Title

Are School-Based Problem-Solving Teams Effective? A Meta-Analysis of Student- and Systems-Level Effects

Journal

Journal of Educational and Psychological Consultation, ahead-of-print(ahead-of-print)

ISSN

1047-4412

Authors

Sims, Wesley A
King, Kathleen R
Preast, June L
et al.

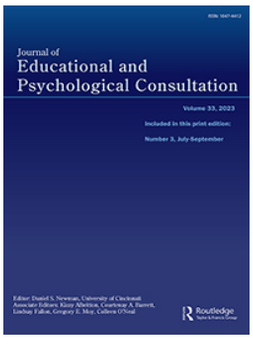
Publication Date

2023

DOI

10.1080/10474412.2023.2232785

Peer reviewed



Routledge
Taylor & Francis Group

Journal of Educational and Psychological Consultation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hepc20

Are School-Based Problem-Solving Teams Effective? A Meta-Analysis of Student- and Systems-Level Effects

Wesley A. Sims, Kathleen R. King, June L. Preast, Matthew K. Burns & Sarah Panameño

To cite this article: Wesley A. Sims, Kathleen R. King, June L. Preast, Matthew K. Burns & Sarah Panameño (2023): Are School-Based Problem-Solving Teams Effective? A Meta-Analysis of Student- and Systems-Level Effects, Journal of Educational and Psychological Consultation, DOI: 10.1080/10474412.2023.2232785

To link to this article: https://doi.org/10.1080/10474412.2023.2232785

Published online: 10 Jul 2023.

JOURNAL OF EDUCATIONAL AND PSYCHOLOGICAL CONSULTATION
https://doi.org/10.1080/10474412.2023.2232785

Are School-Based Problem-Solving Teams Effective? A Meta-Analysis of Student- and Systems-Level Effects

Wesley A. Sims[a], Kathleen R. King[b], June L. Preast[c], Matthew K. Burns[d], and Sarah Panameño[a]

[a]University of California; [b]Palm Spring Unified School District; [c]University of Alabama; [d]University of Missouri

ABSTRACT

Given the widespread use of school-based problem-solving teams (SB PSTs), empirical evidence establishing their effectiveness is essential. Guided by the *Input-Mediator-Output-Input* (IMOI) Framework, this meta-analysis examined SB PST effectiveness generally and differentiated measured effects by targeted outcomes and specific SB PST processes used. Data from 13 studies (14 effects) yielded a large overall weighted estimate of effect ($g = 0.84$). Measured SB PST effects on student outcomes (e.g. academic or behavioral improvement) appeared large ($g = 0.89$), on systems outcomes (i.e., reduction in referrals to special education) were moderate ($g = 0.66$), and on team outcomes (i.e., procedural fidelity) appeared large ($g = 1.00$). Although not a significant moderator of effect, team use of an evidence-based process was associated with larger effects ($g = 1.78$, $k = 8$). Similarly, although non-significant, studies that met rigorous, published standards for research methods showed larger outcome effects ($g = 1.10$, $k = 2$).

ARTICLE HISTORY
Received 15 March 2023
Revised 4 June 2023
Accepted 30 June 2023

School-based problem-solving teams (SB PSTs), across their varied names and targeted outcomes, have played a prominent role in school-based service delivery across the United States for several decades (Rosenfield et al., 2018). The emphasis on prevention and early intervention efforts in schools has resulted in an increased use of SB PSTs (Welch et al., 1999). A national survey found that 86% of states either required or recommended the use of SB PSTs prior to the use of special education services (Truscott et al., 2005). SB PSTs are recognized as a group of at least three educators working collaboratively to drive problem remediation efforts that are not subsumed by special education service provision (Reinke et al., 2018; Rosenfield et al., 2018; Sims et al., 2022). This definition serves to distinguish SB PSTs from other collaborative efforts to address student difficulties (e.g., Behavioral Consultation, multidisciplinary individualized education plan [IEP] teams), which are qualitatively different across composition, students served, and outcomes. For example, consultation

CONTACT Wesley A. Sims wesleys@ucr.edu Graduate School of Education, University of California Riverside, 900 University Ave 1207 Sproul Hall, Riverside, CA, 92521 USA

is often a dyadic relationship rather than a team, and IEP teams work collaboratively to support students within the procedural safeguards afforded students that are or are suspected of being eligible for special education services. Additionally, in some consultation models (i.e., Conjoint Behavioral Consultation [CBC]; Sheridan et al., 2013), participation by parents or legal guardians is expressly required, while parental involvement is encouraged in SB PST activities but not essential. Operationally distinguishing SB PSTs from other collaborative or team-based support mechanisms in schools is essential to evaluating their effectiveness or utility and to avoiding overgeneralization of findings inappropriately.

Empirical research, technical reports, and best practice recommendations portray SB PSTs as effective at the student and system levels. For example, previous meta-analytic SB PST research found a large overall effect ($d = 1.10$), with larger effects noted for student outcomes (e.g., time on-task, $d = 1.15$) than for system outcomes (e.g., referrals to special education, $d = 0.90$; Burns & Symington, 2002). More recent SB PST research that examined fidelity to the problem-solving process (Burns et al., 2008; Newton et al., 2012; Todd et al., 2011) or team member attitudes toward the process (Rosenfield et al., 2014) led to similarly favorable conclusions. However, relative to their widespread use for decades, SB PST effectiveness research appears mixed and limited (Rosenfield et al., 2018). Researchers note variability across SB PST purpose, processes, composition, and targeted outcomes, as well as research oversight as factors confounding evaluations of SB PST effectiveness (Burns & Symington, 2002; Rosenfield et al., 2018). In response, some scholars have called for the application of lessons learned from the broader team science literature, to use "an evidence-based, team science framework to aid in the organization and interpretation of research on PSTs" (Rosenfield et al., 2018, p. 409).

Input-Mediator-Output-Input (IMOI) framework

SB PST literature has largely ignored the influence of distinct factors such as school or team climate, team member roles and responsibilities, or specific team processes used on outcomes (Rosenfield et al., 2018; Santangelo, 2009). In contrast, the Input-Mediator-Output-Input (IMOI) model is a widely accepted, nuanced, and robust framework used to conceptualize and evaluate teaming in broader team science research (Ilgen et al., 2005). The IMOI framework defines *Inputs* as antecedent factors that enable or constrain interactions between team members. Examples of *Inputs* include culture, shared vision, resources, team composition, leadership support, and task structure (Rosenfield et al., 2018). *Mediators* describe how *Inputs* are transformed to *Outputs*, and include two sub-categories: processes and emergent states (Mathieu et al., 2008). Team processes describe actual members' interactions. Emergent states are the cognitive, motivational, and affective states of teams.

Outputs are the results or by-products of the *Mediators*, and may include performance or satisfaction (Ilgen et al., 2005; Mathieu et al., 2008). *Outputs* then drive further team adaptation or evolution (i.e., *Inputs*).

Variation in implementation

Current SB PSTs can vary substantially across problem-solving processes (Rosenfield et al., 2018; Sims et al., 2022) and implementation procedures utilized (Reinke et al., 2018; Rosenfield et al., 2018; Sims et al., 2022). In fact, previous research on SB PSTs found such variation across *Inputs*, *Mediators*, and *Outputs* (e.g., team membership, team roles, and training) that Burns, Vanderwood, et al. (2005) recommended that additional research was needed before SB PSTs could be used as part of high-stakes service delivery decision-making (e.g., SLD identification).

