive years. Schools were then randomized to one of the three intervention conditions within the triads. A priori stratification and matching have been recommended as effective strategies to limit selection-related threats to internal validity (e.g., selection, differential maturation, differential history) in GRTs (Murray, 1998).

Parents provided written consent for their children to participate in the evaluation of the trial outcomes following procedures that were approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board (IRB Protocol # H.33.04.07.22.B).

Sample Demographics: Students

Eligible participants included students enrolled in K-5 classrooms in each of the 9 participating schools at the beginning of the school year. Across the three cohorts, there were a total of 7,024 students enrolled in the participating schools during our baseline, or pretest, fielding period. Of the total eligible, we obtained written parent consent for 79.9% ($N = 5611$); 7.1% refused participation and 12.9% did not respond to the consent request during the fielding period. Of the 5611 enrolled students, 50.4% were male, 89.6% were African American, and 86.5% received free or reduced lunch. The mean grade level was 2.36. The demographic profiles of the students enrolled in the study were comparable to the overall profiles of the participating schools in terms of gender, ethnicity and free and reduced lunch status (FARMS; a proxy for family income), which were 50.81% male, 87.93% African American, and 85.96% FARMS eligible. Figure 1 illustrates the demographics of the enrolled students by intervention condition.

Sample Demographics: Teachers

A total of 331 primary teachers from the 27 participating schools were enrolled in the trial. A teacher was considered a primary teacher for a given classroom if s/he was assigned as the classroom's homeroom teacher at the beginning of the baseline data collection period, and/or if the teacher spent more than 75% of the school year as the homeroom teacher. Of the 331 primary teachers, 118 (35.6%) were in control condition schools, 116 (35.0%) were in PAX Alone condition schools, and 97 (29.3%) were in PATHS to PAX condition schools. The majority of the teachers were female (88.2%) and obtained regular or standard teaching certificates (79.5%). More than half of the teachers (52.0%) held graduate degrees and 60.5% had at least five or more years of teaching experience. Fifty-seven percent of teachers taught early elementary grades, kindergarten through second, and 43% taught grades three through five.

Interventions

We describe below each of the interventions separately and then their integration into PATHS to PAX.

The PAX Good Behavior Game (PAX GBG). The GBG, originally developed by Barrish et al. (1969), allows teachers to

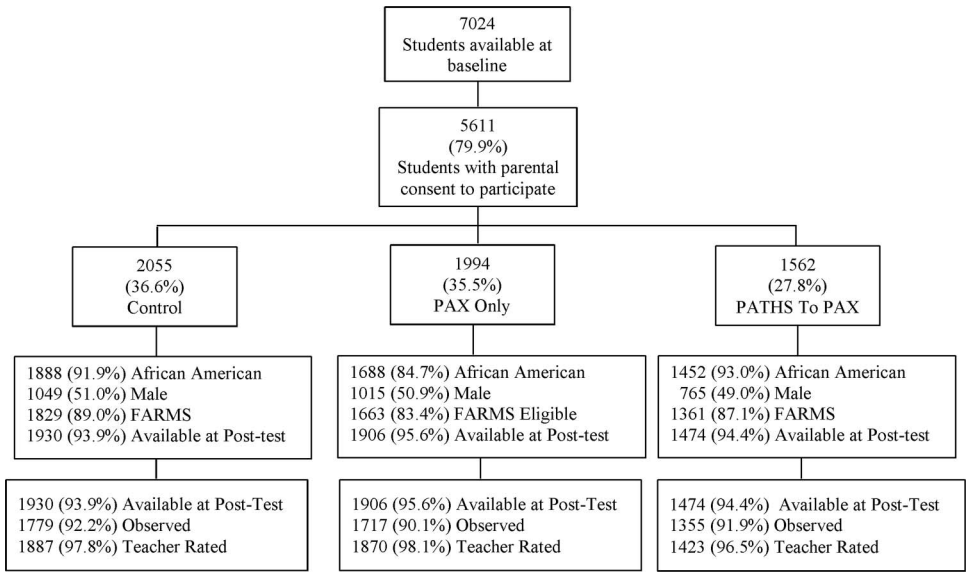

Figure 1. Sample Demographics of Enrolled Students.

utilize social learning principles within a team-based, game-like context to reduce aggressive/disruptive and off-task behavior and, consequently, facilitate academic instruction. In the current study we used the PAX GBG, which represents Dr. Embry's and colleagues' (Embry et al., 2003) efforts to improve the effectiveness of the original GBG and to make it ready for wide scale dissemination. Like the original GBG, the PAX GBG involves group-based rewards several times per day. Rotating student groups or "teams" are reinforced for their collective success in voluntary, group control contextually defined "unwanted" behavior. The team-based nature of the game allows teachers to take advantage of positive peer pressure for improving academic and prosocial student behavior at the individual as well as at the classroom level. Teachers assign students to one of three to four teams. The teacher seeks to evenly match the teams in terms of student behavior to insure all teams have an equal chance of winning the "game". The teams work cooperatively to maintain PAX behavior (which stands for Peace, Productivity, Health, & Happiness) in the classroom. Points are given to the team when a member displays a "spleem," or an infraction of the Game rules, such as, failing to follow the teacher's instructions. At the end of the game period, all teams with three or fewer spleems, or rule infractions, win the game. The students are essentially rewarded for displaying self-control, emotional regulation and group regulation while not attending to or reinforcing the misbehavior of others. The rewards for winning the game are usually nonmaterial and are well within the capacity of teachers or schools to provide.

