

Outcomes of a Small Group Program for Early Elementary Students with Self-Regulation Difficulties:
Limitations of Transportability from Clinic to School

This preprint (dated 3/18/21) has not undergone peer review or any post-submission improvements or corrections. The Version of Record of this article was published October 15, 2021

in *School Mental Health*, and is available online at

<https://link.springer.com/article/10.1007%2Fs12310-021-09480-4>

Desiree W. Murray^{1,2}, Laura J. Kuhn¹, Michael T. Willoughby³,
Doré R. LaForett^{1,2}, Alyson M. Cavanaugh¹

¹Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill, Chapel Hill,
NC

²Currently at Child Trends, Inc., Chapel Hill, NC

³Research Triangle Institute, Durham, NC

Corresponding Author: Desiree W. Murray, desiree.murray@unc.edu

Acknowledgements: This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150169 to the University of North Carolina at Chapel Hill. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. Correspondence concerning this article should be addressed to desiree.murray@unc.edu (orchid ID 0000-0002-1554-1665). We are grateful to the many undergraduate and graduate students, research staff, and the school personnel, students and parents who contributed to the success of this project.

Declarations

Funding: This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150169 to the University of North Carolina at Chapel Hill. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Conflicts of interest/Competing interests: Dr. Murray is a trained mentor in the Incredible Years® Teacher Classroom Management Program and receives compensation from community organizations for providing trainings and consultation. Dr. LaForett is a certified group leader in the Incredible Years® Teacher Classroom Management Program and has received compensation from community organizations for providing trainings. Co-authors Kuhn, Willoughby and Cavanaugh have no relevant financial or non-financial interests.

Availability of data and material: Data collected for this study from which analyses were conducted are publicly available at: <https://dataverse.unc.edu/dataset.xhtml?persistentId=doi:10.15139/S3/FIGIXL>.

Code availability – available from study authors

Compliance with Ethical Standards: All procedures performed with human participants were approved by the UNC Institutional Review Board and adhere to the tenets of the Declaration of Helsinki.

Consent to Participate: Parents, teachers, and other school participants provided written consent/permission for children to participate prior to engaging in any research activities.

Consent for Publication: n/a

Abstract

Several mental health programs have been developed in clinics and transported into schools, which has great potential for increasing access to intervention for students who may not be otherwise served. However, such programs may lack consideration of the complexity and constraints of schools, including the diversity of student needs and backgrounds, raising questions about their effectiveness in this context. This study evaluates the efficacy an evidence-based clinical program – the Incredible Years® Dina Dinosaur School small group treatment program – under such realistic conditions. A total of 138 first and second grade students identified as having self-regulation difficulties were randomized to Business-as-Usual (BAU) or intervention, which included delivery of 34 group and 12 individual recess coaching sessions over 6 months, teacher consultation and inservice presentations, and three parent workshops.

Multi-method outcome measures were collected before and after the intervention and at 6-month follow up, evaluating self-regulation, disruptive behavior, social competence, and academic proficiency. Results indicated few significant main effects, with consistently small effect sizes. Effects were generally larger for self-regulation and oppositional defiant disorder (ODD) outcomes. For students at-risk for ODD, several additional effects were observed at follow up including on emotion regulation ($d = .75$) and academic proficiency ($d = .55$). Findings suggest less impact than clinic-based delivery which typically includes a parent program component, although reduced effects may also be related to lower fidelity in some of the child groups. Results contribute to understanding transportability and have useful implications for school mental health programming.

School-based programs promoting emotional and behavior health in young children have increasingly adopted a self-regulation framework (e.g., Diamond, 2012; Greenberg, 2006), consistent with a growing body of research that has identified self-regulation as a core mechanism underlying a wide range of psychiatric disorders (Gross & Jazaieri, 2014). Self-regulation, which can be defined as *the act of managing cognition*

and emotion to enable goal-directed actions, such as organizing behavior, controlling impulses, and solving problems (Murray et al., 2019), has also been specifically linked to educationally relevant outcomes including school readiness and achievement (Jones et al., 2016; Nigg, 2017; Zhou et al., 2012). Theoretical models integrating the two key components of self-regulation - executive functioning and emotion regulation – have been described specifically in application to school-based intervention (Bailey & Jones, 2019; Ursache et al., 2012). A self-regulation framework may be particularly useful for intervention programs that target a range of emotional and behavioral difficulties in early elementary students.

School-based programs have significant potential for preventing mental health disorders (Farmer et al., 2003; Owens & Murphy, 2004); however, there are several limitations to existing programs. As Weare & Nind (2011) noted in their synthesis of 52 reviews of such programs, “many types of interventions can be effective, sometimes strikingly so, but their effectiveness cannot be relied on”. One factor contributing to this variable efficacy may be the lack of transportability of interventions from clinics to schools. Many programs delivered in schools for children with significant emotional and behavioral difficulties were developed in clinic settings with homogenous populations under highly controlled conditions without adequate attention to features of the school context that may affect delivery (Reddy et al., 2009; Ringeisen et al., 2003). Not surprisingly, there are many questions about the feasibility and acceptability of such programs within day to day school practice (Evans et al., 2014; Gansle, 2005). Indeed, numerous implementation challenges have been identified including resource constraints, logistical difficulties, and competing demands (Forman et al., 2009; Langley et al., 2010). As such, identifying programs that are “relevant and doable in applied settings” has been identified as a priority research area for better addressing school mental health challenges (Evans et al., 2014).

Rationale for Transporting the IY® Dinosaur School Program

One program that holds potential for transportability from clinic to schools for young children is the Incredible Years® Dina Dinosaur small group program (referred to as “Dinosaur School” which is different from the IY® Dina universal classroom program). Dinosaur School was first developed as a clinic-based

program for children aged 3-8 years with or at risk for conduct problems (Webster-Stratton & Reid, 2013). Program content was designed to address deficits in inhibitory control, emotional regulation, and social-cognitive perceptions and social behaviors, consistent with the program's theoretical model (Webster-Stratton, Reid, et al., 2011). Dinosaur School is delivered during 18-20 two-hour clinic sessions, with parents attending concurrent parent training sessions. Two highly trained therapists lead skill building groups with four to six children using developmentally appropriate activities such as video-modeling, art activities, sociodramatic play with puppets, and role plays. Therapists provide explicit feedback and reinforcement (or "coaching") to teach and reinforce skills. Program content includes sequenced lessons addressing school rules, emotional literacy, empathy and perspective-taking, social problem solving, anger management, friendship skills, and communication skills.

Several studies support the efficacy of Dinosaur School with clinical sample, when delivered in combination with the IY® parent program (Larsson et al., 2009; Webster-Stratton & Hammond, 1997; Webster-Stratton, Reid, & Beauchaine, 2011; Webster-Stratton, Reid, & Hammond, 2004), with average effect sizes ranging from .35-.41 for increasing positive peer behaviors and decreasing negative school behaviors and from .6-.8 for social-emotional outcomes rated by parents. However, research on the program's efficacy in school contexts is quite limited. Hutchings et al. (2007) evaluated an abbreviated version of the program in a randomized study of 24 children aged 5-9 years old in one school in Wales and did not find any significant teacher-rated or observed effects, although problem-solving improved in a subsample rated at higher risk for mental health difficulties. In an uncontrolled study of 88 K-3rd grade predominantly Black and Hispanic students in the southwestern U.S. who exhibited disruptive behavior, pre-post teacher ratings indicated improvements in prosocial behavior and decreased conduct problems (Venter et al., 2012). A recent randomized controlled trial (RCT) of Dinosaur School in Turkish schools ($n = 32$) with 4-6 year olds found significant effects on a measure of social problem-solving ($d = .54$) but not on two teacher measures of social-behavioral difficulties (Bayrak & Akman, 2018). Thus, effectiveness of Dinosaur School in schools is unclear and warrants further rigorous evaluation with a larger sample. Moreover, there are no studies of Dinosaur School that include follow up

efficacy data with a control group.

Extensive guidance has been provided for implementation of Dinosaur School, including recommendations regarding dosing, adaptations, and training and technical assistance to support fidelity (Scott, 2008; Webster-Stratton, Reinke, et al., 2011). For schools specifically, such factors include selecting optimal group leaders to deliver the program; providing them with accredited training workshops coupled with ongoing supportive mentoring, consultation, peer, administrative, and facilitative supports; and ongoing program evaluating and monitoring (Webster-Stratton & Herman, 2009). To support its implementation in schools, Dinosaur School can be delivered within multi-tiered systems of support as a Tier 2 (small group) intervention, although its intensity may more closely resemble services provided at Tier 3 (individual interventions).

Challenges to Transportability

One of the key challenges in generalizing Dinosaur School clinic-based research to school-based application is that it has primarily been conducted with small samples that are predominantly White males with conduct problems from generally well-educated families with two residential parents. School-based programs serve an ethnically and economically diverse group of U.S. students, with less than 50% who are White and approximately 50% who are eligible for free/reduced lunch based on incomes at or below 130% of the poverty level (Department of Education, 2019). In contrast, both Webster-Stratton & Hammond (1997) and Webster-Stratton et al. (2004) evaluated the program with 89% and 83% White children (respectively) and with samples that were 74% and 93% male; all children were diagnosed with Oppositional Defiant Disorder (ODD) or Conduct Disorder. The large majority of parents had some college education and were married or partnered. Total sample sizes in each of these studies were relatively small ($n = 49$ for Webster-Stratton et al., 1997; $n = 56$ for Webster-Stratton et al., 2004). In addition, the one school based RCT of Dinosaur School in Turkey, with students identified by teacher ratings, had only 32 participants and limited socio-economic diversity (Bayrak & Akman, 2018).

Examining whether Dinosaur School's efficacy varies by student characteristics is highly relevant for

school practitioners who may be interested in using the program with populations that are more diverse with regard to race/ethnicity, SES, and presenting problems beyond oppositional behavior. A recent review of IY® research did not identify significant moderation by sex, SES or initial severity of conduct problems (Weeland et al., 2017), although other intervention studies suggests that lower SES status and lower parent education may be associated with worse outcomes (Farahmand et al., 2011; Lundahl et al., 2006) and that elementary-aged boys may demonstrate greater reduction in aggression than girls (Stoltz et al., 2013). In addition, children at higher risk for emotional and behavioral difficulties have been found to benefit more from intervention (Conduct Problems Prevention Research Group, 2010; Stoltz et al., 2013).

Finally, despite well-specified guidelines for implementing IY® programs in schools, fidelity has not been systematically examined in relation to efficacy. This seems particularly important given general challenges of school-based implementation (Atkins et al., 2003; Langley et al., 2010), which may be heightened for Dinosaur School which requires extensive pull out time, a high level of clinical skill, and specific physical space needs for the group. Program structure and process have been identified as the two core aspects of program fidelity, which include dosage and quality of delivery, respectively (Mowbray et al., 2003). Both aspects have been found to predict outcomes in school mental health programs, although delivery quality (or process) is believed to be more significant for outcomes (Leiva Bahamondes & Rojas Andrade, 2018).

Study Objectives

Our primary objective was to evaluate effects of the small group Dinosaur School as it is likely to be delivered in schools - with a heterogeneous sample identified based on teacher referral and co-delivered with school personnel. We examined the following research questions: 1) How effective is Dinosaur School for early elementary students with self-regulation difficulties? 2) Are there follow up effects during the next school year? 3) Does program efficacy vary by student characteristics including ODD symptoms, sex, and parent education? and 4) How is fidelity of group implementation related to student outcomes? Based on IY®'s theory of action and prior literature, we anticipated replicating positive effects on self-regulation outcomes, social competence, and ODD behaviors. We also anticipated stronger effects for students with elevated ODD

symptoms than those without. Given inconsistencies in the literature, we did not have specific hypotheses regarding student sex or parent education as moderators.

Method

Participants

Across two cohorts participating during the 2016-2017 and 2017-2018 school years, nine schools were recruited near a medium-sized city in the southeast, including six schools in an urban district and three in rural counties. Of the nine total schools, three participated in both years. A two-phase process was used to identify students. Phase I occurred in the spring of the year prior to intervention. Kindergarten and first grade teachers nominated students “with challenging behaviors or difficulties managing emotions, interacting with peers, and meeting behavioral expectations in the classroom.” Parent consent for child participation was sought for all nominated children, which resulted in a 54% ($n = 230$ of 425 students) consent rate (see Figure 1). In Phase II, three weeks after the start of the new school year, new teachers rated each nominated and consented child on the Total Difficulties scale of the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), a widely used measure of social-emotional difficulties. Children who were rated above the “risk” threshold (≥ 12) on the total score by this second teacher were enrolled in the study ($n = 138$). Students with autism spectrum disorder (by parent or school counselor report), full-time placement in special education classrooms, significant intellectual deficits and non-proficiency in English (per school report) were not eligible. During the intervention year, 45% of the sample was in first grade and 55% in second grade (mean age = 7.2 years).

