A randomized wait-list controlled analysis of the implementation integrity of team-initiated problem solving processes

J. Stephen Newton [a,*], Robert H. Horner [a], Bob Algozzine [b], Anne W. Todd [a], Kate Algozzine [b]

[a] Educational and Community Supports, 1235 University of Oregon, Eugene, OR 97403-1235, USA
[b] BRIC/EDLD/COED, University of North Carolina Charlotte, Charlotte, NC 28223, USA

ARTICLE INFO

ABSTRACT

Members of Positive Behavior Interventions and Supports (PBIS) teams from 34 elementary schools participated in a Team-Initiated Problem Solving (TIPS) Workshop and follow-up technical assistance. Within the context of a randomized wait-list controlled trial, team members who were the first recipients of the TIPS intervention demonstrated greater implementation integrity in using the problem-solving processes during their team meetings than did members of PBIS Teams in the Wait-List Control group. The success of TIPS at improving implementation integrity of the problem-solving processes is encouraging and suggests the value of conducting additional research focused on determining whether there is a functional relation between use of these problem-solving processes and actual resolution of targeted student academic and social problems.

© 2012 Society for the Study of School Psychology. Published by Elsevier Ltd. All rights reserved.

1. Introduction

It has long been advocated that educators use data to monitor student performance (Boudett, City, & Murnane, 2005; Deno, 1985; Individuals with Disabilities Education Improvement Act, 2004; Ysseldyke & Algozzine, 2006). Teams of school personnel who may use data for problem identification and problem solving have been referred to in various ways, including Teacher Assistance Team (Chalfant, Pysh, &

* Corresponding author. Tel.: +1 336 404 3219; fax: +1 541 346 5517.
 E-mail addresses: jnewton@uoregon.edu (J.S. Newton), robh@uoregon.edu (R.H. Horner), rfalgozz@uncc.edu (B. Algozzine), awt@uoregon.edu (A.W. Todd), kmalgozz@uncc.edu (K. Algozzine).
 ACTION EDITOR: Robin Codding

Moultrie, 1979); Instructional Consultation Team (Rosenfield & Gravois, 1996); prereferral intervention team (Graden, Casey, & Christenson, 1985); instructional support team (Kovaleski & Glew, 2006); student assistance team (Rankin & Aksamit, 1994); and School-Based Intervention Team (McDougal, Clonan, & Martens, 2000). In some instances, a school may have multiple teams, each addressing a specific issue (e.g., a literacy team) within a problem-solving or response-to-intervention framework (e.g., Batsche, Curtis, Dorman, Castillo, & Porter, 2007; Brown-Chidsey & Steege, 2005; Fuchs, Mock, Morgan, & Young, 2003). School psychologists are often leaders of such problem-solving teams (e.g., Cowan & Cohn, 2009; Gelzheiser, 2009; National Association of School Psychologists [NASP], 2010; Shriberg, 2007; Tilly, 2008; Ysseldyke, Burns, & Rosenfield, 2009). The extent to which members of such teams actually use data for problem identification and problem solving—and also engage in other related problem-solving processes— is often unclear. Team members' failure to monitor implementation integrity with regard to problem-solving processes, including the use of data to identify problems, is problematic in that assessment of treatment (or implementation) integrity is crucial for determining whether there is a functional relation between changes in the environment and changes in a target behavior (e.g., Baer, Wolf, & Risley, 1968; Billingsley, White, & Munson, 1980).

1.1. Research concerning implementation integrity

Reviews of published research have shown that the failure to gather or report treatment integrity data is widespread (Gresham, Gansle, & Noell, 1993a, 1993b; Gresham, Gansle, Noell, & Cohen, 1993; McIntyre, Gresham, DiGennaro, & Reed, 2007; Peterson, Homer, & Wonderlich, 1982; Sanetti, Gritter, & Dobey, 2011), and a recent study by Cochrane and Laux (2008) demonstrated that the lack of treatment integrity data is specifically an issue for school-based problem-solving teams. Cochrane and Laux created an online survey completed by 806 self-selected holders of the Nationally Certified School Psychologist (NCSP) credential. Of the respondents, 56.2% strongly agreed that treatment integrity is a key factor to consider when evaluating the success of an intervention, but when asked whether, in their experience, school-based problem-solving teams measured treatment integrity, only 1.9% responded "yes."

A team's failure to measure implementation integrity of their problem-solving processes may result in both poorer implementation integrity and poorer student outcomes. In a pilot empirical meta-analysis of research concerning the effect of prereferral intervention teams (PITs) on student outcomes and systemic variables, Burns and Symington (2002) found that the mean effect size for university-based PITs (1.32) was more than twice that of field-based teams (0.54). This difference led Burns and Symington to speculate that the lack of consistent positive findings from field-based studies might be a result of inconsistent implementation integrity by PITs, possibly due to local school districts having to design and implement their own approaches to PITs, as originally suggested in results from a survey conducted by Carter and Sugai (1989) and later replicated by Buck, Polloway, Smith-Thomas, and Cook (2003). In a subsequent study, conducted by Burns, Vanderwood, and Ruby (2005) and focused on pre-referral intervention teams, the authors reviewed published research and concluded that although the PIT model had both a sound theoretical base and data to support its effectiveness, its lack of reliable implementation suggested a "significant flaw" (p. 101).

For the relatively few studies in which the implementation integrity of school-based problem-solving teams has been measured via (a) self-report data provided by team members, (b) data gathered by external observers, or (c) both, results generally show much room for improvement. Flugum and Reschly (1994) used a questionnaire, completed by Iowa school personnel, to determine the extent to which six quality indices of prereferral problem-solving were used. Results indicated that the typical prereferral process did not include a behavioral definition of the target behavior, direct measure of the target behavior in the natural setting prior to intervention (baseline data), step-by-step intervention plan, graphing of intervention results, or direct comparison of post-intervention performance with baseline data. In a similar study, Telzrow, McNamara, and Hollinger (2000) evaluated the implementation integrity of multidisciplinary teams (MDTs) in their use of the Intervention Based Assessment (IBA) approach to problem-solving. Teams provided the researchers with their "best case documentation" for a single student. Telzrow et al. reviewed the documentation and found that the teams used the eight IBA problem-solving components with varying degrees of implementation integrity, with the lowest implementation scores assigned to the following components: hypothesized reason for the problem, evidence that the

intervention was implemented as designed, and direct comparison of the student's post-intervention performance with baseline data. Doll, Haack, Kosse, Osterloh, and Siemers (2005) conducted an evaluation of the extent to which student assistance teams (SATs) adhered to the eight problem-solving components detailed in the Telzrow et al. instrument. Examination of team members' self-assessment ratings for the eight problem-solving components demonstrated that team members rated themselves as unsystematic in their implementation of the components and "did not meet the criteria for high-quality performance for most of the indicators" (p. 137). Evaluators also asked the teams to provide documentation for the single case (student) that members believed to represent their best application of the SAT process. Evaluators' examination and rating of this documentation also revealed that teams were not fully adhering to the problem-solving components.

Similar findings were obtained in research that included data not only from team members' self-reports but also from external observers. McDougal et al. (2000) conducted a study with four schools that used the School-Based Intervention Team (SBIT) prereferral consultation process. At the end of the initial team meeting for each of 47 students, team members completed a SBIT Process Integrity Checklist to rate their use of each of 10 SBIT consultation objectives during the meeting. For 49% of the meetings, one or more external observers also completed the checklist. McDougal et al. found that team members rated their implementation integrity higher than did the observers, with the observers more often rating the following objectives as unmet: identifying antecedents and consequences for the problem, specifying a measure of student progress, and specifying a time for a follow-up meeting.

In summary, research has shown that although treatment (implementation) integrity is crucial for determining functional relations, measurement of treatment integrity is often neglected by both researchers and practitioners. For the relatively few studies that have addressed this issue with school-based problem-solving teams, researchers have found that the following may be of particular concern: (a) lack of access to preexisting standardized training, making it necessary for each local school district to develop and deliver its own training for problem-solving teams; (b) researchers' reliance on self-report data (and best case documentation) for determining implementation integrity, which some research suggests may result in overestimation; and (c) generally low implementation integrity for certain problem-solving processes (e.g., gathering baseline data, graphing intervention results, and comparing post-intervention performance with baseline data). Given these poor implementation integrity findings, additional research focused on developing and examining procedures for improving implementation integrity, including the concerns noted above, is needed. The problem-identification and problem-solving processes of Positive Behavior Interventions and Supports (PBIS) teams (Lewis, Jones, Horner, & Sugai, 2010; Sugai & Horner, 2006, 2009) provide an excellent focus for such research.

1.2. PBIS Teams and team-initiated problem solving

In schools that are implementing School-wide Positive Behavior Interventions and Supports (SWPBIS; Horner, Sugai, & Anderson, 2010), PBIS Teams regularly meet for the purpose of reviewing student data and identifying and addressing students' social and academic problems. In team meetings, data from the School-wide Information System (SWIS: Horner et al., 2008; Irvin et al., 2006; May et al., 2003) can be a particularly useful source of information. SWIS provides a methodology for defining and collecting data about student office discipline referrals and a Web-based computer application that allows for both the entry of hand-collected referral data and the subsequent production of pre-defined and custom reports concerning the referrals (May et al., 2003). To help PBIS Teams become more efficient and effective at using SWIS data for problem solving, we developed the Team-Initiated Problem Solving (TIPS) model (Newton, Algozzine, Algozzine, Horner, & Todd, 2011; Newton, Horner, Algozzine, Todd, & Algozzine, 2009; Newton, Horner, Todd, Algozzine, & Algozzine, 2012; Newton, Todd, Algozzine, Horner, & Algozzine, 2009; Todd et al., 2011).

The TIPS model was developed for use by PBIS Teams that have previously received training in the Tier I (primary, or universal, problem behavior prevention and intervention) level of SWPBIS, as well as training in the use of SWIS. Tier I training results in team members learning to (a) adopt a common approach to school discipline, (b) define a set of school-wide behavioral expectations for students, (c) teach the behavioral expectations to students, (d) establish a continuum of procedures for regular acknowledgement or reward of students' engagement in the behavioral expectations, and (e) develop a continuum of

corrective consequences for responding to students' rule violations. Team members also learn about the importance of accurate and timely data for guiding decisions concerning students' academic and social behavior (Horner et al., 2010; Sugai & Horner, 2009). Subsequent training in the use of SWIS results in team members learning to enter office discipline referral data, produce reports, and use the reports to make decisions. SWIS training is provided by a SWIS facilitator who is employed by a school district, and who has previously participated in a two-and-a-half day SWIS Facilitator Training event conducted by SWIS developers and researchers (Horner et al., 2008).

