STUDY PROTOCOL

Open Access



Jill Locke^{1*}, Nathaniel J. Williams², Aksheya Sridhar¹, Mark G. Ehrhart³, Alex Dopp⁴, Marissa Thirion¹, Christine Espeland^{1,5}, Brandon Riddle⁶, Kelcey Schmitz¹, Kurt Hatch⁵, Lindsey Buehler⁶ and Aaron R. Lyon¹

Abstract

Background Schools need to implement universal student supports that prevent social, emotional, and behavioral difficulties; minimize associated risks; and promote social, emotional, and behavioral competencies. The purpose of this study is to examine the efficacy of the Helping Educational Leaders Mobilize Evidence (HELM) implementation strategy for promoting school-level implementation leadership, implementation climate, and high-fidelity delivery of an evidence-based practice. We will test HELM with an exemplar EBP, Positive Behavioral Interventions and Supports (PBIS). The specific aims of the study are to: 1) experimentally evaluate the effects of HELM versus PBIS training and technical assistance only (control condition); and 2) explore for whom, under what conditions, how equitably, and through which processes HELM works to improve outcomes, as well as its cost-effectiveness.

Methods This study will use a hybrid type 3 effectiveness-implementation trial to provide a rigorous test of the effects of HELM in elementary schools. Schools will be randomly assigned to HELM + PBIS training and technical assistance (n=21 schools; n=210 educators) or PBIS training and technical assistance only (n=21 schools; n=210 educators) in a 1:1 ratio within cohorts using covariate constrained randomization that accounts for degree of prior PBIS exposure (measured using the Tiered Fidelity Inventory at baseline) and school size. A series of mixed effects models (time within educator, educator within school) will test within-subject/between-subject interactions across three timepoints (12 months total) to examine whether HELM will show steeper gains than the control on implementation leadership (primary outcome), implementation climate, PBIS fidelity, and student outcomes. Mediational analyses will test hypothesized mechanisms of change (i.e., implementation leadership and climate) of HELM on PBIS fidelity. Sequential mixed-methods data collection and analyses will further explore how organizational mechanisms are linked to implementation outcomes. Cost-effectiveness analyses will compare costs and outcomes of PBIS training and technical assistance only versus PBIS implementation with HELM.

*Correspondence: Jill Locke jjlocke@uw.edu Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Discussion The nature of leadership support in schools can make the difference between successful and unsuccessful EBP implementation. Testing HELM within the context of PBIS implementation will provide rigorous evidence about whether and how HELM can equitably address important EBP and student outcomes.

Name of the registry clinicaltrials.gov.

Trial Registration Clinical Trials ID: NCT06586723. Date of Registration: August 27, 2024. Prospectively registered. URL of Trial Registry Record: https://clinicaltrials.gov/study/NCT06586723?intr=helm&rank=1

Keywords Schools, Implementation leadership, Implementation climate, Implementation strategy, Prevention

Contributions to the Literature

- Few effective strategies to enhance evidence-based practice implementation in schools address implementation leadership and climate, and those that do have only been evaluated qualitatively.
- Schools have unique organizational characteristics, and it is important to test whether organizational implementation strategies developed in other contexts can be effective in improving implementation outcomes in schools.
- This study will evaluate the effects of a tailored adaptation of the Leadership and Organizational Change for Implementation (LOCI) strategy for schools, entitled Helping Educational Leaders Mobilize Evidence (HELM), on implementation mechanisms (implementation leadership and implementation climate), PBIS fidelity, and student outcomes.
- This study also will explore for whom, under what conditions, and how equitably HELM demonstrates its effects, providing critical information about influence on – and the limitations of – the impact of this organizational implementation strategy.
- Finally, the study will evaluate the processes through which HELM works to improve outcomes, contributing to the emerging literature on implementation mechanisms, as well as its cost-effectiveness.

Introduction

Social, emotional, and behavioral problems occur frequently among elementary school students and dramatically impede student outcomes [1–3]. Numerous evidence-based practices (EBPs) exist to address student social, emotional, and behavioral needs, prevent problems, and ensure academic success [4]. A recent metaanalysis supports the utility of universal (i.e., "Tier 1") interventions in improving student social, emotional, and behavioral functioning [5]. One exemplar universal EBP is Positive Behavioral Interventions and Supports (PBIS; [6–8]), a multi-tiered, problem-solving, and teambased continuum of supports that promotes all students' social, emotional, and behavioral development [8]. Studies indicate that fidelity of PBIS delivery is highly variable across schools and often falls below levels associated with improvements in student functioning [9]. Unfortunately, variable fidelity attenuates the impact of even the most efficacious programs [10] and results that are rarely leveraged to benefit the broader population [11].

Prior research suggests that organizational factors at the level of the school building are associated with successful implementation of EBPs [12–19]. In particular, implementation leadership (i.e., proactive leader behaviors that facilitate EBP use; [20]) and implementation climate (i.e., educator perceptions of whether EBP use is valued, expected, rewarded by the school; [21]) have been repeatedly linked to higher EBP fidelity in elementary schools [13, 17, 22, 23]. Unfortunately, few effective interventions to enhance EBP implementation in schools address these factors and those that do have only been evaluated qualitatively [24].

Helping educational leaders mobilize evidence (HELM)

To fill the gap in implementation strategies that address organizational factors within schools, our research team recently adapted the evidence-based Leadership and Organizational Change for Implementation (LOCI; [25-30]) strategy for use in the education sector. LOCI is informed by the Exploration, Preparation, Implementation, Sustainment (EPIS) framework, which details influences on implementation success over multiple phases [28]. We used a human-centered design framework [29] to enhance LOCI's acceptability, feasibility, contextual appropriateness, usability, and effectiveness for public schools [30, 31]. The redesigned strategy – Helping Educational Leaders Mobilize Evidence (HELM) - aims to improve principals' use of implementation leadership to support the high-fidelity delivery of EBPs that improve child outcomes [32]. HELM was designed so that it can be flexibly applied to support the implementation of any universal EBP.

The HELM theory of change (Fig. 1) model depicts the HELM core components, hypothesized organizational mechanisms, implementation outcomes, and student outcomes. Identification of implementation mechanisms is critical to developing effective and streamlined



Fig. 1 HELM Theory of Change

implementation strategies [27–29, 33–36]. The theory of change posits that HELM will improve school implementation leadership and climate [20, 21], which leads to higher-fidelity implementation of PBIS, which then leads to improved student social, emotional, and behavioral outcomes such as disciplinary referrals [15].

Preliminary HELM studies

Previous research found that implementation leadership and climate are malleable constructs and that changes in implementation leadership due to intervention contribute to improvements in implementation climate, EBP adoption, and observed fidelity [22, 37, 38]. Results from these studies validate key organizational mechanisms in HELM's theory of change.