Promoting alternative thinking strategies (PATHS). PATHS is based on the Affective-Behavioral-Cognitive-Dynamic model of development (Greenberg, Kusche, & Speltz, 1990), which places primary importance on the developmental integration of affect (and emotion language), behavior, and cognitive understanding as they relate to social and emotional competence. PATHS is designed to improve skills in four domains: (a) prosocial friendship skills, (b) emotional understanding and emotional expression skills, (c) self-control/emotion regulation (e.g., the capacity to inhibit impulsive behavior and organize goal-directed activity), and (d) problem-solving skills, including interpersonal negotiation and conflict resolution skills, which, in turn, are expected to improve problem behavior and social-emotional skills. In the current study, the preschool/Kindergarten version of PATHS (Domitrovich, Greenberg, Cortes, & Kusche, 1999) was used in kindergarten classrooms and the Fast Track (CPPRG, 1999) version of the curriculum was used in Grades 1 through 5. Fast Track (CPPRG, 1999) refers to a multisite randomized trial of a multilevel, K-12 intervention designed to prevent conduct disorder. The universal level was confined to the elementary school years and included PATHS as a primary component. The Fast Track version of PATHS contains approximately 90% of the content included in the current published version of the curriculum (Kusche & Greenberg, 1993; CPPRG, 2011) that was tested in prior as well as recent randomized PATHS trials (CPPRG 1999, 2010). Both the preschool and elementary curriculum models are designed to improve cognitive skills related to planning and other tasks that require verbal self-regulation (Greenberg et al., 2004). About 40% of the lessons focus on skills related to understanding and communicating emotions. As a basic step toward self-control, PATHS teaches young children to recognize the internal and external cues of affect and to label them with appropriate terms. In the elementary version, additional lessons help children understand the difference between feelings and behaviors. Appropriate and inappropriate behavioral responses are discussed. Another 30% of the lessons in both the preschool and elementary programs focus on skills related to the increase of positive social behavior (e.g., social participation, pro-social behavior, communication skills). Lessons address making/sustaining friendships, using good manners, taking turns and sharing in games, expressing one's viewpoint, and listening to others. Finally, 30% of the lessons focus on self-control and other steps in social problem solving.

PATHS to PAX. The integration of the PAX GBG and PATHS into what the developers refer to as PATHS to PAX was carried out over 3 years and included as a first step, the developers

training in each other's intervention and then observing the other's program as it was being used by teachers. A second step in the process was to conceptually integrate the two interventions, which formed the rationale for the three components of the intervention which are: (a) lessons, (b) activities and (c) practices. The third step of program development involved streamlining the PAX GBG + PATHS combination into a single intervention that was feasible for teachers to use in the classroom. The fourth step in the program development was to create an effective training model that prepared teachers to use the combined intervention. This initial version of the training model was then implemented in six K-5 elementary schools. Following this initial implementation and with feedback from teachers in the form of focus groups, individual interviews, and anonymous questionnaires, the next generation of the training model was developed, which was utilized in the study described here.

Training and Implementation Support Model

In terms of training, there were initial, 1-day, group-based trainings for teachers in both intervention conditions followed by ½ day booster Sessions 3 months later. Of the 97 teachers assigned to the PATHS to PAX condition, 90.8% and 89.7% attended the initial 1-day PATHS and PAX GBG training days, respectively. Additionally, 72.2% and 67% of the PATHS to PAX condition teachers attended the ½ day PATHS and PAX GBG booster trainings, respectively. Of the 116 teachers assigned to the PAX GBG Alone condition, 91.4% attended the initial 1 day PAX GBG training and 55.2% attended the ½ day PAX GBG booster training. An additional 7.5% of teachers across conditions received some form of training during the trial year, including abbreviated trainings provided on an individual basis by one of our intervention staff. A total of 3.8% of intervention teachers received no training at all during the study year. These teachers refused to participate in the intervention training and posttraining coaching.

The initial full day trainings were followed by weekly face-to-face coaching for the remainder of the school year (i.e., 31 weeks), which included check-ins, modeling, needs assessments, and technical assistance/performance feedback. Coaching was manualized with tailoring based on teacher need. PATHS to PAX teachers received more coaching over time as would be expected given the PATHS to PAX condition included 2 interventions. There were no differences between two intervention conditions (PAX GBG Alone and PATHS to PAX) with respect to the number of training/mentoring refusers ($p > .26$).

Level of implementation: Magnitude. With regard to measurement of the magnitude of intervention implementation, teachers maintained a daily log of the number of PAX GBG games and minutes played and PATHS lesson taught. On average, teachers played the game 154.22 ($SD = 106.46$) times over the school year in the PATHS to PAX condition for 1,583.43 min ($SD = 1,483.14$), whereas they played the game 150.18 ($SD = 94.92$) times for 1,431.84 min ($SD = 1,298.38$) in the PAX GBG Alone condition. The differences between the conditions with respect to the number of games and minutes played were not significant [Games Played, F (14), 0.08, $p > 77$; Minutes Played, F (14), 0.46, $p > .50$)]. On average, teachers completed 71.80% ($SD = 0.27$) of the scheduled PATHS lessons across the school year.

Level of implementation quality: PAX GBG and PATHS fidelity rubric. Four approximately bimonthly observations of intervention teachers' quality of implementation of the PAX GBG and PATHS lessons were carried out over the school year. PAX GBG and PATHS coaches carried out the observations. Interrater reliability was established for approximately 20% of the observations at the onset of each wave of observations. The quality of implementation was coded using a 22-item implementation rubric observation scale developed to assess PAX GBG and PATHS implementation fidelity. All the items are rated on a 5-point scale from zero (*no identified characteristics evident*) to 4 (*most identified characteristics present and implemented as trained*). About a one-third of the items are designed to assess more general aspects of the teacher's classroom management and presentation skills, these items include teacher interpersonal style (e.g., teacher creates a positive and responsive atmosphere), level of punitive discipline (e.g., does not use punitive or shaming techniques), and teacher management and discipline skills (e.g., teacher provides clear structure, expectations and routines and consistent discipline). The remaining items in each implementation fidelity rating scale are based on the "recipe" given to teachers for doing a PATHS lesson and playing the game and implementing the PAX GBG evidence-based kernels. The recipe contains each of the steps necessary to implement the PAX GBG and PATHS intervention protocols. Intraclass correlation coefficients between raters reached or exceeded .80 for all items across each wave of implementation quality assessments. The average overall rating of implementation quality based on the rubric total score for the PAX GBG quality indicators was 3.24 ($SD = 0.57$) out of possible score of 4 for teachers in the PAX GBG Alone condition and 3.20 ($SD = 0.52$) in the PATHS to PAX condition, which were not significantly different [F(17), 0.00, $p = .986$]. The overall quality rating for PATHS implementation was 3.40 ($SD = 0.50$).

Descriptions of Measures of Intervention Outcomes

Independent observations of student behavior. Classroom observations of student behavior were carried out by independent observers on two occasions, one week apart, at pre- and posttest respectively. The observation system was based in part on the one used in the Fast Track study (CPPRG, 1999; Tapp, Wehby, & Ellis, 1995). The behaviors observed included on-task and disruptive and physically and verbally aggressive behaviors. On average, each student was observed for about 5 min at pre- and posttest, respectively.