Purpose of SB PSTs

A lack of clear purpose, both in terms of intended purpose (i.e., *Inputs*) and actual outcomes (i.e., *Outputs*), is frequently noted as a significant barrier to SB PST implementation and effectiveness (Nellis, 2012). In the early 1980s, teams of educators began using procedures first outlined in behavioral consultation to guide prereferral intervention activities across varied student outcome areas. The Prereferral Intervention Teams (PIT; Graden et al., 1985) model described group collaborative efforts broadly, including formats where an individual with more specialized training or expertise helped a general educator better support a student experiencing a challenge at school. The recommendations to support students developed by the PIT were implemented prior to or during the process of the student being evaluated to identify or rule out a special education disability (Zins et al., 1989). The goal of the PIT was to help avoid inappropriate referrals to special education. Thus, the PITs that were common in the 1980s served a different function than the SB PSTs that were developed in response to the use of multitiered systems of support (MTSS) such as positive behavior intervention and supports (PBIS) and response to intervention (RTI; Nellis, 2012; Reinke et al., 2018). In other words, the PIT model used in the 1980s focused on identifying problems, but SB PSTs within a MTSS framework focus on finding solutions to problems (Burns, Vanderwood, et al., 2005).

In a recent survey of more than 3,000 educators in a southwestern state, the most common stated goal of SB PSTs was to increase student academic performance (28% of respondents), followed by decreasing inappropriate special education referrals (21%), providing students needed supports (16%), and providing interventions (12%; Sims et al., 2022). Similarly, Burns, Vanderwood, et al. (2005) identified individual student improvement as the most frequently cited and easily recognized measure of SB PST effectiveness. Systemic outcomes such as

reductions in the number of referrals for evaluations to identify special education disabilities are also noted in available research on SB PST effectiveness (Burns & Symington, 2002; Burns, Appleton, et al., 2005; Gravois & Rosenfield, 2006).

Problem-solving model

Contemporary SB PSTs utilize a problem-solving model to drive team processes (Burns, Appleton, et al., 2005), which is often linked to Kratochwill and Bergan's (1990) four-step problem-solving process. However, some SB PSTs use Deno's (2005) five-step IDEAL model, the Team-Initiated Problem-Solving (TIPS; Todd et al., 2011) five-step model, Tilly's (2008) four-step approach, or Bahr and Kovaleski's (2006) eight-step problem-solving model. Sims et al. (2022) documented educator reports of problem-solving processes used by SB PSTs. Reports overwhelmingly indicated the use of problem-solving processes that did not align with those espoused within SB PST literature. Unfortunately, Burns and Symington previously noted that SB PSTs that utilized a research-based model of problem solving were more effective than those based on no model (2002).

Team composition

SB PST composition can look quite different across the number of members and their areas of expertise depending on the model being used or targeted outcome. Schools in North Carolina and Oregon reported an average of 10.59 ($SD = 4.20$) members on their TIPS teams (Newton et al., 2012), which was larger than the five to seven members typically seen in SB PSTs (Burns et al., 2008). One of the longest running debates for SB PSTs is which school-based professionals should make up the team and, more specifically, the inclusion or role of special education personnel (Reinke et al., 2018). Recent research indicated SB PSTs are typically comprised of varying numbers of members, primarily general education personnel (e.g., teachers, administrators, counselors), with less representation of special education and specialized support personnel (Sims et al., 2022). Underlying this debate are concerns related to member knowledge, skills, and experience. Some argue that SB PSTs should rely exclusively on general education teachers and omit special education personnel, while others posit that special educators should be included based on their specialized skills in assessment, problem-solving, and individualized intervention (Burns, Vanderwood, et al., 2005). TIPS team members receive a two-and-a-half day training (Newton et al., 2012), but most SB PST members receive little to no training (Burns, Vanderwood, et al., 2005).

Study purpose

Rosenfield et al. (2018) reported considerable variability in name, purpose, procedures, and foci of SB PSTs, which results in "a lack of long-term studies showing outcomes over time, and an abundance of snapshots of specific

models, usually completed by the authors of those models, so that a systematic empirical base over time and varying conditions has not been established" (p. 415). Thus, it is imperative to further evaluate SB PST approaches and procedures to better understand the effects of this widely adopted model. The current study had three aims. First, the study sought to evaluate SB PST effectiveness using available empirical research and meta-analytic techniques. Second, although limited by a historic absence of theory-driven SB PST research, this work sought to evaluate the impact of variability in SB PSTs *Inputs*, *Mediators*, and *Outputs* (e.g., process used, fidelity measured, outcome measured) as gleaned within available research on team effectiveness . Third, this work sought to evaluate the effects identified by rigorous, contemporary empirical research evaluating SB PST effectiveness. The following research questions guided the study, a) How effective are SB PSTs overall and by student- and system-level outcome; b) as guided by an IMOI framework, is SB PST effectiveness moderated by identifiable *Inputs*, *Mediators*, and *Outputs* (e.g., explicitly stated problem-solving process, targeted outcome); and c) given noteworthy advances in research standards, methods, and technologies, including those related to meta-analytic research, coupled with changes in the conceptualization and use of SB PSTs, does effect size vary for studies that meet What Works Clearinghouse (WWC;, 2020) standards and for those conducted within the last 10 years?

Method

Meta-analytic methods outlined by Borenstein et al. (2009, 2010) and Cooper (2015) were used to guide study procedures. The study began by specifying search processes, continued with executing coding of identified works, and concluded with data-analysis.

Initial studies search

Studies were identified for possible inclusion in the meta-analysis by searching ProQuest and EBSCO Host databases, which included ERIC, PsycINFO, Academic Search Premier, Open Dissertations, Educators Reference Complete, Academic Search Premier, and Education Full Test. Searches included independent use of primary search terms in combination with the term "school" (e.g., "student assistance team" and "school" or "team-initiated problem-solving" and "school"). Varying independent search terms included "school based problem solving team," "instructional consultation team," "response to intervention team," "prereferral team," "prereferral intervention team," "student support team," "multitiered systems of support team," "team initiated problem solving," "child study team," "Positive Behavior Intervention Support team," "teacher assistance

team," "Tier II team," and "Tier III team." Searches included variations, such as plural forms, acronyms, and hyphenations of several of the primary search terms. For example, searches related to Response to Intervention Teams included "response to intervention team," "response to intervention teams," "RTI team," and "RTI teams."

The search for eligible works concluded in December 2022. No restrictions were placed on the date of publication and unpublished works including dissertations or theses, were eligible for inclusion. Each of the searches was set to include articles that were written in English, peer-reviewed, or was a dissertations or thesis. A total of 38 independent searches across the iterations of 14 primary search terms yielded 565 studies for initial screening, after removing 33 duplicates from the 598 initially identified studies. The screening and coding process is illustrated in Figure 1.

Initial source screening

Titles and abstracts of initially identified studies were reviewed and compared to the following initial inclusion criteria:

(1) The study was conducted in a K-12 school setting.
(2) The study included original outcome data that was sufficient to compute an effect size.
(3) The study included a team of at least three educators that was engaged in a problem-solving process to drive remediation efforts for individual students.
(4) The study examined the effectiveness of SB PSTs using student or systemic outcomes.
(5) The study reported the number of members that served on the SB PST.

The most frequent reasons for article exclusion were for 1) not being school-based (e.g., teams of emergency responders or nurses) or 2) not including quantitative data (i.e., review papers, conceptual articles, and qualitative studies). After screening the titles and abstracts, the 124 remaining studies were evaluated more thoroughly for inclusion in effect size calculations. Additionally, the reference lists for the 124 articles were searched to identify additional articles of potential interest that may have been missed in the database search, as recommended by Cooper (2015). This bibliographic search resulted in the addition of 22 studies. A cursory review of the titles for these 22 works suggested that the terminology or phrasing choices in their titles allowed them to elude being identified by formal search procedures (e.g., "problem solving implementation," "team initiated problem solving"). Ultimately, 146 studies were retained after initial screening procedures.