HELM was iteratively developed using the Discover, Design/Build, Test (DDBT) framework which leverages human-centered design and implementation science to guide adaptation of complex interventions for new users or settings [32]. The initial Discover and Design/Build phases involved 1) focus groups (N=54 educators; 34); 2) expert input (from N=15 implementation researchers and school practitioners) using a nominal group decision making process and "hackathon" solution generation (34); and 3) the Cognitive Walkthrough for Implementation Strategies (CWIS: [39]) method to evaluate HELM usability (i.e., the extent to which an intervention can be used by specified users with effectiveness, efficiency, and satisfaction; N=15 principals; 30). These activities informed iterative adaptations to improve the suitability of HELM for the school context, such as aligning assessment windows to school calendars and prioritizing former school leaders as HELM coaches [31]. Furthermore, a pilot study examined the feasibility of HELM and research procedures for testing it in a large-scale Test phase trial [32]. HELM schools had significantly better implementation leadership and implementation climate (HELM's organizational mechanisms of change) over time than control schools. Students in HELM schools also demonstrated significant increases in positive behavior compared to those in control schools, such as the extent to which their classroom teachers rated them as following directions the first time when asked (Cohen's $f^2=0.84$, a large effect size).

PBIS

The current project aims to test whether HELM can yield school-wide improvement in implementation outcomes and student outcomes when implementing PBIS as an exemplar EBP. PBIS organizes support across multiple tiers of intensity that vary based on the level of student need. The three tiers are: 1) Tier 1: Primary Prevention, where all students receive universal supports to increase social, emotional, and behavioral outcomes to enhance academic success; 2) Tier 2: Secondary Prevention, where some students who have not been successful with Tier 1 support alone and are at an elevated risk for problems receive supplemental support to prevent more challenging behaviors; and 3) Tier 3: Tertiary Prevention, where a few students at high risk or experiencing significant challenges receive individualized support to reduce severity (www.pbis.org). PBIS has demonstrated consistent evidence for its effects including reduced student problem behavior and improved social, emotional, and behavioral functioning in elementary schools [40-44]. PBIS is a school-wide program that integrates: 1) student and school data to identify needed supports; 2) using classroom- and student-level EBPs; and 3) implementing a systems approach designed to support fidelity. In PBIS, the system is the infrastructure (e.g., operational supports to use data) to support educators to successfully implement EBPs, and leadership is an integral part of this system.

Study purpose and aims

The purpose of this study is to test the effectiveness of HELM on school-wide implementation of PBIS (i.e., fidelity) relative to a PBIS training and technical assistance condition. Experimental and exploratory aims are:

Aim 1: Experimentally evaluate the effects of HELM.

Hypothesis 1a. Schools randomized to HELM will demonstrate higher PBIS fidelity compared to schools randomized to PBIS training and technical assistance only (control schools).

Hypothesis 1b. Schools randomized to HELM will demonstrate greater improvement in school context factors that support implementation (i.e., implementation leadership, implementation climate) compared to control schools.

Hypothesis 1c. Students in schools randomized to HELM will demonstrate improved social, emotional, and behavioral and academic outcomes compared to students in control schools.

Aim 2: Explore for whom, under what conditions, how equitably, and through which processes HELM works to improve outcomes, as well as its cost-effectiveness.

Research Question 2a. Are the effects of HELM on PBIS fidelity mediated by improvement in organizational mechanisms of change (i.e., implementation leadership and climate)?

Research Question 2b. Are the effects of HELM on student outcomes mediated via implementation outcomes (i.e., PBIS fidelity)?

Research Question 2c. Are the effects of HELM consistent and equitable across buildings or are they moderated by school contextual factors (e.g., school size, student body racial/ethnic diversity, free/reduced price lunch)?

Research Question 2d. What explains implementation success in schools where HELM's explanatory model does not fit? What factors outside of the hypothesized mechanisms explain implementation success (i.e., high fidelity) in schools where HELM's explanatory model does fit?

Research Question 2e. What are the costs and costeffectiveness of HELM vs. PBIS implementation as usual?

Method

This study will use a hybrid type 3 effectiveness-implementation trial with a standard PBIS training and technical assistance only control condition to provide a rigorous test of the effects of HELM in elementary schools.

Participants

Participating schools (N=42) will be recruited into three cohorts (n=14 schools p/cohort) from districts located in Washington state using a stratified recruitment approach. Schools will be compensated \$400 total to support time for recruitment, retention, and other research activities. Inclusion criteria for districts are: 1) presence of a district-wide commitment to implement PBIS in elementary schools, as evidenced by either having completed a PBIS training in the past two years from certified coaches or a willingness to complete the training at the outset of study participation; 2) data sharing agreements with the project investigative team (including sharing of administrative data regarding students); 3) commitment to help recruit 10 educators per elementary school to complete study assessments (N=420); and 4) no previous HELM exposure. Principals and leadership teams will participate in HELM for schools assigned to that condition. The research team will invite all educators from the school to participate in surveys and PBIS trainings and contacted by email or phone to complete informed consent. We will gather de-identified school administrative data for students in each classroom with no identifiable information and, therefore, active parental permission is not needed.

Randomization. Random assignment will occur at the school building level because both HELM and PBIS are building-level processes. The study methodologist will generate the randomization sequence, with allocation concealed from all other study personnel and participants. We will randomize schools within cohorts with equal probability (1:1) to: 1) PBIS+HELM; or 2) standard PBIS training and technical assistance. We will use covariate constrained randomization that accounts for degree of prior PBIS implementation and school size [45, 46]. Covariate constrained randomization enumerates a large number of possible assignments of the interventions to schools and quantifies the balance across arms with regard to a set of prespecified covariates (i.e. degree of prior PBIS implementation and school size). From a subset of possible assignments that achieve adequate balance, one is randomly chosen as the final allocation of interventions for the study.

Implementation strategies

HELM. Principals, their distributed leadership teams (DLT), and school district-level leaders (e.g., Elementary Education Director, Director of Student Learning, etc.) will participate in HELM. HELM is a 9-month, data-driven organizational and leadership implementation strategy that entails eight core components: 1) Assessment and Feedback. 360° surveys measuring

implementation leadership and climate are administered to principals and educators at three time points. These data will be synthesized into a detailed feedback report, which will be shared with the DLT and used to create a tailored leadership development plan to support implementation coaching throughout the year. 2) Initial Training. A 4-h didactic and interactive training will be provided to principals, their DLT, and district-level leaders and cover developing strategic implementation leadership behavior and building a positive EBP implementation climate in their schools. 3) Leadership Development Plan. During the initial training, principals and their DLT work individually with their HELM coach to review their 360° assessment data and develop goals for improving implementation leadership and climate. 4) Individual Coaching. HELM coaches provide monthly 1-h coaching sessions in person or via Zoom to review progress and update the leadership development plan. The coaching structure includes reflective questions about 1) broader school updates (10 min); 2) EBP implementation (10 min); 3) Leadership Development Plan progress (20 min); 4) barriers to implementation and solution generation (10 min); 5) next steps (5 min); and 6) "what other support is needed" from the district and/ or HELM coaches (5 min). 5) Group Coaching. Coaches offer optional monthly 1-h group coaching calls with all HELM principals and DLT in each Cohort to review progress and share strategies across schools for idea generation and implementation support (this component is optional because there were mixed results in our pilot where some schools found Group Coaching helpful and others felt it was burdensome). 6) Organizational Strategy Development. Two 1-h meetings with district-level leaders are held, one in Fall semester and one in Spring semester of study enrollment to develop and update an organizational Climate Development Plan. This meeting will provide a structured discussion of alignment between school-level and district efforts to support EBP implementation. 7) Professional Learning Collaboratives. Two professional learning collaboratives are held with principals and their DLT to review content (align HELM strategies with principles from the National Educational Leadership Standards and EBP sustainment for the following school year) and share strategies across participants. 8) Graduation. During graduation, principals' and their DLT team's final feedback is reviewed, and progress for the past year is celebrated. Results from the pilot suggest HELM is feasible to deliver, meets the needs of school leaders, and allows school leaders to support EBP use.