The TIPS model is an explicit, sequenced set of procedures for integrating the SWPBIS and SWIS skills of PBIS team members into a conceptually coherent, data-based approach to identifying and solving students' social behavior problems. (Although the model's generic problem-solving processes appear appropriate for also addressing academic problems, our focus to date has been on social behavior.) For example, TIPS can be used to identify and address student social behavior problems at the level of primary prevention or intervention (a.k.a. Tier I or universal supports), secondary intervention (a.k.a. Tier II or targeted interventions), and tertiary intervention (a.k.a. Tier III or intensive interventions). PBIS team members learn to use the TIPS model's problem-solving processes via a TIPS Workshop. Follow-up technical assistance is then provided by a team's PBIS Coach, consistent with the recommendations of Joyce and Showers (2002), who found that even when a workshop for teachers includes (a) presentations that provide theory and rationale for the workshop content, (b) demonstrations or modeling of the skills that teachers are to acquire, and (c) opportunities for teachers to practice the skills, teachers' success at transferring those skills to classrooms is minimal. However, when post-workshop coaching is also provided, a large and dramatic increase in transfer typically occurs, allowing teachers to demonstrate implementation integrity in applying workshop-related skills in their classrooms.

The TIPS model addresses the general limitations noted in the previous review of the literature in that (a) training content and activities are documented in a TIPS training manual (Newton, Todd, Algozzine et al., 2009), which may make it possible for local school districts to adopt and deliver this standardized training rather than developing their own training materials; (b) TIPS research data are gathered by an independent observer who attends team meetings and records data concerning the extent to which team members are actually achieving implementation integrity with regard to the TIPS problem-solving processes, rather than relying on team members' self-report data or best case documentation; and (c) those problem-solving processes for which teams have generally demonstrated low implementation (gathering baseline data, graphing intervention results, comparing post-intervention performance with baseline data) may be easier to implement within TIPS, given that SWIS facilitates data entry and the production of custom reports and graphs.

For school-based problem-solving teams, such as PBIS Teams, assessment of the implementation integrity of a team's problem-solving processes is essential because it is through these problem-solving processes that team members select solutions for resolving problems. In developing TIPS, we recognized that before we can assess whether use of TIPS is functionally related to improvement in student outcomes, we must demonstrate that TIPS training for PBIS team members results in implementation integrity of the problem-solving processes in their meetings.

1.3. Preliminary research on TIPS

As a prelude to the randomized wait-list controlled trial described in this article, we conducted a small-scale study (Todd et al., 2011). PBIS team members from four Title I elementary schools in Oregon participated in the study. The four teams were implementers of SWPBIS and users of SWIS. Prior to each team's participation in the TIPS Workshop, which the teams accessed in the context of a multiple-baseline-across-teams design (e.g., Hersen & Barlow, 1976), an observer attended at least six team meetings and used the Decision Observation, Recording, and Analysis (DORA) instrument (Newton, Todd, Horner, Algozzine, & Algozzine, 2009) to gather data concerning the extent to which a team (a) established "foundational" aspects for conducting an efficient and effective meeting, such as assigning team roles and developing an agenda, and (b) engaged in the problem-solving processes exemplified by the TIPS model: identifying a problem, using data to identify and define the problem with precision, discussing and selecting a solution to address the problem, and developing an action plan for implementing the solution. (Details concerning the DORA data collection instrument and the TIPS Workshop are presented in the

description of the randomized wait-list controlled trial in the Method section.) Following each team's participation in the TIPS Workshop, a researcher attended the team's next two regularly scheduled meetings and provided TIPS "coaching" consisting of (a) providing prompts, as necessary, for team members to engage in the TIPS problem-solving processes; (b) praising correct implementation; (c) providing corrective feedback when necessary; and (d) answering team members' questions. DORA data were then collected for at least two additional team meetings, during which TIPS coaching was not provided. Visual analysis of the DORA data demonstrate a functional relation between TIPS training (i.e., the TIPS Workshop and the two coached meetings) and improvements in team members (a) establishing the foundations for efficient and effect team meetings and (b) implementing the TIPS problem-solving processes (i.e., demonstrating implementation integrity).

Given the promising results of this small-scale study, we were interested in expanding the scope and generality of our research to answer the following broad question: Within the context of a randomized wait-list controlled trial, do PBIS team members who participate in a TIPS Workshop and follow-up technical assistance (the "Immediate" group), demonstrate greater implementation integrity in using the problem-solving processes during team meetings than do PBIS team members who receive the TIPS Workshop and follow-up technical assistance at a later point in time (the "Wait-List Control" group)?

2. Method

2.1. Participants and selection procedures

Participants were the members of 34 PBIS Teams, 22 teams from schools in North Carolina and 12 teams from schools in Oregon. All participating schools met the following criteria for inclusion in the project: (a) is an elementary school; (b) has been implementing SWPBIS for at least 1 year, as assessed via the School-wide Evaluation Tool (SET; Horner et al., 2004), the Benchmarks of Quality (BoQ; Cohen, Kincaid, & Childs, 2007), or the Team Implementation Checklist (TIC; Sugai, Horner, & Lewis-Palmer, 2009); (c) has been a user of SWIS during at least the most recent 6-month period; and (d) has a PBIS team that meets at least once a month. Based on these selection criteria, schools were recruited from a sample of convenience nominated by state PBIS Coaches. PBIS Coaches were school district employees with responsibility for providing PBIS-related training and technical assistance to school teams. The PBIS Coaches sent school principals a recruitment letter, provided by the researchers, describing the research and asking principals about their interest in participating. The first 22 North Carolina schools and the first 12 Oregon schools that expressed interest in the research were selected for participation. All team members provided informed consent for participation.

Data concerning the demographic characteristics of the PBIS Teams and schools in the Wait-List Control group and in the Immediate group are provided in Table 1. No statistically significant differences ($p > .05$) were found between the two groups with regard to number of team members (Wait-List Control group range, 7 to 14; Immediate group range, 6 to 23), number of years implementing SWPBIS (Wait-List Control group range, 1 to 5; Immediate group range, 1 to 5), total school enrollment, percent student enrollment for kindergarten through fifth grade, teacher–pupil ratio, student ethnicity, or student gender. Likewise, there was no statistically significant difference between the two groups regarding the number of schools receiving Title I federal assistance, $X^2 = 1.21$, $df = 1$, $p = .271$. The majority of teams in both groups held their team meetings once a month (Wait-List Control group, 65%; Immediate group, 76%). Five teams in the Wait-List Control group (29%) and three teams in the Immediate group (18%) held their meetings twice a month; one team in each group held its meetings once every 3 weeks. There was no statistically significant difference between the frequency of meetings for teams in the two groups, $X^2 = 0.67$, $df = 2$, $p = .717$.

2.2. Dependent variables, observers, interobserver agreement, and data collection schedule

2.2.1. Problem-solving processes
The primary dependent variables were the problem-solving processes engaged in by team members. We also gathered information concerning the foundational aspects of a team meeting, as we had done in our small-scale preliminary research (Todd et al., 2011).

Table 1
Demographic characteristics of PBIS Teams and schools in Wait-List Control group and Immediate group.

Characteristic	Wait-List Control group		Immediate group		Obtained t
	M	SD	M	SD	
Number of PBIS team members	9.88	2.39	10.59	4.20	− 0.60
Years implementing SWPBIS	2.59	0.94	2.71	1.36	− 0.29
School enrollment (Number)	567.94	198.24	607.12	224.68	− 0.54
School enrollment (Percent)					
Kindergarten	16.64	3.22	16.68	1.99	− 0.04
First	16.89	1.94	16.81	2.32	0.11
Second	16.98	1.60	16.84	1.85	0.58
Third	16.63	1.69	16.71	1.34	− 0.16
Fourth	16.78	2.14	16.59	1.93	0.26
Fifth	16.09	2.36	16.57	2.05	− 0.63
School teacher–pupil ratio[a]	16.39	2.76	16.22	3.61	0.16
Student ethnicity (percent)					
African American	26.86	24.62	23.17	22.16	0.46
American Indian	1.28	1.67	0.65	1.15	1.28
Asian	4.14	4.29	2.59	1.99	1.35
Hispanic	13.36	7.91	13.64	7.68	− 0.10
Caucasian	54.36	31.02	59.96	24.55	− 0.58
Student gender (percent)					
Male	51.54	1.45	50.75	2.21	1.23

Note. No statistically significant differences ($p<.05$) were found between groups.
 [a] Average of number of students in school, divided by number of teachers in school, as reported on state education department websites

To measure the problem-solving processes, we directly observed team meetings at all 34 schools, using the Decision Observation, Recording, and Analysis (DORA) instrument (Algozzine, Newton, Horner, Todd, & Algozzine, in press; Newton, Todd, Horner et al., 2009; Todd, Newton, Horner, Algozzine, & Algozzine, 2009) to record information during the baseline and post-TIPS conditions. (A copy of DORA is available on request from the authors.) We used DORA to record student problems that team members identified during a meeting. A "problem" was defined as a discrepancy between students' actual and desired academic or social behavior that team members consider significant enough to require remediation (e.g., Deno, 2005). In the case of social behavior, a problem might be identified as requiring a solution at the level of (a) the student body (e.g., a Tier I problem involving the vast majority of students in the cafeteria during lunch time), (b) a targeted group of students (e.g., a Tier II problem involving only sixth graders during their recess period on the playground), or (c) an individual student (e.g., a Tier III problem involving a single student's acts of disrespect during math instruction).