As can be seen in Table 1, enrolled children were predominantly male (68%), racially and ethnically diverse (61% Black, 9% Latine 23% White, and 7% multiracial), and predominantly low income (76% receiving free/reduced lunch). A total of 34% lived with both biological parents, 54% with one biological parent, and 12% with other parents/guardians. Seven percent of parents reported their child spoke another language besides English in the home, primarily Spanish. Based on parent report (available for 72% of the sample), 29% had previously been diagnosed with an emotional, behavioral, or learning disorder; the majority (73%) of which were identified as having Attention-Deficit/Hyperactivity Disorder (ADHD). In addition, 22%

were participating in counseling outside of school and 22% were taking medication to help with behavior or mood (only half of whom were also in counseling). By school staff report, the sample also received relatively high rates of educational services including special education and other academic supports as can be seen in Table 1, highlighting the sample's intervention needs.

As can be seen in Figure 1, attrition did not vary by intervention status and was due primarily to student moves. At post-test, 4 students could not be assessed; at follow up during the next school year, 12 students were not assessed (92% retention). There were very few significant differences between students retained and lost on socio-demographic characteristics or outcome variables. At post-test, attriters were rated as more oppositional ($p = .02$), more impaired overall in the classroom ($p = .03$) and as having greater social-behavioral difficulties ($p = .02$ at Time 2 and $p = .04$ at Time 3); they were also an average of 5 months younger.

Measures

Primary outcomes most proximal to intervention targets include measures of self-regulation (inhibitory control, social problem solving, and emotion regulation). Secondary outcomes include disruptive behavior (symptoms of ADHD and ODD, classroom impairment, discipline referrals), and other functional outcomes (grades, peer interactions).¹ We also examined implementation data including measures of fidelity (dosage and quality).

Self-Regulation Outcome Measures

The *Happy-Sad Stroop* (HSS; Lagattuta et al., 2011) requires children to point to a happy face when the examiner says “sad” and visa-versa in 20 trials following four teaching trials, with total number of errors scored (higher scores = worse performance). It is a widely used neurocognitive measure ((Spreen & Strauss, 1991) sensitive to social-emotional interventions in early elementary students (Riggs et al., 2006). Internal consistency based on Cronbach's alpha in the current sample was .70.

¹ Most of our teacher measures have been widely used with racially and ethnically diverse samples with little evidence of differential validity or bias (Mason et al., 2014), although research in this area is limited.

In the *Puzzle Box task* (Eisenberg et al., 1996), children are shown a wooden alphabet puzzle in a large wooden box covered with a cloth (so they cannot see the puzzle) and are asked to complete it in four minutes in order to receive a special prize. Research staff show the child how to cheat (e.g., lifting up the cloth) and then left the room while the child's behavior is video-recorded. Incidents of cheating and off-task behavior are coded and subtracted to obtain the percentage of time persisting. This score consistently predicts parent and teacher ratings of inhibitory control (Cumberland-Li et al., 2004). ICC in the present study was .94.

The *Head Toes Knees Shoulders test* (HTKS; McClelland et al., 2007) has three sections with up to four paired behavioral rules: touch your head/touch your toes and touch your shoulders/touch your knees. Children are instructed to respond according to different rules (e.g., touch their head when told to touch their toes) with increasing cognitive complexity (i.e., head goes with knees, shoulders go with toes) that were designed to reflect the ability to integrate and follow classroom instructions and demands (McClelland & Cameron, 2012). The HTKS demonstrates test-retest stability and is related to teacher rated inhibitory control across time ($r = .27$ and $.21$; (Becker et al., 2014; McClelland et al., 2007). Internal consistency based on Cronbach's alpha in the current sample was .85.

The *Wally Problem-Solving Test* (Webster-Stratton, 1990) includes six colored pictures depicting hypothetical social problem scenarios in a color drawing of a same-sex child (Webster-Stratton & Hammond, 1997; Webster-Stratton & Lindsay, 1999). In 1:1 interview, children are asked to think of as many solutions as they can to each problem (i.e., being teased, wanting a toy you can't have, having one piece of pizza left for you and a friend). Solutions were coded with an adapted coding system based on Lochman & Lampron (1986). Proportion of positive solutions across scenarios was examined, consistent with prior research (Webster-Stratton et al., 2013). ICC based on 20% of double-coded protocols of .91.

The *Emotion Regulation Checklist* (ERC) is a 24-item teacher questionnaire that measures children's emotional lability, negativity, and emotion regulation (Shields & Cicchetti, 1997). The ERC is rated on a 4-point Likert rating scale with higher scores reflecting better regulation. ERC scores are related to academic

competence (Graziano et al., 2007), peer rejection (Kelly et al., 2008), and empathy (Smith, 2001). Internal consistency in this study was good ($\alpha = .81$).

Secondary Social-Behavioral and Academic Outcomes

The *Revised Edition of the School Observation Coding System* (REDSOCS; Jacobs et al., 2000) was used for coding classroom behavior during two 10-minute observations conducted by trained observers during instructional time on different days. Off-task behavior was coded when the child did not attend to the assigned task or exhibited behaviors such as talking to a classmate or being out of seat. Inappropriate behavior was coded for behavior that was distracting or disruptive, such as whining, crying, yelling, and aggression. A percentage of time for each behavior was calculated based on presence across coding intervals, and the two scores were averaged. In this study, ICC's for 20% of double-coded observations was .78 and .86 for Off-Task and Inappropriate, respectively.

The *Coder Observation of Child Adaptation, Revised* (COCA-R; Webster-Stratton et al., 2004) was used to code recess behavior during 25-30 minute sessions using a 0-5 frequency scale (0 = *Almost Never* to 5 = *Almost Always*); higher scores reflect greater difficulties. To strengthen reliability based upon initial piloting, we selected 12 items assessing social competence (drawn from both the *Poor Social Contact* and *Poor Social Health* subscales) with minor wording changes. Items assess behaviors such as the amount of time playing with others; playing appropriate with peers and appearing disliked by classmates. The original measure correlates with teacher ratings of social problems (Webster-Stratton, Reid, et al., 2011). ICC in the present study was .81.

The *Strengths and Weaknesses of ADHD Symptoms and Normal Behavior* (SWAN; Swanson et al., 2012) was used to assess symptoms of ADHD and ODD based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), using a seven-point scale rated from 3 (*far below average*) to -3 (*far above average*), where zero is considered average and higher scores indicate greater difficulties. Separate scores can be calculated for noncompliant and angry/irritable ODD items (Drabick & Gadow, 2012). The SWAN demonstrates strong internal consistency (subscale $\alpha = .89-.97$; Gold et al., 2013; Gomez et al., 2016) and construct validity ($r > .80$) with similar measures (Cornish et al., 2005). Internal consistency (Cronbach's alpha

coefficient) was .93 and .90 for the SWAN Inattentive and Hyperactivity-Impulsivity subscales, respectively, and .84 and .85 for the ODD angry/irritable and non-compliant subscales.

Selected *Strengths and Difficulties Questionnaire* (SDQ; Goodman, 1997) subscales were also used as an outcome. The SDQ has been widely used in research and practice and its reliability and validity are well established (Goodman & Goodman, 2009; Goodman, 2001; Stone et al., 2010) including with a comparable ethnically diverse sample of 1st graders (Hill & Hughes, 2007). It is rated on a scale of 1 (*Not True*) to 3 (*Certainly True*), with higher scores reflecting more difficulties. Internal consistency was .75 for the SDQ HI scale and .72 for the CP subscale.

The *Impairment Rating Scale* (IRS; Fabiano et al., 2006) is a 5-item teacher rating of a child's severity of impairment in peer relations, relationship with teacher, academic progress, classroom functioning, and self-esteem on a 5-point scale ranging from 0 (*no problem; definitely does not need intervention or special services*) to 4 (*extreme problem; definitely needs intervention or special services*). The IRS demonstrates temporal stability and construct validity with elementary samples (Fabiano et al., 2006; Girio-Herrera et al., 2015). Internal reliability in this study was high (Cronbach's alpha = .87).

Academic proficiency was assessed via *report card grades* using students' quarterly report card grades in reading, math, and writing were coded as proficient or not based on a "satisfactory" grade (or numerical grade of 75 or above) to allow for comparison across schools with different grading systems. A summary score reflects the number of academic content domains in which a student was proficient (0-3) for the specific quarter assessed (T1=1st quarter pre-intervention; T2=4th quarter, post intervention; T3=2nd quarter follow up year). Report card grades are considered an ecologically valid measure of students' academic success (Perfect et al., 2014; Rasmussen & Laumann, 2013), and may reflect non-academic factors (Allen & Lamb, 2001) relevant for this study.

Office discipline referrals were recorded by trained school staff and entered into an online data entry site monthly. Referrals were defined as occasions when a student was sent to the office for any disciplinary action. Referrals were analyzed as raw counts.

Fidelity

Fidelity was defined as both dosage (student level session attendance) and group delivery quality. Adequate “dosage” was defined as attendance at 70% or more of the sessions. Delivery quality was rated independently by a nationally certified trainer from Incredible Years, Inc. based on review of three or four randomly selected videos of each group using the *Small Group Therapy Process Checklist* (Webster-Stratton & Reid, 2013). Each group received an average score (1 = *not at all* to 5 = *frequently/extremely well*) based on the average of five items assessing use of recommended group procedures, coleader collaboration, instructional methods, individualizing the program to students, and engagement strategies. Intervention groups were characterized based on quality ratings as either high fidelity (≥ 4 on 1-5 scale) or low fidelity (< 4).

Services

Data were collected at post-test and follow up on services students were receiving in school and their community. School counselors reported on referrals made by teachers for additional evaluation/services, whether students had an Individualized Educational Plan (IEP), and if they received academic and/or counseling supports. Parents reported on whether their children were participating in counseling outside of school and if they were taking medication to help with emotional or behavioral concerns.

Costs

Actual costs required to deliver and support the Dinosaur School program were estimated based on the ingredients method (Levin et al., 2017) from a variety of data sources concurrent with implementation (2015-2018). Resources included salaries of clinical research staff for program delivery and support during 8 months per year as well as student volunteer time, IY® training and consultation fees, program materials, and the opportunity cost of time from school counselors co-delivering the program. Data on “in kind” costs incurred by schools for space, administrative time, etc. were not collected. School staff ($n = 14$) completed surveys asking about actual time spent for program delivery and associated activities during three randomly selected weeks each year. Total estimated costs were then averaged by student.

Procedures

A blocked randomization plan was used. Specifically, for classrooms in which two or more children were enrolled, randomization occurred at the level of individual children within classrooms. All remaining children (who were the only enrolled child in their respective classroom) were combined into a single block and randomization occurred at the level of individual children. Students who were randomized to the intervention condition participated in pull-out intervention groups that were co-led by project staff and school counselors (in total, 13 intervention groups were administered with 4-6 students each). Students randomized to the Business-as-Usual (BAU) condition received services through school and in the community.

Students were observed and assessed at three time points by research assistants who were masked to intervention condition, systematically trained to pre-defined competency standards, and subjected to ongoing reliability review. Pretest and posttest assessments occurred prior to randomization and at the end of active intervention (approximately September and April). A 6-month follow-up visit was conducted in the subsequent school year. Individual child assessments were completed during 45-minute pull-out sessions with trained RAs at school who were blind to randomization status.

Intervention Implementation

A total of 13 intervention groups were provided across two cohorts during pull-out sessions during the school day. As detailed elsewhere (Authors, 2019), we made some key adaptations to the program for delivery within schools in consultation with the developer. The most significant of these were: 1) changes to group structure to accommodate students' class schedule (twice per week for 45 minutes instead of once weekly for 2 hours), 2) addition of "recess coaching" to compensate for less time in groups and promote generalizability, and 3) group co-facilitation by research therapists paired with school counselors or other trained student support staff. In addition, we offered two hours of teacher in-service focused on basic concepts about self-regulation development and positive behavior management; additional teacher consultation addressed individual students' behavioral success in the classroom. Three parent workshops were provided to share highlights of skills students were learning and suggestions for reinforcing these at home in addition to weekly homework activities; materials were translated into Spanish and interpretation provided as needed. Dinner, language interpreters, and

childcare were provided to support attendance in addition to \$25 compensation for each meeting to assist families with transportation or other expenses incurred to attend.