PBIS training. Standard PBIS implementation includes initial training and technical assistance (i.e., booster training, fidelity monitoring, and coaching) – cornerstone

implementation strategies across programs and domains [47] – which all participating schools will receive regardless of condition. Implementation Coaches will help schools form internal school implementation teams comprised of five to six members (e.g., teachers, administrators, community mental health providers, families). Implementation Coaches will provide ongoing support for the schools throughout the school year, inclusive of four booster training events (approximately 4–6 h) that will use progress monitoring data to determine specific skill building during the school year for all PBIS school implementation teams.

Research procedures

Focus Groups. To understand "hypothesis defying residuals" (i.e., schools where implementation leadership and climate are inconsistent with their documented implementation outcomes), schools whose observed change in fidelity from T1 to T3 is greater than 0.5 standard deviation away from their predicted change will be identified for invitation to a focus group, balanced between users and non-users and HELM and control conditions. School-level focus groups with up to 10 participants (approximately 45-60 min) will be conducted at a convenient time for identified educators via Zoom and audio recorded. Participants will be paid \$100 for their time. Recordings will be transcribed prior to coding. The mixed methods design will be sequential in structure (quantitative data collected prior to qualitative data); the functions are sampling (using quantitative data to identify our qualitative sample) and expansion (using qualitative data to provide depth and breadth of understanding of the factors that contribute to implementation outcomes that deviate from our theory of change; i.e., QUAN+QUAL); and the process is connecting (the qualitative dataset will build on the quantitative dataset; 49).

Measures

Educators (principals/administrators, teachers, paraeducators) will complete secure web-based surveys via REDCap at three time points (Table 1). Time points will be baseline (April/May; T1) before the academic year in which schools will receive HELM, winter (Jan/Feb; T2, 8–9 months after baseline), and spring of the subsequent year (April/May; T3, 12 months after baseline). Educators will self-report their demographic characteristics, organizational mechanisms of change (implementation leadership and climate), implementation time and costs, and deidentified classroom-level outcomes. Educators will be compensated \$40 for the completion of study instruments at each data collection timepoint.

Table 1 Study Measures

Construct	Measure Description	Source	Time
School demographics	School size, student body racial/ethnic composition, % free/reduced lunch	А	T1
	Organizational Mechanisms		
Implementation Leadership	<u>School Implementation Leadership Scale</u> (SILS; 15): Adapted from the original ILS [20], the SILS has 24 items loading onto 8 subscales: Proactive, Knowledgeable, Supportive, Perseverant, Communication, Vision, Available, and Distributed Leadership. Subscale internal consistencies range from 0.91 to 0.96, and scores correlate with other leadership measures	P DLT T	T1-T3
Implementation Climate	<u>School Implementation Climate Scale</u> (SICS; [48]): Adapted from the original ICS [21] includes 21 items loading onto 7 subscales: Focus on EBP, Educational Support for EBP, Recognition for EBP, Rewards for EBP, Use of Data to Support EBP, Existing Supports for EBP, and EBP Integration, with good internal consistency estimates (range: 0.81–0.90)	P DLT T	T1-T3
Burnout	<u>Maslach Burnout Inventory</u> (MBI; [49]): Burnout will be measured using the 22-item MBI. Respondents will use a 7-point Likert ("0-Never" to "6-Every Day") to respond to items across three subscales: nine items measure emotional exhaustion (e.g., "I feel used up at the end of my workday"), five items measure depersonalization (e.g., "I feel I treat some students as if they were impersonal objects"), and eight items measure personal accomplishment (e.g., "I deal very effectively with the problems of my students"). Internal consistency was accept- able for burnout (α = .89), emotional exhaustion (α = .91), and personal accomplishment (α = .83). The lower estimate for depersonalization (α = .69) is consistent with extant research using the MBI-ES [50]	P DLT T	Τ1,Τ3
Implementation Citizenship Behavior	Implementation Citizenship Behavior Scale (ICBS; [51]): The School Implementation Citizenship Behavior Scale is a 12-item scale with four subscales. This study will evaluate 6 items loading onto 2 subscales: Helping Others and Keeping Informed. Research demonstrates acceptable reliability (α s = .8892) for all subscales	P DLT T	T1-T3
Coordination	<u>Team Process Scale (TPS): Coordination</u> [52]: Team-related processes will be measured using the 5-item Coordination response scale of the TPS. Respondents will use a 5-point Likert ("0=not at all to 5=to a very great extent") to report the extent to which they perceive their team engages in effective team processes. Research supports the content and construct validity of several versions of this scale, including the 5-item Coordination response scale.	P DLT	T1-T3
Change Fatigue	<u>Change Fatigue</u> [53]: This 6-item validated measure will be used to measure the impact of organizational change on employee outcomes (e.g. well-being, withdrawal, organizational commitment, turnover intentions).	P DLT T	T1,T3
	Implementation Outcomes		
PBIS Fidelity	The <u>Tiered Fidelity Inventory</u> (TFI; [54]) is a 45-item tool used to provide a valid, reliable, and efficient measure of the extent to which school personnel are applying the core features of PBIS. The TFI is a group-assessment completed by a schoolwide system planning team with external facilitation with strong internal consistency (α = .96). The TFI will be collected for both PBIS + HELM and PBIS + IAU conditions across the three time points	P DLT T O	T1-T3
Implementation Cost	Costs of 1) delivering HELM to augment PBIS implementation, as well as 2) IAU, will be cal- culated using activity-based cost metrics. Inputs will include time, supplies, travel, overhead, and costs associated with HELM training/coaching meetings, including pre-work, schedul- ing, and attending meetings; as well as costs associated with PBIS training and delivery in each study condition.	A P DLT T	T1-T3
	Behavioral and Academic Outcomes		
Academic engagement; Prosocial behavior; & Problem behavior	<u>Modified Direct Behavior Rating (DBR)</u> will be used in which teachers observe and rate their classroom's overall academic engagement (i.e., actively or passively participating in the classroom activity, 0–100%), disruptive behavior (i.e., interrupting activities, 0–100%), and prosocial behavior (i.e., following directions, 0–100%; Chafouleas et al., 2012). The DBR is sensitive to behavior change [55, 56] and is more feasible than assessing individual students.	Т	T1-T3
Educational and Behavioral Outcomes	Standardized academic test scores, attendance rates, and disciplinary incidents (office discipline referrals, suspensions, expulsions) will be derived from administrative records	A	T1 -T3
	HELM Manipulation Check		
HELM Fidelity	Coding of video recordings of HELM delivery	0	T1 -T3

Note: A Administrative data, P Principal, DLT distributed leadership team, T Teacher, O Trained observer; T1 = Baseline (Apr/May prior to HELM/PBIS delivery year), T2 = Jan/Feb, 9 months after baseline, T3 = Apr/May, 12 months after baseline

Fidelity. The study's primary outcome is school-level PBIS fidelity, assessed using the Tiered Fidelity Inventory (TFI; [54]). Facilitated by an expert PBIS coach,

school PBIS implementation teams will complete the TFI in both the HELM and control conditions. The TFI is designed to be used 1) for initial assessment to

determine the degree to which a school is using PBISconsistent practices; 2) as a guide for implementation of Tiers 1, 2, and 3 EBPs, and 3) to track PBIS implementation over time. TFI data will be collected at the same time points as educator data.

