5

FIRST STEP TO SUCCESS

A Preventive, Early Intervention for Young Students with Disruptive Behavior Problems

Hill M. Walker
OREGON RESEARCH INSTITUTE
UNIVERSITY OF OREGON

Andy J. Frey
UNIVERSITY OF LOUISVILLE

Jason W. Small
OREGON RESEARCH INSTITUTE

Ed Feil
OREGON RESEARCH INSTITUTE

John R. Seeley
OREGON RESEARCH INSTITUTE
UNIVERSITY OF OREGON

Jon Lee
NORTHERN ARIZONA UNIVERSITY

Annemieke Golly
OREGON RESEARCH INSTITUTE

Shantel Crosby
UNIVERSITY OF LOUISVILLE

Steve Forness
UNIVERSITY OF CALIFORNIA AT LOS ANGELES

PROOF

Original First Step Program Reviews

The original First Step program was positively reviewed by the Institute for Education Sciences (IES) regarding its intervention efficacy for achieving prevention outcomes.[1] The What Works Clearinghouse (WWC) uses a standardized format to evaluate programs as to their efficacy or effectiveness within a number of domains; the review process can take up to two years. In April 2012, the First Step program was certified by IES reviewers as a promising, evidence-based intervention that produces positive outcomes at a moderate level of magnitude. The First Step program has been recommended as an evidence-based intervention and included in more than a dozen reviews of effective early intervention programs for addressing challenging and problem behavior patterns among at-risk, K-3 students. This includes being listed in the Model Program Guide of the US Department of Justice's Office of Juvenile Justice and Delinquency Prevention.

First Step Program Foundations

First Step is a social ecological, tier 2 intervention designed to address the poor fit that often exists between a student's behavioral characteristics and the routines, demands, and social contingencies of school environments (Romer & Heller, 1983). The theoretical foundation and underpinnings of the First Step program are based on a social learning formulation of behavior. Negative, social

DOI: 10.4324/9781003275312-7

9781032230344c01-end_p1-557.indd 74 15-Jun-22 23:00:32

learning contingencies often characterize family ecologies in which coercive, child behavior patterns are inadvertently strengthened via parenting practices (Patterson, 1982). At-risk children and youth growing up in such environs learn coercive strategies and bring them to the schooling process where they eventually produce rejection, social isolation, and punishment from both teachers and peers (Patterson et al., 1992; Reid et al., 2002).

The primary focus of First Step is (1) to reverse these contingency arrangements with parents and teachers and (2) to teach at-risk students an equally efficient but adaptive set of school success skills. Parents, teachers, and peers are involved in the First Step intervention as participants and supportive collaborators. Their intervention roles are coordinated and supervised by a behavioral coach such as a school psychologist or school counselor, school social worker, PBIS implementer, or behavioral specialist.

First Step Goals

First Step seeks to (1) facilitate efforts by at-risk students in achieving the best start possible to their school careers, (2) effectively engage parents in collaborating with schools in developing their child's current and long-term school success, and (3) enhance the bonding and engagement of students with the schooling process (Walker et al., 1997). First Step Next is focused on a series of skills that enhance academic engagement and contribute to satisfactory teacher and peer-related adjustments, which are two of the most critically important relationships that all students must negotiate in their school careers (Walker et al., 2004).

First Step Structure and Implementation Procedures

First Step Next relies upon group dependent contingencies at school and individual contingencies at home to motivate the target student's participation. If the student meets a daily reward criterion at school, a brief free time activity is earned which is shared with classmates thus potentially encouraging their involvement in his or her success. The student also earns a prearranged reward at home for achieving the school reward. A "good day card" is used to communicate this information to parents who are urged to praise the child's school performance.

First Step operates for 30 continuous program days with a daily reward criterion which must be met in order to advance to the next program day. If the reward criterion for a particular day is not met, then the student is recycled to a prior successful program day. Thus, as a rule, approximately 2 to 3 months are required to complete the First Step program which begins with a 20-minute period where the student can earn points on a coach or teacher-administered card for following classroom rules and meeting the teacher's expectations. By the end of the intervention, it has been gradually extended to the full school day or to those periods of the day where the student struggles with teacher expectations.

A visual cueing system, in the form of a green-red point card, is used to record teacher praises and earned points. When the student is following classroom rules, the green side of the card is shown to him or her by either the coach or teacher and then turned to red when s/he is not. A return to appropriate behavior prompts the card to be turned back from red to green. The First Step program and how it operates is explained privately to the student, the cooperating teacher, and the child's parents. Consent to participate is secured from all these parties before the First Step program can begin.

The green-red point card can be used to easily extend the program to recess periods and to class periods other than homeroom. There are three phases to First Step Next: the coach phase; the teacher phase; and a post-intervention, maintenance phase. The coach phase lasts for the first ten program days and is coordinated and actively managed by the behavioral coach. During the remainder of implementation, the teacher has primary responsibility for the program and is supervised by the coach. At program day ten, the homeBase program component is introduced to the parents who are taught and supervised in how to support and encourage school success skills at home. The last ten days of implementation are focused on maintaining achieved program gains and includes four options ranging from

9781032230344c01-end_p1-557.indd 75 15-Jun-22 23:00:34

increased teacher praise and support to a brief, temporary reinstatement of the FS program depending on a student's needs. Longer term, occasional booster shots, based on the work of Paine et al. (1982), have been successfully implemented as a means of helping preserve First Step Next behavioral gains.

The remainder of this chapter provides information on the following topics: (a) Development of the First Step program, (b) Revision of the original First Step program and Creation of First Step Next, (c) Evaluation research on the First Step program, (d) Synthesis of First Step Evidence from Randomized Controlled Trials, (e) Research on the Sustainability of First Step Program Adoption and Usage, (f) Cost analysis of First Step implementation, (g) Implications of First Step findings for managing the classroom context, and (h) Concluding remarks.