Data Analysis

Initially we conducted descriptive analyses, including examining implementation and costs. Prior to conducting outcome analyses, we engaged in data reduction efforts to reduce the number of outcomes as is consistent with educational research guidelines for addressing concerns of multiple comparisons. For outcomes in which the random effect was not statistically significant, the model was re-estimated with only a residual variance term. Each model was estimated twice, once each for posttest and follow-up outcomes.

To address our third question regarding the differential effectiveness of the intervention as a function of student characteristics, we extended the model to include child sex, child ODD risk, and parent education as moderators in separate models. Statistically significant two-way interaction terms (i.e., intervention x moderator) were probed to determine the differential magnitude of intervention effects as a function of each moderator. Finally, we conducted sensitivity analyses to examine fidelity as previously defined, using similar ITT model specifications and HLM modeling to 1) determine if results changed when 6 students with inadequate dosage were dropped, and 2) compare students in high and low fidelity groups to those in BAU. Effect sizes were computed as the difference between the intervention and control group least squares means (adjusted for pretest values) divided by the control group standard deviation.

RESULTS

Baseline Characteristics and Equivalence

As can be seen in Table 2, the sample demonstrated considerable evidence of self-regulation difficulties at baseline. On inhibitory control measures, students persisted on the puzzle box task only 50% of the time (lower than the 63% average in a community sample for this age; Zhou et al., 2007). On the HTKS, first graders averaged 53% correct and second graders averaged 75%, similar to a southwestern US sample that was 50% Hispanic (Hernández et al., 2018). Mean number of errors on the HSS was 5, which is at the 80th% (Lagattuta et al., 2011). Mean ADHD and ODD scores on the SWAN were above average (> 1) and 38% of the sample was

identified as at risk for ODD. The sample was off-task or engaged in inappropriate classroom behavior 28% of the time. Overall impairment on the IRS ($M = 2.53$) was similar to that of sample of children identified as at-risk for emotional-behavioral difficulties (Girio-Herrera et al., 2015) as were SDQ mean scores ($M = 18.65$). A total of 53 students (39%) received a discipline referral, primarily due to noncompliance and defiance (50%), disrupting class (33%) and physical aggression towards peers (32%). The sample was also quite academically impaired, with only 22%, 17%, and 34% considered proficient in reading, writing, and math, respectively. Intervention and BAU groups did not differ significantly on any demographic characteristic or baseline variables with one exception; intervention students were rated as having more difficulties in observed social competence on the COCA.

Service Use

Students in the intervention and BAU groups differed in services they received beyond participation in the Dinosaur School program at the end of the first school year and at follow up (see Table 1). There appeared to be a pattern for those in the comparison group to have received greater services outside of school than did those in the intervention group, including counseling ($X^2 = 3.94$, $p = .06$ at follow up) and medication ($X^2 = 4.86$, $p = .03$), although we do not have data on whether these were pre-randomization differences. On the other hand, those who participated in the intervention were more than twice as likely to receive special education services (10.8%) as BAU students (4.6%), $X^2 = 5.47$ $p = .02$ at follow up; 70% of intervention students began IEP services during the intervention year relative to 42% of comparison students. There was little indication of differences between groups in rates of referrals to the school intervention team or academic support services provided.

Implementation Analyses

Full details on implementation data are provided in Authors (2019), which includes data from a pilot year as well as the two cohorts for whom data were analyzed for this study. Students attended an average of 34.4 ($SD = 1.9$) 45-minute sessions. Student attendance averaged 86% ($SD=19\%$) and recess coaching sessions averaged 20-25 minutes with 12.5 sessions per child ($SD = 3.6$). Each child's teacher also participated in

individual in-person or phone consultation meetings ($M = 5.8$, $SD = 2.0$) and 76% of teachers attended at least one in-service meeting. Average parent attendance was 56.5% across three meetings, with 78% attending at least one meeting (see Authors, in press for additional details on parent engagement methods).

Based on group leader report on session checklists, 90.68% ($SD = 1.56$) of the required content and activities were completed across sessions and groups. An average of 43 ($SD = 13.24$) vignettes were shown, notably fewer (1.26 per session) than the recommended two per session, although this varied considerably across the 13 groups. Average fidelity ratings were 3.90 on a 1-5 scale. Although these fidelity data are considered good, clear guidelines for interpretation are not available.

Costs per Student

Based on school staff report, co-delivering the program required 1 hour and 17 minutes per week to deliver the program ($SD = .64$) and 55 minutes per week ($SD = .51$) to participate in other intervention-related activities. Average intervention costs incurred per student was \$7,850, of which 80% was for staff time, accounting for full-time clinical research staff salaries during each school year, best practice training and supervision, and numerous quality control procedures.

Dinosaur School Effectiveness

Unadjusted means of all outcome variables are presented in Table 2. ITT analyses addressing the effectiveness of Dinosaur School identified only one statistically significant main intervention effect, for the Wally's Problem-Solving Test ($\beta = .37$, $SE = .08$, $p < .0001$). Examination of adjusted means (based on z-scores) indicated that intervention students ($M = .50$, $SE = .06$) generated a higher percentage of positive solutions to social problems than did comparison students ($M = .13$, $SD = .06$) at post-test, with a small effect size ($d = .37$). Effect sizes for all outcomes can be seen in Figure 2, reflecting small intervention effects in the expected directions that did not reach conventional levels of statistical significance.

Follow up Effects

Regarding our second research question addressing follow up effects, significant intervention effects were observed for inhibitory control ($\beta = .18$, $SE = .07$, $p = .01$) and ODD ($\beta = -.41$, $SE = .14$, $p = .01$), $d = .24$

and -.45, respectively. Interestingly, although not statistically significant ($p = .15$), fewer students in the intervention condition met a risk threshold for ODD (i.e., at least four symptoms of ODD) at follow-up than did students in the BAU group (38% vs. 51%). Again, the remaining outcomes consisted of small intervention effects that did not reach conventional levels of statistical significance (see Figure 2).

Moderators of Intervention Effects

As seen in Figure 3, significant interaction effects were found by ODD risk for observed social competence ($\beta = .66$, $SE = .31$, $p = .04$), teacher-rated classroom impairment ($\beta = -.81$, $SE = .29$, $p = .01$), and the number of discipline referrals students received ($\beta = 2.66$, $SE = 1.14$, $p = .02$) at post-test. When the model interaction was significant, we tested the simple effect of BAU versus intervention for ODD risk type (low or high) using model estimated means. Upon probing, there were significantly fewer discipline incidents for the low ODD risk intervention group relative to the low ODD risk BAU group ($d = -.41$). For the high ODD risk group, impairment ratings were lower ($d = -.57$) and social competence was higher ($d = .43$). None of the interaction terms involving child sex or parent education were statistically significant.

At follow up, none of the post-intervention moderator effects remained statistically significant (see Figure 4). However, several new significant moderator effects were found in the expected direction at follow up for problem-solving ($\beta = .81$, $SE = .28$, $p = .02$), emotion regulation ($\beta = -.87$, $SE = .35$, $p < .01$), and academic proficiency ($\beta = -.77$, $SE = .32$, $p = .02$). When we tested the simple effect of BAU versus intervention for ODD risk type, we found all effects occurred for the high ODD risk group with moderate to large effects on these three outcomes ($d = .60 - .75$). Statistically nonsignificant but moderate-sized effect ($d \geq .5$) were also seen for ODD and ADHD outcomes.

Fidelity Analyses

As noted, we were interested in whether results vary in relation to fidelity defined by child group attendance (receiving adequate “dosage”), and group delivery quality (based on video ratings). To examine dosage, we used the previous ITT model specifications but with removal of six children who did not receive the full dose of the intervention. These sensitivity analyses results identified significant effects for problem-solving

($B = .38$, $SE = .08$, $p < .001$) as well as emotion regulation ($B = .27$, $SE = .14$, $p = .05$), similar to the main ITT effects.

To examine quality, we redefined the intervention group into two groups of students; those in groups with low ($n = 44$) and high quality ($n = 25$) and compared each to BAU. Significant effects were found for problem-solving in the low-fidelity group at post- test ($d = .50$), contrary to expectation. At follow up, effects were significant for students in the high-fidelity group for inhibitory control ($d = .41$), ODD ($d = -.72$), emotion regulation ($d = .53$) and inappropriate classroom behavior ($d = -.43$), in the expected direction. These are notably larger than effects seen in ITT analyses for the overall sample.

DISCUSSION

This study examined the effectiveness of the IY® Dinosaur School program, designed for young children with ODD seen in clinics, as delivered to early elementary school students with a range of self-regulation difficulties. We examined effects on a range of self-regulation outcomes as well as social-emotional and academic competence both at end of program delivery and at a 6-month follow up during the next school year. Given that the program was designed for children with conduct problems, we evaluated it's effects on students with and without elevated symptoms of ODD. To further understanding of potential intervention moderators, we also examined whether outcomes varied by parent education and student sex. Finally, we conducted sensitivity analyses to better understand if fidelity of implementation influenced results.

This study extends prior Dinosaur School efficacy research in several important ways. First, it reflects the most rigorous evaluation of this program within a school context to date. It is the first school-based RCT conducted within the U.S. Second, it is the only study of the Dinosaur School program that includes controlled follow up data, which is particularly critical for understanding benefits for young children who may experience developmental improvement over time. Third, we address limitations of prior Dinosaur School research by including a predominantly low-income, Black and Latine sample with a significant number of girls. Finally, we examined fidelity in relation to efficacy to increase scientific understanding of implementation issues in

program transportability. Although this work focuses on one specific social-emotional program, we believe results can inform a broad range of school-based mental health interventions.

Intervention Effects

Main effects of the intervention at immediate post-test were generally small, albeit in the expected direction ($d = .07 - .28$), with only one of 10 statistically significant outcomes. The same general pattern of findings was evident at follow-up ($d = .07 - .24$ with the exception of ODD which was $d = -.45$). Effect sizes were larger for primary self-regulation outcomes (i.e., inhibitory control, emotion regulation, and problem-solving) than for most secondary outcomes, including those that are likely most relevant to school staff (i.e., grades and discipline referrals). The largest sustained effect was seen on ODD symptoms, which are directly targeted in the program's theory of change. In terms of practical significance, the percentage of students who were at risk for ODD decreased over a one-year period from 46% to 38% in the intervention group and increased for the BAU group (47% to 51%), a 13% relative risk decrease, suggesting potential benefits of the intervention in shifting a negative developmental trajectory.

Subgroup analyses indicated that the intervention was equally effective for boys and girls and for families with varying parent educational levels. However, it appeared notably more effective for students at high risk for ODD, as expected given that the Dinosaur School program was originally developed for young children with diagnoses of ODD and/or conduct problems. For those students, effect sizes were consistently larger than for the full sample across almost all outcomes at both post-intervention ($d = .25 - .55$ with one exception) and at follow up ($d = .37 - .84$ exclusive of observational measures, with smaller effects). With regard to impact that may be meaningful to schools, significant effects were observed on teacher ratings of emotion regulation and classroom impairment as well as grades (at follow up).

Relative to other evaluations of the IY® Dinosaur School program as an independent intervention delivered in schools, results were similar to Bayrak and Akman (2018) and Hutchings et al. (2007) in showing the strongest effects on students' abilities to generate positive solutions to social problems on the Wally Problem-Solving Test, which as noted is highly aligned with the intervention. Those two studies also failed to

show significant effects on classroom observations and teacher ratings of social competence with much smaller samples than the present study, although the study authors cited trends in means in the expected direction favoring the intervention group. Relative to the broader literature, effects seen in this study for the Dinosaur School program are comparable to what has been reported for other SEL interventions (i.e., $ES = .26 - .54$), where smaller effects are typically seen for universal samples and moderate sized effects for children at risk for behavior disorders (Gansle, 2005; Reddy et al., 2009; Wilson et al., 2003). Because our study differed in several ways from prior evaluations of the clinic version of Dinosaur School, direct comparisons to results of those studies are not interpretable.