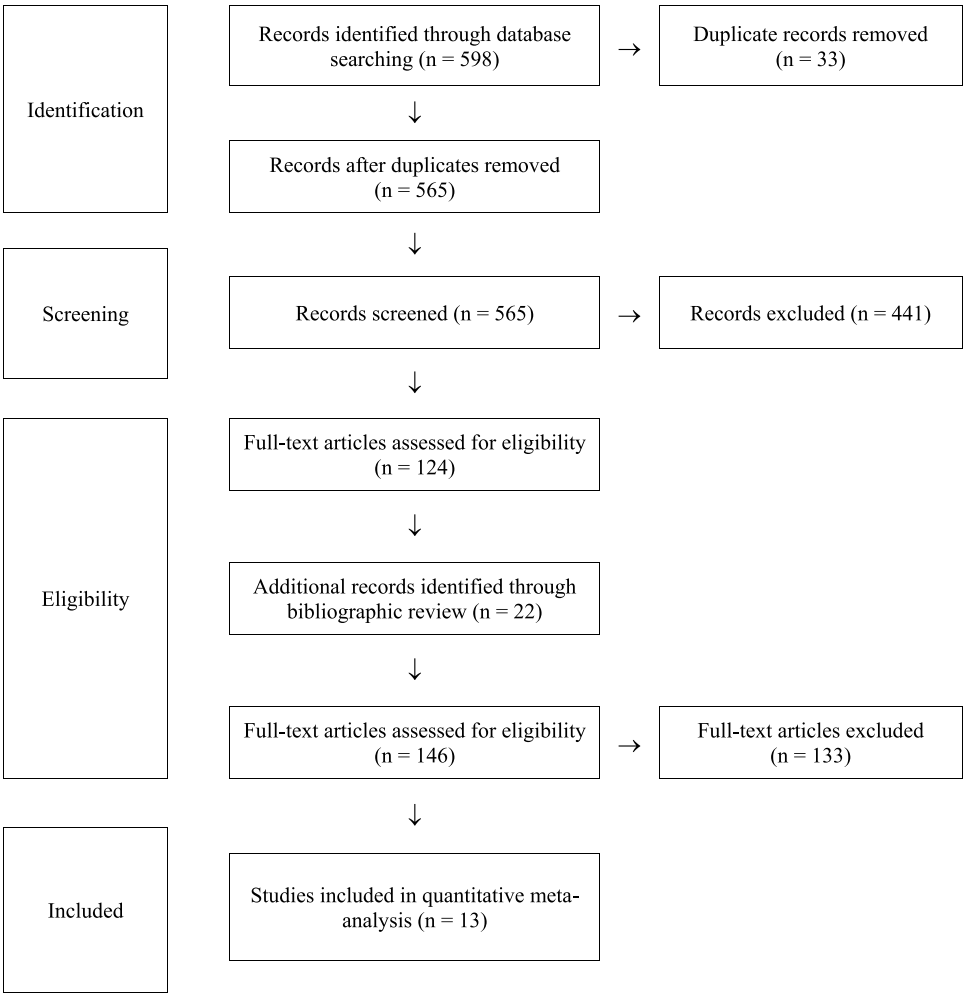


Figure 1. Studies retained across screening and inclusion process. Note. Table adapted from Moher et al. (2009). For more information, visit www.prisma-statement.org.

Eligibility screening

The full text of the studies that were retained after initial screening was thoroughly reviewed according to the inclusion criteria. After a full review of the 146 initially retained studies, 133 articles were excluded due to not meeting at least one of the inclusion criteria. The majority ($n = 42$, 31.6%) of the excluded articles did not report outcome data, but an additional 31 articles (23.3%) reported data that were insufficient for calculating effect sizes (e.g., posttest data only with no control group or presented an outcome number without a mean or standard deviation). Twenty articles (15.0%) were not included because they did not study the effects of SB PST on student or systemic outcomes, and 33 articles (24.8%) were excluded because they did not study an SB PST as defined by having at least three members (e.g.,

Instructional Consultation in a one-on-one consultation model). A total of five (3.8%) articles were excluded for not occurring in a school-aged setting. Finally, two (1.5%) works were excluded because they could not be located despite attempts to contact the primary and secondary authors. Secondary screening resulted in the retention of 13 articles for coding and calculation of effects.

The primary authors of the 13 remaining articles were contacted to identify potentially missed articles that may be in press, unpublished, or missed during the article search. A total of 13 studies ere included in the meta-analysis after contacting authors did not result in the inclusion of any additional works. One study (i.e., Newton et al., 2012) reported two different outcomes (see Table 1), which resulted in two separate effects that were included in the meta-analysis for a total of 14 effects from 13 articles.

Coding and analysis

A 34-item coding scheme was developed to facilitate moderator analyses. Key variables were coded across identifiable *Inputs*, *Mediators*, *Outputs*, and *Inputs,* including problem-solving process used, study outcomes, fidelity monitoring, and study design.

Problem-Solving Process and Procedural Fidelity Monitoring

Problem-solving process variables are the characteristics of the approach used by specific teams to identify, analyze, and remediate student difficulties. Problem-solving process variables included presence of an explicitly stated problem-solving process, ideally a process identified in best practice literature (e.g., Deno, 2005; Tilly, 2008), the number of steps involved in that explicitly stated problem-solving process, and the assessment of fidelity of the intervention and/or team. A total of 8 of the 14 (57.1%) effects used an explicitly stated problem-solving process and 6 (42.9%) did not. Explicitly stating team processes is essential to evaluate procedural fidelity for SB PST teams. As with interventions or treatments, assessing procedural fidelity for team activities is critical to reaching a valid conclusion regarding team effectiveness (Burns et al., 2008). Increased implementation or procedural implementation fidelity has been overwhelmingly associated with increases in treatment effects in intervention research (Barnett et al., 2014; Sanetti & Kratochwill, 2009). Given its importance to promoting improved outcomes (i.e., effective team-based problem-solving) and attributing these outcomes to teams appropriately, studies were coded based on whether they reported procedural fidelity of the problem-solving process used. Of the studies identified for inclusion in study analyses, 6 (42.9%) documented procedural fidelity of team processes and 8 (57.1%) did not.

Table 1. Description of studies included in the meta-analysis.

Study	Design	*n*	Outcomes	Evidence-Based	Team Name	Fidelity Measured	*g*
Bahr et al. (2006)	Between Group	134	Perceived Team Effectiveness	No	Three-step process – Name Not Reported	Not Reported	1.15
Bay et al. (1994)	Between Group	33	Referrals to and placements in special education	No	Teachers Assisting Teachers	Not Reported	0.18
Carter and Sugai (1989)	Within Group	40	Ratings of effectiveness	No	Teacher Assistance Team Five-step process	Yes	0.26
Davidoff (2012)*	Within Group	30	Student response to intervention	No	Not reported	Not Reported	1.69
Gravois and Rosenfield (2006)	Between Group	13	Reductions in referrals to and placements in special education	No	Instructional Consultation Team	Not Reported	0.27
Hartman and Fay (1996)	Between Group	22	Reductions in referrals to and placements in special education	No	Instructional Support Team with four-step process	Not Reported	1.10
Horner et al. (2018).	Between Group	38	Reductions in office discipline referrals and suspensions	Yes	TIPS	Yes	0.76
Kovaleski et al. (1999)	Between Group	1,190	Increases in student task completion, task comprehension, and time on-task	Yes	Instructional Support Team with four-step process	Yes	0.21
Lhospital and Gregory (2009)	Within Group	33	Reductions in teacher stress and students experiencing difficulties	No	Not reported	Not Reported	1.58
McDougal et al. (2000)	Within Group	20	Teacher ratings of helpfulness of the team	No	School-Based Intervention Team	Yes	0.53
McKay and Sullivan (1990)	Within Group	8	Reduction in number of students referred for special education	No	Student Assistance Team	Not Reported	1.22
Newton et al., (2012)	Within Group	35	Reduction in identified problems	Yes	TIPS	Yes	1.06
Newton et al., (2012)	Within Group	35	Satisfaction with team meeting process and results	Yes	TIPS	Yes	1.58
Short and Talley (1996)	Between Group	26	Reduction students referred to and placed into special education	No	Teacher Assistance Teams	Not Reported	0.15

*Unpublished dissertation.

