Not getting better but not getting worse: A cluster randomized controlled pilot trial of a leadership implementation strategy

Jill Locke[1] (iD), Catherine M. Corbin[2] (iD), Roger Goosey[1],
Vaughan K. Collins[1] (iD), Mark G. Ehrhart[3], Kurt Hatch[4],
Christine Espeland[4] and Aaron R. Lyon[1] (iD)

Abstract

Background

Implementation of evidence-based practices (EBPs) in schools is fraught with challenges. Even when EBPs are initiated, deterioration of implementation efforts often hinders their long-term success. School leadership behaviors can influence teachers' EBP implementation. Our study tested an implementation strategy called Helping Educational Leaders Mobilize Evidence (HELM), adapted from the Leadership and Organizational Change for Implementation strategy, to enhance EBP implementation through improvements in school leadership teams' implementation leadership and climate to buffer against the deterioration of implementation efforts. This study explores the impact of HELM on theorized mechanisms of change (i.e., implementation leadership, climate), educator-level factors (i.e., implementation citizenship), and implementation outcomes (i.e., fidelity, initiative stability).

Method

One school district and 10 schools in Washington participated. Five of the schools were randomized to receive the HELM strategy and the remaining five schools received an alternative leadership training as an implementation attention control. Teachers at every school ($n = 341$) received training for an EBP called Positive Greetings at the Door that has been previously demonstrated to reduce student behavior problems. Principals and Assistant Principals ($n = 18$) received the HELM strategy or alternative leadership training. Three district Administrators also participated in HELM as part of the Organizational Strategy Development meetings.

Results

HELM significantly slowed the average decline of implementation leadership (perseverant leadership and communication), three dimensions of implementation climate (recognition, rewards, and existing supports) and total implementation climate, and one dimension of implementation citizenship (keeping informed). No significant effects were found with regard to implementation outcomes (i.e., fidelity, initiative stability).

Conclusions

HELM shows promise in buffering the deterioration of EBP implementation efforts in schools. HELM positively influenced implementation leadership and climate, which are the hypothesized mechanisms for promoting

[1]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA
[2]School of Special Education, School Psychology, and Early Childhood Studies, University of Florida, Gainesville, FL, USA
[3]Department of Psychology, University of Central Florida, Orlando, FL, USA

[4]Educational Administration, University of Washington, Tacoma, WA, USA

Corresponding author:
Jill Locke, Department of Psychiatry and Behavioral Sciences, University of Washington, 6200 NE 74th Street, Suite 100, Seattle, WA 98115, USA.
Email: jjlocke@uw.edu

successful long-term implementation efforts. An appropriately powered trial is needed to determine the efficacy of HELM in the future.

Name of the registry: clinicaltrials.gov
Trial registration number: NCT06340074
Date of registration: March 29, 2024. Retrospectively registered
URL of trial registry record: https://clinicaltrials.gov/study/NCT06340074?intr=helm&rank=

Plain Language Summary

Using practices that have been proven to work in research to support youth social, emotional, and behavioral well-being in schools is hard. Efforts to use these practices decrease across the school year. School leadership behaviors can influence educators' use of these practices. Our study tested an implementation strategy or support designed to help educators use a research-backed practice called Helping Educational Leaders Mobilize Evidence (HELM). HELM was adapted from the Leadership and Organizational Change for Implementation (LOCI) strategy, to enhance school leadership teams' specific leadership behaviors to protect against the decline in implementation effort. This study explores the impact of HELM on outcomes related to practice use. One school district and 10 schools in western Washington State participated. Five of the schools received the HELM strategy and the remaining five schools received an alternative leadership training. Teachers at every school ($n = 341$) received training in a practice called Positive Greetings at the Door (PGD) that has been shown to reduce student behavior problems. Principals and Assistant Principals ($n = 18$) received the HELM strategy or alternative leadership training. Three district administrators also participated in HELM. HELM unexpectedly slowed the average decline of implementation leadership (perseverant leadership and communication), three dimensions of implementation climate (recognition, rewards, and existing supports), total implementation climate, and one dimension of implementation citizenship (keeping informed). No significant effects were found with regard to use of the practice. The HELM strategy shows promise in slowing the deterioration of practice in schools and positively influenced implementation leadership behaviors.