Development of the Original First Step Program

The original First Step program was developed through a four-year, model development grant from the US Office of Special Education Programs to the first author that ran from 1992 to 1996. This development grant involved four sets of collaborative partners representing the University of Oregon (U of O), the Oregon Social Learning Center (OSLC), Eugene School District 4J, and the Oregon Research Institute (ORI). Original First Step was designed as a tier 2, early intervention for achieving secondary prevention goals and outcomes and was jointly based upon a foundation of prior research on families and antisocial children conducted by the Oregon Social Learning Center (OSLC), under the direction of Gerald Patterson and his associates (Patterson, 1982; Patterson et al., 1992), and two decades of school-based screening and intervention research by Walker and his colleagues focused on students with challenging behavior (Walker et al., 1995).

Revision of Original First Step and Creation of First Step Next

Beginning in 2014, a year-long process was begun to update original First Step based on our implementation experience and feedback from end user consumers during the first decade of the program's existence. This process is described in Walker et al. (2018) and resulted in the development of First Step Next (Walker et al., 2015). First Step Next combines the original (K-3) First Step program with the First Step preschool version for 3–5 year-olds (see Feil et al., 2016) and now provides a seamless intervention for use from PreK and Kindergarten through grade two.

The 2014–15 First Step Next revision was guided by three goals: (a) to standardize the First Step Next program procedures and components and forge them into a single unified module, (b) to make the program more user friendly for implementers, especially cooperating parents, and (c) to increase the intervention's efficacy by adding new components and updating existing ones based on consumer feedback (see Walker et al., 2018).

The following developments spurred the revision. Over several decades of existence, end users had created a number of differing First Step innovations and adaptations that resulted in competing First Step Next versions which were circulating. Further, coaches indicated they needed more information about how to individualize the program more effectively. The coach phase also needed to be expanded to allow younger students and those with more intractable problems additional time to acquire new skills. First Step coaches found that for some parents, requiring them to directly teach school success skills at home proved to be a burden and resulted in a low-quality implementation effort leading to elimination of this requirement in the revised intervention. This and other valuable consumer and coach feedback was integrated into the revision process. The resulting First Step Next program is considered more streamlined, less complex, easier to implement with integrity, and more acceptable to consumers than the original.

The revised First Step Next program also has a more balanced academic and social-emotional focus and includes the addition of seven new Super Student Skills lessons, more robust maintenance

9781032230344c01-end_p1-557.indd 76 15-Jun-22 23:00:34

and trouble-shooting options, and additional supplemental materials including parent and teacher workbooks, coloring books, behavioral skills charts, new and revised implementation forms, and new program task demonstration videos. The FSN super student skills are: *follows directions*, *be cool*, *be a team player*, *fix-ups are OK*, *ask for help in the right way*, *do your best work*, *and play well with others*. In addition, the homeBase component of First Step Next, as noted above, relieves parents of the responsibility for directly teaching school success skills at home as in original First Step. First Step Next is now considered a parent-supported rather than a parent-school intervention due to the changed implementation role of parents.

Trial Testing of First Step Intervention Components

A series of informal, quasi-experimental studies was used to trial test components of the First Step intervention during its initial developmental phase. The First Step screening procedures were derived from research conducted by Walker and colleagues on the use of teacher rankings, ratings, and behavioral observations in the universal screening and identification of students having challenging forms of behavior (Walker et al., 1988, 1990). The school intervention procedures were based upon an adaptation of the *Contingencies for Learning Academic and Social Skills (CLASS) Program* for acting out children previously developed and researched by Hops, Walker, and colleagues (Hops & Walker, 1988). Two investigators (Reid and Kavanagh) from the Oregon Social Learning Center designed an adaptation of their parent training-intervention model for antisocial children and youth that enabled parents to directly teach their child school success skills at home. These three modules (screening, school intervention, parent support and involvement) were then implemented, in combination, for a selected number of cases representing the K-3 grade range and examined as to feasibility, the occurrence of logistical problems, and identification of critical implementation-delivery issues. Finally, literature searches and reviews of published programs in these three domains were conducted to identify any needed enhancements that might improve the program's efficacy, consumer acceptability to end user professionals, and ease of delivery prior to publishing a manual and user's guide to enable implementation, replication, and adoption of First Step.

Evaluation Research on Original First Step and First Step Next

The empirical knowledge base on First Step and its revision has been largely developed over the past two decades by the Oregon School Study Group consisting of researchers from the Oregon Research Institute, the University of Louisville, the University of Oregon, Northern Arizona University, and UCLA. Our evaluation research was informed by the hierarchy evidence standard as opposed to the threshold standard of scientific evidence (see Drake et al., 2004) which allows for differing designs connected to investigations of a variety of research questions. A mix of single case and group designs was used to successfully investigate important questions in constructing and validating the First Step intervention (Walker et al., 2014b). Single case designs provided a robust information yield in relation to their required implementation effort—particularly as deployed in the early stages of our research and development process. We found their sensitivity to both behavioral process variables (during the intervention) and outcome variables (after the intervention) were instrumental to our First Step development efforts (Carter & Horner, 2007, 2009; Horner et al., 2005). Group experimental designs were used to address questions pertaining to First Step internal and external validity.

First Step Efficacy and Effectiveness Studies

There have been five studies of FS efficacy and three studies on the effectiveness of the original FS program version. Overall, our efficacy studies were carefully controlled by the FSN developers and

9781032230344c01-end_p1-557.indd 77 15-Jun-22 23:00:34

produced robust effects. Our effectiveness studies produced less robust effects as expected (see Flay et al., 2005 for a further discussion of this distinction). Two of the efficacy studies and one of the effectiveness studies are briefly reviewed following.

Nelson et al. (2009) included original First Step as a selected intervention in a complex study investigating the impact of a 3-tiered behavioral intervention conducted over a 4-year period that focused on 407 students in grades K-3 drawn from 7 elementary schools. Students from one of four longitudinal cohorts participated in this study. There were 153 universal intervention students, 173 selected First Step students who received the tier 2 intervention, and 81 students who received the indicated MultiSystemic Therapy intervention. The demographic breakdown of the study sample was as follows: 130 girls and 277 boys; 35% students of color; and 61% who qualified for free and reduced lunch.