Observers received 2 weeks of training in the observational system employed. The bulk of the training consisted of the coding of videotapes of student behavior and live observations in classrooms during which agreement with gold standard observers was established. Observer agreement with a gold standard observer was checked on a weekly basis over the course of the pre- and posttest observations. Data on agreement with the gold standard observers was obtained for approximately 15% of the observations at pre- and posttest, respectively. Due to the low base rate of observed disruptive and aggressive (physical and verbal) behaviors, these three codes were collapsed into one for the purpose of calculating interobserver agreement. Behaviors were observed in 10 second

intervals and were recorded as present if they occurred at least once during a 10-s interval. The percent agreement (agreements/agreements + disagreements) for on-task behavior across the 6 waves of observations was 95.4%, whereas it was 70.2% for the aggressive/disruptive behavior composite.

A single Total Problem Behavior score was derived for the outcome analyses, which was based on the number of intervals in which the student engaged in off-task, disruptive and/or aggressive behaviors (physical and/or verbal aggression). The per interval Total Problem Behavior score could range from 0–4, with 0 signifying that none of the problem behaviors occurred and 4 signifying all of the problems behaviors were observed at least once during a 10-s interval. The Total Problem Score used in the analyses was the average score across the 10-s intervals the student was observed.

Teacher Observation of Classroom Adaptation-Revised (TOCA-R). The TOCA-R (Werthamer-Larsson, Kellam, & Wheeler, 1991) requires teachers to rate the child's adaptation to classroom task demands over the last three weeks. Adaptation is rated by teachers on a six-point frequency scale (1 = *almost never* to 6 = *almost always*). The domains include authority acceptance (or oppositional defiant/conduct problems) and readiness to learn (or attention-concentration problems). The mean of the teacher ratings across the items making up each of these subscales was used in the outcome analyses. The authority acceptance (oppositional defiant/conduct problems behavior) subscale includes items such as, breaks rules and talks backs to teachers, (items were reversed coded so that a higher score translated to less frequent problem behaviors). The readiness to learn (attention-concentration problems) subscale consists of items such as ready to learn, stays on task, and concentrates. The coefficient alphas for the authority acceptance (oppositional defiant/conduct problems) and readiness to learn (attention-concentration problems) subscales exceeded .80, respectively.

The Social Health Profile Social Competence Scale (SHP SCS, CPPRG, 1999). In addition to the TOCA-R, teachers completed the SHP Social Competence and Emotion Regulation subscales, which were used in the evaluation of the Fast Track intervention. The Social Competence subscale items include resolves peer problems, understands others, suggests without bossiness, whereas controls temper in a disagreement, appropriate expression of needs/feelings, thinks before acting, can calm down are examples of the items included in the Emotion Regulation subscale. Like the TOCA-R, items are rated by teachers on a 6-point frequency scale from almost never to almost always observed by the teacher over the last 3 weeks. Also consistent with the TOCA-R, we used the subscale means in our outcome analyses. The coefficient alphas for the Social Competence and Emotion Regulation subscales were .82 and .80, respectively, in the present study.

Results

Analytic Plan

An intent to treat approach was employed. That is, student outcome data for all intervention teachers were included regardless of teacher level of implementation of the PAX GBG and/or PATHS. Linear mixed model analysis of covariance was used to evaluate the effects of the intervention on the study outcomes with adjustment for their pretest levels. School was included as a random effect, given randomization was at the school level. Planned comparisons were carried out between the intervention and control conditions and the two intervention conditions. We also tested for interactions between intervention status and gender, ethnicity, free lunch status, grade level, cohort, and pretest level of the study outcomes, respectively. For the purpose of the intervention x grade level interactions, the grade levels were collapsed into two categories, K-2 and 3–5, to increase the statistical power to detect interactions in line with McClelland and Judd (1993). Consistent with Bauer and Curran (2005), interactions were probed using the Johnson-Neyman technique (Johnson & Neyman, 1936) to determine the *regions of significance* within which the simple slope was significantly different (at $p \leq .05$) from zero with respect to the pretest levels of the outcome (which were mean-centered for interpretability). Effect sizes were calculated using Cohen's *d.* In the case of interactions, we chose an arbitrary value within the Johnson and Neyman (1936) regions of significance to calculate the effect size.

Preliminary Analyses

There were no significant differences between conditions at pretest in terms of student gender, $F(2, 24) = 0.078$, $p = .925$; student ethnicity, $F(2, 24) = 0.337$, $p = .717$; number of FARMS eligible students, $F(2, 24) = 0.713$, $p = .550$; or school size, $F(2, 24) = 1.90$, $p = .171$. In addition, there were no differences between conditions in terms of the pretest levels of the study outcome variables or the percentage of students missing a posttest outcome measure. Finally, no differences in terms of baseline characteristics (i.e., gender, ethnicity, FARMS, and school size) were found between those with complete versus pretest only data in terms of classroom observations and teacher ratings, respectively. Consistent with the finding of no differences, only those students with both pretest and posttest data in terms of classroom observations and teacher ratings, respectively, were included in the outcome analyses.

Outcome Analyses

Table 1 includes the adjusted posttest means for the 3 conditions, their standard deviations and the F-statistics and *p*-levels for the main or interaction effects. In the presence of an intervention x pretest level of outcome, the F-statistic reported is for the interaction. As noted above, Johnson and Neyman (1936)'s regions of significance analysis was used to probe the significant interactions to determine the pretest levels of the outcomes (Z) at which the slopes of the outcome posttest score (Y) on intervention condition (X) were significant a $p \leq .05$. As noted above, the pretest levels of the outcomes (Z) were mean-centered for these analyses to aid interpretability. The boundaries of the regions of significance are given in Table 2 for each the significant interactions. Note that the Johnson and Neyman (1936) regions of significance analysis can generate ranges of significance above and below the mean in a single analysis as reflected in Table 2. Also note that the regions of significance are based on model implied estimates, which do not necessarily translate into the range of observed values. Values that are out of the range of possible pretest mean-centered scores are identified in the Table 2 note.