Each problem recorded on DORA provided an opportunity to record additional information concerning team members' problem-solving processes for that problem. Table 2 provides a detailed description of the DORA variables, as well as the procedures we used to transform data recorded on DORA into the following scores:

- a Foundations score, providing a summary measure of "structural elements" of a team's meeting
- a Problem Precision score, providing a measure of whether a team identified and defined a problem in detail
- a Use of Quantitative Data score, providing a measure of whether team members used quantitative data to identify and define a problem
- a Solution score, providing a measure of the extent to which team members selected solutions from specific intervention categories to address a problem
- an Action Plan score, providing a measure of the extent to which team members demonstrated accountability for resolving a problem
- a Thoroughness of Problem Solving score, providing a summary measure of the extent to which team members demonstrated (a) problem precision, (b) use of quantitative data, (c) solution selection, and (d) development of an action plan for identified problems.

Table 2
Description of DORA dependent variables and scoring.

DORA variable and description	Number of items	DORA scoring
Foundations—"structural elements" deemed important for conducting effective and efficient meetings (e.g., meeting starts on time; previous meeting minutes are available; agenda is available; team members serve as Facilitator, Minute Taker, and Data Analyst; and quantitative data are available for decision making)	12	Percentage of 12 items demonstrated by the team members
Problem precision—for each problem identified, the extent to which team members also identified five precision elements: what (the specific problem behavior), who (the individual student, or group of students, engaging in the problem behavior), where (the school locations in which the problem is occurring), when (the times of day when the problem is occurring), and why (a hypothesis about environmental conditions that may be setting the occasion for, and maintaining occurrences of, the problem behavior)	5	Percentage of the five precision elements identified by the team members
Use of quantitative data—for each problem identified, whether team members used quantitative data (e.g., SWIS data) to identify a problem or to refine the definition of an identified problem to achieve greater precision	1	Score of 100%, if the team used quantitative data; score of 0% if the team did not use quantitative data
Solution—for each problem identified, the extent to which team members selected solutions from the following five categories: prevention, teaching, reward, correction, extinction	5	Percentage of the five solution categories from which team members selected a solution action for implementation
Action plan—for each problem identified, the extent to which team members demonstrated accountability for resolving the problem by indicating (a) who will complete a solution-related task; (b) the date by which the solution task will be completed; and (c) the goal that, when met, will provide evidence that team members' implementation of the solution tasks has resolved the problem to their satisfaction	3	Percentage of the three accountability items demonstrated by team members
Thoroughness of problem solving—for each problem identified by team members, the extent to which team members demonstrated Problem Precision, Use of Quantitative Data, Solution, and Action Plan	NA	Average of the scores for Problem Precision, Use of Quantitative Data, Solution, and Action Plan

Technical characteristics (i.e., validity and reliability evidence) of data gathered using DORA have been documented (Algozzine et al., in press). Content validity analyses reflected correspondence between the TIPS problem-solving processes assessed via DORA and the critical features of five widely-accepted problem-solving and decision-making models (Boudett et al., 2005; Bransford & Stein, 1984; Deno, 1985; Gilbert, 1978; Hamilton et al., 2009), as well as a high degree of congruence between DORA's assessed problem-solving processes and the content reflected in the Rational Problem Solving subscale of the Social Problem-Solving Inventory—Revised (D'Zurilla & Nezu, 1990; D'Zurilla, Nezu, & Maydeu-Olivares, 2002). DORA demonstrated good interobserver agreement (e.g., Cooper, Heron, & Heward, 2007) for measures gathered at three PBIS team meetings during early pilot testing (Newton et al., 2012), with average agreements of 85% for Foundations (range, 73% to 100%), 50% for problems identified (range, 0% to 100%), 80% for Problem Precision (range, 80% to 80%), 84% for Use of Quantitative Data (range, 67% to 100%), 78% for Solution (range, 67% to 89%), 85% for Action Plan (range, 80% to 90%), and 85% for overall Thoroughness (range 82% to 87%). Likewise DORA demonstrated good interobserver agreement for measures gathered at four meetings of PBIS Teams participating in the previously described small-scale research study (Todd et al., 2011), with agreement for Foundations scores averaging 94% (range, 72% to 100%) and agreement for Thoroughness scores averaging 88% (range, 50% to 100%).

2.2.2. Observers

Data for the study were collected by the first and fourth authors, two research assistants, three doctoral students, and an independent evaluator with a master's degree in education. Observers received training via a DORA Data Collection Workshop, which was conducted by the two authors who served as observers. The workshop involved use of standardized procedures and materials to ensure consistency in training across North Carolina and Oregon. Observers received an overview of the TIPS model and reviewed the DORA Data Collection Protocol and DORA instrument. They read 30 DORA "training vignettes" (Todd et al.,

2009) that sampled the range of problem-solving processes engaged in by a hypothetical PBIS team, recorded associated information on a DORA instrument, and received feedback on their performance after the trainer reviewed their DORAs. Finally, observers watched a 15-minute segment of a video depicting a hypothetical PBIS team meeting (Liaupsin, Scott, & Nelson, 2006), recorded information about the meeting on a DORA instrument, and again received performance feedback after the trainer reviewed their DORAs. Following participation in the DORA Data Collection Workshop, each observer individually accompanied the trainer to a regular meeting of two PBIS Teams, and both the trainer and the observer independently collected DORA data. If the DORAs completed by the trainer and the observer achieved agreement of 85% or better (as computed using the procedures described in the interobserver agreement section that follows), the observer met criterion for gathering future DORA research data. All observers met this criterion.

2.2.3. Interobserver agreement

Once observers met the training criterion and began collecting research data for the study, we assessed interobserver agreement for DORA measures by having two observers independently complete a DORA at 31 of the 99 (31%) team meetings (proportionally distributed across meetings held prior to a team's participation in the TIPS Workshop and meetings held after the team's participation in the workshop). In calculating observers' agreement concerning the Foundations variable, we examined the 12 Foundation items on the observers' DORA instruments and compared the content on a discrete trial (item-by-item) basis by (a) adding the number of items both observers agreed the team demonstrated at the meeting (e.g., agenda is available) to the number of items both observers agreed the team did not demonstrate; (b) dividing that total by 12, and (c) multiplying the quotient by 100% (e.g., Cooper et al., 2007; Page & Iwata, 1986). The average interobserver agreement for the Foundations variable across the 31 DORAs was 97% (range, 83% to 100%).

In calculating observers' agreement on problems identified, we examined the 31 DORAs to determine what descriptions of student problems (e.g., instances of aggression on the playground) the two observers wrote on their DORAs as a record of problems identified by team members. We then (a) divided the number of problems that both observers agreed the team identified (i.e., identical problem descriptions); by (b) the number of problems that both observers agreed the team identified, plus the number of problems that only one observer indicated the team identified; and (c) multiplied the quotient by 100%. The average interobserver agreement for problems identified across the 31 DORAs was 90% (range, 0% to 100%). (There were six instances in which one observer's DORA noted that the PBIS team identified a problem, but the second observer's DORA did not include the problem.)

Finally, to calculate observers' agreement on the Problem Precision, Use of Quantitative Data, Solution, and Action Plan variables, we restricted the analysis to the 49 problems that both observers agreed team members identified across the 31 DORAs. We compared the content of the observers' DORAs for the individual items associated with a given TIPS variable on a discrete trial (problem-by-problem) basis by (a) adding the number of the variable's items that both observers agreed the team achieved for an identified problem (e.g., the what, who, where, when, and why items of the Problem Precision variable) to the number of items that both observers agreed the team did not achieve; (b) dividing that total by the number of items for the variable (e.g., five); and (c) multiplying the quotient by 100%. The average interobserver agreement for the variables was as follows: Problem Precision, 90% (range, 60% to 100%); Use of Quantitative Data, 96% (range, 0% to 100%); Solution, 93% (range, 40% to 100%); and Action Plan, 89% (range, 0% to 100%).

The lower values in the range for percent of interobserver agreement were largely a function of the limited number of items associated with each variable. The Problem Precision variable had 5 items, allowing for 6 possible percent agreement scores. Across the 49 problems that both observers agreed team members identified, there were 4 problems for which the observers' DORAs showed agreement on the presence or absence of 3 of the 5 items, producing a lower range value of 60%. For the Solution variable, which also had 5 associated items, there was 1 problem for which the observers' DORAs demonstrated agreement on the presence or absence of 2 of the 5 items, producing a lower range value of 40%. The Action Plan variable had 3 items, allowing for 4 possible percent agreement scores. Across the 49 problems that both observers agreed team members identified, there was 1 problem for which the observers' DORAs demonstrated agreement on the presence or absence of none of the 3 associated items, producing a lower

range value of 0%. Finally, the Use of Quantitative Data variable included only a single "item." Observers simply noted whether the team members used quantitative data to identify the problem or to refine the definition of an identified problem to achieve greater precision. This procedure resulted in percent interobserver agreement of either 0% or 100% for a given problem. Across the 49 problems that both observers agreed team members identified, there were 2 problems for which 1 observer's DORA indicated that team members used quantitative data and the other observer's DORA indicated that team members did not use quantitative data, thus producing a lower range value of 0% interobserver agreement for these 2 problems.

2.2.4. Data collection schedule

Table 3 provides the DORA data collection schedule and the research design. Prior to the TIPS Workshop for the PBIS Teams that had been randomly assigned to the Immediate group, the observers attended a team meeting at all 34 schools and collected DORA data (O_1 in Table 3) to determine the extent to which teams may have already been implementing problem-solving processes exemplified by TIPS. Following this first round of data collection, team members from the 17 teams in the Immediate group attended the TIPS Workshop (X_1).

After the team members in the Immediate group attended the workshop and were scheduled to have received follow-up technical assistance from their PBIS Coaches at their next two team meetings, observers once again attended a team meeting at all 34 schools and collected a second round of DORA data (O_2). If a team's PBIS Coach also attended this meeting, the coach did not provide TIPS-related technical assistance. This rule ensured that the DORA data collection reflected the behavior of the team members in the absence of any prompting or feedback from the coach. Coaches attended 28% of these meetings. (Note that except for requesting coaches to provide technical assistance at a team's first two post-TIPS Workshop meetings, we advised the coaches simply to adhere to their typical schedules for attending team meetings.)