Thus, the three intervention levels of universal, selected, and indicated were represented respectively by the Behavior and Academic Support and Enhancement (BASE) intervention, First Step, and MultiSystemic Therapy (MST); dependent measures included a range of social skills, problem behavior, and academic performance measures. Employing a series of 2-level linear growth analyses, the authors concluded that the 3-tiered behavior model used in the study achieved hypothesized outcomes in social and behavioral domains but not in academic performance areas. The authors reported that First Step students showed significant gains on the study measures in these two domains and that gains were maintained at one- and two-year follow-up assessments.

Walker and his colleagues reported a large-scale, randomized controlled trial of FS that was funded by a 4-year grant from the Institute of Education Sciences (IES). The study was conducted in the Albuquerque Public Schools (APS) with a diverse sample of students in grades 1–3 (over 70% of the participants were students of color). Results of this study were reported in Walker et al. (2009) which describes an efficacy trial of the FS program. Year 1 of this study involved planning, student participant screening and identification, and recruitment of schools, teachers, and coaches, along with staff training. Years 2 and 3 were intervention years and involved two cohorts of 99 and 101 participants respectively who were randomly assigned to intervention and control conditions. Year 4 of this study was focused on follow-up and maintenance activities, training additional APS staff, data analysis, report writing and dissemination efforts. The APS sample of 200 cases was 73% male with 57% Hispanic and 24% Caucasian students. The remainder of the sample was distributed across four other underrepresented groups that included 7% African-American students. Eighty eight percent of the sample came from English-speaking households and more than 70% were eligible for free or reduced lunch.

School outcome measures examined social skills, classroom behavior (adaptive and maladaptive), academic performance, and oral reading fluency. These measures were in the form of teacher ratings and also direct performance assessments as well as behavioral observations of participating students. Teachers completed the SSRS rating scale measure of social skills (i.e. Social Skills Rating System, Gresham & Elliott (1990). Academic engaged time (AET) was recorded with a stopwatch in all experimental and control classrooms using the *Systematic Screening for Behavior Disorders* observation system (Walker et al., 2014a).

Pre and post assessments showed relatively strong effects for social, adaptive, and maladaptive behavior and AET domains but no effects on direct academic performance measures as measured by standardized achievement tests. However, the academic performance subscale of the Social Skills Rating System (Gresham & Elliott, 1990) was sensitive to the FS intervention with an effect size of .66. For maladaptive behavior, effect sizes ranged from .62 to .73; for adaptive behavior and social skills, effect sizes ranged from .54 to .87; and for academic performance, they ranged from .13 to .66.

This study was the first large-scale, randomized test of the FS program within a diverse, large urban school district. The relatively robust intervention effects achieved on our behavioral measures demonstrate the short-term benefits of the program. However, follow-up and maintenance effects were disappointing. One year after the end of intervention, most of the gains for FS students in the two

9781032230344c01-end_p1-557.indd 78 15-Jun-22 23:00:34

cohorts had decayed with no differences detectable between experimental and control participants. This result prompted the design and implementation of a classroom-wide (i.e. universal) FS maintenance plan designed to reestablish and stabilize the initial program gains. However, its outcomes were judged to have been only modestly effective in meeting this goal.

First Step Effectiveness Study

The FS program has been the focus of one large-scale, national effectiveness study funded by the Institute of Education Sciences and reported by Sumi et al. (2013). This study involved a total of 48 schools, randomly assigned to intervention or control conditions with 142 intervention students and 144 control or comparison students. The implementation sites were in Illinois, West Virginia, Florida, California, and Oregon. In this study, the FS developers and associated investigators were much less involved in the FS implementation protocol than in previously reported First Step Next efficacy studies (see Walker et al., 1998, 2009). Their role was primarily limited to providing initial staff training in FS implementation procedures and in coordinating and the collection of study measures. Also, this study involved five FS implementation sites as opposed to one participating site as in the prior FS efficacy studies—thus substantially increasing its complexity and logistics. The dependent measures used in the study were identical to those reported in Walker et al. (2009).

Findings of the Sumi et al. (2013) effectiveness study generally replicated outcomes of the large-scale efficacy study reported by Walker et al. (2009) but effect sizes were of lower magnitude than in the Walker et al. study. Outcomes favored FS intervention students on 8 of 10 dependent measures. As in prior FS research, direct academic performance of oral reading fluency on the Woodcock Johnson Diagnostic Reading III Inventory (Schrank et al., 2004) were not positively impacted by exposure to the intervention but student behavior on all other teacher, parent and observational measures were. Effect sizes for the Sumi et al. study ranged from .11 to .67. In the efficacy study reported by Walker et al. (2009), they ranged from .57 to .87. Such reduced effects are expected when comparing effectiveness and efficacy studies (see Flay et al., 2005; Weisz & Jensen, 2001) where much tighter control is maintained over implementation procedures, dosage levels, and the trouble shooting of problems that arise during the intervention.

On the issue of implementation, large amounts of definitional, conceptual, and empirical work remains to be conducted on the relationship between the quality of implementation and intervention outcomes beginning with the quality of fidelity measures used to predict such outcomes. Sheridan et al. (2009), in a study of fidelity measurement in consultation, noted that the relationship between implementation fidelity and intervention outcomes may be more a function of adherence to meaningful intervention protocols over time (i.e., dosage). We have evidence from our First Step research that although good fidelity is associated with good outcomes, a less than perfect implementation of First Step Next does not insure poor outcomes. For example, in the Walker et al. (2009) efficacy study, the canonical correlation between our measure of fidelity and intervention outcomes was .50. In the Sumi et al. (2013) effectiveness study, there was a similar finding regarding this relationship with First Step.