Outcomes reported

In addition to the student and system outcomes used by Burns and Symington (2002), a category of team outcomes was added. These three categories were used to assign the studies to groups. Outcome measures that were placed into the student group included individual intervention effects ($k = 5$, 35.7%). Team outcomes ($k = 4$, 28.6%) included team effectiveness, team efficiency,

and team satisfaction. Systemic variables ($k = 5$, 35.7%) included referrals for special education evaluation, eligibility for special education, and documentation of problems remediated.

Study design

The study design variables included those that characterized the methodology used by each study. Those variables included the following: number of study participants, independent variable, identification of pre- and post-test design (if used), and identification of control/experimental group design (if used). Studies that used pre- and post-test designs without a control group were identified as within-group designs ($k = 7$, 50%), and those that used a control group and compared the effects between a treatment and control group were designated as between-group design ($k = 7$, 50%). Additionally, the study design was coded with the What Works Clearinghouse (WWC, 2020) standards for research design. Studies that used a randomized design with evidence for attrition that was equal between groups and demonstrated baseline equivalence were rated as meeting standards ($k = 2$, 14.3%), studies that did not use a randomized design or used a randomized design with differential attrition but demonstrated baseline equivalence for the two groups were rated as meets standards with reservations ($k = 2$, 14.3%), and studies that did not use a randomized design and did not demonstrate baseline equivalence were rated on not meeting standards ($k = 10$, 71.4%).

Age of the study

Given continuous advances in empirical research, and that the conceptualization and role of SB PSTs has changed since the proliferation of multi-tiered service delivery systems, the age of each study was considered as a potential moderating factor of SB PST effectiveness. Studies that were completed or published in 2011 or later were coded as "within 10 years," and those that were published in 2010 or earlier were coded as "older than 10 years." A total of four effects (29%) were from studies that were completed or published in the past 10 years and 10 (61%) were from studies that were older than 10 years. Although there is no universally established standard for reviewing or citing research, an unwritten rule of "within 10 years" is recommended as a starting point when considering available literature within many disciplines (Adair & Vohra, 2003; Chiang et al., 2015). Use of the most recent literature available is considered best practice as it should reflect more contemporary perspectives, conclusions, and recommendations for a given subject. In addition to the literature itself, the use of recent literature should also reflect advances in research methodologies, instruments, and analytic approaches. To this point, the first significant revision of WWC research standards was published in 2011. This first significant update to the standards that influenced

severalmethodological choices for this study and aligned well with the study's aims and timeline.

Interobserver agreement

Interobserver agreement (IOA) evaluations across study procedures followed guidelines outlined by Ahn et al. (2012). In the early phases of article screening for inclusion, IOA was not assessed. Rather, if a screener was unsure whether a study met inclusion criteria, the article was retained for a more stringent evaluation (see Figure 1 - Screening). At the *Eligibility* stage, 53 studies (36.3%) of the 146 retained works were evaluated for final inclusion by multiple evaluators. The number of times that both raters agreed that a work should be included or excluded was divided by the total number of studies and equaled 87%. The study authors reviewed seven studies where there was disagreement, reached a consensus for inclusion or exclusion, and recorded the work. The 13 studies retained for analysis were independently evaluated and coded by three of the study authors and then discussed to reconcile any disagreement.

Data analyses

A random-effects meta-analysis was conducted due to the relative lack of synthesis across team outcomes and effect sizes within empirical research examining SB PSTs (Borenstein et al., 2010; Hedges & Olkin, 2014). Treatment effects were estimated using Hedges *g* (Hedges, 1981), which adjusted for the size of the samples when computing pooled standard deviations. Hedges *g* is a standardized mean difference statistic used with small samples (Hedges, 1981) and is recommended by WWC (2020) standards for meta-analytic research. Each outcome measure is adjusted to correct for the slight upward bias of Hedges *g* when estimating population effects (Hedges, 1981). Hedges *g* was computed for between-group designs by using the means and standard deviations of the control and treatment groups, but it was computed for within-group designs using the means and standard deviations of the pre- and post-test data. Effect sizes, confidence intervals, estimates of homogeneity, and estimates of publication bias were computed using Microsoft Excel spreadsheets from Suurmond et al. (2017).

All but one study contributed one effect size to the meta-analysis. If a study reported multiple outcomes, only the outcome that was relevant to study research questions was included for analysis. For example, Newton et al. (2012) reported two outcomes that were relevant to the study, and both were included as separate effects, which resulted in 14 effects from 13 studies.

Publication bias

Publication bias refers to the increased likelihood of publication for studies with significant effects, which can positively bias the results of a meta-analysis (Borenstein et al., 2010). The likelihood of an overestimation of the average true effect due to publication bias was tested through visual inspection of a funnel plot created with Microsoft Excel spreadsheets (Suurmond et al., 2017). The distribution of individual effect sizes by their standard errors is displayed on a funnel plot and the distribution indicates the precision of the effect size estimate (see Figure 2). Symmetry in the funnel plot suggests a lack of publication bias, and asymmetry suggests a potential for publication bias. An Egger's (Egger et al., 1997) regression, which regresses the effect sizes on the inverse of their variances, was also conducted. A nonsignificant regression suggests a lack of publication bias.

Moderating factors

The second and third research questions sought to evaluate the impact of moderating factors identified in retained studies on SB PST effectiveness (e.g., process used, procedural fidelity, study rigor). As anticipated and consistent with the critique provided by Rosenfield et al. (2018), few Input or Mediator variables were explicitly included in empirical SB PST works. From an IMOI framework perspective, explicit statement of the SB PST process used and inclusion of an implementation fidelity monitoring mechanism were the only discernible Input variables identified for use in study analyses. Absent from these works were other typical *Inputs* noted in team literature, including team composition, organizational health, training, and administrative support. However, the moderating impacts of outcome measures and study design were examined. Effect sizes for each were computed and were weighted according to the inverse of the variance. Weighting an effect in a random-effects meta-analysis with the inverse of the variance is preferable because it is proportional to sample size, but provides a more precise estimate that minimizes the variance of the combined effects (Borenstein et al., 2010). Estimates of effect for g between 0.2 and 0.49 are considered small, 0.5 and 0.79 are considered medium, and 0.8 and higher are considered large (Cohen, 1988).

To compare the observed variance in the distribution of effect sizes to what would be expected from sampling error, a test of homogeneity (Q test) was conducted. Variability in effect sizes between studies that is greater than what would be expected by sampling error alone is indicated by a significant Q statistic, and significantly heterogeneous effect sizes suggest that the data were potentially influenced by moderating variables. An I^2 indicates the proportion of the variation of the observed effect sizes in relation to the proportion of the variation in the true effect sizes or the proportion of variance beyond sampling error (Borenstein et al., 2017). The larger the I^2, the more

likely the data to approximate the variability of true effect sizes, and 75% or larger is considered a high level.

The median weighted effects were reported for each level of the moderating variable because the number of effects for some levels of these variables was quite small and too small to report a parametric mean. Confidence intervals of 95% of median effect sizes were compared to determine the presence of interval overlap. Intervals that did not overlap were considered reliably different at the $p < .05$ criterion. Moreover, confidence intervals that did not include 0 suggested that the effect size was reliably greater than 0 to $p < .05$ criterion.

Because the meta-analysis included both between- and within-group designs, the relative effects from each type of study were evaluated with a nonparametric analysis to determine the extent to which they were comparable (Burns et al., 2012). Thus, the review of moderating factors began by conducting a Mann-Whitney U nonparametric test of two independent sets of data to determine if the effects from the two types of designs were significantly different.

Results

Data from the 14 effects were used to answer the research questions (see Figure 3). The individual studies are summarized in Table 1. First, the data from the 14 effects were examined for a potential publication bias with the funnel plot shown in Figure 2. A visual analysis of the funnel plot found a symmetry that suggested a lack of publication bias, and the Q statistic was not significant $Q = 17.79$, $p = .17$. Egger's (Egger et al., 1997) test for publication bias was conducted, which was not significant (Intercept = −0.26, 95% CI = −2.04 to 1.52, $t = -0.32$, $p = .78$), and did not indicate a need to adjust the estimates of effect due to publication bias.