From this point forward, the research designs in North Carolina and Oregon slightly differed (see Table 3). In North Carolina, observers collected a third round of DORA data (O_3) at all 22 schools, and then we provided the TIPS Workshop (X_2)—and the coaches provided the follow-up technical assistance—to the 11 PBIS Teams in the Wait-List Control group. In Oregon, observers collected a third round of DORA data (O_3) for the six teams in the Immediate group, and then we provided the TIPS Workshop—and the coaches provided the follow-up technical assistance—to the six teams in the Wait-List Control group. It was only after the Oregon teams in the Wait-List Control group had experienced the TIPS intervention that the third round of DORA data were gathered for them. This approach allowed for a small-scale assessment of the extent to which any effects of the TIPS intervention on the 17 teams in the Immediate group were replicated with the six teams in Oregon's Wait-List Control group. Although our original design did not include this assessment of replication effects with the teams in the Wait-List Control group, after several months of the project had elapsed, we made a decision to alter the design in this fashion for only the six teams in the Oregon. Including only the Oregon Wait-List Control group teams in this effort was largely a matter of practicality, given that the North Carolina research team (a) was working with a greater number of teams, (b) had already set the dates for all data collection visits and for delivering the TIPS Workshop for

Table 3
Design for randomized wait-list controlled trial.

North Carolina								
[a]R (N = 11) Immediate group	[b]O_1	[c]X_1	[d]TA	TA	O_2	O_3			
R (N = 11) Wait-List Control group	O_1				O_2	O_3	[e]X_2	TA	TA
Oregon								
R (N = 6) Immediate group	O_1	X_1	TA	TA	O_2	O_3			
R (N = 6) Wait-List Control group	O_1				O_2	X_2	TA	TA	O_3

Note. [a] R = random assignment.
[b] O_i = ith observation of PBIS Team meeting and collection of DORA data.
[c] X_1 = TIPS Workshop for teams in Immediate group.
[d] TA = follow-up technical assistance delivered to PBIS Teams by PBIS Coach.
[e] X_2 = TIPS Workshop for teams in Wait-List Control group.

teams in the Wait-List Control group, and (c) included only two observers. Thus, the North Carolina research team lacked the flexibility and resources of the Oregon research team. Finally, the third round of DORA data collection (O_3) for the Immediate group teams in North Carolina and Oregon allowed for an assessment of the extent to which any effects of the TIPS intervention that were present at O_2 were maintained at O_3.

2.3. Research design

We used a randomized, wait-list controlled trial (Coalition for Evidence-Based Policy, 2003; Horner et al., 2009; Myers & Dynarski, 2003), which controls for threats to internal validity (e.g., Campbell & Stanley, 1963; Gall, Gall, & Borg, 2007; Shadish, Cook, & Campbell, 2002). Due to the relatively small number of teams involved in the project, we did not use simple random assignment of teams to groups, which could have resulted in the design becoming unbalanced with regard to (a) sample size, (b) a key variable correlated with an outcome variable, or (c) both. To achieve a good balance of teams between the two groups, we instead employed blocked randomization (e.g., Matthews, 2006; Torgerson & Torgerson, 2008), using the six possible sequences of a block size of four (e.g., ABAB and AABB), to randomly assign teams to the two groups. This technique resulted in a balanced design, with 17 teams randomly assigned to the Immediate group and 17 teams randomly assigned to the Wait-List Control group. The research design is depicted in Table 3 for both the North Carolina and Oregon teams.

2.4. Independent variables

The independent variables consisted of (a) the TIPS Workshop (Newton, Todd, Algozzine et al., 2009) and (b) TIPS follow-up technical assistance provided to members of the participating teams. The TIPS Workshop was conducted by the first and fourth authors, hereinafter referred to as the TIPS trainers; follow-up technical assistance was provided by a team's PBIS Coach. PBIS Coaches were school district employees with the pre-existing responsibility of providing PBIS-related training and technical assistance to school teams. They agreed to also provide the TIPS-related technical assistance to the teams participating in the TIPS research project, as their schedules and other duties permitted. As described in a following section, the TIPS trainers provided a TIPS Workshop for the PBIS Coaches prior to the TIPS Workshop that was held for the PBIS team members, in order to ensure that the coaches were prepared to assist team members in completing workshop activities and in implementing the TIPS problem-solving processes during a team's first two post-TIPS Workshop meetings.

2.4.1. TIPS Workshop

The 5-hour TIPS Workshop was conducted in North Carolina and in Oregon. Content for the workshop sessions was referenced to the components of the TIPS model: establish problem-solving foundations, identify problems, develop hypothesis, discuss and select solutions, develop and implement action plan, and evaluate and revise action plan. Participating schools were reimbursed for the cost of hiring substitute teachers for up to five of a school's team members. Under this arrangement, each school sent five team members to the workshop.

The format for each of the workshop sessions was identical, and was designed to ensure that participants acquired the associated skills. The TIPS trainer first provided the content for a session (e.g., how to identify student problems using data from the school's SWIS database), interspersing questions that participants answered aloud to demonstrate their understanding. Next the participants completed a related activity, with technical assistance from their PBIS Coach as necessary, allowing them to demonstrate acquisition of the session-related skills under simulated circumstances (e.g., reviewing graphs and reports of SWIS data for a hypothetical school and identifying problems), followed by performance feedback from their coach and the TIPS trainer. Finally, the participants demonstrated acquisition of the skills in a manner directly relevant for their school/team (e.g., using a laptop computer each team brought to the workshop, accessing their school's SWIS database, and identifying a student problem at their school), again followed by performance feedback from their coach and the TIPS trainer.

The session on *establishing problem-solving foundations* focused on important "structural" aspects of a team meeting, including assigning team roles, developing an agenda, and electronically displaying agenda

items and a record of the team's decisions and tasks. A particular emphasis was placed on the role of a Data Analyst who reviews student data, leads team members in a discussion of possible new student problems, and provides members with an update on the extent to which implementation of team-developed solutions (i.e., interventions) for student problems identified at previous meetings is having the desired effect. The session ended with participants using a checklist to record related tasks to complete before their next meeting (e.g., assigning the role of Data Analyst to a team member).

The sessions on *identifying problems and developing hypotheses* focused on how to use data (e.g., SWIS data) to identify student problems. A problem was defined as (a) a discrepancy between the current status of a student behavior and its expected or desired status (e.g., Bransford & Stein, 1984) that (b) team members judged to be significant enough to address (e.g., Deno, 2005). Participants learned to review data in a structured manner that resulted in their writing a "precise problem statement" that defined the what, where, when, who, and why elements of an identified problem. The session ended with participants writing one precise problem statement on a trainer-provided form (the TIPS Meeting Minutes and Problem-Solving Action Plan form) after having reviewed their school's SWIS data.

The session on *discussing and selecting solutions* focused on how to develop solution actions (interventions) that are (a) consistent with the hypothesis about why a problem is occurring and (b) derived from the broad "solution strategy" options of preventing the problem behavior, defining and teaching appropriate alternative behavior, rewarding or reinforcing instances of the appropriate behavior, withholding rewards or reinforcers for the problem behavior (i.e., extinction), and taking non-rewarding or non-reinforcing corrective action if the problem behavior occurs. The session ended with participants discussing and recording at least one possible solution action for addressing the problem described in their school's precise problem statement, and agreeing to continue the discussion at their next team meeting as necessary.

The session on *developing, implementing, evaluating, and revising an action plan* focused on teaching team members to achieve accountability for implementing their "solution actions" by recording the name of the person who would implement (or coordinate implementation of) each solution action and the date by which each solution action would be implemented. Participants also learned to write a goal statement that empirically defined resolution of the problem to team members' satisfaction, and to use data at their meetings for progress monitoring concerning the effectiveness of the implemented solution actions.

2.4.2. TIPS Workshop for PBIS Coaches and follow-up technical assistance to PBIS Teams

Prior to the TIPS Workshop for PBIS team members, the TIPS trainers conducted a TIPS Workshop for the PBIS Coaches to ensure they were prepared to provide technical assistance to team members during the workshop activities. The TIPS trainer provided an overview of the rationale, objectives, and content for each workshop session; described each of the workshop activities in detail; asked the coaches to complete the activities; and provided assistance, as necessary, and performance feedback until the coaches expressed confidence that they could provide technical assistance during the workshop.

The TIPS trainer then provided the coaches with a TIPS Technical Assistance Data Collection Form (available from the authors) that noted their post-TIPS Workshop technical assistance responsibilities in checklist format. The responsibilities involved contacting their team's Facilitator, Minute Taker, and Data Analyst prior to the team's first post-TIPS Workshop meeting and offering and providing technical assistance related to that role. Next, coaches were to provide technical assistance at a team's first two post-TIPS Workshop meetings. Because the coaches had substantial experience at providing PBIS-related training and technical assistance, the TIPS trainer described these responsibilities in general terms: (a) providing the least amount of prompting required to ensure team members progressed through the components of the TIPS model as time allowed; (b) providing corrective feedback and assistance, as needed, if team members omitted component processes or were unable to complete them; (c) providing praise as team members independently completed component processes; and (d) answering team members' questions. We regarded this follow-up technical assistance as crucial, given the previously-cited findings by Joyce and Showers (2002), and given the results of previous studies demonstrating the effectiveness of performance feedback at increasing treatment and implementation integrity (e.g., Burns, Peters, & Noell, 2008; Codding, Feinberg, Dunn, & Pace, 2005; DiGennaro & Martens, 2007; Duhon, Mesmer, Gregerson, & Witt, 2009; Moore et al., 2002; Noell et al., 2005; Rodriguez, Loman, & Horner, 2009; Sterling-Turner, Watson, & Moore, 2002). The TIPS Workshop for PBIS Coaches ended with a brief

question-and-answer period, followed by an agreement from coaches to provide the technical assistance as their schedules and other duties permitted, and to provide the TIPS trainer with their completed TIPS Technical Assistance Data Collection Forms.

2.5. Implementation integrity of TIPS independent variables

The independent variables consisted of (a) the TIPS Workshop and (b) TIPS follow-up technical assistance. Delivery of the TIPS Workshop with implementation integrity was assured, given that the TIPS trainers who conducted the TIPS Workshop in North Carolina and Oregon presented the material from a standardized set of "scripts," presentation slides, and workshop activities contained in the TIPS training manual (Newton, Todd, Algozzine et al., 2009).