PATHS to PAX versus control. As can be seen in Table 1, after adjustment for the pretest levels of the respective outcomes,

Table 1
Adjusted Post-Test Means, SDs and Mixed Model F-Statistics, and Significance Levels for PATHS to PAX Vs Control and PAX GBG Versus Control Contrasts

	PATHS to PAX VS CONTROL			PAX GBG VS CONTROL			PATHS to PAX VS PAX GBG		
Construct	P2P Mean (*SD*)	Control Mean (*SD*)	*F*-Statistic	GBG Mean (*SD*)	Control Mean (*SD*)	*F*-Statistic	P2P Mean (*SD*)	GBG Mean (*SD*)	*F*-Statistic
Teacher ratings									
Readiness to learn	4.33 (2.41)	4.17 (2.76)	5.32^{2*}	4.26 (2.55)	4.16 (2.63)	1.25	4.28 (1.93)	4.21 (2.14)	3.76^{2*}
Social competence	4.10 (2.74)	3.92 (3.14)	10.89^{2***}	4.06 (2.43)	3.90 (2.51)	3.71	4.05 (2.52)	4.02 (2.76)	5.89^{2*}
Emotion regulation	4.10 (2.41)	3.98 (2.72)	5.23^{2*}	4.08 (2.39)	3.97 (2.46)	1.71	4.06 (2.34)	4.05 (2.60)	.05
Authority acceptance	4.79 (1.97)	4.79 (2.25)	7.71^{2**}	4.80 (2.06)	4.81 (2.12)	.03	4.79 (1.90)	4.78 (2.10)	.01
Classroom observations									
Total problem behavior	.158 (.62)	.210 (.71)	4.56^{1*}	.162 (.78)	.210 (.79)	5.28^{2*}	.158 (.54)	.162 (.61)	.033

[1] Main Effect. [2] Interaction Effect.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

the planned contrasts between the PATHS to PAX and Control conditions yielded a main effect for the independent observations of Total Problem Behavior scale (composite index of off-task and aggressive/disruptive behavior). Relative to the Control condition at posttest, the PATHS to PAX condition manifested a significantly lower score on the Total Problem Behavior scale. The size of the effect was small as reflected in a Cohen's *d* of 0.08 (Cohen, 1992).

Although no significant (all *p*-levels exceeded .10) PATHS to PAX x gender, ethnicity, grade level, cohort, and free lunch status interactions, respectively, were found, the PATHS to PAX x pretest-level interactions for each of the 4 teacher-rated constructs reached significance. The regions of significance analyses for the PATHS to PAX by pretest level of Readiness to Learn, Authority Acceptance, Social Competence, and Emotion Regulation (See Table 2) interactions revealed that students who were rated at the lowest levels of each of these constructs at pretest benefitted the most from the PATHS to PAX intervention at posttest. The effect sizes at −1 *SD* below the mean of pretest scores were all in the small range (Readiness to Learn, 09; Authority Acceptance, 0.03; Social Competence, .09; Emotion Regulation, 0.07).

Of note, the regions of significance analysis for teacher-rated authority acceptance also revealed that pretest scores at ≥2.3 were associated with negative slopes between posttest scores and intervention condition. That is, PATHS to PAX students at the high end of authority acceptance at pretest had significantly worse posttest scores relative to their control counterparts.

PAX GBG versus control. We did not find any main effects for the PAX GBG versus Control contrasts. However, the PAX GBG versus Control intervention x pretest level of the outcome interaction (see Table 1) for the classroom observation Total Problem Behavior Scale score was significant. None of the interactions between the PAX GBG and gender, ethnicity, grade level, cohort, and free lunch status, respectively, proved significant (all *p*-levels exceeded .10).

Table 2
Mean Centered Pre-Test Values at Which the Slope of the Outcome on Intervention Condition is Significant at $p \leq .05$

Intervention contrasts	Pre−test scores at or below	Pre−test scores at or above
Intervention contrasts		
PATHS to PAX vs Control		
Outcomes		
Teacher-rated readiness to learn	−.25 (3.88)	19.36*
Teacher-rated social competence	−.25 (3.63)	6.70*
Teacher-rated emotion regulation	−1.19 (2.77)	14.30*
Teacher-rated authority acceptance	−2.59 (2.18)	2.32*
Intervention contrasts		
PAX GBG vs Control		
Outcomes		
Classroom observation of total problem behavior	−3.21*	.05 (.26)
Intervention contrasts		
PATHS to PAX vs PAX GBG		
Outcomes		
Teacher-rated readiness to learn	−1.61 (2.62)	6.21*
Teacher-rated social competence	−3.63*	5.54*

Note. Values outside of the parentheses are mean centered values of the moderator, Z. Values within parentheses are the actual scale values when the value is within the range of possible scale values.
* At or beyond the range of possible scale values.

With respect to the classroom observation Total Problem Behavior Scale, the regions of significance analysis (see Table 2) revealed that relative to their control counterparts, the PAX GBG resulted in a significant reduction in the Total Problem Behavior scale score at posttest for those students whose pretest scores on the Total Problem Behaviors scale were at the upper end of the distribution. The size of the effect at 1 *SD* above the pretest mean was .05.

PAX GBG alone versus PATHS to PAX. Only two of the PAX GBG Alone versus PATHS to PAX comparisons were significant and both involved significant intervention x pretest level of the outcomes interactions (see Table 1). Students at the lower end of the pretest distribution (see Table 2) of teacher—rated Readiness to Learn and Social Competence subscales tended to show higher scores on each of these subscales at posttest if they were in the PATHS to PAX condition in contrast to their PAX GBG Alone counterparts. The sizes of the effects at −1 *SD* below the mean were small (Readiness to Learn, 0.06; Social Competence, 0.04).

Exploration of grade-level variation in intervention response. Although the grade level x intervention and grade level x intervention by baseline level of the outcome behaviors interactions were not significant for any of the study outcomes, an examination of grade level variation in intervention response seemed justified scientifically given the lack of study of this issue in prior GBG and PATHS trials. As noted above, to maximize statistical power to test for such effects, grade levels K-2 and 3–5 were collapsed, respectively, into two groups. As with the whole sample, mixed model analyses of covariance and post hoc probing of interactions using regions of significance analyses were then carried out for each of the study outcomes separately for the K-2 and 3–5 groups.

With respect to the K-2 PATHS to PAX versus Control condition mixed model analyses of covariance, significant intervention x pretest level of outcome interactions were found for teacher-rated Authority Acceptance, $F = 13.18$, $p < .001$ and independent observations of Total Problem Behavior, $F = 5.04$, $p < .05$. Regions of significance analyses of the Authority Acceptance interaction revealed that on average students at the lower end of the Authority Acceptance distribution at pretest (>1.34 units below the Mean) in the PATHS to PAX condition had significantly higher Authority Acceptance scores at posttest than their Control group counterparts. As to the Total Problem Behavior interaction, the PATHS to PAX students at the upper end of the Total Problem Behavior distribution at pretest (>0.39 units above the mean) had on average significantly lower Total Problem Behavior scores at posttest than Control group students with similar pre-Test Total Problem Behavior levels.