Introduction

Organizationally focused implementation strategies are needed to improve the implementation and sustainment of evidence-based practices (EBPs) in schools. The Leadership and Organizational Change for Implementation (LOCI) intervention is an organizational implementation strategy that has been shown to improve implementation leadership (i.e., leader behaviors that facilitate implementation; Aarons, Farahnak, et al., 2014) that in turn, fosters a conducive implementation climate (shared perceptions that EBP use is expected, supported, and rewarded; Ehrhart et al., 2014) in organizations (Aarons, Farahnak, et al., 2015; Aarons, Moulin, et al., 2017; Aarons, Sklar, et al., 2024; Skar et al., 2022; Williams et al., 2023). Across public service sectors including community mental health and public school settings, several studies have corroborated Aarons, Farahnak, et al.'s (2014) theory of change that first-level leaders can influence provider fidelity to EBPs by contributing to the development of a focused EBP implementation climate within different organizations (Aarons, Farahnak, et al., 2015; Aarons, Moulin, et al., 2017; Brookman-Frazee & Stahmer, 2018; Skar et al., 2022; Williams, Benjamin-Wolk, et al., 2020; Williams, Ehrhart, et al., 2024; Williams, Hugh, et al., 2022; Williams, Marcus, et al., 2023).

Recently, we used a human-centered design framework for complex interventions, the Discover, Design/Build, Test (DDBT) framework (Lyon et al., 2019) to adapt LOCI for use in the education sector. Adaptation occurred in partnership with school leaders and staff to enhance LOCI's acceptability, feasibility, contextual appropriateness, usability, and effectiveness for public elementary schools. Applying DDBT, we (a) facilitated a series of focus groups ($N = 54$ educators); (b) gathered expert input ($N = 15$ implementation researchers and school practitioners) using a nominal group decision-making process (Locke et al., 2024) and "hackathon" solution generation; and (c) applied the Cognitive Walkthrough for Implementation Strategies (CWIS; Collins et al., 2024) method to evaluate usability (i.e., the extent to which an intervention can be used by specified users with effectiveness, efficiency, and satisfaction; $N = 15$ principals). Results at each step provided avenues to improve the suitability/fit of LOCI for the school context, such as aligning assessment windows to school calendars and prioritizing former school leaders as coaches (Locke et al., 2024). Usability ratings from CWIS (scale 0–100) were acceptable to good ($M = 77.8–87.5$) and redesign solutions included doing a "walk and talk," where coaches virtually walk the building and discuss the school climate with principals (Collins et al., 2024). Our redesigned strategy—

Helping Educational Leaders Mobilize Evidence (HELM) —is intended to improve school leadership team's use of implementation leadership to foster a positive implementation climate to support the high-fidelity use of EBPs to improve student outcomes.

HELM is designed to be "EBP agnostic" and focuses on supporting school leadership teams to engage in collaborative distributed leadership practices (Bolden, 2011; Tian et al., 2016) as a component of implementation leadership. Distributed leadership is a common approach in schools where teachers and staff work closely with school administrators (i.e., principals, assistant principals) to make decisions and support innovative change efforts within the building (Camburn et al., 2003; Diamond & Spillane, 2016; Hickey et al., 2022). Distributed leadership structures create the conditions that promote initiative stability, defined as continued organizational commitment (e.g., maintaining EBP priorities, protecting time) and consistency (e.g., regular discussions about EBP importance) surrounding an implementation initiative despite changes such as staff turnover (i.e., principal and staff), shifts in policy, and/or budget changes (Heck & Hallinger, 2009; Jantzi & Leithwood, 2000). However, distributed leadership can only be achieved when formal leaders (e.g., principals) create space for it. As a result, some schools may already have distributed leadership structures and practices in place while others do not. HELM accounts for this variability by supporting principals through data-based feedback and coaching to either (a) refine and improve an *existing* distributed leadership structure to support implementation goals more strategically or (b) initiate *new* distributed leadership structures to advance EBP implementation. HELM targets both formal and informal leaders within buildings as part of the distributed leadership team and addresses variable district-level leadership structures (i.e., site-based vs. centralized control). It is important to test whether this adaptation and replication of LOCI in schools activates its intended mechanisms of change when considering critical contextual factors (e.g., school calendar year, distributed leadership structures).