Overall effectiveness

To address the first research question, the data were analyzed with a random-effects approach. The effect sizes for each study were calculated using spreadsheets as outlined by Suurmond et al. (2017) and are displayed in Table 1 and Figure 3. The data resulted in a large overall weighted estimate of effect of $g = 0.84$ (95% CI = 0.54 to 1.14) with a confidence interval that did not include 0. The resulting Q of 84.42 ($p < .01$) and I^2 of 84.60 (95% CI = 75.64 to 90.27) suggested acceptable heterogeneity to conduct moderator analyses.

Next, the effects of studies that used a between-group design ($k = 7$) were compared to those that used a within-group design ($k = 7$) before examining potential moderating factors. The weighted effect for a between-group design was 0.63 (95% CI = 0.21 to 1.04), and the within-group studies resulted in

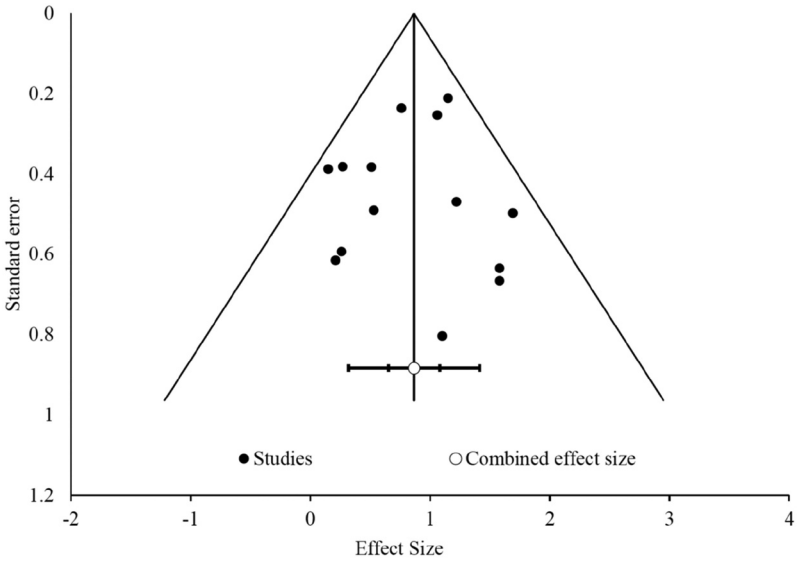


Figure 2. Funnel plot to examine potential publication bias.

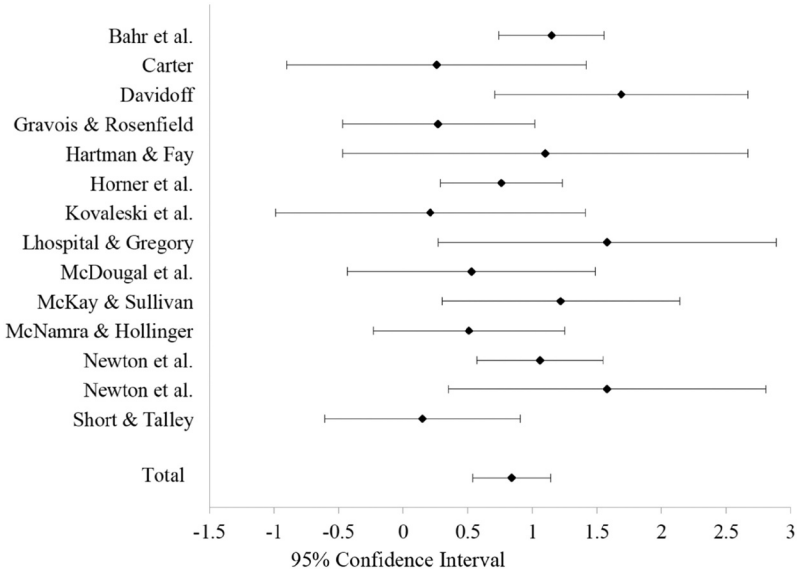


Figure 3. Effects for studies included in the meta-analysis.

a weighted effect of 1.09 (95% CI = 0.58 to 1.60). The Mann–Whitney U resulted in mean ranks of 5.43 for between-group designs and 9.57 for within-group designs, which equaled U = 10.00, $z = 1.85$, $p = .064$. The analysis was not significant, and there was considerable overlap between the two ranges. Therefore, the data were combined across the two designs to address the research questions.

Moderating factors

The second and third research questions addressed the effects associated with various *Inputs, Mediators*, and *Outcomes* identified within available SB PST literature. As anticipated, consistent with criticisms noted in prior research, documentation of potential team *Inputs* and *Mediators* in available SB PST literature was limited. For example, none of the included studies described team composition, a common *Input* identified in IMOI and team science literature (Rosen et al., 2015). One salient *Input,* an explicitly stated problem-solving process, was identified and used in these analyses. Similarly, a single *Mediating* variable was evident in reviewed research, procedural (i.e., team processes) fidelity monitoring. Varying *Output* variables (i.e., targeted outcomes) were identified in the reviewed literature (e.g., student outcomes, team satisfaction).

Explicitly stated problem-solving process

First, the reviewed studies were differentiated based on their use of an explicitly stated problem-solving process, a team *Inputs* variable. Eight effects were from studies that reported an explicit problem-solving process, which resulted in a weighted effect of 1.78 (95% CI = 0.04 to 3.52). Six effects were from studies that did not state a problem-solving process, which resulted in a weighted effect size of $g = 0.85$ (95% CI = 0.24 to 1.44). Although the confidence intervals overlapped and neither contained 0, and the effect size for studies that explicitly stated the problem-solving process used was much larger than those that did not.

Procedural fidelity monitoring

The effect of measuring procedural fidelity (i.e., problem-solving process adherence), a team *Mediators* variable, was also examined. Six effects were from studies that reported fidelity data, all of which reported procedural adherence rates above 90%. The resulting effect size of 0.71 (95% CI = 0.24 to 1.18) suggested a moderate to large effect for which the range did not include 0. Eight effects were from studies that did not report fidelity and resulted in a weighted effect size of 0.95 (95% CI = 0.63 to 1.28), which was a large effect with a range that did not include 0.

Targeted outcome

Five effects from studies that measured student outcomes (i.e., IMOI *Outputs*), including academic and/or behavioral improvement, resulted in a large weighted effect size of 0.89 (95% CI = 0.28 to 1.49). System outcomes, including referrals to and actual placements in special education ($k = 5$), resulted in a moderate effect size of 0.66 (95% CI = 0.15 to 1.18). Finally, four studies examined team outcomes (e.g., members rating team effectiveness) that

yielded a large effect of 1.00 (95% CI = 0.41 to 1.59). Although there was some variability, the ranges for the three effect sizes overlapped and suggested moderate to large effects.

Study design

The final research question inquired about the influence of more rigorous research approaches on SB PST effectiveness. Only two studies met WWC standards, which resulted in a large weighted effect of $g = 1.10$ (95% CI = 0.76 to 1.43). The two studies that met standards with reservations had a small effect $g = 0.21$ (95% CI = −0.34 to 0.77) with a confidence interval that included 0 and that overlapped with the effect size for the studies that met standards, but by only one point. Ten of the effects were derived from studies that did not meet standards (71.4%), which resulted in a large effect $g = 0.91$ (95% CI = 0.58 to 1.25) that overlapped considerably with the other two ranges.

Recency of study

Four of the effects were from studies conducted in the last decade (i.e., since 2011), which resulted in a large weighted $g = 2.30$ (95% CI = 0.63 to 3.86). The remaining 10 effects from studies conducted before 2011 (i.e., 2010 or earlier) resulted in a large weighted effect of $g = 1.06$ (95% CI = −0.45 to 2.56). Although the two ranges overlapped, the confidence interval for the effect size for studies that were older than 10 years included 0.