In our view, these outcomes speak to the structure and power of applied interventions as key factors in accounting for intervention outcomes that produce meaningful behavior change. In addition to fidelity, a huge panoply of variables likely account for intervention outcomes such as the design and characteristics of the intervention, the severity of target student behavior problems, staff training and motivation, classroom ecology, supervision and consultation with implementers, parent and administrative support, and so on. We look forward to research that clarifies and isolates the role of these variables in future studies of intervention outcomes.

Although not expected, we argue that the implementation fidelity and intervention outcomes noted above can be considered a potential strength of the First Step Next program. That is, given the not infrequent, semi-chaotic ecologies and behavioral challenges that exist in many of today's

9781032230344c01-end_p1-557.indd 79 15-Jun-22 23:00:34

classrooms and schools, if a targeted intervention such as First Step Next can be implemented, albeit with a lower level of fidelity and still produce acceptable student outcomes, it can be viewed as a positive result. Less than adequate implementation fidelity may actually be a more likely event than satisfactory fidelity in many "real world" applications of behavioral interventions. Because there is likely to be a less than ideal implementation of First Step Next in many instances, this outcome becomes a potentially noteworthy finding. However, regardless of the magnitude of achieved intervention effects, our research consistently shows that decay of FS behavioral outcomes begins occurring soon after termination of the intervention unless systematic maintenance procedures are in place.

Overall, the studies conducted on the original FS program from its development until the 2014–15 revision show it producing relatively robust outcome effects. However, as noted, the durability of achieved behavioral gains in these studies has been disappointing and is a challenge to be addressed going forward (Walker et al., 2014b).

First Step Next Research Following the 2014–15 Program Revision. Substantive changes were made to the original FS program in the 2014–15 revision that resulted in First Step Next. Following this revision, a question arose as to whether these changes would negatively influence the FSN program's outcomes and efficacy as compared to original First Step Next. Two recent IES-funded studies, conducted respectively with preschool and elementary student samples, provide evidence that they do not. These studies are briefly described below.

Feil et al. (2021) recently reported a validation study of FSN conducted in four preschool settings. One hundred and sixty students at risk for school failure, and their teachers, were randomized to the First Step Next intervention and control conditions. Results for the three prosocial outcome measures had Hedges' *g* effect sizes that ranged from 0.34 to 0.91 and favored the First Step Next intervention condition. For the problem behavior domain, Hedges' *g* effect sizes ranged from 0.33 to 0.63 again favoring the intervention condition.

Frey et al. (in press) conducted a comparative efficacy study of First Step Next where they identified a sample of 379 student-teacher-parent triads in which identified students were at elevated risk for disruptive behavior and school failure. Triads were randomly assigned to a school only condition, a home only condition, and combined condition along with a business as usual condition. Across prosocial, problem behavior and academic domains, results showed substantial support for the school only and combined conditions compared to business as usual. Modest support was obtained for the home only condition. Effect sizes in the school only and combined conditions averaged moderate to large magnitudes across domains and were small for the home only condition ranging from 0.09 to 0.26. Using the WWC Improvement Index, the mean percentile gains for the school only, combined, and home only conditions were respectively 18.9, 18.5, and 7.5.

First Step Next Applications to Attention Deficit Hyperactive Disorder and Other Psychiatric Disorders. An important condition affecting school performance and adjustment is Attention Deficit Hyperactive Disorder (ADHD). Attention Deficit Hyperactive Disorder has long been established as a co-occurring condition with conduct disorder (August et al., 1996). To date, there have been no formal, large-scale trials of First Step focused exclusively on students with disruptive behavior and ADHD. However, using data from the larger randomized efficacy trial (N=200) reported by Walker et al. (2009), Seeley et al. (2009) conducted a series of analyses for 42 participants (23 intervention and 19 control) who met eligibility criteria for co-occurring ADHD in order to determine FS outcomes for this subpopulation.

Students were identified for this ADHD subsample using Conner's DSM symptom cutoff for teacher-reported ADHD symptomatology. As part of the baseline assessment measures for the larger Albuquerque Public Schools (APS) study of 200 cases, intervention and control teachers completed an 18-item version of Conner's ADHD/DSM-IV scale (CADS-T; Conners, 1997). Thus, 21% (n=42) of our larger APS sample met ADHD diagnostic criteria. As compared with the larger sample, the ADHD subsample had significantly fewer females and significantly more Hispanic students. When this study was conducted, 1% of the students with ADHD in the control condition and 3% of students with ADHD in the intervention

9781032230344c01-end_p1-557.indd 80 15-Jun-22 23:00:34

condition were receiving medication. Dependent measures were grouped into five areas or domains as follows: (a) school-based measures, (b) ADHD and disruptive behaviors and symptoms, (c) social functioning, (d) academic functioning, and (e) home-based outcome measures.

When the ADHD subsample was compared to the larger, remaining sample of students with disruptive behavior but without ADHD (n=158), the ADHD subsample had significantly more problematic profiles on measures of hyperactivity, attention, oppositional defiant disorder and maladaptive behavior. They also had less favorable profiles on measures of adaptive behavior and social skills. There were no differences between them in the areas of academic functioning and home-based measures completed by parents (i.e. SSRS problem behavior and social skills).

An overall multivariate model was tested for the four post-test behavioral symptom measures. Intervention students with ADHD had large overall gains compared to control students with ADHD on each of these measures with effect sizes ranging from d=.74 to 1.32. In the social functioning domain, effect sizes for these same students were (d=.80) for adaptive behavior and (d=1.01) for social skills. In the academic domain, there was a medium level difference between the groups on two academically related measures. These were the AET observational measure (d=.76) and teacher ratings on the SSRS academic competence subscale (d=.58). Parent ratings on the SSRS problem behavior subscale were in the predicted direction (d=.60) but there were no significant differences on the SSRS social skills scale (d=.31).