Relevance for Schools

In terms of the transportability of the IY® small group Dinosaur School program from clinic to schools, we experienced clear implementation challenges that resulted in fewer hours of session delivery and lower fidelity than expected (especially for showing videos), despite several adaptations made to enhance fit and feasibility. As detailed in Authors (2019) this may be related to a variety of dynamic and interactive factors such as highly challenging student behaviors that were sometimes negatively reinforced by peers, limitations in group leader management skills, and school discipline policies that interfered with fully implementing positive behavioral approaches such as ignoring. Additionally, implementation may have been negatively impacted in schools where teacher stress levels were high, where there were fewer resources, and a lack of universal positive behavior supports. As expected, fidelity *did* matter with regard to student outcomes, particularly at follow up when students in high fidelity groups demonstrated better self-regulation and less inappropriate and oppositional behavior than students who received BAU services. Our experiences and findings reflect the broader literature on implementation of evidence-based school mental health programs (Evans et al., 2014; Forman et al., 2009; Langley et al., 2010), and suggest that caution and study are needed in transporting EBPs across contexts.

With regard to the effectiveness of the Dinosaur School program, results suggest benefits on self-regulation and disruptive behavior for those students at risk for ODD, at a level that is likely to be observable to

school staff (i.e., medium to large effect sizes). However, there was little evidence of impact on other outcomes relevant to school functioning, such as discipline referrals or on observed classroom and playground behavior. Costs to deliver the Dinosaur School program were estimated at \$7,850 as implemented within this RCT, which may be higher than costs to deliver in practice, particularly over multiple years when training costs would decrease. This amount exceeds the annual \$5,700 educational costs of ADHD, which includes special education, retention, and discipline referrals (Robb et al., 2011; converted to 2018 costs). However, cumulative costs for Conduct Disorder across 13 years of schooling have been estimated to be over \$100,000 (converted to 2018 costs; Foster et al., 2006). Thus, to the extent that future intensive educational services and crime costs are prevented, the program might be considered cost-effective. Schools will certainly need to consider program costs within the context of multiple competing resource demands, ideally with long-term cost benefits in mind.

With regard to other intervention approaches that might be considered for young children with significant self-regulation difficulties, particularly those who experience socio-economic adversity and who have been historically marginalized, we recommend integrating ecological approaches that address both the school and family context as well as support children's resilience and skill development. This is, in fact, the rationale for the Incredible Years® program series that includes extensive trainings for teachers (42 hours) as well as parents (36-40 hours). It is certainly possible that implementation of all three components may enhance program outcomes, as seen in Webster-Stratton et al. (2004), although implementation and costs of such a comprehensive approach would likely be very difficult for schools to support. School-community partnerships could provide additional supports for children and families, although there are many challenges with financing evidence-based programs in community mental health as well (Stewart et al., 2015).

It is also important to note that, consistent with the premise of multi-tiered interventions in schools (Stephan et al., 2015), providing integrated supports at both universal and targeted levels should be more effective for students with self-regulation difficulties than targeted interventions like the Dinosaur School program alone. In our study, we observed that many schools lacked consistent school-wide positive behavior support systems (Tier 1), which may be necessary for successful implementation of Tier 2 programs. Future

research is needed to determine the extent to which Tier 2 programs are dependent on Tier 1 programs and inform how best to combine multi-tiered school mental health interventions (Fazel et al., 2014). One such research opportunity would be to evaluate the universal Tier 1 classroom Dina program in combination with the Tier 2 small group Dinosaur School program.

Finally, there are a few general clinical implications of this work for schools. The first is that students should be selected carefully for groups to ensure that a Tier 2 group-based intervention is appropriate and that Tier 3 individualized interventions are not indicated. Attempting to serve students with an inappropriate level of service may decrease their own benefit as well as risk negative impact for other children. In addition, it is clear that programs adopted by schools should be utilized for the populations for which they were intended. In this study, including children with self-regulation difficulties beyond oppositionality did not result in significant positive effects like it did for those students with oppositional behavior. And as noted, attempting to adapt programs developed in clinics and fit them into school contexts may not be a useful approach. Instead we recommend designing interventions specifically for delivery within schools that can consider the resources, training, and constraints of schools up front, while simultaneously leveraging the school environment to support skill generalization and promote positive peer interactions (e.g., Masia Warner et al., 2007).

Limitations

Interpretation of findings should be considered in the context of several limitations. First, we enrolled a smaller sample than intended due to recruitment and implementation challenges that required an unplanned pilot year. Second, our heterogeneous sample reflects students who may be referred for school-based supports for disruptive and dysregulated behavior but precluded direct comparisons to prior randomized trials of this program that have focused on children with specific ODD behaviors and diagnoses. And as we saw, this heterogeneity likely undermined the overall intervention effects. Third, it is clear that some of our groups were not conducted with the ideal fidelity, which may be related to a number of factors discussed further in Authors (2019), and that may contribute to decreased overall effects.

Finally, overall effect sizes may have been impacted by almost one-third of the BAU students utilizing community-based services (counseling and medication), which we suspect may be related to parents' seeking additional supports when their child was not randomized to intervention. However, we cannot determine this given that community service use was not assessed at baseline, a recommendation we would make for others doing school-based mental health research. In addition, intervention students were also more likely to receive new special education services, which is a confound for assessing intervention effects. Although other services beyond a specified intervention typically cannot be controlled in clinical trials for ethical reasons, it is certainly important to assess to facilitate interpretation of intervention effects.

Conclusion

For early elementary students with a broad range of self-regulation difficulties the Dinosaur School program appears to have small to moderate effects on self-regulation related outcomes and oppositional behavior, with some evidence of maintenance of effects although this was inconsistent across outcomes. Results were notably more positive and sustained for those children at high risk for ODD, which suggests that this should be the targeted population for this intervention. However, there was considerably more room for improvement suggesting that more comprehensive programs involving parents, teachers, and/or multi-tiered intervention may be needed. For school-based mental health intervention, this may require community partnerships and significant resources that can be invested to prevent costly longer-term negative educational and economic consequences for young children at risk for conduct problems.

References

- Allen, J. D., & Lambating, J. (2001, November 21). Validity and reliability in assessment and grading: Perspectives of preservice and inservice teachers and teacher education professors. *American Educational Research Association*.
- Atkins, M. S., Graczyk, P., Frazier, S. L., & Adil, J. A. (2003). Toward a new model for school-based mental health: Accessible, effective, and sustainable services in urban communities. *School Psychology Review*, 12, 503–514.
- Bailey, R., & Jones, S. M. (2019). An Integrated Model of Regulation for Applied Settings. *Clinical Child and Family Psychology Review*, 22(1), 2–23. <https://doi.org/10.1007/s10567-019-00288-y>
- Bayrak, H. U., & Akman, B. (2018). Adaptation of the “incredible years child training program” and investigation of the effectiveness of the program. *Educational Sciences: Theory & Practice*. <https://doi.org/10.12738/estp.2018.2.0259>
- Becker, D. R., McClelland, M. M., Loprinzi, P., & Trost, S. G. (2014). Physical Activity, Self-Regulation, and Early Academic Achievement in Preschool Children. *Early Education and Development*, 25(1), 56–70. <https://doi.org/10.1080/10409289.2013.780505>
- Conduct Problems Prevention Research Group. (2010). The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology*, 78(2), 156–168. <https://doi.org/10.1037/a0018607>
- Cornish, K. M., Manly, T., Savage, R., Swanson, J., Morisano, D., Butler, N., Grant, C., Cross, G., Bentley, L., & Hollis, C. P. (2005). Association of the dopamine transporter (DAT1) 10/10-repeat genotype with ADHD symptoms and response inhibition in a general population sample. *Molecular Psychiatry*, 10(7), 686–698. <https://doi.org/10.1038/sj.mp.4001641>
- Cumberland-Li, A., Eisenberg, N., & Reiser, M. (2004). Relations of young children’s agreeableness and

resiliency to effortful control and impulsivity. *Social Development*, 13(2), 193–212.

<https://doi.org/10.1111/j.1467-9507.2004.000263.x>

Department of Education, N. C. for E. S. (2019). *Fast facts: Back to school statistics*.

<https://nces.ed.gov/fastfacts/display.asp?id=898>

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Current Directions in Psychological Science : A Journal of the American Psychological Society*, 21(5), 335–341.

<https://doi.org/10.1177/0963721412453722>

Drabick, D. A. G., & Gadow, K. D. (2012). Deconstructing oppositional defiant disorder: clinic-based evidence for an anger/irritability phenotype. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(4), 384–393. <https://doi.org/10.1016/j.jaac.2012.01.010>

Eisenberg, N., Fabes, R. A., Guthrie, I. K., Murphy, B. C., Maszk, P., Holmgren, R., & Suh, K. (1996). The relations of regulation and emotionality to problem behavior in elementary school children. *Development and Psychopathology*, 8(01), 141–162.

Evans, S. W., Stephan, S. H., & Sugai, G. (2014). Advancing research in school mental health: introduction of a special issue on key issues in research. *School Mental Health*, 6(2), 63–67. <https://doi.org/10.1007/s12310-014-9126-8>

Fabiano, G. A., Pelham, W. E., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., Onyango, A. N., Kipp, H., Lopez-Williams, A., & Burrows-Maclean, L. (2006). A practical measure of impairment: psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child and Adolescent Psychology*, 35(3), 369–385. https://doi.org/10.1207/s15374424jccp3503_3

Farahmand, F. K., Grant, K. E., Polo, A. J., & Duffy, S. N. (2011). School-based mental health and behavioral programs for low-income, urban youth: A systematic and meta-analytic review. *Clinical Psychology:*

Science and Practice, 18(4), 372–390. <https://doi.org/10.1111/j.1468-2850.2011.01265.x>

Farmer, E. M. Z., Burns, B. J., Phillips, S. D., Angold, A., & Costello, E. J. (2003). Pathways into and through mental health services for children and adolescents. *Psychiatric Services*, 54(1), 60–66.

<https://doi.org/10.1176/appi.ps.54.1.60>

Fazel, M., Hoagwood, K., Stephan, S., & Ford, T. (2014). Mental health interventions in schools 1: Mental health interventions in schools in high-income countries. *The Lancet. Psychiatry*, 1(5), 377–387.

[https://doi.org/10.1016/S2215-0366\(14\)70312-8](https://doi.org/10.1016/S2215-0366(14)70312-8)

Forman, S. G., Olin, 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(1), 26–36.

<https://doi.org/10.1007/s12310-008-9002-5>

Foster, E. M., Jones, D., & Group, C. P. P. R. (2006). Can a costly intervention be cost-effective?: An analysis of violence prevention. *Archives of General Psychiatry*, 63(11), 1284–1291.

<https://doi.org/10.1001/archpsyc.63.11.1284>

Gansle, K. A. (2005). The effectiveness of school-based anger interventions and programs: A meta-analysis. *Journal of School Psychology*, 43(4), 321–341. <https://doi.org/10.1016/j.jsp.2005.07.002>

Girio-Herrera, E., Dvorsky, M. R., & Owens, J. S. (2015). Mental health screening in kindergarten youth: a multistudy examination of the concurrent and diagnostic validity of the Impairment Rating Scale.

Psychological Assessment, 27(1), 215–227. <https://doi.org/10.1037/a0037787>

Gold, A. B., Ewing-Cobbs, L., Cirino, P., Fuchs, L. S., Stuebing, K. K., & Fletcher, J. M. (2013). Cognitive and behavioral attention in children with math difficulties. *Child Neuropsychology*, 19(4), 420–437.

<https://doi.org/10.1080/09297049.2012.690371>

Gomez, R., Vance, A., & Gomez, R. M. (2016). Maternal ratings of the ADHD symptoms: Subtypes versus severity in clinic-referred children and adolescents. *Journal of Attention Disorders*, 20(5), 414–423.

<https://doi.org/10.1177/1087054713514606>

Goodman, A., & Goodman, R. (2009). Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health. *Journal of the American Academy of Child and Adolescent Psychiatry*.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*.

Goodman, Robert. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38(5), 581–586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>

Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The Role of Emotion Regulation and Children's Early Academic Success. *Journal of School Psychology*, 45(1), 3–19.

<https://doi.org/10.1016/j.jsp.2006.09.002>

Greenberg, M. T. (2006). Promoting resilience in children and youth: Preventive interventions and their interface with neuroscience. *Annals of the New York Academy of Sciences*, 1094, 139–150.

<https://doi.org/10.1196/annals.1376.013>

Gross, J. J., & Jazaieri, H. (2014). Emotion, Emotion Regulation, and Psychopathology: An Affective Science Perspective. *Clinical Psychological Science*, 2(4), 387–401. <https://doi.org/10.1177/2167702614536164>

Hernández, M. M., Eisenberg, N., Valiente, C., Spinrad, T. L., Johns, S. K., Berger, R. H., Silva, K. M., Diaz, A., Gal-Szabo, D. E., & Thompson, M. S. (2018). Self-regulation and academic measures across the early elementary school grades: Examining longitudinal and bidirectional associations. *Early Education and Development*, 29(7), 914–938.