The implementation integrity with which PBIS Coaches delivered technical assistance to team members as they engaged in workshop activities was monitored by the TIPS trainers, who simply asked team members to raise their hands when they had completed a workshop activity. The coaches' provision of post-TIPS Workshop technical assistance to the PBIS Teams in the Immediate group was assessed by reviewing their completed TIPS Technical Assistance Data Collection Forms. For 16 of the 17 teams in the Immediate group, the team's coach returned a completed TIPS Technical Assistance Data Collection Form. (During the study, one of the coaches was assigned to another position within the school district. Although her TIPS-related responsibilities were assigned to another coach, a TIPS Technical Assistance Data Collection Form for one of that coach's teams was never completed.) On the 16 completed TIPS Technical Assistance Data Collection Forms we received, coaches reported that they contacted 15 of the 16 Facilitators (94%), 12 of 15 Data Analysts (80%), and 11 of 15 Minute Takers (73%) at least once prior to the team's first post-TIPS Workshop meeting. (One of the 16 teams was reported as neither having a preexisting Data Analyst nor having assigned someone to act in that capacity by the time the coach was ready to initiate contact with the Data Analyst. The same was true of the Minute Taker position for another of the 16 teams.)

The data also revealed that coaches asked 9 of the 15 contacted Facilitators (60%) whether they were prepared to assume the specific responsibilities of the position, and asked the same question of 8 of the 12 contacted Data Analysts (67%) and 8 of the 11 contacted Minute Takers (73%). Of those who were asked about their preparedness to assume the responsibilities of the position, the coaches subsequently provided related technical assistance to 6 of the 9 Facilitators (67%), 4 of the 8 Data Analysts (50%), and 3 of the 8 Minute Takers (38%). (Note, however, that we do not know whether some or all of the team members who did not receive technical assistance from the coaches simply declined the offer. We failed to construct the TIPS Technical Assistance Data Collection Form so as to take this possibility into account.)

Data from the 16 completed TIPS Technical Assistance Data Collection Forms revealed that coaches attended the first post-TIPS Workshop meeting for 15 of the 16 PBIS Teams in the Immediate group (94%). When asked to report how much prompting, in general, these teams required from the coach in order to implement components of the TIPS model during this meeting, the coaches reported that 3 teams required *none* (20%), 7 required *a little* (47%), 4 required *a moderate amount* (27%), 0 required *a lot* (0%), and 1 team was reported as not having implemented components of the TIPS model in spite of the coach's prompting (7%). When asked about their attendance at the 16 teams' second post-TIPS Workshop meeting, coaches provided data indicating that they attended 13 of the meetings (81%). When asked to report how much prompting the teams required, coaches reported that 3 teams required *none* (23%), 8 required *a little* (62%), 1 required *a moderate amount* (8%), 0 required *a lot* (0%), and 1 team was reported as not having implemented the components of the TIPS model in spite of the coach's prompting (8%). This team was the same one that was reported as not having implemented components of the TIPS model at its first post-TIPS Workshop meeting.

3. Results

Table 4 provides summarized data from the first two DORA observations conducted at meetings of all 17 teams in the Wait-List Control group and at meetings of all 17 teams in the Immediate group. There were no statistically significant differences ($p > .05$) between the baseline mean scores (O_1) for the teams in the Wait-List Control group and the baseline mean scores for the teams in the Immediate group for any

Table 4
DORA means, standard deviations, and ANCOVA results, by TIPS variable, for PBIS Teams in Wait-List Control group and Immediate group.

| DORA variable | Wait-List Control group ($N=17$) | | | | Immediate group ($N=17$) | | | | ANCOVA[a] | |
| | Baseline (O_1) | | Baseline (O_2) | | Baseline (O_1) | | Post-TIPS (O_2) | | | |
	M	SD	M	SD	M	SD	M	SD	$F(1,31)$	ES
Foundations	60.78	17.12	63.24	9.80	65.69	13.14	78.92	11.07	17.54[**]	1.46
Problem precision	36.37	32.78	30.00	35.18	44.04	33.31	72.31	34.81	12.35[**]	1.18
Use of quantitative data	31.37	46.35	29.41	46.97	46.47	45.38	72.55	44.46	7.09[*]	0.94
Solution	12.16	13.54	7.06	11.60	10.59	10.97	33.57	27.15	13.57[**]	1.24
Action plan	11.11	19.25	10.78	19.49	15.03	18.11	55.03	27.76	27.96[**]	1.80
Thoroughness of problem solving	22.75	19.99	19.31	21.57	29.03	21.43	58.37	27.38	19.78[**]	1.55

Note. ES = effect size, computed as Hedges' g, with correction factor applied (Immediate group mean − Wait-List Control group mean)/SD_{pooled}. There were no statistically significant differences between the baseline mean scores (O_1) for the two groups.
[a] ANCOVA using baseline values as covariate.
[*] $p<.05$.
[**] $p<.01$.

of the five individual DORA variables (i.e., Foundations, Problem Precision, Use of Quantitative Data, Solution, Action Plan) or for the overall measure of Thoroughness of Problem Solving. Thus, the randomization process succeeded in eliminating any initial differences between the groups on these variables. Likewise, there was not a statistically significant difference between the baseline mean for number of student problems identified during team meetings by members of teams in the Wait-List Control group ($M=1.18$, $SD=1.24$) and the baseline mean for number of student problems identified during team meetings by members of teams in the Immediate group ($M=1.71$, $SD=1.65$, $t(32)=-1.06$, $p=.298$).

Controlling for the initial baseline scores, the ANCOVA analyses presented in Table 4 revealed statistically significant differences ($p<.05$) between the second baseline mean DORA scores for the teams in the Wait-List Control group (O_2) and the first post-TIPS Workshop mean DORA scores for the teams in the Immediate group (O_2), across all five of the individual DORA variables, as well as for the overall measure of Thoroughness of Problem Solving. Teams in the Immediate group consistently received the higher mean DORA scores. For example, the teams in the Wait-List Control group achieved a mean DORA score of 30.00 on the Problem Precision variable at the time of their second baseline measure, indicating that they implemented an average of 30% of the five elements required to add precision to the definition of a broadly-identified problem (i.e., what, who, where, when, and why). In contrast, the teams in the Immediate group achieved a mean DORA score of 72.31 on the Problem Precision variable at the time of their first post-TIPS Workshop measure, indicating that they implemented an average of 72% of the five elements required to add precision to the definition of a broadly-identified problem. The practical significance of this difference was large (Cohen, 1988), $g=1.18$, with teams in the Immediate group achieving a mean DORA score on Problem Precision that was a little more than 1 standard deviation greater than the mean DORA score for the teams in the Wait-List Control group. Effect sizes for these data (and for the data displayed in all subsequent tables) were computed as Hedges' g, with the correction factor applied to reduce bias (Hedges, 1981, 1982; Kline, 2005). As noted in the final column of Table 4, effect sizes for all DORA variables were large (Cohen, 1988), ranging from 0.94 for Use of Quantitative Data to 1.80 for Action Plan.

The results show that within the context of the randomized wait-list controlled trial the PBIS team members in the Immediate group implemented a significantly greater percentage of the problem-solving processes after having participated in the TIPS Workshop and follow-up technical assistance than did the team members in the Wait-List Control group who had not yet received the TIPS intervention, a finding furthered strengthened by the large effect size found for each TIPS variable.

The third round of DORA data collection (O_3) for the Immediate group teams in North Carolina and in Oregon (see Table 3) allowed for an assessment of the extent to which the effects of the TIPS intervention demonstrated at O_2 were maintained at O_3. (The average time elapsed between the O_2 and O_3 observations for the Immediate group teams was 1.5 months [range, 0.50 to 2.33]). The Table 5 data show

Table 5
Post-TIPS and maintenance DORA means, standard deviations, mean contrasts, and effect sizes for PBIS Teams in the Immediate group (n = 17).

| | Occasion | | | | | |
| | Post-TIPS (O_2) | | Maintenance (O_3) | | | |
DORA variable	M_2	SD_2	M_3	SD_3	$M_2 - M_3$	ES
Foundations	78.92	11.07	78.43	15.88	0.49	0.04
Problem precision	72.31	34.81	58.35	38.42	13.96	0.37
Use of quantitative data	72.55	44.46	61.18	48.72	11.37	0.24
Solution	33.57	27.15	22.27	19.72	11.30	0.48
Action plan	55.03	27.76	46.41	42.49	8.62	0.24
Thoroughness of problem solving	58.37	27.38	47.05	32.19	11.32	0.37

Note. ES = effect size, computed as Hedges' g, with correction factor applied (O_2 Mean − O_3 Mean)/SD_{pooled}.

that the DORA mean scores for each of the five individual DORA variables and for the overall measure of Thoroughness of Problem Solving were lower at the time of the third observation, with mean contrasts showing decreases ranging from 0.49 for Foundations to 13.96 for Problem Precision. However, pairwise comparisons (using Bonferroni adjustment, corrected α = .01) indicated that these differences were not statistically significant, and that the 95% confidence interval for each of the TIPS variable differences ranged above and below 0.0. Similarly, the effect size data (range, 0.04 to 0.48) show that the decreases for all of the variables were small (Cohen, 1988), with the exception of the Solution variable, which rounded to a medium effect size.

Table 6 provides data concerning the small-scale assessment of the extent to which effects of the TIPS intervention on the performance of teams in the Immediate group were replicated with the six teams in Oregon's Wait-List Control group. Statistically significant differences ($p < .05$) between baseline O_2 means and post-TIPS O_3 means were found for all of the individual DORA variables with the exception of Use of Quantitative Data,[1] and for the overall measure of Thoroughness of Problem Solving, with the higher mean scores following the TIPS intervention. Effect sizes were large, ranging from 0.84 to 4.75. Follow-up pairwise comparisons (using Bonferroni adjustment, corrected α = .01) indicated that baseline to baseline (O_2–O_1) contrasts were not statistically significantly different, and that baseline to post-TIPS (O_3–O_2) contrasts were statistically significantly different.