The K-2 PAX GBG versus Control condition analyses yielded a main effect for Authority Acceptance, $F = 4.28$, $p < .05$ and an interaction, $F = 5.42$, $p < .05$ for the Total Problem Behavior scale. Regarding the former, students in the PAX GBG condition were rated as higher on Authority Acceptance at posttest than students in the Control condition (PAX GBG, MN = 4.96, *SD* = 2.06 vs. Control, MN = 4.88, *SD* = 2.12, ES = .04). Like with the PATHS to PAX Total Problem Behavior interaction, post hoc probing revealed that PAX GBG students at the higher end (>0.27 units above the mean) of the Total Problem Behavior distribution at pretest had significantly lower post-Test Total Problem Behavior scores than their counterparts in the Control condition.

The Grades 3–5 analyses only yielded significant effects for the PATHS to PAX versus Control condition comparisons. Significant intervention x pretest level of the outcome interactions were found for teacher-rated Readiness to Learn, $F = 3.86$, $p < .05$ and Social Competence, $F = 6.91$, $p < .01$, whereas main effects were found for teacher-rated Emotion Regulation, $F = 4.09$, $p < .05$ and the independent observations of Total Problem Behavior ($F = 5.03$, $p < .01$). In the case of each of the interactions, PATHS to PAX students at the lower end of the Readiness to Learn (>1.70 units below the mean) and Social Competence (>0.46 units below the mean) pretest scores had on average significantly higher scores on these subscales at posttest than their Control group counterparts. The main effect analysis revealed that the PATHS to PAX condition had significantly higher Emotion Regulation (PATHS to PAX, MN = 3.96, *SD* = 2.41 vs. Control, MN = 3.82, *SD* = 2.72, ES = .07) and significantly lower Total Problem Behavior scores at posttest than the Control condition (PATHS to PAX, MN = .14, *SD* = .59 vs. Control, MN = .22, *SD* = .65, ES = .13).

Discussion

The primary aim of this study was to determine the relative efficacy with respect to the magnitude and range of effects of the PAX GBG alone versus its combination with PATHS in reducing the early risk behaviors for substance abuse, antisocial behavior, depression and poor academic and occupational attainment in adolescence and young adulthood. These early risk behaviors included attention/concentration problems, aggressive/disruptive behavior, and deficits in emotion regulation and social competence. As noted, it was hypothesized that 1) the size and 2) breadth of effects of the combination of the GBG and PATHS—PATHS to PAX—versus the control condition would be superior to that of the PAX GBG alone versus control comparison. The basis for these hypotheses was that the teachers in the combined condition could use the PAX GBG to increase student attention and on task behavior during the PATHS lessons, which should, in turn, improve acquisition of the PATHS skills taught. The emotional regulation and social problem-solving skills acquired within the context of the PATHS lessons should then facilitate generalization of the on task and prosocial behaviors seen during the playing of the Game across time and settings.

Some evidence was found to support the hypothesis of the superiority of the PATHS to PAX condition. The evidence was, however, in terms of the greater diversity of effects on the targeted risk behaviors and not necessarily their magnitude as discussed below. More specifically, whereas the PATHS to PAX versus control condition comparisons yielded significant effects for all 4 teacher-rated constructs and independently observed Total Problem Behavior, the only significant PAX GBG versus control condition comparison was for the Total Problem Behavior measure. With 2 exceptions, all the significant intervention effects as hypothesized for the PATHS to PAX versus control and the PAX GBG versus control condition consisted of intervention x pretest level interactions. This is typical in universal preventive intervention trials as pointed out earlier, given those participants who manifest the targeted risk behaviors are most likely to benefit from the interventions (Greenberg & Abenavoli, 2016).

With respect to the broader array of effects hypothesis and in line with the explicit focus of PATHS on social competence and emotion regulation, those PATHS to PAX condition students at the lower end of the pretest distributions of these teacher-rated constructs had significantly higher scores on both at posttest relative to their control condition counterparts. Moreover, on average, PATHS to PAX condition students had significantly greater posttest scores than control group students in terms of teacher-rated readiness to learn and authority acceptance. These findings were again confined to those PATHS to PAX and control students at the lower end of the pretest distribution of these constructs. No such effects were found for the PAX GBG alone versus control condition comparisons for any of these teacher-rated constructs. Finally, the PATHS to PAX versus control group comparison yielded a main effect, wherein the PATH to PAX condition's independently observed Total Problem Behavior score was significantly lower at posttest than the control condition. Importantly, however, a PAX GBG Alone x baseline level of observed off-task/aggressive/disruptive behavior was found, with those students at the upper end of the distribution at baseline benefiting the most at posttest.

A number of potential moderators of intervention impact were examined besides the intervention x pretest level of the outcomes interactions. Among these potential moderators was age/grade level. However, none of the intervention x grade level interactions carried out reached conventional levels of significance. Importantly, the lack of statistically significant findings may have been due to the fact that number of schools was not sufficient to provide an adequately powered test of these interactions. Nevertheless, grade level variation in intervention response was explored given there have been no studies to date of variation in GBG and PATHS intervention outcomes as a function of age and/or grade level. As noted in the introductory section of this article, in line with the concept of developmental cascades as reflected in Patterson and colleagues' (1992) model of the development of antisocial behavior over the life-course, one might expect poorer response to interventions targeting aggressive/coercive behavior in the late elementary school years given it may become increasingly more difficult to overcome the cumulative consequences of the history of social transactions between children and their parents, teachers and peers over time and across contexts (e.g., family, school, peer group, and neighborhood; Masten & Cicchetti, 2010). Alternatively, it may be the case that late elementary school students may be more advanced in terms of their social–cognitive development than their early elementary school counterparts and, as such, may be more likely to comprehend and translate into action the PATHS socioemotional learning concepts being taught to them. The exploratory analyses of grade level variation in intervention response yielded evidence in support of the contention that aggressive/disruptive/coercive behavior might prove more malleable in K-2 children than in children in Grades 3–5. Moreover, the finding that PATHS to PAX teachers rated their Grades 3–5 students higher on Readiness to Learn, Social Competence, and Emotion Regulation than their control group counterparts, but not on Authority Acceptance, supports the contention that the grades 3–5 children may have been more advanced in terms of social–cognitive development and, in turn, better able to comprehend and internalize the PATHS concepts. Importantly, like the intervention effects for the sample as a whole, the size of the grade-level intervention effects were small. Moreover, these results should be interpreted with caution as the analyses were exploratory in nature. Future trials of school-based universal interventions such as the PAX GBG and PATHS should be adequately powered to test for variation in intervention response as a function of age/grade level. The knowledge obtained may serve to inform the next generation of preventive interventions.