HELM fidelity checklist. This checklist will be used as a manipulation check to document HELM delivery (based on observations of recorded HELM trainings). HELM coaches will complete a standardized measure of dates specific steps were completed.

Qualitative Focus Groups. We will develop a systematic, comprehensive semi-structured focus group guide that draws from the EPIS framework to examine multilevel (i.e., intervention, individual, inner, and outer settings) determinants that explain what processes facilitated or hindered implementation [28]. We will generate questions that explore the most salient implementation determinants and mechanisms, and how implementation leadership and climate may interact with other relevant characteristics of the setting [57, 58].

Administrative Data. At the end of each school year, deidentified academic records by classroom will be requested for all participating schools to extract students' attendance, discipline (office disciplinary referrals), and achievement (grades, standardized test scores). To assist in modeling, these data will be collected retrospectively for the year prior to HELM participation as well as each cohort's HELM year.

Cost Assessments. Activity-based costing will be used to comprehensively estimate the costs of HELM and the comparison condition [59, 60]. We will measure cost from the payor (i.e., school system) perspective, since the primary costs and associated decision-making are within the implementing school district [61]. We will isolate costs of HELM as an implementation strategy, but since HELM may have secondary impacts on PBIS costs – e.g., by increasing educator engagement in implementation – we will measure and monetize differences between study conditions in costs related to teachers' participation in PBIS training, consultation, and delivery.

We will include open-ended items in each survey that ask about unexpected resources needed for HELM and/ or PBIS. Using a sequential qual \rightarrow QUANT development function, we will rapidly analyze responses on an ongoing basis so that we can immediately incorporate any newly identified cost categories into future surveys for quantitative measurement [62, 63]. To assign cost values to categories that are measured in non-monetary values (e.g., trainer time delivering HELM, teacher time spent on PBIS), we will use other data sources such as school district records or project expense reports.

Data analysis

Power Analysis. The planned sample, including attrition, will provide sufficient power to test the primary hypotheses regarding the direct effects of HELM on school-, educator/classroom-, and student-level outcomes, assuming small to medium minimum detectable effect sizes of 0.23 to 0.65. These effect sizes are reasonable, likely, and clinically meaningful, based on prior research, our pilot trial, and standard interpretations of effects sizes for implementation strategies (Williams NJ, Ehrhart MG, Aarons GA, Esp S, Sklar M, Carandang K, Vega NR, Brookman-Frazee L, Marcus SC: Increasing fidelity to measurement-based care in youth mental health through improved organizational leadership and focused implementation climate: A process evaluation within a randomized trial, in preparation) [64-66]. Power analyses were conducted using Power Analysis and Sample Size Software [67]. They account for clustering of timepoints within educators/classrooms within schools (as applicable, based on the analysis), and assume final samples of 42 schools, 10-12 educators/classrooms per school (as applicable and after accounting for attrition), ICCs at the school level of 0.1 to 0.35 (depending on the outcome and as consistent with prior research), within-subjects correlations of timepoints equivalent to 50% of the variance in posttests explained by pretests (consistent with prior research), and three timepoints [68-71]. We will enroll at least 42 schools (1 extra per condition) to address potential concerns about attrition.

Anticipated statistical power for our mediation and moderation analyses vary by outcome and analytic model; however, they generally align with HELM's anticipated effects and the effects of potential moderators based on prior research [22, 38, 72]. Making similar assumptions as above where applicable, and assuming a two-mediator, serial mediation model (Fig. 2), our sample has adequate power to detect large effect sizes of 0.9 for the paths from HELM to implementation leadership and from implementation leadership to implementation climate, and a medium effect size of 0.49 for the path from implementation climate to PBIS fidelity. Minimum detectable effect sizes for analyses testing moderators of HELM's effects on PBIS fidelity and on student outcomes range from small to medium.

Data Analytic Approach. We will explore for baseline equivalence between conditions on all school, teacher, and student variables following Institute of Educational Sciences guidelines, with effect sizes between 0.05 and 0.25 indicating that statistical adjustment for nonequivalent baseline characteristics is required [73]. Non-equivalent baseline characteristics will be included as covariates at the school, educator, or student level, as appropriate. Following best practice guidelines, variables used for



covariate constrained randomization will be included as covariates in all models [46, 74]. Data missing at random will be modeled using full information maximum likelihood estimation in mixed effects modeling or multiple imputation for other analyses. Although all analyses will use an intent-to-treat approach, we will examine the robustness of condition assignment through descriptive analyses of scores on the HELM fidelity tools. Schools that achieve 80% or higher of the maximum possible score on all HELM fidelity criteria will be considered having received a full dose of HELM training and coaching.

All analyses will use an intent-to-treat approach in which units are analyzed based on condition assignment regardless of HELM implementation success [75]. Because the trial includes outcomes at multiple levels, the specific analytic approach will vary depending on the outcome. However, in general, analyses will employ 2- or 3-level mixed effects models reflecting data collection time points nested within educator/classroom nested within schools. We will test for significantly large ICCs at the district levels to determine if statistical nesting is necessary. Standard model-building procedures will be used [76, 77], including fitting a null model to facilitate calculation of variance accounted for in later models, fitting unconditional growth models (e.g., linear, quadratic, piecewise) to determine the optimal functional form for time, and finally, fitting models that include the variable for condition (and the condition by time interaction, as appropriate) along with covariates. Iterative models with possible covariates will be tested. Covariates not significantly contributing to the model at p < 0.10 based on likelihood ratio tests will be removed. We will obtain estimates of whether there were statistically significant differences between conditions on rate of change over time (i.e., slope), and whether there are statistically significant condition differences in average score on each outcome at T2 and T3. Models will be generalized, with appropriate link functions (e.g., log-link, Poisson) applied based on distributional form (e.g., dichotomous, zero-inflation). Estimations will be fit using full maximum likelihood. Models will be assessed for possible violations of assumptions. Inference will be evaluated relative to p < 0.05. For hypotheses/research questions with multiple DVs, we will adjust for familywise error using Benjamini and Hochberg's False Discovery Rate [78].

Qualitative Analysis. Certain codes will be conceptualized during the protocol guide development and driven by the EPIS framework (i.e., deductive approach) and others will be developed through reading an initial subset of transcripts (i.e., inductive approach). Themes will provide a way of understanding the most salient factors that impact implementation and extend beyond the existing HELM organizational mechanisms and theory of change [79, 80]. EPIS-driven directed coding will include system levels (i.e., intervention, individual, inner setting, outer setting) as initial "parent nodes." After a stable set of codes is developed, a consensus process will be used in which all reviewers independently code and compare their coding to arrive at consensus judgments through open dialogue [81–83].