Hill, C. R., & Hughes, J. N. (2007). An examination of the convergent and discriminant validity of the strengths and difficulties questionnaire. *School Psychology Quarterly : The Official Journal of the Division of School Psychology, American Psychological Association*, 22(3), 380–406. <https://doi.org/10.1037/1045-3830.22.3.380>

- Hutchings, J., Bywater, T., Daley, D., & Lane, E. (2007). A pilot study of the Webster Stratton Incredible Years Therapeutic Dinosaur School programme. *Clinical Psychology Forum*, 170, 21–24.
- Jacobs, J. R., Boggs, S. R., Eyberg, S. M., Edwards, D., Durning, P., Querido, J. G., McNeil, C. B., & Funderburk, B. W. (2000). Psychometric properties and reference point data for the revised edition of the school observation coding system. *Behavior Therapy*, 31(4), 695–712. [https://doi.org/10.1016/S0005-7894\(00\)80039-8](https://doi.org/10.1016/S0005-7894(00)80039-8)
- Jones, S. M., Zaslow, M., Darling-Churchill, K. E., & Halle, T. G. (2016). Assessing early childhood social and emotional development: Key conceptual and measurement issues. *Journal of Applied Developmental Psychology*, 45, 42–48. <https://doi.org/10.1016/j.appdev.2016.02.008>
- Kelly, B. M., Schwartz, D., Gorman, A. H., & Nakamoto, J. (2008). Violent victimization in the community and children's subsequent peer rejection: the mediating role of emotion dysregulation. *Journal of Abnormal Child Psychology*, 36(2), 175–185. <https://doi.org/10.1007/s10802-007-9168-6>
- Lagattuta, K. H., Sayfan, L., & Monsour, M. (2011). A new measure for assessing executive function across a wide age range: children and adults find happy-sad more difficult than day-night. *Developmental Science*, 14(3), 481–489. <https://doi.org/10.1111/j.1467-7687.2010.00994.x>
- Langley, A. K., Nadeem, E., Kataoka, S. H., Stein, B. D., & Jaycox, L. H. (2010). Evidence-based mental health programs in schools: Barriers and facilitators of successful implementation. *School Mental Health*, 2(3), 105–113. <https://doi.org/10.1007/s12310-010-9038-1>
- Leiva Bahamondes, L., & Rojas Andrade, R. (2018). Validación de la Escala de Problemas Emocionales y del Comportamiento Youth Self Report (YSR/11-18) en una muestra de adolescentes chilenos. *Terapia Psicológica*, 36(1), 27–36. <https://doi.org/10.4067/s0718-48082017000300023>
- Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2017). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. SAGE publications.

- Lochman, J. E., & Lampron, L. B. (1986). Situational social problem-solving skills and self-esteem of aggressive and nonaggressive boys. *Journal of Abnormal Child Psychology*, 14(4), 605–617.
<https://doi.org/10.1007/BF01260527>
- Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: moderators and follow-up effects. *Clinical Psychology Review*, 26(1), 86–104. <https://doi.org/10.1016/j.cpr.2005.07.004>
- Masia Warner, C., Fisher, P. H., Shrout, P. E., Rathor, S., & Klein, R. G. (2007). Treating adolescents with social anxiety disorder in school: an attention control trial. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 48(7), 676–686. <https://doi.org/10.1111/j.1469-7610.2007.01737.x>
- Mason, B. A., Gunersel, A. B., & Ney, E. A. (2014). CULTURAL AND ETHNIC BIAS IN TEACHER RATINGS OF BEHAVIOR: A CRITERION-FOCUSED REVIEW. *Psychology in the Schools*, 51(10), 1017–1030. <https://doi.org/https://doi.org/10.1002/pits.21800>
- McClelland, M. M., & Cameron, C. E. (2012). Self-Regulation in Early Childhood: Improving Conceptual Clarity and Developing Ecologically Valid Measures. *Child Development Perspectives*, 6(2), 136–142. <https://doi.org/https://doi.org/10.1111/j.1750-8606.2011.00191.x>
- McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, 43(4), 947–959. <https://doi.org/10.1037/0012-1649.43.4.947>
- Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity Criteria: Development, Measurement, and Validation. *American Journal of Evaluation*, 24(3), 315–340.
<https://doi.org/10.1177/109821400302400303>
- Murray, D. W., Rosanbalm, K., Christopoulos, C., & Meyer, A. L. (2019). An Applied Contextual Model for Promoting Self-Regulation Enactment Across Development: Implications for Prevention, Public Health and Future Research. *The Journal of Primary Prevention*, 40(4), 367–403. <https://doi.org/10.1007/s10935->

019-00556-1

- Nigg, J. T. (2017). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 58(4), 361–383. <https://doi.org/10.1111/jcpp.12675>
- Owens, J. S., & Murphy, C. E. (2004). Effectiveness research in the context of school-based mental health. *Clinical Child and Family Psychology Review*, 7(4), 195–209. <https://doi.org/10.1007/s10567-004-6085-x>
- Perfect, M. M., Levine-Donnerstein, D., Archbold, K., Goodwin, J. L., & Quan, S. F. (2014). The contribution of sleep problems to academic and psychosocial functioning. *Psychology in the Schools*, 51(3), 273–295. <https://doi.org/10.1002/pits.21746>
- Rasmussen, M., & Laumann, K. (2013). The academic and psychological benefits of exercise in healthy children and adolescents. *European Journal of Psychology of Education*, 28(3), 945–962. <https://doi.org/10.1007/s10212-012-0148-z>
- Reddy, L. A., Newman, E., De Thomas, C. A., & Chun, V. (2009). Effectiveness of school-based prevention and intervention programs for children and adolescents with emotional disturbance: A meta-analysis. *Journal of School Psychology*, 47(2), 77–99. <https://doi.org/10.1016/j.jsp.2008.11.001>
- 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(1), 91–102. <https://doi.org/10.1007/s11121-005-0022-1>
- Ringeisen, H., Henderson, K., & Hoagwood, K. (2003). Context matters: Schools and the “research to practice gap” in children’s mental health. *School Psychology Review*, 32(2), 153–168. <http://www.scopus.com/inward/record.url?eid=2-s2.0-0043156189&partnerID=40&md5=16ea6b27c52fe862cd476f60a01e9305>

- Robb, J. A., Sibley, M. H., Pelham Jr, W. E., Foster, E. M., Molina, B. S. G., Gnagy, E. M., & Kuriyan, A. B. (2011). The Estimated Annual Cost of ADHD to the U.S. Education System. *School Mental Health*, 3(3), 169–177. <https://doi.org/10.1007/s12310-011-9057-6>
- Scott, S. (2008). An update on interventions for conduct disorder. *Advances in Psychiatric Treatment*, 14(1), 61–70.
- Shields, A., & Cicchetti, D. (1997). Emotion regulation among school-age children: the development and validation of a new criterion Q-sort scale. *Developmental Psychology*, 33(6), 906.
- Smith, M. (2001). Social and Emotional Competencies: Contributions to Young African-American Children's Peer Acceptance. *Early Education & Development*, 12(1), 49–72. https://doi.org/10.1207/s15566935eed1201_4
- Spren, O., & Strauss, E. (1991). *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary*. Oxford University Press.
- Stephan, S. H., Sugai, G., Lever, N., & Connors, E. (2015). Strategies for integrating mental health into schools via a multitiered system of support. *Child and Adolescent Psychiatric Clinics*, 24(2), 211–231.
- Stewart, R. E., Adams, D. R., Mandell, D. S., Hadley, T. R., Evans, A. C., Rubin, R., Erney, J., Neimark, G., Hurford, M. O., & Beidas, R. S. (2015). The Perfect Storm: Collision of the Business of Mental Health and the Implementation of Evidence-Based Practices. *Psychiatric Services*, 67(2), 159–161. <https://doi.org/10.1176/appi.ps.201500392>
- Stoltz, S., Deković, M., van Londen, M., Orobio de Castro, B., & Prinzie, P. (2013). What Works for Whom, How and under What Circumstances? Testing Moderated Mediation of Intervention Effects on Externalizing Behavior in Children. *Social Development*, 22(2), 406–425. <https://doi.org/10.1111/sode.12017>
- Stone, L. L., Otten, R., Engels, R. C. M. E., Vermulst, A. A., & Janssens, J. M. A. M. (2010). Psychometric

properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: A review. *Clinical Child and Family Psychology Review*, 13(3), 254–274.

<https://doi.org/10.1007/s10567-010-0071-2>

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., Clevenger, W., Wasdell, M., McCleary, R., Lakes, K., & Wigal, T. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and the SWAN rating scales. *The International Journal of Educational and Psychological Assessment*, 10(1), 51–70.

<https://www.ncbi.nlm.nih.gov/pubmed/26504617>

Ursache, A., Blair, C., & Raver, C. C. (2012). The promotion of self-regulation as a means of enhancing school readiness and early achievement in children at risk for school failure. *Child Development Perspectives*, 6(2), 122–128. <https://doi.org/10.1111/j.1750-8606.2011.00209.x>

Venter, J., Cardinal, A. M., Shaw, M., Wiskerchen, L., DeSimone, S., & Ledesma, A. (2012). School-based intervention for k-2nd graders with disruptive behavior disorders. *Paper Presented at the Southwest Interdisciplinary Research Center, Arizona State University*.

Weare, K., & Nind, M. (2011). Mental health promotion and problem prevention in schools: What does the evidence say and what are the messages for Europe? *Health Promotion International*, 26, i29–i69.

Webster-Stratton, C. (1990). *Wally Game: A problem solving test*. .

Webster-Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology*, 65(1), 93–109. <https://doi.org/10.1037/0022-006X.65.1.93>

Webster-Stratton, C., & Herman, K. C. (2009). Disseminating incredible years series early-intervention programs: Integrating and sustaining services between school and home. *Psychology in the Schools*, n/a-n/a. <https://doi.org/10.1002/pits.20450>

- Webster-Stratton, C., & Lindsay, D. W. (1999). Social competence and conduct problems in young children: Issues in assessment. *Journal of Clinical Child Psychology*, 28(1), 25–43.
- Webster-Stratton, C., & Reid, J. (2013). Tailoring the Incredible Years Parent, Teacher, and Child Interventions for Young Children with ADHD. In J. Ghuman & H. Ghuman (Eds.), *ADHD in Preschool Children: Assessment and Treatment* (1st ed., pp. 113–131). Oxford University Press.
- Webster-Stratton, C., Reid, M. J., & Beauchaine, T. (2011). Combining parent and child training for young children with ADHD. *Journal of Clinical Child and Adolescent Psychology*, 40(2), 191–203.
<https://doi.org/10.1080/15374416.2011.546044>
- Webster-Stratton, C., Reid, M. J., & Beauchaine, T. P. (2013). One-year follow-up of combined parent and child intervention for young children with ADHD. *Journal of Clinical Child and Adolescent Psychology*, 42(2), 251–261. <https://doi.org/10.1080/15374416.2012.723263>
- Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology*, 33(1), 105–124. https://doi.org/10.1207/S15374424JCCP3301_11
- Webster-Stratton, C., Reinke, W. M., Herman, K. C., & Newcomer, L. L. (2011). The Incredible Years Teacher Classroom Management Training: The Methods and Principles That Support Fidelity of Training Delivery. *Online Submission*, 40(4), 509–529.
- Weeland, J., Chhangur, R. R., van der Giessen, D., Matthys, W., de Castro, B. O., & Overbeek, G. (2017). Intervention Effectiveness of The Incredible Years: New Insights Into Sociodemographic and Intervention-Based Moderators. *Behavior Therapy*, 48(1), 1–18. <https://doi.org/10.1016/j.beth.2016.08.002>
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: a meta-analysis. *Journal of Consulting and Clinical Psychology*, 71(1), 136–149.
<https://doi.org/10.1037/0022-006X.71.1.136>

- Zhou, Q., Chen, S. H., & Main, A. (2012). Commonalities and differences in the research on children's effortful control and executive function: A call for an integrated model of self-regulation. *Child Development Perspectives*, 6(2), 112–121. <https://doi.org/10.1111/j.1750-8606.2011.00176.x>
- Zhou, Q., Hofer, C., Eisenberg, N., Reiser, M., Spinrad, T. L., & Fabes, R. A. (2007). The developmental trajectories of attention focusing, attentional and behavioral persistence, and externalizing problems during school-age years. *Developmental Psychology*, 43(2), 369.