These outcomes were comparable in magnitude to school-home interventions that are directly focused on ADHD students without disruptive behavior (see Pfiffner et al., 2007). They also compare favorably to interventions that are focused on externalizing problem behavior among at-risk students (Conduct Problems Prevention Research Group, 2011). The ADHD sample showed a robust response to the original First Step intervention suggesting that it may have a positive impact on students with other externalizing problems and disorders.

A replication of the Seeley (2009) ADHD study was subsequently conducted by Feil et al. (2016) with preschoolers co-occurring ADHD. This study involved a subsample from a relatively large RCT study of First Step for 126 preschoolers (Feil et al., 2014). Using a similar diagnostic measure for co-occurring ADHD, as in the study just described for elementary students, we retrospectively identified 45 preschoolers with ADHD symptoms, 26 from the intervention condition and 19 from the control or usual care condition. Preschoolers with ADHD who received the First Step Next intervention had statistically significant improvement compared to control preschoolers with ADHD on all teacher and parent-reported outcomes. For teacher-reported improvements in symptoms, Hedges' g effect sizes ranged from 0.80 to 1.4. As in the elementary-aged study above, parent-reported outcomes were smaller with effect sizes of 0.77 for symptoms and 0.65 for social functioning.

As there is some evidence that very young children with Autism Spectrum Disorder (ASD) may also be at risk for conduct or oppositional disorders (Kim et al., 2012) and that anxiety disorders have a high co-occurrence with disruptive behavior disorders (Franz et al., 2014), we used the same preschool sample from Feil et al. (2014) to also study subsamples having risk for ASD and anxiety disorders. In the ASD study, we used cutoffs on the autism diagnostic measure from the Early Childhood Inventory (ECI) (Gadow & Sprafkin, 2000) to retrospectively identify a subsample of 17 preschoolers with ASD who received the FS intervention and 17 with co-occurring ASD and who were in the usual care condition (Frey et al., 2015). Hedges' g effect sizes ranged from 0.65 to 1.12 in favor of the FS subsample in teacher-reported reductions in symptoms and from 0.29 to 1.18 in teacher-reported improvements in social functioning. Effect sizes in parent-reported improvements in social functioning ranged from 0.48 to 0.57 in favor of the FS subsample. All but one of the 11 outcome measures was statistically significant.

We then used the anxiety disorders diagnostic measure from the ECI to retrospectively identify a subsample of 38 preschoolers with anxiety disorders who were respectively assigned to the First Step intervention or to a usual care condition (19 per condition). Seven of 9 outcome measures were

9781032230344c01-end_p1-557.indd 81 15-Jun-22 23:00:34

statistically significant in favor of the subsample assigned to First Step (Seeley et al., 2018). Hedges' *g* effect sizes ranged from 0.59 to 0.79 for disruptive behavior and 0.60 to 0.85 for social functioning improvements as reported by teachers; parent reported effect sizes on the same measures were 0.66 and 0.6 respectively. For both teacher and parent-reported outcomes for internalizing behavior, however, effect sizes were smaller at 0.42 and 0.23 suggesting that anxiety symptoms continued to be problematic. This was an expected finding as original FS was not designed to address internalizing anxiety problems; the achieved effects in this domain are likely a result due to disruptive students who also had co-occurring anxiety disorders.

Since the above studies were quite limited due to their small sample sizes, we are in the process of retrospectively identifying larger subsamples of children with co-occurring psychiatric disorders to also include depression. A recently completed RCT of nearly 400 elementary teacher-student-parent triads by Frey et al. (in press) will allow us to partially address this goal.

Synthesis of Original First Step and First Step Next Evidence from Randomized Controlled Trials

School professionals are well aware of the positive impact of early intervention and prevention efforts for successfully reducing disruptive behaviors and the likelihood of poor developmental outcomes (Hawkins et al., 1999). Following its development and initial testing in a small-scale RCT (Walker et al., 1998), the First Step intervention, as we have noted herein, was validated in a large-scale study conducted in the diverse Albuquerque School District (Walker et al., 2009) and via a national effectiveness study conducted by Sumi et al. (2013).

The Oregon School Research Group recently reviewed and synthesized five randomized controlled trials conducted between 2009–21 that included efficacy and effectiveness studies involving both FS and its revised FSN version along with several subsample analyses demonstrating impact across a range of co-occurring disorders and diverse target populations. Collectively, these studies show that the First Step intervention has resulted in small to large effect sizes and statistically significant improvements, compared to students randomized to control conditions, on multiple indicators of prosocial and problem behavior. A manuscript describing this best evidence synthesis is presently in press to Remedial and Special Education. This article provides professional consumers with accessible synthesized, First Step information that may facilitate program evaluation and adoption decisions.

Research on the Sustainability of First Step Next Program Usage

Flay et al. (2005) defined a set of evaluation standards for applied interventions which specify that (1) intervention programs should demonstrate replicable efficacy and effectiveness, (2) in addition to being judged as effective, they should meet criteria ensuring that agencies will adopt them, and (3) adopting agencies and organizations have the capacity and resources to make effective use of them. Loman et al. (2010) reported a comprehensive, long-term study of the sustainability of First Step that investigated these questions. These investigators systematically examined variables associated with the First Step program's usage and implementation over a ten-year period following First Step training and adoption by 29 Oregon schools located in 13 Oregon school districts. A survey, called the First Step Evaluation Tool was developed by the study's authors to investigate these questions. Each participating school was represented by one school-level respondent (i.e. personnel formally trained in First Step) and was included in the sample based on the following criteria: (a) being an elementary school with K-3 grades, (b) having personnel who received formal training in First Step, (c) having personnel who at one time implemented First Step in their school, and (d) having school personnel with past and present knowledge of First Step implementation in their school.