Cost and Cost-Effectiveness Analyses. We will use the CostOut program to complete the cost analysis for HELM and PBIS implementation-as-usual. CostOut specifies ingredients for each intervention condition, assigns prices (national and user-inputted local values), and calculates costs based on the units per ingredient used [59, 84]. We will use CostOut to generate descriptive statistics describing typical costs (i.e., means, standard deviations) for HELM and PBIS. We will calculate total costs for each condition, and incremental costs of HELM over PBIS training and technical assistance only. We also will provide cost breakdowns to help administrators understand the budget implications of HELM, including personnel versus other direct expenses and start-up versus maintenance costs. Finally, we will conduct sensitivity analyses to examine the robustness of our cost estimates by identifying areas of uncertainty in measuring the units

and prices for our ingredients, and then calculating costs across a range of plausible values [85, 86].

We also will use CostOut to calculate the cost-effectiveness of HELM versus PBIS [87]. This will involve calculating a series of incremental cost-effectiveness ratios for each student outcome (social, emotional, behavioral, academics) and PBIS implementation outcome (fidelity). Finally, we will again use sensitivity analyses to examine the robustness of our cost-effectiveness results, both across the ranges of costs examined and across plausible effectiveness estimates (i.e., 95% CIs for student and implementation outcomes).

Discussion

One out of 5 elementary students exhibit social, emotional, and behavioral difficulties that hinder their academic success which are linked to later negative outcomes, including substance use problems, unemployment, houselessness, and contact with the legal system [1, 88, 89]. Despite the promise of EBPs like PBIS to promote student social, emotional, and behavioral functioning, their routine use in schools is limited, reducing their large-scale impact on student outcomes [90, 91]. The majority of schools in the United States are attempting to implement multi-tiered frameworks such as PBIS, yet evidence suggests that implementation in educational settings is typically absent, inconsistent, or incomplete [91, 92]. School leadership plays a pivotal role in the successful implementation of EBPs that effectively reduce social, emotional, and behavioral outcomes [13, 93]. The presence or absence of strong leadership support in schools can make the difference between successful implementation and abandonment. However, leadership is not specifically targeted in standard PBIS training [93].

Research on leadership-focused implementation strategies have almost exclusively focused on outpatient mental health clinics or clinics focused on treating substance use disorder [25-27, 27-31] and has not been widely evaluated in public schools. There is some preliminary evidence to suggest HELM positively impacts implementation leadership, implementation climate, and implementation citizenship and may potentially buffer the decline in EBP implementation efforts that naturally occurs over the school year [94]. This study will be the first full-scale test of HELM's efficacy in public schools, which is a very different service setting with unique organizational characteristics, systems, and processes. It is important to test whether the type of leadershipfocused training and climate development that worked in outpatient mental health clinics or clinics focused on treating substance use disorder also can work in public schools.

HELM has the potential to reduce the substantial waste in time and resources resulting from ineffective universal EBP implementation, including inadequate uptake and low fidelity. HELM is closely aligned with the needs and priorities of educators working in educational settings due to its applied focus and emphasis on strategic implementation behaviors. Demonstrating the effectiveness of HELM on PBIS fidelity will address a highly prevalent barrier to improved population health: abandonment or low-fidelity delivery of effective interventions focusing on social, emotional, and behavioral functioning in schools. Because HELM was developed and tested with school partners, it has the potential to be highly usable and scalable, which has significant implications for low-resource community contexts in which lay providers deliver services. A unique aspect of HELM is that it emphasizes the engagement of DLT in schools, which is a common leadership model in schools but not in outpatient mental health clinics or clinics focused on treating substance use disorder. As of February 2025, no participants have been enrolled.

Limitations

We recognize that it will be difficult to identify school districts that have not engaged in some level of PBIS implementation. However, for maximum generalizability and educational impact, HELM needs to be able to support both initial implementation of an EBP as well as improve implementation efforts that are already in process. Schools with variable levels of PBIS implementation will allow us to examine this effect.

Abbreviations

CWIS	Cognitive Walkthrough for Implementation Strategies
DBR	Direct Behavior Ratings
DDBT	Discover, Design/Build, Test
DLT	Distributed Leadership Teams
EBP	Evidence-based practice
EPIS	Exploration, Preparation, Implementation, Sustainment
ES	Effect Sizes
HELM	Helping Educational Leaders Mobilize Evidence
ICBS	Implementation Citizenship Behavior Scale
IES	Institute of Education Sciences
LOCI	Leadership and Organizational Change for Implementation
MBI-ES	Maslach Burnout Inventory – Educators Survey
MDES	Minimum Detectable Effect Sizes
ML-PA	Multilevel path analysis
MLR	Maximum likelihood estimator
PASS	Power Analysis and Sample Size Software
PBIS	Positive Behavioral Interventions and Supports
REDCap	Research Electronic Data Capture
TFI	Tiered Fidelity Inventory
TPS	Team Process Scale

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13012-025-01429-4.

Supplementary Material 1.

Acknowledgements

We are grateful for the support and collaboration from our school district partners.

Authors' contributions

JL is the principal investigator of this study, generated the idea and designed the study.

ARL is the co-principal investigator. Authors NJW, MGE, and AD are co-investigators and supported the conceptualization of the study. JL, NJW, AS, MT, and ARL were the primary writers of the manuscript and approved all changes. NJW, AD, JL,and ARL drafted all study analyses in the manuscript. Authors BR, CE, KS, KH, LB are core contributors to the research study in terms of protocol and school-based teams development and have provided input into the design of the study. All authors were involved in developing, editing, reviewing, and providing feedback for this manuscript and have given approval of the final version to be published.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This study was previously reviewed and is funded by the Institute of Education Sciences (grant #: R305A240030) in the amount of \$1,399,987. The content is solely the responsibility of the authors and does not necessarily represent the view of the funder.

Data availability

The application described in this manuscript is freely available. Please contact the lead author for more information.

Declarations

Ethics approval and consent to participate

The University of Washington Institutional Review Board approved this study; Study No. 00020559.

Consent for publication

Not applicable.

Competing interests

All Authors declare that there is no conflict of interest.

Author details

¹ University of Washington, Box 357920, Seattle, WA 98115, USA. ²Boise State University, 1910 W. University Dr, Boise, ID 83725, USA. ³University of Central Florida, Pictor Lane, Orlando, FL, USA. ⁴RAND, 1776 Main Street, Santa Monica, CA 90401, USA. ⁵School of Education/Educational Administration, University of Washington Tacoma, Campus, Box 358435, Tacoma, WA 98402, USA. ⁶Educational Service District 105, 33 S 2nd Ave, Yakima, WA 98902, USA.