Figure 1. Consort Table

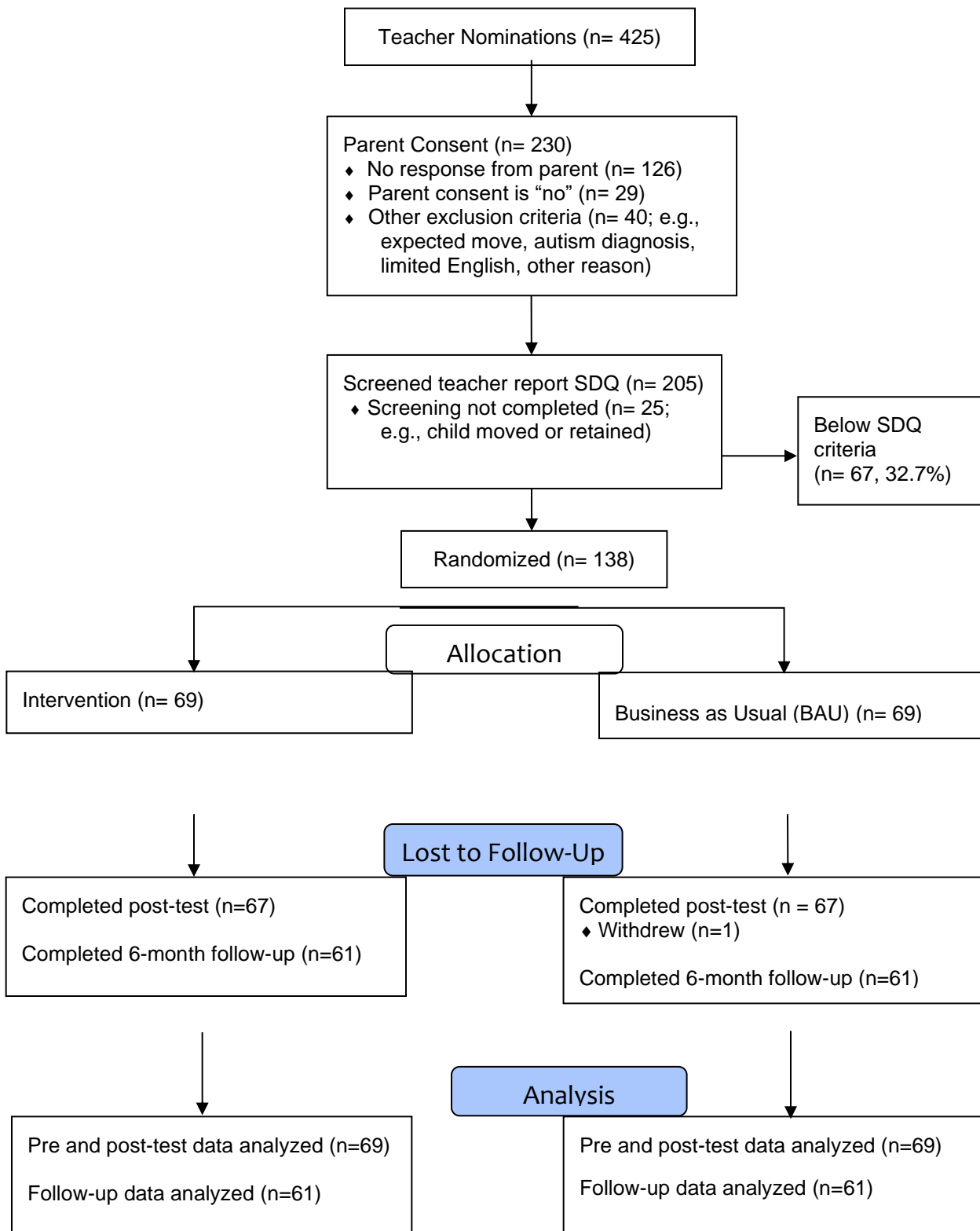
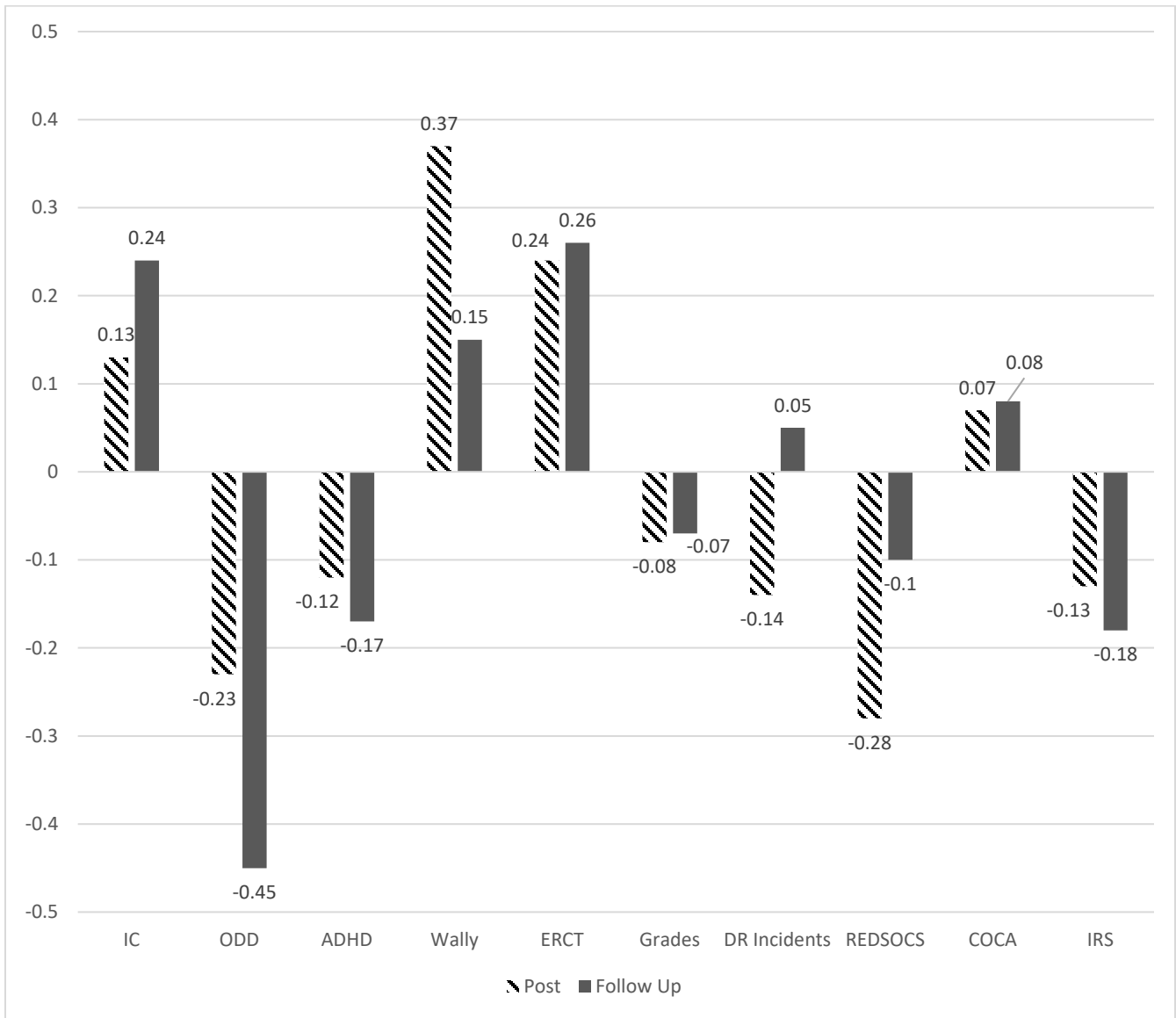


Figure 2

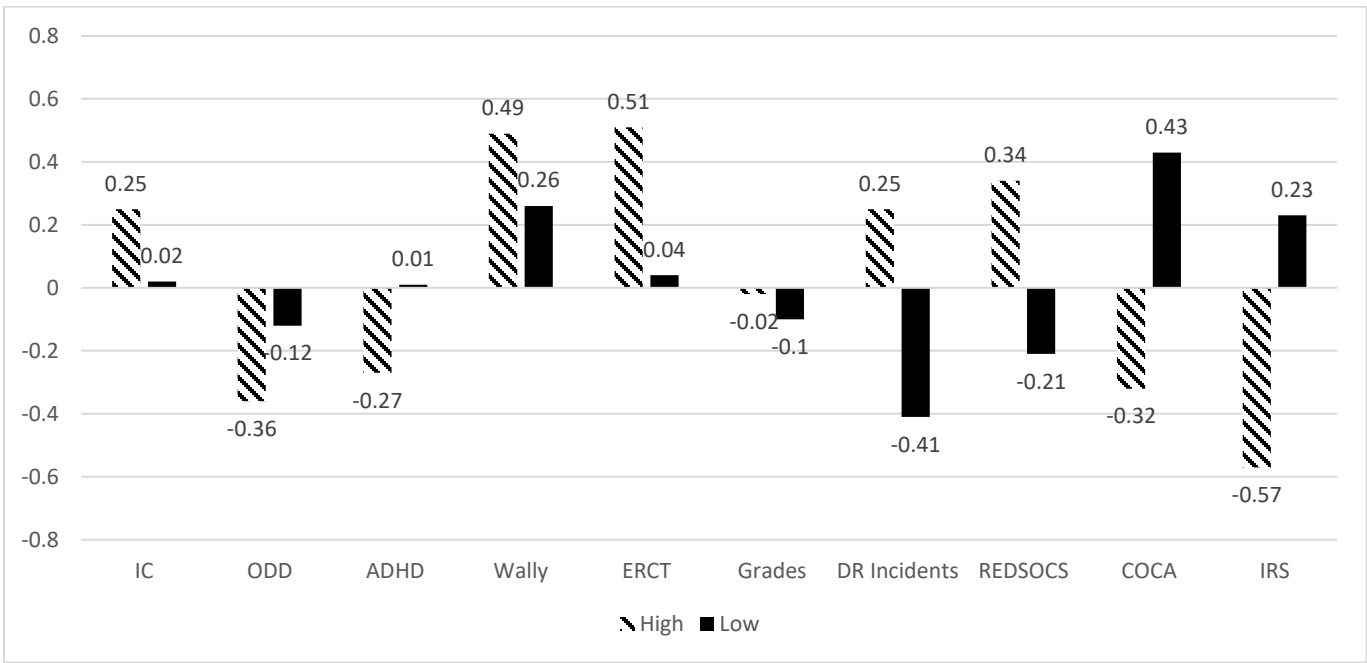
Effect Sizes (d) by Outcome for ITT Analysis at Post-Test and Follow Up



Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised Edition of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale

Figure 3

Effect Sizes (d) by Outcome Comparing High and Low ODD Risk to BAU at Post-Test

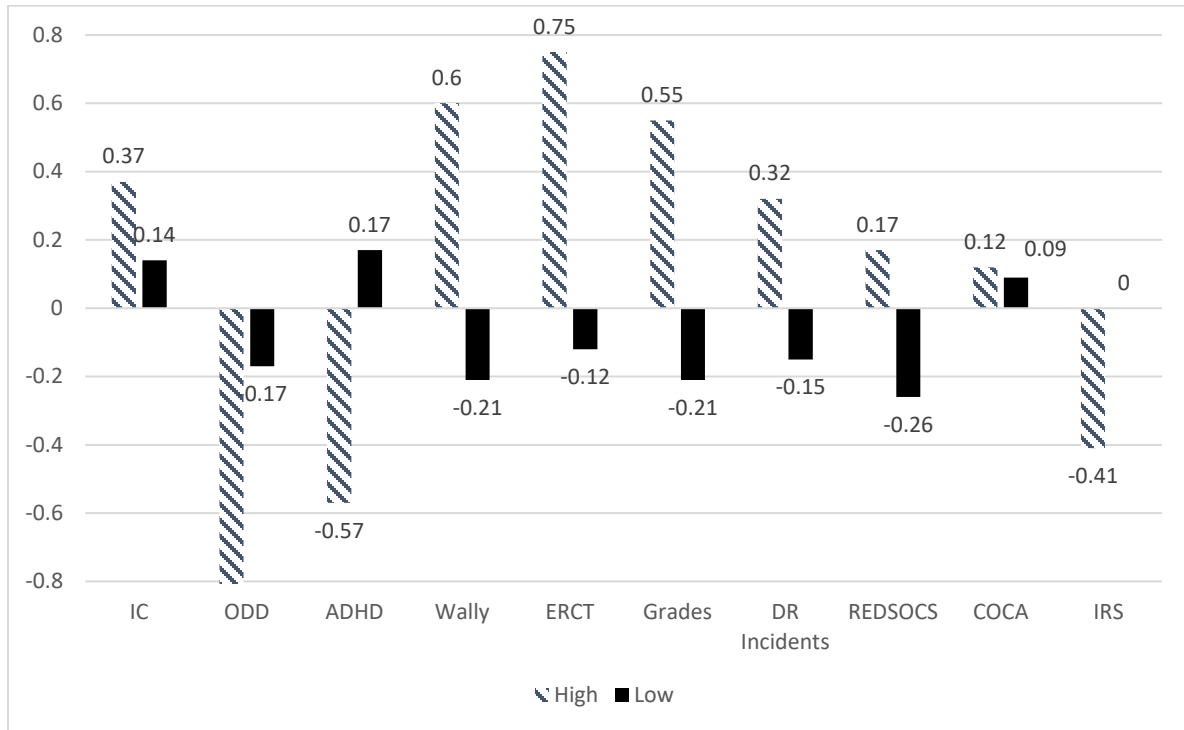


Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised EDITION of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale.