Results showed that 28% of adopting school districts continued to implement the First Step program up to 10 years after initial implementation. Six critical features were reported by school

9781032230344c01-end_p1-557.indd 82 15-Jun-22 23:00:34

personnel as associated with First Step sustainability: (a) dedicated resources, (b) training and orientation activities, (c) district-level coordination, (d) selection of students who are a good fit for the program, (e) highly qualified coaches, and (f) administrative support. This was a well-designed and conducted study that produced valuable information on the differences between First Step sustaining and non- sustaining schools.

Cost Analysis of First Step Next Implementation

The long-term societal costs of disruptive behavior disorders are well known. However, there is a dearth of available information on the costs of interventions that are used to impact them. Frey et al. (2019) conducted a comprehensive cost analysis of First Step Next with 40 preschool and kindergarten students. The per child cost to implement the program with 29 triads in two cohorts was $4,330. The incremental cost per additional student was $2,970—a difference of $1,360. In a recent efficacy study involving 379 teacher-parent-student triads, we examined whether either First Step Next (FSN) alone or First Step Next plus a brief, home visitation intervention (i.e., homeBase) to support parents is more cost-effective in treating disruptive behavior problems among elementary aged students (Frey et al., in press). Intervention costs were estimated using an activities-based Ingredients Method (Levin & McEwan, 2001) and comparative cost effectiveness analyses involved calculating incremental cost-effectiveness ratios (ICERs). The average cost per student of $3,387 and the additional cost of serving one more student once the intervention was in place ($2,538) were similar to the costs identified in Frey et al. (2019) for preschoolers. Intervention effectiveness was defined as a student moving from the borderline range into the normative range or from the clinical range into the borderline or normative range at post-intervention. The combined intervention was always more cost-effective among three co-occurring conditions (i.e., ADHD, CD, and co-occurring ADHD and CD).

Implications of First Step and First Step Next Findings for Classroom Management

There have been just over two decades of research and development conducted on First Step and its 2015 revision since it was published in 1997. This research has produced a number of important findings, empirical as well as experiential, that have implications for managing the classroom context. They are listed and briefly described below.

- First Step reduces conflict between teachers and students.
- Target students during First Step Next implementation experience a surge in their popularity with peers as the target student is instrumental in earning a preferred activity in which classmates all share.
- Sprague and Perkins (2009) found positive spillover effects from First Step to (a) other students with problem behavior in the same classroom, (b) with the teacher and (c) on the classroom's climate.
- Horner and colleagues found that First Step works well for students whose problem behavior is motivated primarily by attention seeking and less well for students motivated by escape from or avoidance of situations they perceive as aversive.
- First Step works best when there is a positive universal intervention, such as PBIS, already in place. Students who fail to respond to the universal intervention are likely good candidates for a Tier 2 intervention such as First Step.
- Target students with strong parental support tend to do better in the program.
- Target students who are younger (i.e., preschoolers) or whose problems are more severe tend to need a longer exposure to the program in order to see positive results.

9781032230344c01-end_p1-557.indd 83 15-Jun-22 23:00:34

- Behavioral gains achieved via First Step do not automatically maintain over the long term unless accompanied by regular monitoring and the provision of ongoing supports.
- It appears that First Step can be implemented with substantially less than perfect fidelity and still produce acceptable intervention outcomes.
- First Step produces positive changes in academically related behavior (i.e., AET and Teacher Ratings of Academic Performance) but not on standardized achievement tests.

Concluding Remarks

First Step is a well-researched early intervention and its accumulated evidence suggests that, on average, it produces moderate to strong effects, even when not implemented as well as desired or expected. There is solid evidence that the First Step program works effectively for a diverse array of Pre-K, and K-2 students and that it does not require specialized accommodations for such students to experience success. A recent What Works Clearinghouse review of a preschool, efficacy validation study of First Step Next by Feil et al. (2021) found that the study met WWC group design standards without reservations. Results of this review are available on the WWC website. However, it is also clear that while assigning the program moderate to high satisfaction ratings, some teachers see it as too much work in exchange for the benefits they perceive FS as delivering. Teacher comments suggest they believe the program pulls them away from teaching academics and other essential tasks as opposed to making these tasks easier. However, parents of target students tend to give the program high marks and slightly more positive satisfaction ratings than do teachers. A consistent finding from our research is that the First Step program does not impact direct academic performance measures but does impact important academically related measures such as academic engagement and academic competence-enablers (i.e., forms of student behavior that support direct academic performance such as being organized for instruction, working on assigned tasks, cooperating, responding to teacher directives, and so forth). It is gratifying to see the positive effects of First Step with ADHD students and those having autism and anxiety disorders.

The challenge for programs like First Step is to find ways to sustain their positive effects over the long term but without incurring unreasonable costs (time, effort, materials) in doing so. It has been argued that behavior change is a two-part process. That is, there is one set of procedures for producing it and determining efficacy-effectiveness and another set of procedures governing its durability and generalization. Part two of this change process continues to be laden with challenges that remain to be solved. It appears that First Step is robust in addressing part one of the behavior change process—at least in the relative short term (i.e., within a school year). We will address the issue of inducing longer-term intervention outcomes as an important component of a continuing program of First Step Next research.

Note

1 The authors are indebted to Ed Sabornie for his editorial assistance during the production of this manuscript. Portions of the material contained herein in the Evaluation Research section were previously included in Walker et al. (2014b). Thanks to Kristina Hulegaard for her assistance in finding URLs.

References

August, G., Realmuto, G., MacDonald, A., Nugent, S., & Crosby, R. (1996). Prevalence of ADHD and comorbid disorders among elementary school children screened for disruptive behavior. *Journal of Abnormal Child Psychology*, *24*, 571–595. https://doi.org/10.1007/BF01670101

Carter, D. & Horner, R. (2007). Adding functional behavioral assessment to First Step Next to Success: A case study. *Journal of Positive Behavior Intervention*, *9*(4), Fall, 229–238. https://doi.org/10.1177/10983007070090040501

PROOF

9781032230344c01-end_p1-557.indd 84 15-Jun-22 23:00:34

Carter, D., & Horner, R. (2009). Adding function-based behavioral support to First Step Next to Success: Integrating individualized and manualized practices. *Journal of Positive Behavior Intervention*, *11*(1), 22–34. https://doi.org/10.1177/1098300708319125

Conduct Problems Prevention Research Group (2011). The effects of the Fast Track preventive intervention on the development of conduct disorder across childhood. *Child Development*, January–February, *82*(1), 331–345. https://doi.org/10.1111/j.1467-8624.2010.01558.x

Connors, C. (1997). *Connors' rating scales—revised technical manual*. Multi-Health Systems, Inc.