4. Discussion

To help PBIS Teams become more efficient and effective at using data for problem solving, we developed the Team-Initiated Problem Solving (TIPS) model. This randomized controlled trial was conducted to determine whether PBIS team members who participate in a TIPS Workshop and follow-up technical assistance, demonstrate greater implementation integrity in the use of problem-solving processes during team meetings than do PBIS team members who receive the TIPS Workshop and follow-up technical assistance at a later point in time. The results indicate that the PBIS Teams in the Immediate group achieved statistically-significantly different (and higher) mean DORA scores on all of the individual DORA variables (i.e., Foundations, Problem Precision, Use of Quantitative Data, Solution, and Action Plan) and on the overall Thoroughness of Problem Solving score than did the PBIS team members in the Wait-List Control group. That is, team members in the Immediate group demonstrated greater implementation integrity. Further, the effect sizes derived from the mean contrasts for each individual TIPS variable, and for overall Thoroughness of Problem Solving, were large. Maintenance of these effects was encouraging. Although the mean DORA scores for the teams in the Immediate group were lower at the time of the Maintenance observation, these differences were not statistically significant and, with one exception, the reduced effect sizes were small. Finally, Oregon teams in the Wait-List Control group

---

[1] Contrasts for this variable were similar to those for other variables but statistical comparisons were restricted by zero variance in post-TIPS (O_3) data.

Table 6
DORA means, standard deviations, and ANOVA results for PBIS Teams in the Oregon Wait-List Control group ($n = 6$).

DORA variable	Occasion							
	Baseline (O_1)		Baseline (O_2)		Post-TIPS (O_3)		ANOVA	
	M	SD	M	SD	M	SD	$F(2,10)$	ES
Foundations	62.50	20.24	65.28	8.19	93.06	6.27	10.00**	3.52
Problem precision	43.06	24.32	53.33	32.04	96.67	8.16	9.68**	1.71
Use of quantitative data	55.56	50.18	66.67	51.64	100.00	0.00	2.65	0.84
Solution	16.11	10.84	8.33	13.29	46.67	21.60	9.85**	1.97
Action plan	9.26	12.99	8.33	13.94	88.89	17.21	94.72**	4.75
Thoroughness of problem solving	31.00	16.12	34.17	17.87	83.06	8.41	29.69**	3.23

Note. ES = effect size, computed as Hedges' *g*, with correction factor applied (O_3 Mean $- O_2$ Mean)$/SD_{pooled}$.
** $p < .01$.

replicated the effects by achieving improved implementation integrity of the problem-solving processes following their participation in the TIPS Workshop and follow-up technical assistance.

Some similarities and differences between our methodology and results and the methodology and results of previously-cited studies are worth noting. Whereas our study adopted an experimental approach, seeking to improve the implementation integrity of team problem-solving processes, most previous research in this area has been descriptive, simply assessing the status of a team's implementation integrity at a single point in time (Doll et al., 2005; Flugum & Reschly, 1994; McDougal et al., 2000; Telzrow et al., 2000). Like McDougal et al. (2000) we used observers to gather real-time, direct observation data concerning the problem-solving processes engaged in by team members during their regular team meetings. This stands in contrast to the more common methodologies of gathering implementation integrity data via (a) team members' written responses to a questionnaire or a checklist (Doll et al., 2005; Flugum & Reschly, 1994); or (b) best case documentation, chosen and provided by team members (Doll et al., 2005; Telzrow et al., 2000). Results from the cited descriptive studies showed much room for improvement in the implementation integrity achieved by teams, and examination of the baseline data in our study showed that this was also true of the participating PBIS Teams.

Prior to becoming participants in our study, PBIS Teams had received training in the Tier I (primary, or universal, problem behavior prevention and intervention) level of School-wide Positive Behavior Interventions and Supports, as well as training in the use of the School-wide Information System, but the training did not place those content areas within a problem-solving context. Prior to participating in the descriptive studies, teams had received varied training. Researchers reported that (a) teams had received training provided by multiple regional coordinators as part of a state-wide initiative, but with training varying in content due to the absence of any explicit state-sanctioned training model (Telzrow et al., 2000); (b) teams had received individual training sessions at various points in time, provided both by leaders of a school district and by national experts in problem-solving consultation, but with no mention of the extent to which coordination of content across these training sessions may have been attempted (Doll et al., 2005); (c) teams received training in the use of a specific model (the School-Based Intervention Team model) provided by members of a SBIT planning team responsible for development and diffusion of the project (McDougal et al., 2000); or (d) the majority of participating teams "had not received any training in quality interventions" (Flugum & Reschly, 1994, p. 13), with implementation integrity data from the study providing baseline information for an anticipated statewide staff development and training effort.

The achievement of implementation integrity of problem-solving processes by teams is an important first step in an associated logic model. First, it is through implementation of the problem-solving processes that team members arrive at solutions for resolving targeted problems. And second, it is assumed that solutions developed via the problem-solving processes, when also used with implementation integrity, will resolve the targeted problems. Outside of the context of school-based problem-solving teams, this assumption about the positive relationship between the level of treatment integrity and improved outcomes has some support from research. Descriptive research has found that the level of treatment integrity is an important determiner of core curriculum outcomes for students (O'Donnell, 2008) and of

prevention and health outcomes such as level of physical health and development, academic performance, and drug use for children and adolescents (Durlak & DuPre, 2008). There are also experimental studies in which researchers have systematically manipulated treatment integrity and demonstrated a functional relation between higher levels of treatment integrity and improvement in mathematics performance (Noell, Gresham, & Gansle, 2002) and in child compliance (Wilder, Atwell, & Wine, 2006). However, it should be noted that in a chapter reviewing research on the relationships among the consultation process, treatment integrity, and outcomes, Noell (2008) pointed out that although degradation in treatment integrity has generally been associated with poorer outcomes, there has also been research in which modest decreases in treatment integrity had no significant effect on some outcomes or with some participants (e.g., Gansle & McMahon, 1997; Holcombe, Wolery, & Snyder, 1994; Vollmer, Roane, Ringdahl, & Marcus, 1999).

4.1. Implications for training and technical assistance

Successfully establishing the first step in the logic model is likely to require providing teams with both training and follow-up technical assistance. Our training and research efforts were tightly integrated in the sense that we developed (a) the TIPS Workshop; (b) the follow-up technical assistance provided by a team's PBIS Coach; and (c) the DORA instrument that observers used to record data concerning the extent to which team members succeeded in both establishing foundational aspects for conducting an efficient and effective meeting, and implementing the problem-solving processes exemplified by the TIPS model during their meetings. Collectively these three elements were designed for the explicit purpose of improving (and verifying) the implementation integrity of problem-solving processes exhibited by PBIS Teams. Integration of training, technical assistance, and data collection was not apparent in the descriptive studies.

Ensuring that training is followed by technical assistance is important in increasing the likelihood that participants will transfer problem-solving skills acquired during training sessions to the settings in which the skills are to be applied (Joyce & Showers, 2002). Two experimental studies have directly focused on the effect of providing technical assistance to problem-solving teams, although the technical assistance was not linked to prior training. Bartels and Mortenson (2005) used an observer to gather data concerning the extent to which problem-solving teams at three middle schools employing the Instructional Consultation model (Rosenfield & Gravois, 1996) engaged in 15 problem-solving processes (steps) during team meetings. (The authors did not provide information concerning what training team members may have received in use of the Instructional Consultation model prior to the study.) During a baseline condition, two of the three teams demonstrated low percentages of implementation integrity. This was followed by (a) a performance feedback condition during which each team's school psychologist was shown a graph of the percentage of steps previously implemented by the team, prior to the team's next meeting; and (b) a subsequent checklist condition, during which each team member was provided with a copy of the 15-step checklist prior to each meeting. Neither experimental condition produced substantial, sustained improvement in implementation integrity of problem-solving processes across team meetings, although the checklist condition produced the better results. Providing team members with a checklist of problem-solving processes appears similar to our providing team members with the TIPS Meeting Minutes and Problem-Solving Action Plan form, which listed components of the TIPS model in a format designed both to "prompt" teams through the problem-solving processes during their meetings and to provide team members with a form that resulted in a written record of the team meeting (e.g., problems identified, solutions selected, persons responsible for implementation of solutions, etc.).

In the second experimental study, Burns et al. (2008) also examined the effect of performance feedback on the implementation integrity of problem-solving processes achieved by teams. Participants were problem-solving teams from three elementary schools. (Problem-solving training for teams in the school district had occurred 10 years before the study was conducted, but staff turnover had resulted in an inconsistent and unknown number of trained team members in each school, and team members at the three participating schools experienced only brief refresher information at the beginning of the school year that consisted of a review of the problem-solving team process and a reminder that a problem-solving team manual and forms were available from the school district.) An observer used a 20-item checklist of problem-solving processes to gather data during baseline and performance feedback

conditions in the context of a multiple-baseline design across the three school teams. Each team demonstrated a low percentage of implementation integrity for the checklist items (measuring problem-solving processes) during baseline. During the performance feedback condition, a researcher (a) presented each team member with a graph showing the checklist items implemented by team members in previous meetings, and (b) discussed individual checklist items to reinforce correct implementation and to highlight items that were not implemented. This brief feedback session was conducted just before the beginning of each team meeting. The performance feedback condition resulted in an immediate and improved change in level in the percentage of problem-solving processes implemented. However, in noting the need for additional research in the area of implementation integrity, the authors stated that during the performance feedback condition team members still did not implement the problem-solving processes of monitoring student progress, assessing the effectiveness of the intervention, or measuring treatment integrity of the intervention

In summary, research suggests that each of the following may be important in the effort to improve the achievement of implementation integrity of problem-solving processes by teams: (a) training specifically focused on both establishing foundational elements for conducting an efficient and efficient team meeting and implementing problem-solving processes, particularly when these elements are embedded in an explicit problem-solving model such as TIPS; (b) follow-up technical assistance that complements the training content and allows team members to practice workshop skills in the context of their regular meetings, with help as needed, such as that provided by the PBIS Coaches; and (c) materials that can serve as prompts for guiding team members through the problem-solving processes during meetings, such as the TIPS Meeting Minutes and Problem-Solving Action Plan form. Further, it will be important for an observer to collect implementation integrity data in the context of subsequent team meetings to verify team members' use of the problem-solving processes, such as we did when using the DORA instrument; or as an aid to performance feedback in which team members are presented with a graph showing problem-solving processes completed and omitted during meetings (e.g., Burns et al., 2008).