ODD, ADHD, DR, REDSOCS, COCA, and IRS are scored such that lower scores are better.

Figure 4

Effect Sizes (d) by Outcome Comparing High and Low ODD Risk to BAU at Follow-Up

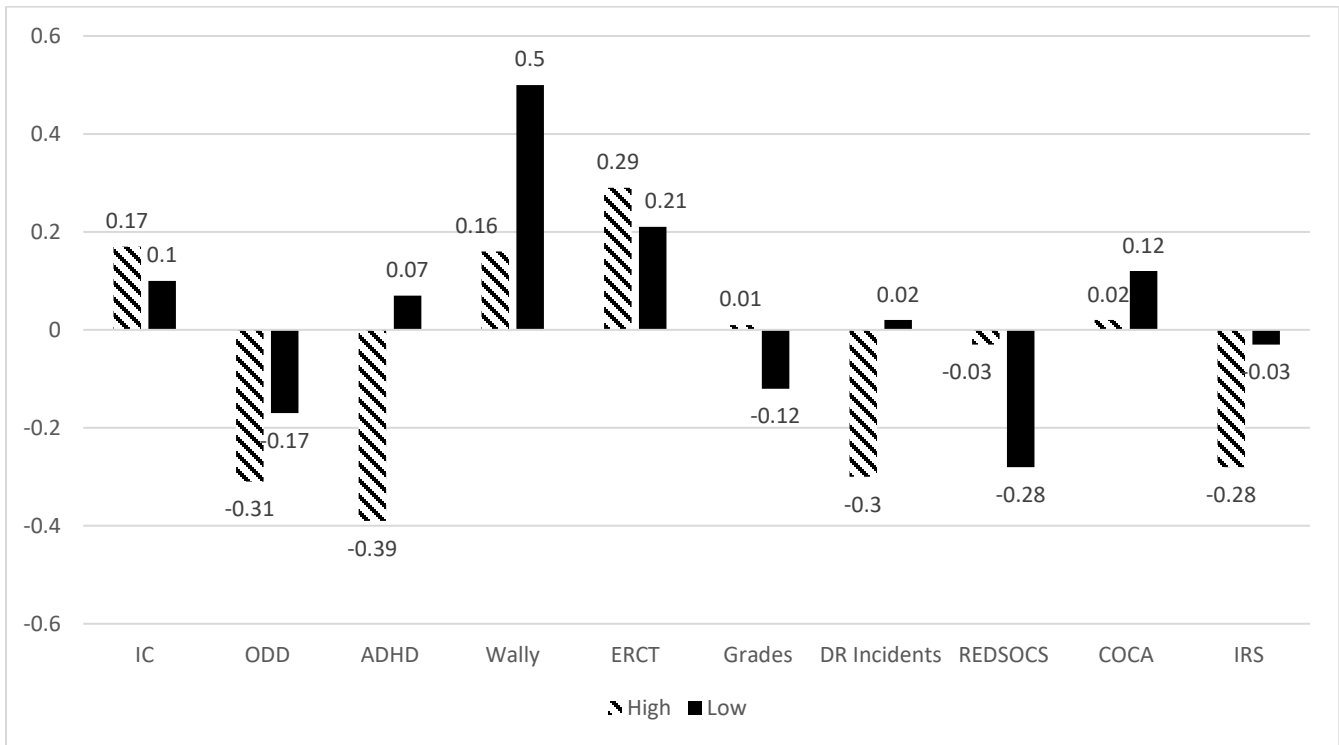


Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised EDITION of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale.

ODD, ADHD, DR, REDSOCS, COCA, and IRS are scored such that lower scores are better.

Figure 5

Effect Sizes (d) by Outcome Comparing High and Low Fidelity to BAU at Post-Test

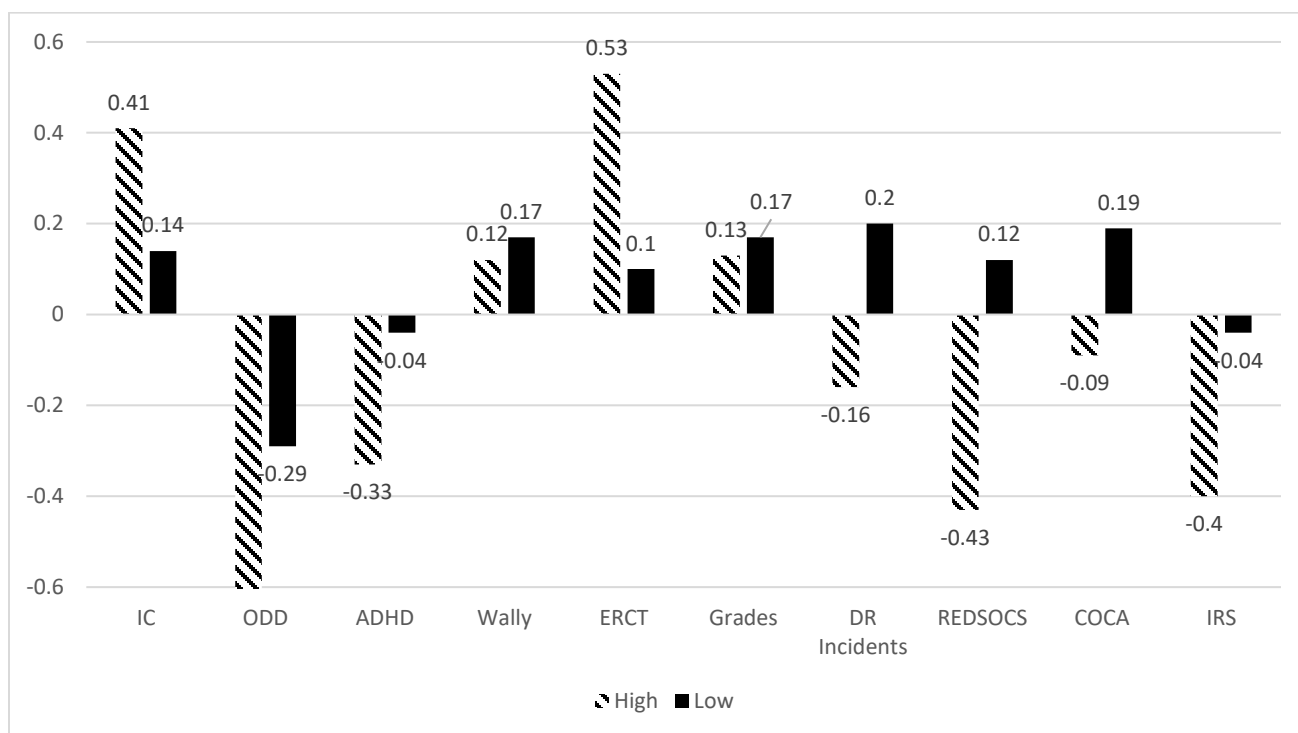


Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised Edition of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale.

ODD, ADHD, DR, REDSOCS, COCA, and IRS are scored such that lower scores are better.

Figure 6

Effect Sizes (d) by Outcome Comparing High and Low Fidelity to BAU at Follow-Up

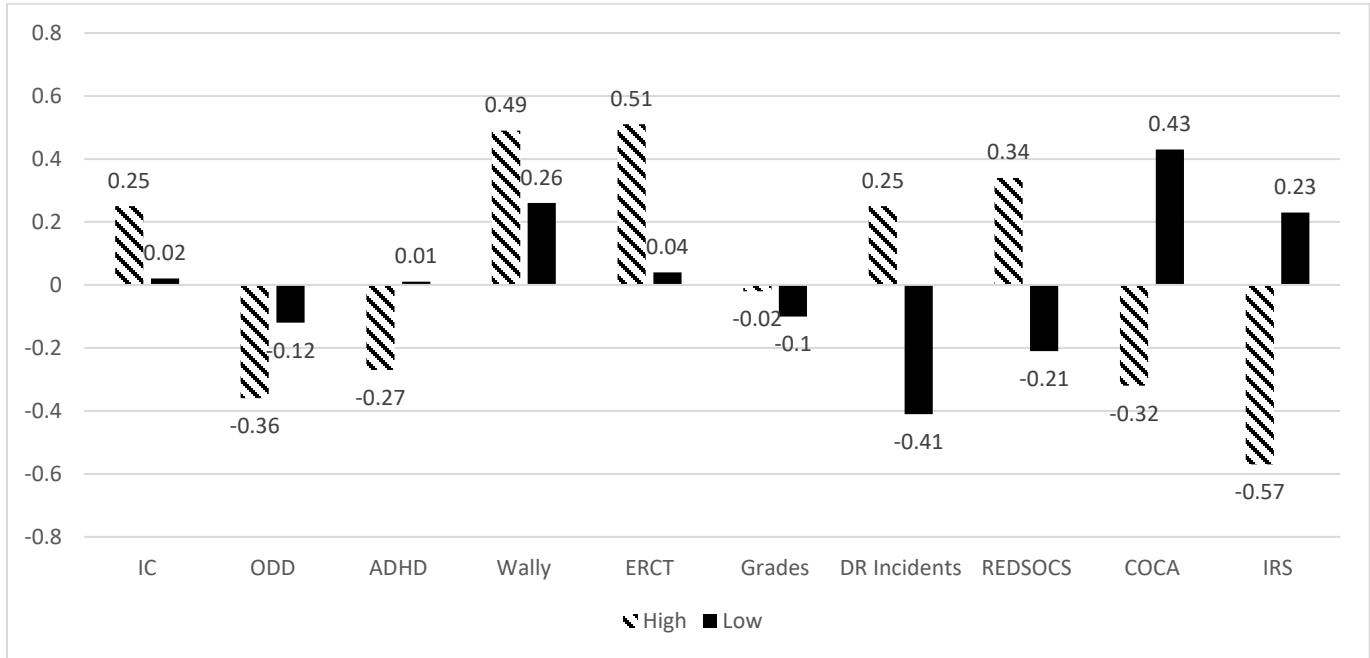


Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised EDITION of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale.

ODD, ADHD, DR, REDSOCS, COCA, and IRS are scored such that lower scores are better.