Drake, R., Latimer, E., Leff, S., McHugo, G., & Burns, B. (2004). What is evidence? *Child and Adolescent Psychiatric Clinics of North America*, *13*, 717–728. https://doi.org/10.1016/j.chc.2004.05.005

Feil, E., Frey, A., Walker, H., Seeley, J., Golly, A., & Forness, S. A. (2014). The efficacy of a home/school intervention for preschoolers with challenging behaviors: A randomized controlled trial of Preschool First Step Next to Success. *Journal of Early Intervention*, *36*(3), 151–170. https://doi.org/10.1177/1053815114566090

Feil, E., Small, J., Seeley, J., Walker, H., Golly, A., Frey, A., & Forness, S. (2016). Early intervention for preschoolers at risk for Attention-Deficit/Hyperactivity Disorder: Preschool First Step Next. *Behavioral Disorders*, *41*, 95–106. https://doi.org/10.17988/0198-7429-41.2.95

Feil, E., Walker, H., Frey, A., Seeley, J., Small, J., Golly, A., Lee, J., & Forness, S. (2021). Efficacy validation of the revised First Step Next program: A randomized controlled trial. *Exceptional Children*, *87*(2), 183–198. https://doi.org/10.1177/0014402920924848

Flay, B., Biglan, A., Boruch, R., Castro, F., Gottfredson, D., Kellam, S., Moscicki, E., Schinke, S., Valentine, J., & Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness, & dissemination. *Prevention Science*, *6*(3), September, 151–175. https://doi.org/10.1007/s11121-005-5553-y

Franz, L., Angold, A., Copeland, W., Costello, E., Towe-Goodman, N., & Egger, H. (2013). Preschool anxiety disorders in pediatric primary care: Prevalence and comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*, 1294–1303. https://doi.org/10.1016/j.jaac.2013.09.008

Frey, A., Kuklinski, M., Bills, K., Small, J., Forness, S. R., Walker, H., Feil, E., & Seeley, J. (2019). Comprehensive cost analysis of First Step Next for preschoolers with disruptive behavior disorder: Using real-world intervention data to estimate costs to scale. *Prevention Science*, *20*, 1219–1232 https://doi.org/10.1007/s11121-019-01035-z

Frey, A., Small, J., Feil, E., Seeley, J., Walker, H., & Forness, S. R. (2015). First Step Next: Application to preschoolers at risk of developing autism spectrum disorders, *Education and Training in Autism and Developmental Disabilities*, *50*(4), 397–407.

Frey, A., Small, J., Seeley, J., Walker, H., Feil, E., Forness, S. R., Lee, J., Lissman, D., Crosby, S., & Golly, A. (in press). First Step Next and homeBase: A comparative efficacy study of children with disruptive behavior. *Exceptional Children*. https://urldefense.com/v3/__https://doi.org/10.1177/0014402920924848.

Frey, A., Small, J., Walker, H., Mitchell, B., Feil, E., Lee, J., & Forness, S. (in press). First Step Next: A synthesis of replication studies from 2009-2021. *Remedial & Special Education*.

Gadow, K., & Sprafkin, J. (2000). *Early childhood inventory-4 screening manual*. Checkmate Plus.

Gresham, F. & Elliott, S. (1990). *Social Skills Rating System. SSRS*. (test and manual), American Guidance Service.

Hawkins, D., Catalano, R., Kosterman, R., Abbott, R., & Hill, K. (1999). Preventing adolescent health risk behaviors by strengthening protection during childhood. *Archives of Pediatric and Adolescent Medicine*, *153*, 226–234. https://doi.org/10.1001/archpedi.153.3.226

Hops, H. & Walker, H. (1988). *CLASS: Contingencies for Learning Academic and Social Skills*. Educational Achievement Systems.

Horner, R., Carr, E., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*(2) 165–179. https://doi.org/10.1177/001440290507100203

Kim, J., Freeman, S., Paparella, T, & Forness, S. (2012). Five year followup of preschoolers with autism and comorbid psychiatric disorders. *Behavioral Disorders*, *38*(1), Nov., 57–70. https://doi.org/10.1177/019874291203800105

Levin, M., & McEwan, J. (2001). *Cost effectiveness analysis: Methods and applications*. Sage.

Loman, S., Rodriguez, B., & Horner, R. (2010). Sustainability of a focused intervention package: First Step to Success in Oregon. *Journal of Emotional Disorders*, *18*(3), 178–191. https://doi.org/10.1177/1063426610362899

Nelson, R., Hurley, K., Synhorst, L., Epstein, M., Stage, S., & Buckley, J. (2009). The child outcomes of a behavior model. *Exceptional Children*, *76*(1), 7–30. https://doi.org/10.1177/001440290907600101

Paine, S., Hops, H., Walker, H., Greenwood, C., Fleischman, D., & Guild, J. (1982). Repeated treatment effects: A study of maintaining behavior change in socially withdrawn children. *Behavior Modification* https://doi.org/10.1177/01454455820062002

Patterson, G. R. (1982). *Coercive family process (Vol. 3): A social learning approach*. Castalia.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Castalia.

Pfiffner, L., Mikami, A., Huang-Pollack, C., Easterlin, B., Zalecki, C., & McBurnett, K. (2007). A randomized, controlled trial of integrated home-school behavioral treatment for ADHD, predominantly intensive type.