It is important to point out that the greater number of significant effects found for the PATHS to PAX versus the PAX GBG alone condition in terms of teacher-rated constructs may in part be due to expectancy effects associated with the greater amount of training and coaching in the PATHS to PAX condition. It is also important to point out that all but one of the intervention outcomes was in the expected direction. The lone exception was the PATHS to PAX versus control contrast in terms of teacher-rated authority acceptance. As noted in the results section, two contrasting effects were found. As hypothesized, those students at the low end of the distribution of pretest levels of authority acceptance had significantly higher scores at posttest than their control group counterparts. But those PATHS to PAX students with higher scores at pretest had lower posttest scores than their control counterparts. One possible explanation for this finding is regression to the mean. Alternatively, as opposed to a statistical artifact, it may be that PATHS to PAX teachers placed a greater focus over the school year on those students exhibiting problems with authority acceptance at the beginning of the school year, which, in turn, resulted in less attention to the better behaving students and, subsequently, the decrease in their level of authority acceptance as measured at posttest. However, this was not apparent for any of the other teacher-rated constructs. Nor have there been similar findings reported in the original randomized trials of the Good Behavior Game using this same teacher rating of authority acceptance (Dolan et al., 1993; Ialongo et al., 1999). The explanation for this finding may have to await future trials of PATHS to PAX.

Despite the larger number of significant PATH to PAX intervention effects relative to the PAX GBG alone, the contrasts between the PATHS to PAX and PAX GBG alone conditions only yielded significant differences for 2 comparisons, reflecting the relatively modest intervention effect sizes found for both intervention conditions. In previous randomized trials of the GBG (Dolan et al., 1993; Ialongo et al., 1999), the effect sizes found were modest as well. In large part, modest effect sizes are the rule rather than the exception in universal preventive intervention trials, given, as noted above, only those individuals who demonstrate some level of risk will likely benefit from the interventions.

One factor that may explain the modest intervention effects found in the present study is that the intervention condition teachers and students were only exposed to the interventions for one year. The reason for doing so was to ensure that a sufficient number of schools could be included in the evaluation given randomization was at the level of the school. Too few schools would have jeopardized the generalizability of our results and undermined statistical power. The trade-off, however, may have been weaker intervention effects. This certainly may be the case for PATHS given it is conceived as a pre-K to Grade 5 intervention. That is, children are expected to complete 7 years of PATHS lessons, given it is during this age span that key developmental milestones in terms of emotion knowledge, self-regulation, and social problem solving are typically achieved. As such, the size of the effect for the PATHS to PAX condition may have been

attenuated given students only received 1 year of PATHS lessons. This same logic could be applied to the PAX GBG condition: multiple years of exposure would have likely yielded larger effect sizes.

Another possible explanation for the modest intervention effects found in the present student for the PAX GBG and PATHS to PAX was teacher level of implementation. Like in Ialongo et al. (1999), there was considerable variation in the level of implementation with respect to the number and minutes of games played. This variation occurred despite the fact that the training and coaching protocol reflected the "best practices" identified in the extant literature on optimizing implementation of evidence-based interventions via training and coaching. These include (a) program content based on a well-articulated theory of cause (e.g., Gottfredson, 1997), (b) program standardization, including manualized intervention materials (e.g., Blakely et al., 1987), (c) provision of training to implementers (e.g., Dusenbury, Brannigan, Falco, & Hansen, 2003), (d) ongoing consultation to implementers (Gorman-Smith, Beidel, Brown, Lochman, & Haaga, 2003); (e) a training and consultation phase long enough to insure a thorough and working knowledge of the core program principles and their translation into practice (e.g., Rose & Church, 1998), and (f) a process for monitoring of implementation which is linked to professional development (Scott & Martinek, 2006).

Importantly, Domitrovich et al. (2015), utilizing the present study sample of teachers, examined the predictors of the number and minutes of the PAX games played. They found that teachers who reported higher levels of emotional exhaustion and perceived less of a fit between the game and their "teaching style" at pretest played fewer games across the school year. None of the other factors assessed at the teacher level including organizational health, motivation to implement, openness to innovation, coach-teacher alliance, administrator support or behavior management self-efficacy were significantly related to the number of games played. This was also the case for school level characteristics, including school size and the number of children receiving free or reduced lunch—a proxy for family income. However, not surprisingly, Musci, Pas, Bradshaw, and Ialongo (2018) found that classroom-levels of authority acceptance and emotional exhaustion were significantly related at pretest, such that, the lower the classroom-level of authority acceptance, the higher the level of emotional exhaustion. Although Musci et al. were not able to draw causal inferences, given authority acceptance and emotional exhaustion were assessed concurrently, this finding points to the need for the assessment of classroom level of authority acceptance and teacher emotional exhaustion early in the school year in order to identify the teachers in need of additional help with managing aggressive/disruptive behaviors in their classrooms.

As noted above, the Washington State Institute for Public Policy (2017) reported that the GBG and PATHS both yielded in excess of 60% return on investment in terms of intervention training, coaching/mentoring, and material costs. The fact that only a handful of significant differences were found between the PAX GBG Alone and PATHS to PAX in terms of student outcomes suggest that combination of the GBG and PATHS may not be justified from an economic standpoint. However, this does not take into account the benefits of PATHS to PAX relative to the PAX GBG Alone on teachers' perception of self-efficacy and personal accomplishment as reported in Domitrovich et al. (2016). These benefits may translate into fewer teachers leaving teaching for other occupations and fewer working days lost due to work-stress related health problems for those who remain in teaching. Unfortunately, Domitrovich et al. (2016) do not report on such outcomes and so the economic benefits of the combination of PATHS to PAX relative to the PAX GBG or PATHS alone remain questionable.

Limitations

With respect to study limitations, the size of the trial did result in a number of trade-offs besides the 1-year length of the intervention. These include the breadth, frequency, and duration of the assessments. More frequent teacher ratings and classroom observations would have allowed a more reliable assessment of the course of student behavior, particularly for low base rate events, such as physical and verbal aggression. Longer term follow-up would have also provided a test of the generalization of intervention effects over time, whereas observations beyond the classroom, including in the hallways and in the cafeteria, would have yielded data relevant to the generalization of intervention effects across multiple school contexts. Peer assessments would have provided an additional and important method of determining intervention effects with respect to social competence. In contrast to these limitations, clear strengths of the present study include the randomized design, the relatively large number of schools, teachers and students, and the intensive teacher training and coaching model employed.