Received: 7 February 2025 Accepted: 31 March 2025 Published online: 24 April 2025

References

- Copeland WE, Alaie I, Jonsson U, Shanahan L. Associations of childhood and adolescent depression with adult psychiatric and functional outcomes. J Am Acad Child Adolesc Psychiatry. 2021. https://doi.org/10. 1016/j.jaac.2020.07.895.
- Meade J. Mental health effects of the COVID-19 pandemic on children and adolescents: a review of the current research. Pediatr Clin. 2021;68:5.
- McIntosh CE, Stone GE. Introduction to the special issue: How COVID-19 has affected students' health, achievement, and mental health. Psychol Sch. 2023. https://doi.org/10.1002/pits.22820.
- Cipriano C, Naples LH, Zieher A, Durlak J, Eveleigh A, Funero M, Chow J. The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. Child Dev. 2023;1181–204.
- 5. Goldberg JM, Sklad M, Elfrink TR, Schreurs KM, Bohlmeijer ET, Clarke AM. Effectiveness of interventions adopting a whole school approach to

enhancing social and emotional development: A meta-analysis. Eur J Psychol Educ. 2018. https://doi.org/10.1007/s10212-018-0406-9.

- Horner RH, Sugai G. School-wide PBIS: An example of applied behavior analysis implemented at a scale of social importance. Behav Anal Pract. 2015. https://doi.org/10.1007/s40617-015-0045-4.
- Eber L, Barrett S, Perales K, Jeffrey-Pearsall J, Pohlman K, Putnam R, Splett J, Weist MD. Advancing education effectiveness: Interconnecting school mental health and school-wide PBIS, Volume 2: An implementation guide. Center for Positive Behavior Interventions and Supports. Eugene, Oregon: University of Oregon Press; 2019.
- Sugai G, Horner RH, Dunlap G, Hieneman M, Lewis TJ, Nelson, Scott T, Liaupsin C, Sailor W, Turnbull AP, Turnbull HP, Wickham D, Wilcox B, Reuf M. (2000). Applying positive behavior support and functional behavioral assessment in schools. Journal of Positive Behavior Interventions. 2000; https://doi.org/10.1177/10983007000200302
- Pas ET, Johnson SR, Debnam KJ, Hulleman CS, Bradshaw CP. Examining the relative utility of PBIS implementation fidelity scores in relation to student outcomes. Remedial and Special Education. 2019. https://doi. org/10.1177/0741932518805192.
- Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol. 2008. https://doi. org/10.1007/s10464-008-9165-0.
- 11. National Research Council (US) and Institute of Medicine (US) Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults Research Advances and Promising Interventions. Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities. M. E. O'Connell & K. Warner, Eds. National Academies Press. 2009
- Locke J, Beidas RS, Marcus S, Stahmer A, Aarons GA, Lyon AR, Cannuscio C, Barg F, Dorsey S, Mandell DS. A mixed methods study of individual and organizational factors that affect implementation of interventions for children with autism in public schools. Implement Sci. 2016. https://doi. org/10.1186/s13012-016-0501-8.
- Locke J, Lawson GM, Beidas RS, Xie M, Aarons GA, Spaulding C, Seidman M, Oh C, Frederick LK, Mandell DS. Individual and organizational factors that affect implementation of evidence-based practices for children with autism in public schools. Implement Sci. 2019;14:39.
- Williams N, Frank H, Frederick L, Beidas R, Mandell DS, Aarons GA, Green P, Locke J. Organizational culture and climate profiles: Relationships with fidelity to three evidence-based practices for autism in elementary schools. Implement Sci. 2019;14:15.
- Lyon AR, Corbin CM, Brown EC, Ehrhart MG, Locke J, Davis C, Picozzi E, Aaron GA, Cook CR. Leading the charge in the education sector: Development and validation of the school implementation leadership scale (SILS). Implement Sci. 2022. https://doi.org/10.1186/s13012-022-01222-7.
- Lyon AR, Cook CR, Brown EC, Locke J, Davis C, Ehrhart M, Aarons GA. Assessing organizational implementation context in the education sector: Confirmatory factor analysis of measures of implementation leadership, climate, and citizenship. Implement Sci. 2018;13:1.
- Corbin CM, Hugh ML, Ehrhart MG, Locke J, Davis C, Brown EC, Cook CR, Lyon AR. Teacher perceptions of implementation climate related to feasibility of implementing schoolwide positive behavior supports and interventions. Sch Ment Heal. 2022. https://doi.org/10.1007/ s12310-022-09528-z.
- Corbin C, Lyon A, Collins VK, Ehrhart MG, Goosey R, Locke J. The incremental association of implementation leadership and school personnel burnout beyond transformational leadership. Sch Psychol. 2024:269–79.
- Zhang Y, Cook C, Fallon L, Corbin C, Ehrhart M, Brown E, Locke J, Lyon A. The interaction between general and strategic leadership and climate on their multilevel associations with implementer attitudes toward universal prevention programs for youth mental health: A crosssectional study. Adm Policy Ment Health. 2023. https://doi.org/10.1007/ s10488-022-01248-5.
- Aarons GA, Ehrhart MG, Farahnak LR. The implementation leadership scale (ILS): Development of a brief measure of unit level implementation leadership. Implement Sci. 2014. https://doi.org/10.1186/1748-5908-9-45.
- Ehrhart MG, Aarons GA, Farahnak LR. Assessing the organizational context for EBP implementation: The development and validity testing of the implementation climate scale (ICS). Implement Sci. 2014. https://doi.org/ 10.1186/s13012-014-0157-1.