9781032230344c01-end_p1-557.indd 85 15-Jun-22 23:00:34

Journal of the American Academy of Child and Adolescent Psychiatry, *46*(8), 1041–1050. https://doi.org/10.1097/chi.0b013e318064675f

Reid, J. B., Patterson, G. R., & Snyder, J. J. (Eds.). (2002). *Antisocial behavior in children and adolescents: A developmental analysis and the Oregon Model for Intervention*. American Psychological Association.

Romer, D., & Heller, T. (1983). Social adaptation of mentally retarded adults in community settings: A social-ecological approach. *Applied Research in Mental Retardation*, *4*, 303–314. https://doi.org//10.1016/0270-3092(83)90031-0

Schrank, F., Mather, N., & Woodcock, R. (2004). *Woodcock-Johnson IIIr Diagnostic Reading Battery*. (test and manual), (2004). Riverside Insights.

Seeley, J., Small, J., Feil, E., Frey, A., Walker, H., Golly, A., & Forness, S. (2018). Effects of the First Step Next intervention on preschoolers with disruptive behavior and comorbid anxiety problems. *School Mental Health*, *10*, 243–253. https://doi.org/10.1007/s12310-017-9226-3

Seeley, J. R., Small, J. W., Walker, H. M., Feil, E. G., Severson, H. H., Golly, A. M., & Forness, S. R. (2009). Efficacy of the First Step to Success intervention for students with ADHD. *School Mental Health*, *1*, 37–48. https://doi.org/10.1007/s12310-008-9003-4

Sheridan, S., Swanger-Gange', M., Welch, G., Kwon, K., & Garbacz, A. (2009). Fidelity measurement in consultation: Psychometric issues and preliminary examination. *School Psychology Review*, *38*(4), 476–495.

Sprague, J., & Perkins, K. (2009). Direct and collateral effects of the First Step to Success program. *Journal of Positive Behavior Interventions*, *11*(4), 208–221. https://doi.org/10.1177/1098300708330935

Sumi, C., Woodbridge, M., Javitz, H., Thornton, P., Wagner, M., Rouspil, K., Yu, J., Seeley, J., Walker, H., Golly, A., Small, J., Feil, E., & Severson, H. (2013). Assessing the effectiveness of First Step to Success: Are short-term results the First Step Next to long-term behavioral improvements? *Journal of Emotional and Behavioral Disorders*, *21*(1). https://doi.org/10.1177/1063426611429571

Walker, H. M. (2004). Commentary: Use of evidence-based interventions in schools: Where we've been, where we are, and where we need to go. *School Psychology Review*, *33*(3), 398–407. https://doi.org/10.1080/02796015.2004.12086256

Walker, H. M., Colvin, G., & Ramsey, E. (1995). *Antisocial behavior in school: Strategies and best practices*. Brooks/Cole, Inc.

Walker, H., Feil, E., Frey, A., Small, J., Seeley, J., Golly, A., Crosby, S., Lee, J., Forness, S. R., Sprick, M., Coughlin, C., & Stiller, B. (2018). First Step Next: An updated version of the First Step to Success early intervention program. *Perspectives on Early Childhood Psychology and Education*, *3*(1), 89–109.

Walker, H. M., Kavanagh, K., Stiller, B., Golly, A., Feil, E., & Severson, H. H. (1997). *First Step to Success: Helping young children overcome antisocial behavior (an early intervention program for grades K-3)*. Sopris West.

Walker, H. M., Kavanagh, K., Stiller, B., Golly, A., Severson, H. H., & Feil, E. G. (1998). First Step to Success: An early intervention approach for preventing school antisocial behavior. *Journal of Emotional and Behavioral Disorders*, *6*(2), 66–80. https://doi.org/10.1177/106342669800600201

Walker, H. M., Seeley, J. R., Small, J., Severson, H. H., Graham, B., Feil, E. G., Serna, L., Golly, A. M., & Forness, S. R. (2009). A randomized controlled trial of the First Step to Success early intervention: Demonstration of program efficacy outcomes in a diverse, urban school district. *Journal of Emotional and Behavioral Disorders*, *17*(4), 197–212. https://doi.org/10.1177/1063426609341645

Walker, H. M., Severson, H., & Feil, E. (2014a) (2nd Ed.). *Systematic Screening for Behavior Disorders. SSBD*. (test and manual). Ancora Publishing, Inc.

Walker, H. M., Severson, H., Seeley, J., Feil, E., Small, J., Golly, A., Frey, A., Lee, J., Sumi, W.C., Woodbridge, M., Wagner, M., & Forness, S. (2014b). The evidence base of the First Step to Success early intervention for preventing emerging antisocial behavior patterns. In H.M. Walker & F.M. Gresham (Eds.), *Handbook of evidence-based practices for emotional and behavioral disorders: Applications in schools* (pp. 518–537). Guilford.

Walker, H. M., Severson, H., Stiller, B., Williams, G., Haring, N., Shinn, M., & Todis, B. (1988). Systematic screening of pupils in the elementary age range at risk for behavior disorders: Development and trial testing of a multiple gating model. *Remedial and Special Education*, *9*(3), 8–14. https://doi.org/10.1177/074193258800900304

Walker, H. M., Severson, H. H., Todis, B. J., Block-Pedego, A. E., Williams, G. J., Haring, N. G., & Barckley, M. (1990). Systematic screening for behavior disorders (SSBD): Further validation, replication and normative data. *Remedial and Special Education*, *11*(2), 32–46. https://doi.org/10.1177/074193259001100206

Walker, H., Stiller, B., Coughlin, C., Golly, A., Sprick, M. & Feil, E (2015) *First Step Next* (2nd Ed.). Ancora Publishing.

Weisz, J., & Jensen, A. (2001). Efficacy and effectiveness of psychotherapy with children and adolescents. *European Child and Adolescent Psychiatry*, *10*, 112–118. https://doi.org/10.1007/s007870170003

9781032230344c01-end_p1-557.indd 86 15-Jun-22 23:00:34