Future Directions

In terms of future directions, follow-up of the study population through adolescence and into young adulthood would allow us to test the hypothesis that the effects of PATHS to PAX are more enduring than the PAX GBG Alone. This hypothesis is consistent with the more explicit focus on socioemotional learning and development of problem solving and conflict resolution skills in PATHS. Randomized trials of interventions targeting teacher stress in order to improve teacher implementation of PATHS and the PAX GBG would also be worth pursuing.

Importantly, we are in the midst of a study where we examine the benefits of the integration of PATHS to PAX with an indicated intervention component, the Incredible Years Child and Parent program (Webster-Stratton & Reid, 2010) on student outcomes. Such an approach would be consistent with the Metropolitan Area Child Study (2002) Research Group's findings, wherein the integration of universal and indicated teacher, student, and parent intervention components in the participating elementary schools produced the greatest benefit in terms of student social and behavior outcomes. Our rationale for expecting an added benefit from the combination of PATHS to PAX with an indicated and treatment intervention is that the latter interventions are more likely to be effective with the nonresponders to PATH to PAX. As noted earlier, these nonresponders represent approximately 50% of the most aggressive/disruptive students in the classroom as found in Petras et al. (2008, 2011). We also hypothesize that the improvement in the behavior of the nonresponders via the indicated or treatment intervention will increase teacher perception of the effectiveness of PATHS to PAX, which should lead to lower levels

of emotional exhaustion and increased PATHS to PAX implementation. This increased implementation should then result in improved student outcomes overall, not just for the nonresponders.

Of course, the scalability of such combinations of universal and indicated interventions is an important question, let alone each intervention on its own. In this regard, it is important to point out that both the PAX GBG and PATHS are now currently in use in school systems across the US, Canada and in at least 8 European and 2 Asian countries. That said, as is the case in K-12 instruction (Castles, Rastle, & Nation, 2018), mental health services (Becker, Smith, & Jensen-Doss, 2013), and medical care (Hisham, Ng, Liew, Hamzah, & Ho, 2016), the use of evidence-based practices such as the GBG and PATHS appears to be the exception as opposed to the rule. One barrier to wider dissemination and implementation of evidence-based practices in the fields of education, mental health, and medicine is a cost-effective method to train and mentor large numbers of teachers and providers in such practices. Importantly, Becker at al. (2014) offer preliminary evidence of the effectiveness of training K-5 teachers in the PAX GBG and PATHS via online methods. More specifically, teachers implemented the PAX GBG and PATHS at the same levels of frequency and quality regardless of whether they participated in an online or in person training. A larger scale study, which also includes online mentoring/coaching and the measurement of student outcomes, is necessary to confirm and extend Becker, Bohnenkamp, Domitrovich, Keperling, and Ialongo's (2014) findings.

References

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,* 119–124. http://dx.doi.org/10.1901/jaba.1969.2-119

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research, 40,* 373–400. http://dx.doi.org/10.1207/s15327906mbr4003_5

Becker, E. M., Smith, A. M., & Jensen-Doss, A. (2013). Who's using treatment manuals? A national survey of practicing therapists. *Behaviour Research and Therapy, 51,* 706–710. http://dx.doi.org/10.1016/j.brat.2013.07.008

Becker, K., Bohnenkamp, J., Domitrovich, C., Keperling, J., & Ialongo, N. (2014). Online training for teachers delivering evidence-based preventive interventions. *School Mental Health, 6,* 225–236. http://dx.doi.org/10.1007/s12310-014-9124-x

Blakely, C. H., Mayer, J. P., Gottschalk, R. G., Schmitt, N., Davidson, W. S., Roitman, D. B., & Emshoff, J. G. (1987). The fidelity-adaptation debate: Implications for the implementation of public sector social programs. *American Journal of Community Psychology, 15,* 253–268. http://dx.doi.org/10.1007/BF00922697

Bradshaw, C. P., Zmuda, J. H., Kellam, S. G., & Ialongo, N. S. (2009). Longitudinal impact of two universal preventive interventions in first grade on educational outcomes in high school. *Journal of Educational Psychology, 101,* 926–937. http://dx.doi.org/10.1037/a0016586

Brown, C. H. (1993). Statistical methods for prevention trials in mental health. *Statistics in Medicine, 12,* 289–300.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19,* 5–51. http://dx.doi.org/10.1177/1529100618772271

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. http://dx.doi.org/10.1037/0033-2909.112.1.155

Conduct Problems Prevention Research Group (CPPRG). (1999). Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology, 67,* 648–657. http://dx.doi.org/10.1037/0022-006X.67.5.648

Conduct Problems Prevention Research Group (CPPRG). (2010). The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting & Clinical Psychology, 78,* 156–168. http://dx.doi.org/10.1037/a0018607

Conduct Problems Prevention Research Group (CPPRG). (2011). The effects of the fast track preventive intervention on the development of conduct disorder across childhood. *Child Development, 82,* 331–345. http://dx.doi.org/10.1111/j.1467-8624.2010.01558.x

Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., . . . Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology, 14,* 317–345. http://dx.doi.org/10.1016/0193-3973(93)90013-L

Domitrovich, C. E., Bradshaw, C. P., Berg, J. K., Pas, E. T., Becker, K. D., Musci, R., . . . Ialongo, N. (2016). How do school-based prevention programs impact teachers? Findings from a randomized trial of an integrated classroom management and social-emotional program. *Prevention Science, 17,* 325–337. http://dx.doi.org/10.1007/s11121-015-0618-z

Domitrovich, C. E., Greenberg, M. T., Cortes, R. C., & Kusche, C. (1999). *The Preschool PATHS Curriculum*. Deerfield, MA: Channing Bete Publishing Company.

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,* 1064–1074. http://dx.doi.org/10.1007/s11121-015-0557-8

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,* 405–432. http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18,* 237–256. http://dx.doi.org/10.1093/her/18.2.237

Embry, D. D., Staatemeier, G., Richardson, C., Lauger, K., & Mitich, J. (2003). *The PAX good behavior game* (1st ed.). Center City, MN: Hazelden.