- Williams NJ, Hugh ML, Cooney DJ, Worley J, Locke J. Testing a theory of implementation leadership and climate across autism evidencebased behavioral health intervention of varying complexity. Behav Ther. 2022a:900–12.
- Meza RD, Beidas RS, Ehrhart MG, Mandell DS, Dorsey S, Frederick L, Oh C, Locke J. Discrepancies and agreement in perceptions of implementation leadership: Associations with dosage of school-based evidence-based practices for children with autism. Adm Policy Ment Health. 2019;518–29.
- 24. Baffsky R, Ivers R, Cullen P, Wang J, McGillivray L, Torok M. Strategies for enhancing the implementation of universal mental health prevention programs in schools: A systematic review. Prev Sci. 2023;24:2.
- Aarons GA, Ehrhart MG, Farahnak LR, Hurlburt MS. Leadership and organizational change for implementation (LOCI): A randomized mixed method pilot study of a leadership and organization development intervention for evidence-based practice implementation. Implement Sci. 2015. https://doi.org/10.1186/s13012-014-0192-y.
- Aarons GA, Ehrhart MG, Moullin JC, Torres EM, Green AE. Testing the leadership and organizational change for implementation (LOCI) intervention in substance abuse treatment: A cluster randomized trial study protocol. Implement Sci. 2017. https://doi.org/10.1186/s13012-017-0562-3.
- Skar AMS, Braathu N, Peters N, Bækkelund H, Endsjø M, Babaii A, Borge RH, Wentzel-Larsen T, Ehrhart MG, Sklar M, Brown CH, Aarons GA, Egeland KM. A stepped-wedge randomized trial investigating the effect of the leadership and organizational change for implementation (LOCI) intervention on implementation and transformational leadership, and implementation climate. BMC Health Serv Res. 2022. https://doi.org/10. 1186/s12913-022-07539-9.
- Williams NJ, Marcus SC, Ehrhart MG, Sklar M, Esp S, Carandang K, Vega N, Gomes A, Brookman-Frazee L, Aarons GA. Randomized trial of an organizational implementation strategy to improve measurement-based care fidelity and youth outcomes in community mental health. J Am Acad Child Adolesc Psychiatry. 2024:991–1004.
- 29. Williams NJ, Ehrhart MG, Aarons GA, Esp S, Sklar M, Carandang K, Vega NR, Brookman-Frazee L, Marcus SC. Improving measurement-based care implementation in youth mental health through organizational leadership and climate: A mechanistic analysis within a randomized trial. Implement Sci. 2024;19:29.
- Aarons GA, Sklar M, Ehrhart MG, Roesch S, Moullin JC, Carandang K. Randomized trial of the Leadership and Organizational Change for Implementation (LOCI) strategy in substance use treatment clinics. J Subst Use Addict Treat. 2024;165.
- Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Administration and Policy in Mental Health and Mental Health Services Research. 2011. https://doi.org/10.1007/s10488-010-0327-7.
- Lyon AR, Munson SA, Renn BN, Atkins DC, Pullmann MD, Friedman E, Areán PA. Use of human-centered design to improve implementation of evidence-based psychotherapies in low-resource communities: Protocol for studies applying a framework to assess usability. JMIR Research Protocols. 2019. https://doi.org/10.2196/14990.
- Collins VK, Corbin CM, Locke JJ, et al. Centering School Leaders' Expertise: Usability Evaluation of a Leadership-Focused Implementation Strategy to Support Tier 1 Programs in Schools. Sch Ment Heal. 2024. https://doi.org/ 10.1007/s12310-024-09635-z.
- Locke J, Corbin C, Collins V, Ehrhart M, Lyon, A. Helping Educational Leadership Mobilize Evidence (HELM): The iterative redesign of the Leadership for Organizational Change for Implementation (LOCI) intervention for use in schools. Implement Res Pract. 2024;5.
- Lewis CC, Klasnja P, Powell BJ, Lyon AR, Tuzzio L, Jones S, Walsh-Bailey C, Weiner B. From classification to causality: Advancing understanding of mechanisms of change in implementation science. Front Public Health. 2018. https://doi.org/10.3389/fpubh.2018.00136.
- Williams NJ. Multilevel mechanisms of implementation strategies in mental health: Integrating theory, research, and practice. Adm Policy Ment Health. 2016:783–98.
- Williams NJ, Benjamin-Wolk C, Becker-Haimes EM, Beidas RS. Testing a theory of strategic implementation leadership, implementation climate, and clinicians' use of evidence-based practice: A 5-year panel analysis. Implement Sci. 2020;15:1.
- Williams NJ, Becker-Haimes EM, Schriger S, Beidas RS. Linking organizational climate for evidence-based practice implementation to observed

clinician behavior in patient encounters: a lagged analysis. Implementation Science Communications. 2022;3:64.

- Lyon AR, Coifman J, Cook H, McRee E, Liu FF, Ludwig K, Dorsey S, Koerner K, Munson SA, McCauley E. The cognitive walkthrough for implementation strategies (CWIS): A pragmatic method for assessing implementation strategy usability. Implementation Science Communications. 2021. https://doi.org/10.1186/s43058-021-00183-0.
- Bradshaw C, Reinke W, Brown L, Bevans K, Leaf P. Implementation of school-wide positive behavioral interventions and supports (PBIS) in elementary schools: Observations from a randomized trial. Educ Treat Children. 2008:1–26.
- Bradshaw C, Mitchell MM, Leaf PJ. Examining the effects of School-Wide Positive Behavioral Interventions and Supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. J Posit Behav Interv. 2010; 133–48.
- Childs KE, Kincaid D, George HP, Gage NA. (2016). The relationship between school-wide implementation of Positive Behavior Intervention and Supports and student discipline outcomes. Journal of Positive Behavior Interventions. 2016; https://doi.org/10.1177/1098300715590398
- Horner RH, Sugai G, Smolkowski K, Eber L, Nakasato J, Todd AW, Esperanza J. A randomized, wait-list controlled effectiveness trial assessing schoolwide positive behavior support in elementary schools. J Posit Behav Interv. 2009. https://doi.org/10.1177/1098300709332067.
- Horner RH, Sugai G, Anderson CM. Examining the evidence base for school-wide positive behavior support. Focus on Exceptional Children. 2010; https://journals.ku.edu/focusXchild/article/view/6906/6254
- Ivers NM, Halperin IJ, Barnsley J, Grimshaw JM, Shah BR, Tu K, Upshur R, Zwarenstein M. Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. Trials. 2012. https://doi.org/ 10.1186/1745-6215-13-120.
- Li F, Lokhnygina Y, Murray DM, Heagerty PJ, DeLong ER. An evaluation of constrained randomization for the design and analysis of group-randomized trials. Stat Med. 2016;35:10.
- Lyon AR, Pullmann MD, Walker SC, D'Angelo G. Community-sourced intervention programs: Review of submissions in response to a statewide call for "promising practices." Administration and Policy in Mental Health and Mental Health Services Research. 2017. https://doi.org/10.1007/ s10488-015-0650-0.
- Thayer AJ, Cook CR, Davis C, Brown EC, Locke J, Ehrhart MG, Aarons GA, Picozzi E, Lyon AR. Construct validity of the school-implementation climate scale. Implementation Research and Practice. 2022. https://doi. org/10.1177/26334895221116065.
- 49. Maslach C, Jackson SE, Leiter MP. Maslach Burnout Inventory Manual. 3rd ed. Mountain View, CA: CPP, Inc.; 1996.
- Schaufeli WB, Bakker AB, Hoogduin K, Schaap C, Kladler A. On the clinical validity of the maslach burnout inventory and the burnout measure. Psychol Health. 2001. https://doi.org/10.1080/08870440108405527.
- Ehrhart MG, Aarons GA, Farahnak LR. Going above and beyond for implementation: the development and validity testing of the Implementation Citizenship Behavior Scale (ICBS). Implementation Sci. 2015. https://doi. org/10.1186/s13012-015-0255-8.
- Mathieu JE, Luciano MM, D'Innocenzo L, Klock EA, LePine JA. (2020). The Development and Construct Validity of a Team Processes Survey Measure. Organizational Research Methods. 2020; https://doi.org/10.1177/10944 28119840801.
- Bernerth JB, Walker HJ, Harris SG. Change fatigue: Development and initial validation of a new measure. Work Stress. 2011. https://doi.org/10. 1080/02678373.2011.634280.
- Algozzine B, Barrett S, Eber L, George H, Horner R, Lewis T, Putnam B, Swain-Bradway J, McIntosh K, Sugai, G. School-wide PBIS tiered fidelity inventory. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. www.pbis.org, 2019.
- Chafouleas SM, Sanetti LMH, Kilgus SP, Maggin D. (2012). Evaluating sensitivity to behavioral change using direct behavior rating single-item scales. Exceptional Children. 2012; https://doi.org/10.1177/0014402912 07800406
- Cook CR, Zhang Y. Monitoring class-wide behavior outcomes in response to proactive classroom management strategies. National Association of School Psychologists Annual Conference 2015, Orlando, FL. 2015
- 57. Birken SA, Powell BJ, Presseau J, Kirk MA, Lorencatto F, Gould NJ, Haines E. Combined use of the consolidated framework for implementation

research (CFIR) and the theoretical domains framework (TDF): A systematic review. Implement Sci. 2017;12:1.