4.2. Limitations

4.2.1. External validity

The study's results are tempered by some limitations, including the issue of external validity. Participants were members of elementary school teams that were implementing School-wide Positive Behavior Interventions and Supports (SWPBIS) and using the School-wide Information System (SWIS) to collect data and produce reports about student office discipline referrals. The schools were recruited by PBIS Coaches from a sample of convenience; the first 22 North Carolina schools and the first 12 Oregon schools that expressed interest were selected for participation. Given the selection criteria, the fact that the teams were not randomly selected, and the relatively small number of teams, the question of generalizability of the findings is a legitimate one. The study's results suggest the appropriateness of future research examining the effects of the TIPS intervention when implemented with (a) a larger and randomly selected sample of PBIS Teams, including PBIS Teams in middle schools and high schools, and (b) problem-solving teams in schools that use behavior support systems and data management systems other than SWPBIS and SWIS.

4.2.2. Interobserver agreement

Although the average interobserver agreement for all of the DORA variables exceeded 85%, some of the lower-range values suggest a need to improve the procedures whereby observers are trained. One improvement in training would be to revise the DORA Data Collection Workshop so that observers use DORA to record data (a) when reviewing a greater number of written training vignettes that sample the range of problem-solving processes, and (b) when viewing one or more videotapes of actual PBIS team meetings. A second improvement in training would involve requiring observers to use DORA to record data at an additional number of actual PBIS team meetings, accompanied by a trainer who also records data, while continuing to adhere to the criterion of a resulting interobserver agreement of 85% or better before the observer is allowed to collect research data.

4.2.3. Additional issues

DORA variable scores achieved by the teams in the Immediate group were lower at the time of the Maintenance observation that was conducted, on average, 1.5 months following the previous observation (and an average of 2.93 months following the end of TIPS training). Although the difference in scores was not statistically significant for any of the DORA variables, there is the possibility that such differences could have clinical significance. This finding suggests that, when feasible, it may be advisable for PBIS Coaches either to provide follow-up technical assistance at team meetings until a team achieves a high performance criterion or to deliver "booster" technical assistance and feedback sessions beyond the two coached meetings of the current training protocol.

Two final limitations should also be considered. First, because our decision to assess the extent to which the improvement in implementation integrity achieved by the teams in the Immediate group would also be achieved by teams in the Wait-List Control group was made after the project had begun, this replication effort focused solely on the 6 Oregon Wait-List Control group teams. We excluded the 11 North Carolina teams due to limited flexibility and resources. In the future, a more inclusive assessment of replication of TIPS effects should be considered during the initial development of the research design. Second, concerning the offer of PBIS Coaches to provide technical assistance to a team's Facilitator, Minute Taker, and Data Analyst prior to the team's first post-TIPS Workshop meeting, we do not know whether some or all of the team members who did not receive this technical assistance simply declined the offer. This limitation can be remedied by revising the form to ensure that we learn (a) whether technical assistance was offered; (b) whether the offer of technical assistance was accepted or rejected; and, in cases where the offer of technical assistance was rejected, (c) the reason the offer was rejected (e.g., the team's Facilitator felt competent to execute TIPS-related responsibilities during meetings without further assistance).

4.3. Conclusion

The present results indicate that (a) prior to TIPS training, the PBIS Teams were not demonstrating high levels of implementation integrity in the use of highly recommended processes for data-based problem solving, as exemplified by the TIPS model, during their meetings; (b) after TIPS training, the teams demonstrated improved implementation integrity of the problem-solving processes; and (c) the teams maintained better-than-baseline implementation integrity, although there was some decrement in performance.

The results encourage future research on implementation integrity of problem-solving processes in several areas. The external validity of this study's findings should be tested by conducting studies that (a) involve a larger and randomly selected sample of PBIS Teams, (b) PBIS Teams in middle schools and high schools, and (c) problem-solving teams in schools that use behavior support systems and data management systems other than SWPBIS and SWIS. Another important and logical next step would involve extending or adapting aspects of the TIPS model to more explicitly focus on academic problems experienced by students, given that TIPS is an operationalization of problem-solving processes that are remarkably consistent, even when conceptualized and implemented in varying contexts (e.g., Bergan & Kratochwill, 1990; Bransford & Stein, 1984; Burns, Deno, & Jimerson, 2007; Carroll & Johnson, 1990; Huber, 1980; Moody, 1983; Tilly, 2008). It will also be important to obtain a better understanding of variables that improve maintenance of gains in implementation integrity, perhaps achieved via a tight link between training and subsequent use of highly structured performance feedback of the type employed by Burns et al. (2008).

Of prime importance, of course, is the need for research that extends the problem-solving logic model, by assessing not only the extent to which a team demonstrates implementation integrity in its use of problem-solving processes, but also the extent to which (a) the solutions derived from the problem-solving processes are used with treatment integrity, and (b) the subsequent rate of a targeted problem behavior demonstrates resolution or reduction of the problem. Demonstration of such functional relations would express the ultimate merit of the use of data-based problem-solving processes.

Acknowledgements

of the Wake County North Carolina Public School System and the Clackamas, Douglas, and Linn–Benton–Lincoln Education Service Districts of Oregon, in making this research possible.

References

Algozzine, B., Newton, J. S., Horner, R. H., Todd, A. W., & Algozzine, K. (in press). Development and technical characteristics of a team decision-making assessment tool: Decision observation, recording, and analysis (DORA). Journal of Psychoeducational Assessment.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*, 91–97.

Bartels, S. M., & Mortenson, B. P. (2005). Enhancing adherence to a problem-solving model for middle-school pre-referral teams: A performance feedback and checklist approach. *Journal of Applied School Psychology*, *22*, 109–123.

Batsche, G. M., Curtis, M. J., Dorman, C., Castillo, J. M., & Porter, L. J. (2007). The Florida problem-solving/response to intervention model: Implementing a statewide initiative. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 378–395). New York, NY: Springer Science.

Bergan, J. R., & Kratochwill, T. R. (1990). *Behavioral consultation and therapy.* New York, NY: Plenum.

Billingsley, F., White, O. R., & Munson, R. (1980). Procedural reliability: A rationale and an example. *Behavioral Assessment*, *2*, 229–241.

Boudett, K. P., City, E. A., & Murnane, R. J. (Eds.). (2005). *Data wise: A step-by-step guide to using assessment results to improve teaching and learning.* Cambridge, MA: Harvard Education Press.

Bransford, J. D., & Stein, B. S. (1984). *The IDEAL problem solver: A guide for improving thinking, learning, and creativity.* New York, NY: W. H. Freeman and Company.

Brown-Chidsey, R., & Steege, M. W. (2005). *Response to intervention: Principles and strategies for effective practice.* New York, NY: Guilford Press.

Buck, G. H., Polloway, E. A., Smith-Thomas, A., & Cook, K. W. (2003). Prereferral intervention processes: A survey of state practices. *Exceptional Children*, *69*, 349–360.

Burns, M. K., Deno, S. L., & Jimerson, S. R. (2007). Toward a unified response-to-intervention model. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 428–440). New York, NY: Springer.

Burns, M. K., Peters, R., & Noell, G. H. (2008). Using performance feedback to enhance implementation fidelity of the problem-solving team process. *Journal of School Psychology*, *46*, 537–550.

Burns, M. K., & Symington, T. (2002). A meta-analysis of prereferral intervention teams: Student and systemic outcomes. *Journal of School Psychology*, *40*, 437–447.

Burns, M. K., Vanderwood, M. L., & Ruby, S. (2005). Evaluating the readiness of pre-referral intervention teams for use in a problem solving model. *School Psychology Quarterly*, *20*, 89–105.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally.

Carroll, J. S., & Johnson, E. J. (1990). *Decision research: A field guide.* Newbury Park, CA: Sage.

Carter, J., & Sugai, G. (1989). Survey on prereferral practices: Responses from state departments of education. *Exceptional Children*, *55*, 298–302.

Chalfant, J. C., Pysh, M. V., & Moultrie, R. (1979). Teacher assistance teams: A model for within-building problem solving. *Learning Disability Quarterly*, *2*, 85–96.

Coalition for Evidence-Based Policy (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user-friendly guide.* Washington, DC: Author.

Cochrane, W. S., & Laux, J. M. (2008). A survey investigating school psychologists' measurement of treatment integrity in school-based interventions and their beliefs about its importance. *Psychology in the Schools*, *45*, 499–507.

Codding, R. S., Feinberg, A. B., Dunn, E. K., & Pace, G. M. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis*, *38*, 205–219.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, R., Kincaid, D., & Childs, K. E. (2007). Measuring school-wide positive behavior support implementation: Development and validation of the benchmarks of quality. *Journal of Positive Behavior Interventions*, *9*, 203–213.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson.

Cowan, K., & Cohn, A. (2009). School psychology awareness: A positive approach. *Communiqué*, *38*, 32–33.

D'Zurilla, T. J., & Nezu, A. M. (1990). Development and preliminary evaluation of the social problem-solving inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*, 156–163.

D'Zurilla, T. J., Nezu, A. M., & Maydeu-Olivares, A. (2002). *Social problem-solving inventory—revised (SPSI—R): Technical manual.* North Tonawanda, NY: Multi-Health Systems.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*, 219–232.

Deno, S. L. (2005). Problem-solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10–40). New York, NY: Guilford Press.

DiGennaro, F. D., & Martens, B. K. (2007). A comparison of performance feedback procedures on teachers' treatment implementation integrity and students' inappropriate behavior in special education classrooms. *Journal of Applied Behavior Analysis*, *40*, 447–461.

Doll, B., Haack, K., Kosse, S., Osterloh, M., & Siemers, E. (2005). The dilemma of pragmatics: Why schools don't use quality team consultation practices. *Journal of Educational and Psychological Consultation*, *16*, 127–155.

Duhon, G. J., Mesmer, E. M., Gregerson, L., & Witt, J. C. (2009). Effects of public feedback during RTI team meetings on teacher implementation integrity and student academic performance. *Journal of School Psychology*, *47*, 19–37.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*, 327–350.

Flugum, K. R., & Reschly, D. J. (1994). Prereferral interventions: Quality indices and outcomes. *Journal of School Psychology*, *32*, 1–14.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice*, *18*, 157–171.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Boston, MA: Allyn & Bacon, Inc..