Gorman-Smith, D., Beidel, D., Brown, T. A., Lochman, J., & Haaga, A. F. (2003). Effects of teacher training and consultation on teacher behavior towards students at high risk for aggression. *Behavior Therapy, 34,* 437–452. http://dx.doi.org/10.1016/S0005-7894(03)80029-1

Gottfredson, D. C. (1997). School-based crime prevention. In L. W. Sherman, D. C. Gottfredson, D. L. MacKenzie, J. Eck, P. Reuter, & S. D. Bushway (Eds.), *Preventing crime: What works, what doesn't, what's promising: A Report to the United States Congress (n.p.)*. Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Greenberg, M. T., & Abenavoli, R. (2016). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness, 10,* 40–67. http://dx.doi.org/10.1080/19345747.2016.1246632

Greenberg, M. T., & Abenavoli, R. (2017). Universal Interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness, 10,* 40–67. http://dx.doi.org/10.1080/19345747.2016.1246632

Greenberg, M. T., & Kusche, C. A. (2006). Building social and emotional

competence: The PATHS Curriculum. In S. R. Jimerson, & M. J. Furlong (Eds.), *Handbook of school violence and school safety: From research to practice* (pp. 395–412). Mahwah, NJ: Erlbaum.

Greenberg, M. T., Kusche, C. A., Cook, E. T., & Quamma, J. P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS Curriculum. *Development and Psychopathology, 7,* 117–136. http://dx.doi.org/10.1017/S0954579400006374

Greenberg, M. T., Kusche, C., & Riggs, N. (2004). The PATHS Curriculum: Theory and research on neuro-cognitive development and school success. In J. Zins, R. Weissberg, & H. Walber (Eds.), *Building school success on social and emotional learning* (pp, 170–188). New York, NY: Teachers College Press.

Greenberg, M. T., Kusche, C. A., & Speltz, M. (1990). Emotional regulation, self-control and psychopathology: The role of relationships in early childhood. In D. Cicchetti, & S. Toth (Eds.), *Rochester symposium on developmental psychopathology,* Vol. 2. New York, NY: Cambridge University Press.

Hisham, R., Ng, C. J., Liew, S. M., Hamzah, N., & Ho, G. J. (2016). Why is there variation in the practice of evidence-based medicine in primary care? A qualitative study. *British Medical Journal Open, 6,* e010565. http://dx.doi.org/10.1136/bmjopen-2015-010565

Ialongo, N., Poduska, J., Werthamer, L., & Kellam, S. (2001). The distal impact of two first-grade preventive interventions on conduct problems and disorder in early adolescence. *Journal of Emotional and Behavioral Disorders, 9,* 146–160. http://dx.doi.org/10.1177/106342660100900301

Ialongo, N. S., Werthamer, L., Kellam, S. G., Brown, C. H., Wang, S., & Lin, Y. (1999). Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression, and antisocial behavior. *American Journal of Community Psychology, 27,* 599–641. http://dx.doi.org/10.1023/A:1022137920532

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their applications to some educational problems. *Statistical Research Memoirs, 1,* 57–93.

Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, P., . . . Wilcox, H. C. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95*(Suppl. 1), S5–S28. http://dx.doi.org/10.1016/j.drugalcdep.2008.01.004

Kusche, C. A., & Greenberg, M. T. (1993). *The PATHS (Promoting Alternative Thinking Strategies) curriculum*. Deerfield, MA: Channing-Bete.

Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology, 22,* 491–495. http://dx.doi.org/10.1017/S0954579410000222

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114,* 376–390. http://dx.doi.org/10.1037/0033-2909.114.2.376

Metropolitan Area Child Study. (2002). A cognitive-ecological approach to preventing aggression in urban settings: Initial outcomes for high-risk children. *Journal of Consulting and Clinical Psychology, 70,* 179–194. http://dx.doi.org/10.1037//0022-006X.70.1.179

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York, NY: Oxford University Press.

Musci, R., Pas, E., Bradshaw, C., & Ialongo, N. (2018). *Teacher and student predictors of teacher implementation of the PAX GBG at the individual and classroom level*. Manuscript in preparation.

Patterson, G. R., Reid, J., & Dishion, T. (1992). *A social learning approach: IV. Antisocial boys*. Eugene, OR: Castalia.

Petras, H., Kellam, S., Brown, C. H., Muthén, B., Ialongo, N., & Poduska, J. (2008). Developmental courses leading to antisocial personality disorder and violent and criminal behavior: Effects by young adulthood of a universal preventive intervention in first- and second-grade classrooms. *Drug & Alcohol Dependence. 95S,* S45–59. http://dx.doi.org/10.1016/j.drugalcdep.2007.10.015

Petras, H., Masyn, K., & Ialongo, N. (2011). The developmental impact of two first grade preventive interventions on aggressive/disruptive behavior in childhood and adolescence: An application of latent transition growth mixture modeling. *Prevention Science, 12,* 300–313. http://dx.doi.org/10.1007/s11121-011-0216-7

Riggs, N. R., Greenberg, M. T., Kusché, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS Curriculum. *Prevention Science, 7,* 91–102. http://dx.doi.org/10.1007/s11121-005-0022-1

Rose, D. J., & Church, R. J. (1998). Learning to teach: The acquisition and maintenance of teaching skills. *Journal of Behavioral Education, 8,* 5–35. http://dx.doi.org/10.1023/A:1022860606825

Rutter, M. (1988). Epidemiological approaches to developmental psychopathology. *Archives of General Psychiatry, 45,* 486–495. http://dx.doi.org/10.1001/archpsyc.1988.01800290106013

Scott, T. M., & Martinek, G. (2006). Coaching Positive Behavior Support in school settings. *Journal of Positive Behavior Interventions, 8,* 165–173. http://dx.doi.org/10.1177/10983007060080030501

Tapp, J., Wehby, J., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instruments & Computers, 27,* 25–31. http://dx.doi.org/10.3758/BF03203616

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,* 1156–1171. http://dx.doi.org/10.1111/cdev.12864

Washington State Institute for Public Policy. (December 2017). *Benefit-cost technical documentation*. Olympia, WA: Author.

Webster-Stratton, C., & Reid, M. J. (2010). The incredible years parents, teachers and children training series: A multifaceted treatment approach for young children with conduct problems. In J. Weisz, & A. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd ed.). New York, NY: Guilford Publications.

Werthamer-Larsson, L., Kellam, S., & Wheeler, L. (1991). Effect of first-grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology, 19,* 585–602. http://dx.doi.org/10.1007/BF00937993

Received January 5, 2018
Revision received December 14, 2018
Accepted January 18, 2019 ■