Discussion

As the use of multi-tiered systems of support in schools has widened, SB PSTs have served as a critical mechanism by which prevention and early intervention services were matched to student needs. Although widely used, relatively little research has examined the efficacy and effectiveness of SB PST. In 2002, Burns and Symington found prereferral team practices, an early form of SB PST, to be effective overall. Large effect sizes were evident for these teams generally across student- and systems-level outcomes. Despite these findings supporting intervention team use, the authors noted several concerns related to research in this area. Most notably, relatively few studies had examined this widely used approach. Sixteen years later, Rosenfield et al. (2018) continued the call for a more explicit study of the components, influential factors, and outcomes of SB PST. In an attempt to address questions around SB PST effectiveness first raised two decades ago, the current meta-analysis sought to evaluate the overall effectiveness of SB PST, to evaluate the impact of variability in SB PSTs *Inputs*, *Mediators*, and *Outputs* as gleaned within available research on SB PST effectiveness, and to evaluate SB PST effects

identified in rigorous, contemporary empirical research. Ultimately, this work sought to stimulate further empirical inquiry targeting SB PST practices, processes, and outcomes using an IMOI framework.

To address the study goals, rigorous random-effects meta-analytic procedures were used to analyze 14 effects that examined SB PST effectiveness. The first research question sought to examine the overall effectiveness of SB PSTs. An overall large effect size suggests that SB PSTs are an efficacious practice. Although the identified effect size appeared smaller, these findings appear to agree with prior research that found large effect sizes for SB PSTs across overall ($d = 1.10$), student ($d = 1.15$), and systems ($d = 0.90$) outcomes (Burns & Symington, 2002).

This work also sought to examine effects for SB PST as influenced by moderating factors identified in available empirical SB PST research. Consistent with the critique by Rosenfield et al. (2018), few if any of the retained studies identified any theoretical framework, like the IMOI Framework, beyond "problem-solving" to guide their work. As such, discerning a significant number of *Inputs* or *Mediators* within the eligible studies was challenging. However, as noted, several varying outcomes (i.e., *Outputs*) including student improvement, special education evaluation referrals, and team satisfaction were evident within the included studies. Use of an explicitly stated process espoused in the literature was a one frequent discernable *Input* for SB PST was evident in these included works. The current results indicated a higher effect for teams with a stated process. Although the two effects were not significantly different, the effect size for studies that explicitly stated the problem-solving process used was twice as large as those that did not. These results appear to be consistent with earlier findings suggesting that university-driven SB PSTs were more effective than field-based SB PSTs (Burns & Symington, 2002). Endorsement of a specific problem-solving process was not explicitly documented in the earlier meta-analysis, but it stands to reason that university-guided studies would be more likely to be aligned with an explicitly stated problem-solving process than field-based studies. Whether explicitly stated or not, the core components of a problem-solving orientation to address student difficulties are relatively consistent (Tilly, 2008). Therefore, any steps included in any process should share at least some commonality. Without additional inquiry, reasons behind the noted difference in effect size between teams that use a process explicitly identified in the literature and those that do not would be speculation, which again illustrates the need for additional work to determine aspects of SB PSTs that influence effectiveness.

Unexpected findings were noted relative to the inclusion of the use of procedural fidelity measures as part of team processes, a *Mediator* within an IMOI framework. Findings suggest teams that did not document procedural fidelity appeared somewhat more efficacious than those that did. This finding is potentially problematic given the breath of research suggesting that

implementation fidelity improves outcomes in most applications (Burns et al., 2008; Kovaleski et al., 1999; McNamara & Hollinger, 2000; Ruby et al., 2011). With this finding in mind, rather than endorse SB PST processes that do not include fidelity measures, additional research incorporating measures of fidelity should be conducted, and should include varied outcome measures consistent with research. Conclusions from these findings should also be tempered as an absence of procedural fidelity documentation does not guarantee a) teams did not adhere strictly to a team process or b) that procedural fidelity was not accounted for. This is to say that assessing and reporting procedural fidelity is not a requirement for implementing a problem-solving process with fidelity. However, assessing and reporting procedural fidelity should be considered best-practice within applied implementation and research (WWC, 2020).

Effects associated with various outcomes (i.e., *Outputs*) appeared largely consistent with those identified in prior literature and included individual student improvement across varied areas of performance, including student achievement and behavior, and referrals for special education eligibility evaluation. Perceptions of team success were also identified as an outcome within some studies. Findings appear to support SB PST effectiveness when examining the influence of targeted SB PST outcomes, such as student academic achievement and behavior. However, the largest effect size was noted in studies evaluating the outcome of SB PST as team member ratings of team effectiveness. This is not surprising given the potentially subjective nature of this outcome measure. It is reasonable to think team members find their work of contributing to SB PST activities as valuable, productive, and effective. A large effect size was also noted for studies using student performance outcomes. Studies documenting effects on system outcomes, such as referrals to and placements in special education, indicated a moderate effect size for SB PST, both of which were consistent with previous research (Burns & Symington, 2002; Rosenfield et al., 2018). Although not without limitations, student and system outcome measures may be preferable outcome measures because they are viewed as more objective than ratings of perceived effectiveness and are likely more indicative of positive outcomes than team self-ratings of effectiveness.

The current study also examined SB PST effects as documented in more rigorous and contemporary research. To this end, effect sizes were calculated for studies that aligned with WWC standards and for those conducted since 2011 (i.e., "within the last 10 years"). While the apparent effect of SB PSTs appeared positive for studies adhering to some level of WWC standards (e.g., fully met, met with reservations), the relatively few studies eligible for inclusion in study analyses were troubling. Only two of the effects were from studies that met WWC (2020) standards and an additional two met standards with reservations. Studies that met standards resulted in large effects, and those that

met standards with reservations resulted in a small combined effect. The somewhat larger effect size noted for pre-post designs also suggests a need for additional experimental research in this area. Additionally, while limited generally, there may be reason to be optimistic based on recent research. Research completed within the past 10 years led to a large weighted effect (g = 2.30). Unfortunately, this optimism must be tempered as approximately two-thirds of the effects eligible for study inclusion were from works that were more than 10 years old. Collectively, this clearly indicates that additional empirical research using more contemporary and rigorous, experimental methodologies, analyses, and technologies consistent with WWC standards to further evaluate SB PST effects is warranted.

Limitations

As noted previously, the results of this study should be interpreted while accounting for limitations associated with the research design. First, ideally, this meta-analysis of SB PST would have included additional empirical work on this topic. However, like many meta-analyses, this study was limited by those works that were identified through a thorough search, those that were written in English, and those that were accessible. Similarly, publication bias is also an important factor to consider in this or any meta-analytic work. Although every attempt was made to include unpublished works, including theses, dissertations, technical reports, and presentations as well as other unpublished works completed by SB PST researchers, and although a visual analysis of a funnel plot and an Egger regression did not suggest publication bias, publication bias may persist.

It should also be noted that the effects from studies that used a between-group design (k = 7) were combined with those from within-group designs (k = 7). The difference between the two weighted effects, 0.63 and 1.04, respectively, was not significant (U = 10.00, z = 1.85, p = .064), but the magnitude for within-group designs was 1.5 times that of between-group designs. It was necessary to combine the two data sets because of the small number of studies that met the inclusion criteria. The effect of combining the two sets of data is unknown and readers should consider that when interpreting the data. Moreover, the computation of g and the interpretation of the resulting effect sizes assumed independence of the data and did not account for potential within-study dependence in effect size estimates. The effect of potential dependence within the data on the outcomes of the meta-analysis is also unknown.