Gansle, K. A., & McMahon, C. M. (1997). Component integrity of teacher intervention management behavior using a student self-monitoring treatment: An experimental analysis. *Journal of Behavioral Education*, *7*, 405–419.

Gelzheiser, L. M. (2009). Preparing for the future of school psychology: A special educator's view. *Journal of Educational and Psychological Consultation*, *19*, 259–266.

Gilbert, T. F. (1978). *Human competence: Engineering worthy performance.* New York, NY: McGraw-Hill.

Graden, J. L., Casey, A., & Christenson, S. L. (1985). Implementing a prereferral intervention system: I. The model. *Exceptional Children*, *51*, 377–384.

Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis*, *26*, 257–263.

Gresham, F. M., Gansle, K. A., Noell, G. H., & Cohen, S. (1993). Treatment integrity of school-based behavioral intervention studies: 1980–1990. *School Psychology Review*, *22*, 254–272.

Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E., Supovitz, J. A., & Wayman, J. C. (2009). *Using student achievement data to support instructional decision making (NCEE 2009–4067).* Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/practiceguide.aspx?sid=12

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107–128.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, *92*, 490–499.

Hersen, M., & Barlow, D. H. (1976). *Single-case experimental designs: Strategies for studying behavior change.* Elmsford, NY: Pergamon Press.

Holcombe, A., Wolery, M., & Snyder, E. (1994). Effects of two levels of procedural fidelity with constant time delay on children's learning. *Journal of Behavioral Education*, *4*, 49–73.

Horner, R. H., Sugai, G., & Anderson, C. M. (2010). Examining the evidence base for school-wide positive behavior support. *Focus on Exceptional Children*, *42*, 1–14.

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., et al. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*, *11*, 133–144.

Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B. (2004). The school-wide evaluation tool (SET): A research instrument for assessing school-wide positive behavior supports. *Journal of Positive Behavior Interventions*, *6*, 3–12.

Horner, R., Todd, A., Sampson, N., Dickey, C., May, S., Amedo, M., et al. (2008). *School-wide information system facilitator training manual.* Eugene, Oregon: Educational and Community Supports, University of Oregon.

Huber, G. P. (1980). *Managerial decision making.* Glenview, IL: Scott, Foresman.

Individuals with Disabilities Education Improvement Act (2004). *H.R. 1350, 108th Congress.*

Irvin, L. K., Horner, R. H., Ingram, K., Todd, A. W., Sugai, G., Sampson, N. K., et al. (2006). Using office discipline referral data for decision making about student behavior in elementary and middle schools: An empirical evaluation of validity. *Journal of Positive Behavior Interventions*, *8*, 10–23.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Kline, R. B. (2005). *Beyond significance testing: Reforming data analysis methods in behavioral research.* Washington, DC: American Psychological Association.

Kovaleski, J. F., & Glew, M. C. (2006). Bringing instructional support teams to scale: Implications of the Pennsylvania experience. *Remedial and Special Education*, *27*, 16–25.

Lewis, T. J., Jones, S. E. L., Horner, R. H., & Sugai, G. (2010). School-wide positive behavior support and students with emotional/behavioral disorders: Implications for prevention, identification, and intervention. *Exceptionality*, *18*, 82–93.

Liaupsin, C. J., Scott, T. M., & Nelson, C. M. (2006). *Functional behavioral assessment and intervention planning: A simplified team process.* [CD-ROM]. Sopris West: Frederick, Co.

Matthews, J. N. S. (2006). *Introduction to randomized controlled clinical trials* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

May, S., Arid, W., Jr., Todd, A. W., Horner, R. H., Glasgow, A., & Sugai, G. (2003). *Schoolwide information system.* Eugene, Oregon: Educational and Community Supports, University of Oregon.

McDougal, J. L., Clonan, S. M., & Martens, B. K. (2000). Using organizational change procedures to promote the acceptability of prereferral intervention services: The school-based intervention team project. *School Psychology Quarterly*, *15*, 149–171.

McIntyre, L. L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in the Journal of Applied Behavior Analysis 1991–2005. *Journal of Applied Behavior Analysis*, *40*, 659–672.

Moody, P. E. (1983). *Decision making: Proven methods for better decisions.* New York, NY: McGraw-Hill.

Moore, J. W., Edwards, R. P., Sterling-Turner, H. E., Riley, J., DuBard, M., & McGeorge, A. (2002). Teacher acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis*, *35*, 73–77.

Myers, D., & Dynarski, M. (2003). *Random assignment in program evaluation and intervention research: Questions and answers.* Washington, DC: Institute of Education Sciences.

National Association of School Psychologists (2010). *Model for comprehensive and integrated school psychological services.* Author: Washington, DC. Retrieved from http://www.nasponline.org/standards/2010standards/2_PracticeModel.pdf

Newton, J. S., Algozzine, B., Algozzine, K., Horner, R. H., & Todd, A. W. (2011). Building local capacity for training and coaching data-based problem solving with positive behavior intervention and support teams. *Journal of Applied School Psychology*, *27*, 228–245.

Newton, J. S., Horner, R. H., Algozzine, R. F., Todd, A. W., & Algozzine, K. M. (2009). Using a problem-solving model to enhance data-based decision making in schools. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of positive behavior support* (pp. 551–580). New York, NY: Springer.

Newton, J. S., Horner, R. H., Todd, A. W., Algozzine, B., & Algozzine, K. M. (2012). A pilot study of a problem-solving model for team decision making. *Education and Treatment of Children*, *35*, 25–49.

Newton, J. S., Todd, A. W., Algozzine, K. M., Horner, R. H., & Algozzine, B. (2009). *Team-initiated problem solving training manual.* Eugene, Oregon: Educational and Community Supports, University of Oregon.

Newton, J. S., Todd, A. W., Horner, R. H., Algozzine, B., & Algozzine, K. (2009). *Decision observation, recording, and analysis: DORA.* Unpublished instrument. Eugene, Oregon: Educational and Community Supports, University of Oregon.

Noell, G. H. (2008). Research examining the relationships among consultation process, treatment integrity, and outcomes. In W. P. Erchul, & S. M. Sheridan (Eds.), *Handbook of research in school consultation* (pp. 323–341). New York, NY: Erlbaum.

Noell, G. H., Gresham, F. M., & Gansle, K. A. (2002). Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance. *Journal of Behavioral Education*, *11*, 51–67.

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

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, *78*, 33–84.

Page, T. J., & Iwata, B. A. (1986). Interobserver agreement: History, theory, and current methods. In A. Poling, & R. W. Fuqua (Eds.), *Research methods in applied behavior analysis: Issues and advances* (pp. 99–126). New York, NY: Plenum Press.

Peterson, L., Homer, A. L., & Wonderlich, S. A. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis*, *15*, 477–492.

Rankin, J. L., & Aksamit, D. L. (1994). Perceptions of elementary, junior high, and high school student assistance team coordinators, team members, and teachers. *Journal of Educational and Psychological Consultation*, *5*, 229–256.

Rodriguez, B. J., Loman, S. L., & Horner, R. H. (2009). A preliminary analysis of the effects of coaching feedback on teacher implementation integrity of First Steps to Success. *Behavior Analysis in Practice*, *2*, 11–21.

Rosenfield, S. A., & Gravois, T. A. (1996). *Instructional consultation teams: Collaborating for change.* New York, NY: Guilford Press.

Sanetti, L. M. H., Gritter, K. L., & Dobey, L. M. (2011). Treatment integrity of interventions with children in the school psychology literature from 1995 to 2008. *School Psychology Review*, *40*, 72–84.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin Company.

Shriberg, D. (2007). The school psychologist as leader and change agent in a high-stakes era. *Journal of Applied School Psychology*, *23*, 151–166.

Sterling-Turner, H. E., Watson, T. S., & Moore, J. W. (2002). The effects of direct training and treatment integrity on treatment outcomes in school consultation. *School Psychology Quarterly*, *17*, 47–77.

Sugai, G., & Horner, R. H. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, *35*, 245–259.

Sugai, G., & Horner, R. H. (2009). Defining and describing schoolwide positive behavior support. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of positive behavior support* (pp. 307–326). New York, NY: Springer.

Sugai, G., Horner, R. H., & Lewis-Palmer, T. (2009). *The team implementation checklist (v. 3.0).* Unpublished instrument. Eugene, Oregon: Educational and Community Supports, University of Oregon.

Telzrow, C. F., McNamara, K., & Hollinger, C. L. (2000). Fidelity of problem-solving implementation and relationship to student performance. *School Psychology Review*, *29*, 443–461.

Tilly, W. D., III (2008). The evolution of school psychology to a science-based practice: Problem solving and the three-tiered model. In J. Grimes, & A. Thomas (Eds.), *Best practices in school psychology V* (pp. 17–36). Bethesda, MD: National Association of School Psychologists.

Todd, A. W., Horner, R. H., Newton, J. S., Algozzine, R. F., Algozzine, K. M., & Frank, J. L. (2011). Effects of team-initiated problem solving on decision making by schoolwide behavior support teams. *Journal of Applied School Psychology*, *27*, 42–59.

Todd, A. W., Newton, J. S., Horner, R. H., Algozzine, B., & Algozzine, K. M. (2009). *Decision observation, recording, and analysis (DORA) training manual.* Educational and Community Supports. Eugene, Oregon: University of Oregon.

Torgerson, D. J., & Torgerson, C. J. (2008). *Designing randomised trials in health, education, and the social sciences: An introduction.* New York, NY: Palmgrave Macmillan.

Vollmer, T. R., Roane, H. S., Ringdahl, J. E., & Marcus, B. A. (1999). Evaluating treatment challenges with differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis*, *32*, 9–23.

Wilder, D. A., Atwell, J., & Wine, B. (2006). The effects of varying levels of treatment integrity on child compliance during treatment with a three-step prompting procedure. *Journal of Applied Behavior Analysis*, *39*, 369–373.

Ysseldyke, J., & Algozzine, B. (2006). *Effective assessment for students with special needs: A practical guide for every teacher.* Thousand Oaks, CA: Corwin.

Ysseldyke, J., Burns, M. K., & Rosenfield, S. (2009). Blueprints on the future of training and practice in school psychology: What do they say about educational and psychological consultation? *Journal of Educational and Psychological Consultation*, *19*, 177–196.