- Carey RN, Connell LE, Johnston M, Rothman AJ, de Bruin M, Kelly MP, Michie S. Behavior change techniques and their mechanisms of action: a synthesis of links described in published intervention literature. Ann Behav Med. 2019;53:8.
- Hollands FM, Pratt-Williams J, Shand R. Cost analysis standards & guidelines 1.0. Cost Analysis in Practice (CAP) Project. https://capproject.org/ resources, 2020.
- 60. Levin HM, McEwan PJ, Belfield C, Bowden AB, Shand R. Economic evaluation in education: Cost-effectiveness and benefit-cost analysis. SAGE publications; 2017.
- Eisman AB, Kilbourne AM, Dopp AR, Saldana L, Eisenberg D. Economic evaluation in implementation science: Making the business case for implementation strategies. Psychiatry Res. 2020; 283.
- 62. Hamilton AB, Finley EP. Qualitative methods in implementation research: An introduction. Psychiatry Res. 2019. https://doi.org/10.1016/j.psychres. 2019.112516.
- 63. Taylor B, Henshall C, Kenyon S, Litchfield I, Greenfield S. Can rapid approaches to qualitative analysis deliver timely, valid findings to clinical leaders? A mixed methods study comparing rapid and thematic analysis. BMJ Open. 2018;8:10.
- Lau R, Stevenson F, Ong BN, Dziedzic K, Treweek S, Eldridge S, Murray E. Achieving change in primary care—effectiveness of strategies for improving implementation of complex interventions: systematic review of reviews. BMJ Open. 2015;5:12.
- Baskerville NB, Liddy C, Hogg W. Systematic review and meta-analysis of practice facilitation within primary care settings. The Annals of Family Medicine. 2012;10:1.
- Grimshaw J, Eccles M, Thomas R, MacLennan G, Ramsay C, Fraser C, Vale L. Toward evidence-based quality improvement: Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966–1998. J Gen Intern Med. 2006;21:S2.
- 67. PASS Power Analysis and Sample Size Software. NCSS, LLC. Kaysville: ncss. com/software/pass; 2022.
- McIntosh K, Massar MM, Algozzine RF, George HP, Horner RH, Lewis TJ, Swain-Bradway J. Technical adequacy of the SWPBIS tiered fidelity inventory. J Posit Behav Interv. 2017;19:1.
- James AG, Noltemeyer A, Ritchie R, Palmer K, University M. Longitudinal disciplinary and achievement outcomes associated with school-wide PBIS implementation level. Psychol Sch. 2019;56:9.
- Zhu P, Jacob R, Bloom H, Xu Z. Designing and analyzing studies that randomize schools to estimate intervention effects on student academic outcomes without classroom-level information. Educational Evaluation and Policy Analysis. 2012; <u>doi/https://doi.org/10.3102/0162373711</u> 423786
- Hedges LV, Hedberg EC. Intraclass correlations and covariate outcome correlations for planning two-and three-level cluster-randomized experiments in education. Eval Rev. 2013;37:6.
- Williams NJ, Ramirez N, Esp S, Watts A, Marcus SC. Organization-level variation in therapists' attitudes toward and use of measurement-based care. Adm Policy Ment Health. 2022c:927–42.
- What Works Clearinghouse. 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. 2020.
- Li F, Turner EL, Heagerty PJ, Murray DM, Vollmer WM, DeLong ER. An evaluation of constrained randomization for the design and analysis of group-randomized trials with binary outcomes. Stat Med. 2017;36:24.
- Gupta SK. Intention-to-treat concept: A review. Perspect Clin Res. 2011;2:3.
- 76. Raudenbush SW, Bryk AS. Hierarchical linear models: Applications and Data Analysis Methods. Sage Publications; 2002.
- 77. Singer JD, Willett JB. Applied longitudinal data analysis: Modeling change and event occurrence. USA: Oxford University Press; 2003.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol). 1995;57:1.
- 79. Glaser BG, Strauss AL. Discovery of Grounded Theory: Strategies for Qualitative Research. Routledge; 2017.

- 80. Strauss AC, Corbin J. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Sage Publications; 1990.
- DeSantis L, Ugarriza DN. The concept of theme as used in qualitative nursing research. West J Nurs Res. 2000:351–72.
- Hill CE, Thompson BJ, Williams EN. A guide to conducting consensual qualitative research. Counseling Psychology. 1997. https://doi.org/10. 1177/0011000097254001.
- 83. Hill CE, Knox S, Thompson BJ, Williams EN, Hess SA, Ladany N. Consensual qualitative research: An update. J Couns Psychol. 2005;52:2.
- Hollands FM, Hanisch-Cerda B, Levin HM, Belfield CR, Menon A, Shand R, Pan Y, Bakir I, Cheng H. CostOut - the CBCSE cost tool kit. New York: Teachers College, Columbia University, Center for Benefit-Cost Studies of Education. Retrieved from www.cbcsecosttoolkit.org, 2015.
- Briggs AH, Gray AM. Handling uncertainty in economic evaluations of healthcare interventions. BMJ. 1999:635–8.
- Levin HM, Belfield C. Guiding the Development and Use of Cost-Effectiveness Analysis in Education. J Res Educ Eff. 2015:400–18.
- Palinkas LA, Horwitz SM, Chamberlain P, Hurlburt MS, Landsverk J. Mixedmethods designs in mental health services research: A review. Psychiatr Serv. 2011;62:3.
- Chang X, Jiang X, Mkandarwire T, Shen M. Associations between adverse childhood experiences and health outcomes in adults aged 18–59 years. PLoS ONE. 2019;14:2.
- Grattan RE, Tryon VL, Lara N, Gabrielian SE, Melnikow J, Niendam TA. Risk and resilience factors for youth homelessness in western countries: A systematic review. Psychiatr Serv. 2022. https://doi.org/10.1176/appi.ps. 202000133.
- Wilson DB, Gottfredson DC, Najaka SS. School-based prevention of problem behaviors: A meta-analysis. J Quant Criminol. 2001. https://doi.org/ 10.1023/A:1011050217296.
- 91. Evans SW, Weist MD. Commentary: Implementing empirically supported treatments in the schools: What are we asking? Clin Child Fam Psychol Rev. 2004. https://doi.org/10.1007/s10567-004-6090-0.
- Ringwalt CL, Vincus A, Ennett S, Johnson R, Rohrbach LA. Reasons for teachers' adaptation of substance use prevention curricula in schools with non-white student populations. Prev Sci. 2004. https://doi.org/10. 1023/B:PREV.0000013983.87069.a0.
- McIntosh K, Girvan EJ, Horner RH, Smolkowski K. Education not incarceration: A conceptual model for reducing racial and ethnic disproportionality in school discipline. Journal of Applied Research on Children: Informing Policy for Children at Risk. 2015. https://doi.org/10.58464/2155-5834. 1215.
- Locke J, Corbin CM, Goosey R, Collins VK, Ehrhart MG, Hatch K, Espeland C, Lyon AR. Not getting better but not getting worse: A cluster randomized controlled pilot trial of a leadership implementation strategy. Implementation Research and Practice. 2025. https://doi.org/10.1177/ 26334895241312405.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.