Lastly, consistent with concerns noted by Rosenfield et al. (2018), largely absent from the works retained for this study was the use of a formal theoretical framework to guide SB PST research. Broader teaming research has embraced the IMOI Framework; however, school-based work on this topic

continues to take a limited, problem-solving process-driven perspective when conducting empirical SB PST research. This approach, as reflected in these findings, has resulted in the evaluation of the factors that influence SB PST that has been limited to [targeted] outcome and relatedly, process used (e.g., academic team process, behavior team process).

Implications for practice and research

The findings of this study have several potential implications for practice and research. The study findings support continued and expanded use of SB PSTs, based on thier positive impact on a variety of outcomes. Use of SB PSTs results in improved individual student outcomes while decreasing student referral for evaluations to determine eligibility for special education services. Additional recommendations focus on evaluating the effectiveness of SB PSTs. Whether in empirical research or applied practice, SB PST users should collect objective outcome data as part of their core processes. Specifically, these outcomes should focus on student outcomes, including referral of students to and affirmative eligibility determinations for special education services (Rosenfield et al., 2018). While team member perceptions of team functioning are an important part of evaluating team functioning, it may result in an overestimation of actual effectiveness (e.g., impact for students).

Further recommendations specific to research focus on expanding the breadth and depth of work in this area. First, given their widespread use and apparent effectiveness, additional inquiry is needed to address apparent inconsistencies across SB SPT practices, processes, and outcomes, which should be studied with rigorous experimental, quasi-experimental, and single-case designs. Outcome measures could include objective measures noted in this and other SB PST literature, such as student outcomes, referrals for special education eligibility evaluation, and special education eligibility determination rates. Next, future research should take a more nuanced, theory-informed (e.g., IMOI Framework) approach when evaluating SB PST functioning, efficacy, or effectiveness. Researchers should endeavor to identify and evaluate the influence of [neglected] *Inputs*, like organizational health, team composition, and training, as well as *Mediators*, like team member interactions, shared goals, and psychological safety, on team *Outputs*. Future SB PST research should focus on both overall effectiveness and the individual elements, factors, and states that influence SB PST outcomes. Such an approach will likely require the development and validation of additional measures that objectively assess *Inputs* and *Mediators* that have largely been neglected in SB PST research to this point. This may be supported by drawing from the larger team science literature in other fields. As part of this increased attention, researchers should adopt rigorous methodological approaches that align

with contemporary recommendations that will situate findings to meaningfully contribute to the evidence base supporting SB PST use. This would include clearly and explicitly documenting IMOI-based factors within a given study and reporting results in a manner that allows systematic replication as well as inclusion of findings in future SB PST meta-analytic work. Lastly, beyond effectiveness outcomes, research could explore efficiency outcomes and cost–benefit analyses.

Conclusion

Although a wide-spread educational practice, SB PSTs have received relatively limited empirical scrutiny. Available evidence supporting their use, including this study, suggests that SB PST use results in noteworthy positive impact on important student- and system-level outcomes. Despite such findings, many important questions about factors impacting team effectiveness persist. From a broader team science perspective, questions related to the influence of specific *Inputs* and *Mediators* on *Outputs* should be emphasized as SB PST research continues. Given the frequency of their use in practice, additional rigorously designed research to examine the effectiveness of SB PSTs appears warranted.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Dr. Wesley Sims, is an Assistant Professor and Practicum Coordinator in the School Psychology program at the University of California, Riverside. His research interests include improving educator service delivery practices within tiered service delivery systems, implementation science in educational settings, and assessment development and validation. Dr. Sims began his career as a practicing School Psychologist where he served a variety of schools and populations, and garnered extensive experience facilitating school-wide and individualized support services within tiered service delivery models.

Dr. Kathleen R. King, provides psycho-educational support services as a School Psychologist to the students and families of Palm Springs Unified School District. Her work and research centers on prevention, early detection, and intervention in the area of behavior. Her research focuses on the development and validation of behavioral observation and screening measures within a broader problem-solving/MTSS framework. In addition to her work targeting student behavior, Dr. King is involved in efforts to evaluate and improve teacher practices in the applied setting.

Dr. June L. Priest, is an Assistant Professor of School Psychology at the University of Alabama. Her work focuses on identifying and implementing systemic change to support the academic and behavioral needs of students. She uses a primary prevention, multi-tiered systems of

support approach to educational decision-making and programming. As such, she has studied the impact of academic interventions to address behavioral issues, the influence of psychosocial factors related to bullying and victimization, the importance of using data to inform academic and behavioral intervention choices, and the attributes of effective teacher teams. Her work is driven by her early career work as a practicing School Psychologist.

Dr. Matthew K. Burns, is a Professor of Special Education and Director of the Center for Collaborative Solutions for Kids, Practice, and Policy (SKiPP). Dr. Burns works closely with schools to study and implement response-to-intervention models, intensive reading and math interventions, school-based teams, and generally supporting students with and without disabilities whose needs are not being met. Dr. Burns was also a practicing school psychologist and special education administrator before shifting the focus of his work to research and training.

Sarah Panameño, B.S. is a doctoral student in the School Psychology program at University of California, Riverside. She earned her Bachelor's degree from the University of California, Riverside, majoring in psychology with a minor in education. Her research interests include best practices for supporting emerging bilingual students in schools, implementation science in schools, and school-based problem-solving teams, and data-driven decision-making.

ORCID

June L. Preast http://orcid.org/0000-0001-8759-3221
Matthew K. Burns http://orcid.org/0000-0001-9946-5729

References

***Indicates a study used in the meta-analysis.**

Adair, J. G., & Vohra, N. (2003). The explosion of knowledge, references, and citations: Psychology's unique response to a crisis. *American Psychologist*, *58*(1), 15–23. https://doi.org/10.1037/0003-066X.58.1.15

Ahn, S., Ames, A. J., & Myers, N. D. (2012). A review of meta-analyses in education: Methodological strengths and weaknesses. *Review of Educational Research*, *82*(4), 436–476. https://doi.org/10.3102/0034654312458162

Bahr, M. W., & Kovaleski, J. F. (2006). The need for problem-solving teams: Introduction to the special issue. *Remedial and Special Education*, *27*(1), 2–5. https://doi.org/10.1177/07419325060270010101

Bahr, M. W., Walker, K., Hampton, E. M., Buddle, B. S., Freeman, T., Ruschman, N., Sears, J., McKinney, A., Miller, M., & Littlejohn, W. (2006). *Creative problem solving for general education intervention teams: A two-year evaluation study. *Remedial and Special Education*, *27*(1), 27–41. https://doi.org/10.1177/07419325060270010401

Barnett, D., Hawkins, R., McCoy, D., Wahl, E., Shier, A., Denune, H., & Kimener, L. (2014). Methods used to document procedural fidelity in school-based intervention research. *Journal of Behavioral Education*, *23*(1), 89–107. https://doi.org/10.1007/s10864-013-9188-y

Bay, M., Bryan, T., & O'Connor, R. (1994). *Teachers assisting teachers: A prereferral model for urban educators. *Teacher Education and Special Education*, *17*(1), 10–21. https://doi.org/10.1177/088840649401700103

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. https://doi.org/10.1002/jrsm.12

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1002/9780470743386

Borenstein, M., Higgins, J. P., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Burns, M. K., Appleton, J. J., & Stehouwer, J. D. (2005). Meta-analytic review of responsiveness-to-intervention research: Examining field-based and research-implemented models. *Journal of Psychoeducational Assessment*, *23*(4), 381–394. https://doi.org/10.1177/073428290502300406

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*(5), 537–550. https://doi.org/10.1016/j.jsp.2008.04.001

Burns, M. K., & Symington, T. (2002). A meta-analysis of prereferral intervention teams: Student and systemic outcomes. *Journal of School Psychology*, *40*(5), 437–447. https://doi.org/10.1016/S0022-44050200106-1

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*(1), 89–105. https://doi.org/10.1521/scpq.20.1.89.64192

Burns, M. K., Zaslofsky, A. F., Kanive, R., & Parker, D. C. (2012). Meta-analysis of incremental rehearsal using phi coefficients to compare single-case and group designs. *Journal of Behavioral Education*, *21*(3), 185–202. https://doi.org/10.1007/s10864-012-9160-2

Carter, J., & Sugai, G. (1989). *Survey on prereferral practices: Responses from state departments of education. *Exceptional Children*, *55*(4), 298–302. https://doi.org/10.1177/001440298905500402

Chiang, I. C. A., Jhangiani, R. S., & Price, P. C. (2015). Reviewing the research literature. In *Research methods in psychology-2nd Canadian edition*. BC Campus. https://opentextbc.ca/researchmethods/chapter/reviewing-the-research-literature/

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.