Figure 3

Effect Sizes (d) by Outcome Comparing High and Low ODD Risk to BAU at Post-Test

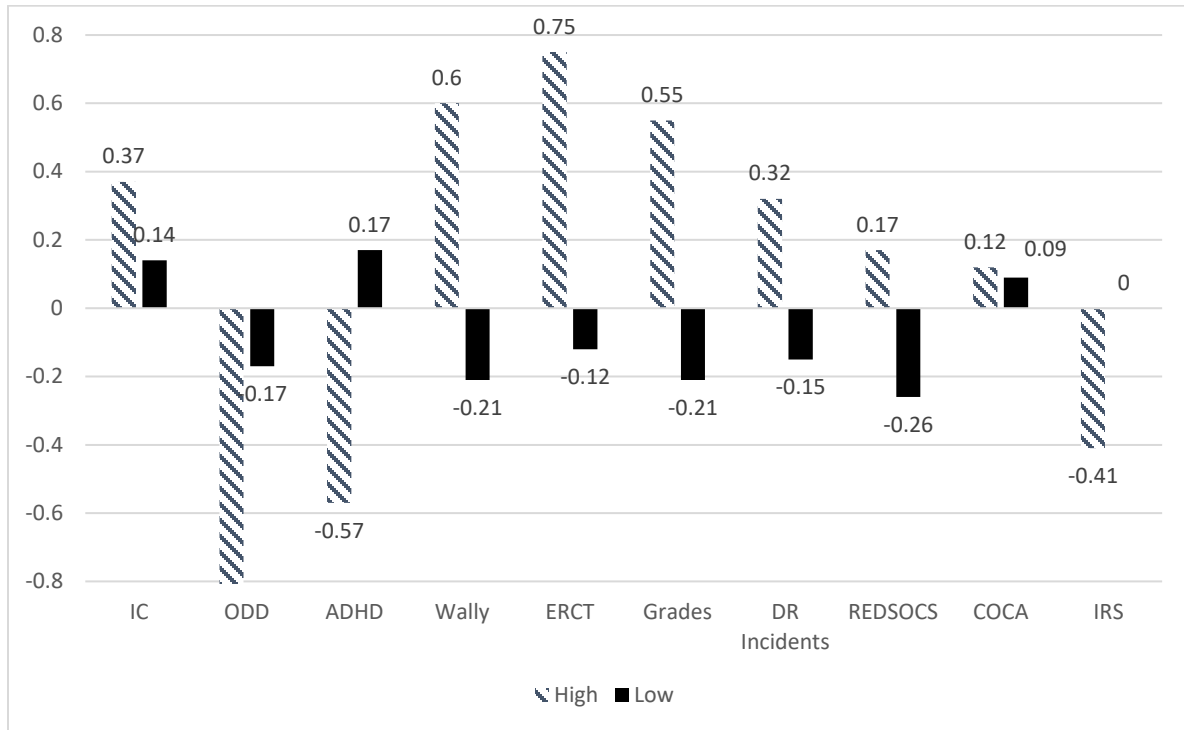


Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised Edition of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale.

ODD, ADHD, DR, REDSOCS, COCA, and IRS are scored such that lower scores are better.

Figure 4

Effect Sizes (d) by Outcome Comparing High and Low ODD Risk to BAU at Follow-Up

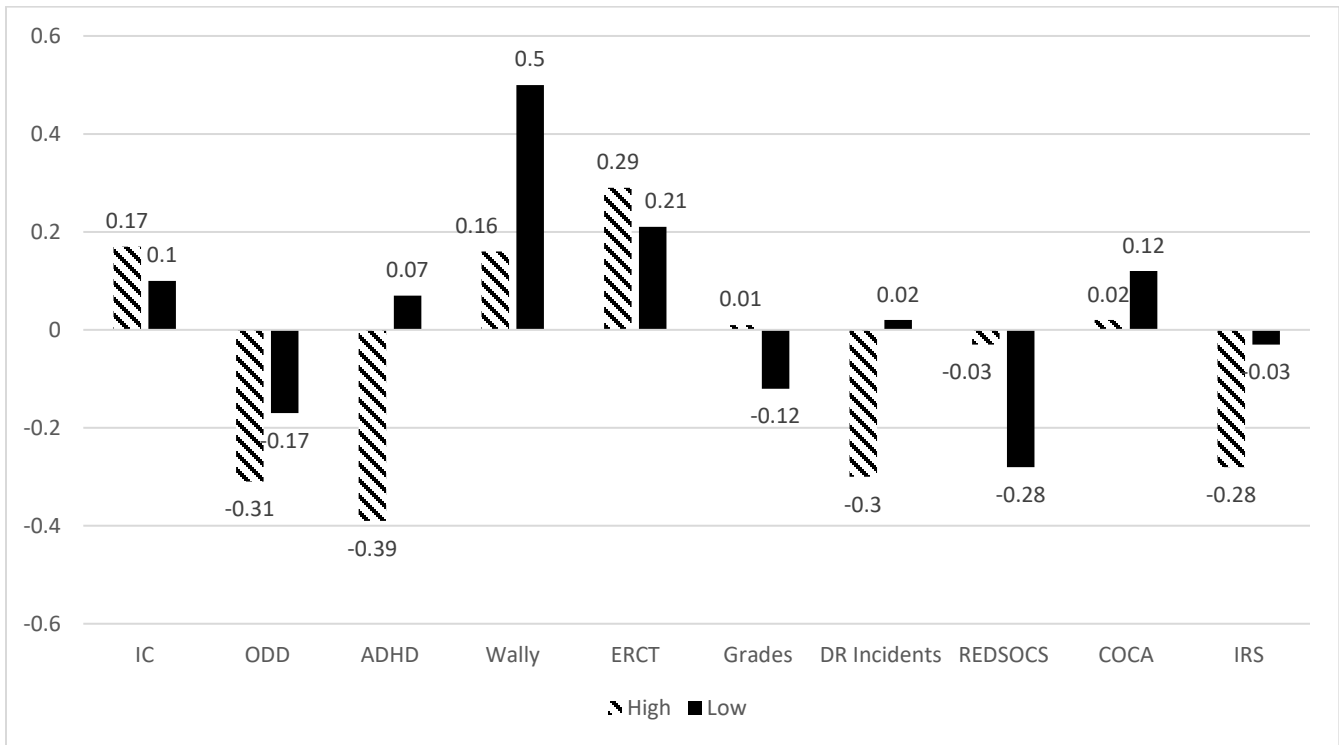


Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised EDITION of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale.

ODD, ADHD, DR, REDSOCS, COCA, and IRS are scored such that lower scores are better.

Figure 5

Effect Sizes (d) by Outcome Comparing High and Low Fidelity to BAU at Post-Test

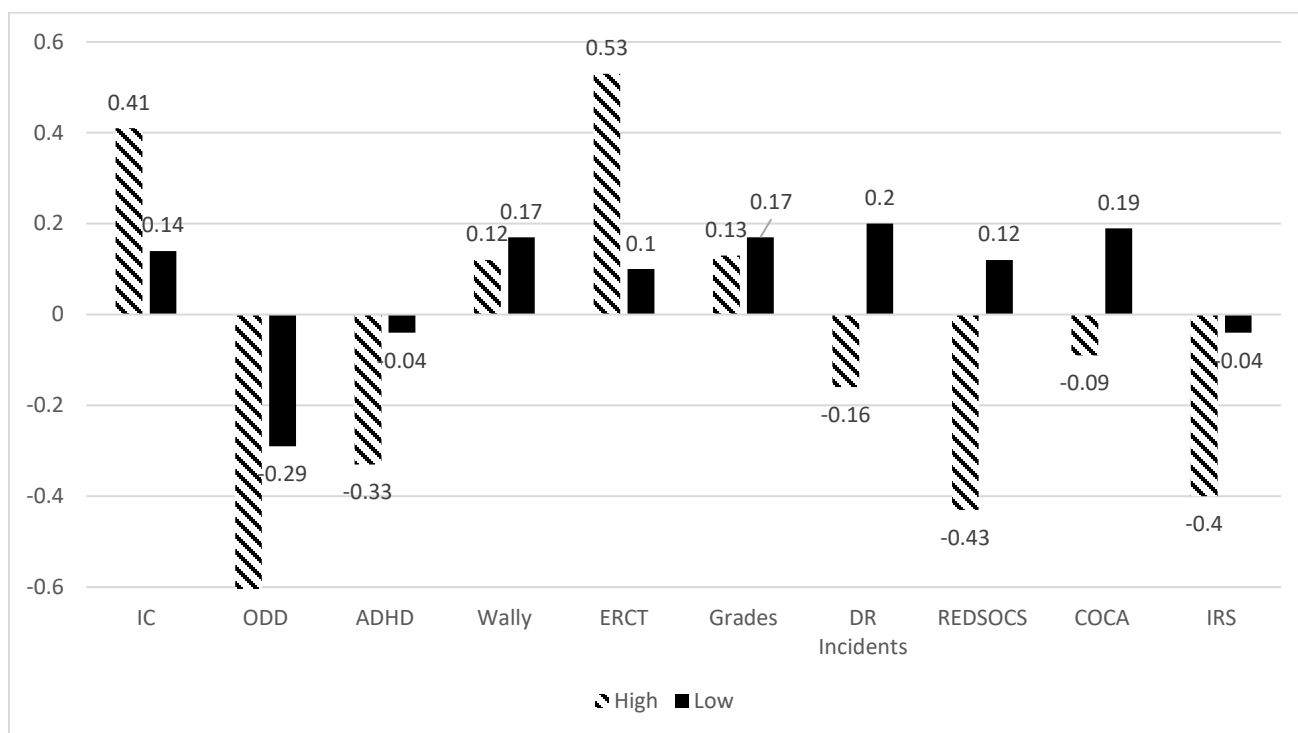


Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised Edition of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale.

ODD, ADHD, DR, REDSOCS, COCA, and IRS are scored such that lower scores are better.

Figure 6

Effect Sizes (d) by Outcome Comparing High and Low Fidelity to BAU at Follow-Up



Note. BAU = Business as Usual; IC = Inhibitory control; ODD = Oppositional Defiant Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Wally = Wally Problem-Solving Test; ERCT = Emotion Regulation; DR = Discipline Referral; REDSOCS = Revised EDITION of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; IRS = Impairment Rating Scale.

ODD, ADHD, DR, REDSOCS, COCA, and IRS are scored such that lower scores are better.

	Study Group											
	BAU (n=69)						Intervention (n=69)					
	Pre		Post		Follow up		Pre		Post		Follow up	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<i><u>Primary Self-Regulation Outcomes</u></i>												
Head Toes Knees Shoulders	37.75	16.06	43.37	14.11	46.95	13.14	40.19	14.12	48.25	8.49	50.77	6.21
Happy Sad Stroop	5.36	3.33	4.00	2.82	3.61	2.63	4.75	3.07	3.66	2.36	3.07	2.39
Puzzle Task - % Persistence	51.70	28.58	53.46	36.05	50.56	35.16	48.35	31.31	54.52	32.73	57.56	34.37
Emotion Regulation Checklist	2.63	0.40	2.73	0.44	2.69	0.43	2.69	0.40	2.82	0.43	2.83	0.41
Wally - % Positive Solutions	0.87	0.11	0.88	0.08	0.88	0.11	0.87	0.13	0.93	0.06	0.90	0.07
<i><u>Social-Behavioral and Academic Outcomes</u></i>												
REDSOCS - % Off-task	31.03	18.73	32.90	18.88	26.70	17.34	32.52	18.78	27.87	16.25	25.85	17.04
REDSOCS - % Inappropriate	23.27	19.67	25.34	20.74	20.06	17.36	25.81	18.21	20.86	16.71	16.50	17.69
COCA – Social Health	0.66	0.83	0.91	0.83	0.76	0.67	0.90	0.79	1.05	1.07	0.77	0.76
COCA – Social Contact	0.60	0.92	0.87	1.14	0.52	0.81	1.12	1.30	1.08	1.32	0.89	1.27
SWAN – Noncompliant	1.21	1.04	1.09	1.08	1.23	1.11	1.14	1.08	0.85	1.18	0.70	1.25
SWAN - Angry Irritable	1.28	1.10	1.24	1.15	1.23	1.21	1.15	1.17	0.90	1.13	0.81	1.30
SWAN - Inattentive	1.43	1.00	1.17	0.96	1.27	0.87	1.49	0.93	1.11	0.98	1.12	1.07
SWAN – Hyperactive/Impulsive	1.28	0.92	1.07	0.95	0.96	0.92	1.38	0.83	1.06	1.01	0.90	1.02
SDQ – Conduct Problems	4.28	2.55	3.91	2.57	4.34	2.77	4.30	2.57	3.83	2.61	3.30	2.29
SDQ – Hyperactivity	7.97	2.26	6.74	2.72	6.95	2.45	8.25	1.97	7.29	2.31	7.02	2.72
IRS	2.50	0.92	2.19	0.96	2.31	1.02	2.56	0.85	2.27	1.07	2.11	1.16
Academic Proficiency	0.67	0.93	0.88	1.05	0.55	0.87	0.77	0.94	0.95	1.16	0.78	0.97
Discipline Referrals	--	--	2.00	3.97	1.26	1.97	--	--	1.79	4.11	1.75	3.83

Table 2. *Unadjusted Means of Outcome Variables (N=138)*

Note. BAU = Business as Usual; Wally = Wally Problem-Solving Test; REDSOCS = Revised EDition of the School Observation Coding System; COCA = Coder Observation of Child Adaptation; SWAN = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior; SDQ = Strengths and Difficulties Questionnaire; IRS = Impairment Rating Scale.

Sample sizes ranged from $N = 64-67$ at post-test; $N = 51-61$ at follow up.