Cooper, H. (2015). *Research synthesis and meta-analysis: A step-by-step approach*. SAGE Publications.

Davidoff, L. J. (2012). **A case study of the use of response to intervention in a public school district* [Doctoral dissertation]. Rutgers, The State University of New Jersey. ProQuest Dissertations & Theses. https://search.proquest.com/pqdtglobal/docview/1033507105/94E71CCF5D51419APQ/1?accountid=14472

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

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Graden, J. L., Casey, A., & Christenson, S. L. (1985). Implementing a prereferral intervention system: Part I. The model. *Exceptional Children*, *51*(5), 377–384. https://doi.org/10.1177/001440298505100502

Gravois, T. A., & Rosenfield, S. A. (2006). *Impact of instructional consultation teams on the disproportionate referral and placement of minority students in special education. *Remedial and Special Education*, *27*(1), 42–52. https://doi.org/10.1177/07419325060270010501

Hartman, W. T., & Fay, T. A. (1996). *Cost-effectiveness of instructional support teams in Pennsylvania. *Journal of Education Finance*, *21*(4), 555–580.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.2307/1164588

Hedges, L. V., & Olkin, I. (2014). *Statistical methods for meta-analysis*. Academic press. https://doi.org/10.1016/C2009-0-03396-0

Horner, R. H., Newton, J. S., Todd, A. W., Algozzine, B., Algozzine, K., Cusumano, D., & Preston, A. (2018). *A randomized waitlist controlled analysis of team-initiated problem solving professional development and use. *Behavioral Disorders*, *43*(4), 444–456. https://doi.org/10.1177/0198742917745638

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, *56*(1), 517–543. https://doi.org/10.1146/annurev.psych.56.091103.070250

Kovaleski, J. F., Gickling, E. E., Morrow, H., & Swank, P. R. (1999). *High versus low implementation of instructional support teams a case for maintaining program fidelity. *Remedial and Special Education*, *20*(3), 170–183. https://doi.org/10.1177/074193259902000308.

Kratochwill, T. R., & Bergan, J. R. (1990). *Behavioral consultation in applied settings: An individual guide*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4757-9395-6

Lhospital, A. S., & Gregory, A. (2009). *Changes in teacher stress through participation in pre-referral intervention teams. *Psychology in the Schools*, *46*(10), 1098–1112. https://doi.org/10.1002/pits.20455

Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*(3), 410–476. https://doi.org/10.1177/0149206308316061

McDougal, J. L., Moody Clonan, S., & 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*(2), 149. https://doi.org/10.1037/h0088783

McKay, B., & Sullivan, J. (1990). **Effective collaboration: The student assistance team model*. Paper presented at the Annual Convention of the Council for Exceptional Children, (ERIC Document Reproduction Service No. ED 322695).

McNamara, K., & Hollinger, C. L. (2000). Fidelity of problem-solving implementation and relationship to student performance. *School Psychology Review*, *29*(3), 443–461. https://doi.org/10.1080/02796015.2000.12086029

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & PRISMA Group*. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269.

Nellis, L. M. (2012). Maximizing the effectiveness of building teams in response to intervention implementation. *Psychology in the Schools*, *49*(3), 245–256. https://doi.org/10.1002/pits.21594

Newton, J. S., Horner, R. H., Algozzine, B., Todd, A. W., & Algozzine, K. (2012). *A randomized wait-list controlled analysis of the implementation integrity of team-initiated problem solving processes. *Journal of School Psychology*, *50*(4), 421–441. https://doi.org/10.1016/j.jsp.2012.04.002

Reinke, W. M., Sims, W., Cohen, D., & Herman, K. C. (2018). Problem solving within an RTI framework: Roles and functions of effective teams. In *Handbook of response to intervention and multi-tiered systems of support* (pp. 96–119). Routledge. https://doi.org/10.4324/9780203102954-8

Rosen, M. A., Dietz, A. S., Yang, T., Priebe, C. E., & Pronovost, P. J. (2015). An integrative framework for sensor-based measurement of teamwork in healthcare. *Journal of the American Medical Informatics Association*, *22*(1), 11–18. https://doi.org/10.1136/amiajnl-2013-002606

Rosenfield, S., Newell, M., Zwolski, S., Jr., & Benishek, L. E. (2018). Evaluating problem-solving teams in K–12 schools: Do they work? *American Psychologist*, *73*(4), 407–419. https://doi.org/10.1037/amp0000254

Rosenfield, S., Silva, A., & Gravois, T. (2014). Bringing instructional consultation to scale. In W. P. Erchul and S. M. Sheridan (Eds.), *Handbook of Research in School Consultation* (pp. 248–275). Routledge. https://doi.org/10.4324/9780203133170

Ruby, S. F., Crosby-Cooper, T., & Vanderwood, M. L. (2011). Fidelity of problem solving in everyday practice: Typical training may miss the mark. *Journal of Educational & Psychological Consultation*, *21*(3), 233–258. https://doi.org/10.1080/10474412.2011.598017

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, *38*(4), 445.

Santangelo, T. (2009). Collaborative problem solving effectively implemented, but not sustained: A case for aligning the sun, the moon, and the stars. *Exceptional Children*, *75*(2), 185–209. https://doi.org/10.1177/001440290907500204

Sheridan, S. M., Kratochwill, T. R., & Bergan, J. R. (2013). *Conjoint behavioral consultation: A procedural manual*. Springer Science & Business Media.

Short, R. J., & Talley, R. C. (1996). *Effects of teacher assistance teams on special education referrals in elementary schools. *Psychological Reports*, *79*(3_suppl), 1431–1438. . https://doi.org/10.2466/pr0.1996.79.3f.1431

*

Sims, W. A., King, K. R., Wicoff, M., Mandracchia, N., Womack, T., & Anazagasty, J. M. (2022). School-based problem-solving teams: Educator-reported implementation trends and outcomes. *Contemporary School Psychology*, 1–15. https://doi.org/10.1007/s40688-022-00412-w

Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison, and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods*, *8*(4), 537–553. https://doi.org/10.1002/jrsm.1260

Tilly, W. D. (2008). The evolution of school psychology to science based practice: Problem solving and the three-tiered model. In A. Thomas, J. Grimes (Eds.), *Best Practices in School Psychology* (2nd ed., Vol. 1, pp. 17–36). 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 school wide behavior support teams. *Journal of Applied School Psychology*, *27*(1), 42–59. https://doi.org/10.1080/15377903.2011.540510

Truscott, S. D., Cohen, C. E., Sams, D. P., Sanborn, K. J., & Frank, A. J. (2005). The current state(s) of prereferral intervention teams: A report from two national surveys. *Remedial and Special Education*, *26*(3), 130–140. https://doi.org/10.1177/07419325050260030101

Welch, M., Brownell, K., & Sheridan, S. M. (1999). What's the score and game plan on teaming in schools? A review of the literature on team teaching and school-based problem-solving teams. *Remedial and Special Education*, *20*(1), 36–49. https://doi.org/10.1177/074193259902000107

What Works Clearinghouse. (2020). *What Works Clearinghouse Standards handbook*, Version 4.1. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/handbooks .

Zins, J. E., Graden, J. L., & Ponti, C. R. (1989). Prereferral intervention to improve special services delivery. *Special Services in the Schools*, *4*(3–4), 109–130. https://doi.org/10.1300/J008v04n03_07