Implementation mechanisms are the processes by which implementation strategies produce desired effects (Lewis, Frank, et al., 2024; Lewis, Klasnja, et al., 2018; Williams, 2016). Consistent with past research on LOCI, HELM is intended to most directly influence implementation leadership behaviors. In schools, implementation leadership often is distributed across members of a building-level leadership team (Aarons et al., 2024; Locke et al., 2019; Skar et al., 2022; Williams et al., 2024). Implementation leadership is the most significant driver of implementation climate (Aarons, Moullin, et al., 2017; Aarons, Sklar, et al., 2014; Klein et al., 2001; Williams et al., 2020, 2022) which, in turn, effectively enhances implementation citizenship behaviors like helping colleagues overcome implementation-related challenges (Corbin, Ehrhart, et al., 2024; Corbin, Zhang, et al., 2024; Ehrhart et al., 2015; Lyon et al., 2018) and increases

the reach of prioritized innovations (Aarons et al., 2024). To this end, the objective of this pilot study was to explore the impact of HELM on theorized mechanisms of change (i.e., implementation leadership, implementation climate), educator-level factors (i.e., teacher implementation citizenship), and implementation outcomes (i.e., fidelity, initiative stability). Specifically, we hypothesized that schools in HELM would have greater implementation leadership, climate, citizenship, fidelity, and initiative stability than in the control condition.

Method

Participants

Ten elementary schools (K-5; one school was kindergarten only) in one school district located in WA State participated in this study. Participants were included if they were: a district leader ($n = 3$), principal or assistant principal ($n = 18$; two schools did not have assistant principals), or other school personnel ($n = 341$) during the 2022–2023 school year. School personnel included 244 general education teachers, 50 paraeducators, 21 specialists (e.g., reading), 10 guidance counselors, nine other personnel (e.g., library assistant, nurse), six office/administrative staff, and one custodial/maintenance worker. One implementation attention control (IAC) and two HELM teachers ($n = 3$) were excluded from analyses because they were associated with two different schools during the pilot trial; the analytic sample included 241 teachers (see Table 1 for teacher counts by treatment condition and time). Table 2 presents demographic information for school leaders, teachers, and all participating school personnel. School administrators identified as women (83%), white (94%), non-Hispanic (100%), and had an average age of 46 years ($SD = 5.2$; range = 35–54 years). One principal was in their first year of principalship while others had 4 ($n = 2$), 6 ($n = 4$), and 10–17 ($n = 3$) years of experience with an average of 5 years ($SD = 4.1$; range = 1–17 years). All principals had obtained a master's degree (100%) and had previous experience implementing EBPs in school settings (100%). Teachers largely identified as women (88%) with approximately 95% identifying as white, 2.2% as More than One Race, 1.4% as American Indian or Alaska Native, and 1.4% as Asian. Approximately, 63% of teachers had a master's degree, 36% had a bachelor's, and less than

Table 1

Count of Teacher Respondents by Treatment Condition and Time

	Time 1	Time 2	Time 3
HELM	128	116	107
IAC	113	94	91
Total	241	210	198

Note. HELM = Helping Educational Leaders Mobilize Evidence; IAC = implementation attention control.

Table 2
Sample Demographics

	Principals (n = 18) %	Teachers (n = 241) %	School personnel (n = 341) %
Gender			
Man	16.7	10.4	9.5
Woman	83.3	88.4	89.3
Prefer not to disclose	0.0	1.2	1.2
Ethnicity			
Hispanic	0.0	7.0	6.5
Race			
American Indian or Alaska Native	0.0	1.4	2.9
Asian	0.0	1.4	2.6
Black or African American	0.0	0.0	0.3
Middle Eastern or North African	0.0	0.0	0.0
Native Hawaiian or Other Pacific Islander	0.0	0.0	0.3
White	94.4	95.0	88.0
Biracial	0.0	2.2	0.0
Prefer not to disclose	5.6	0.0	5.8
Age (average)	46.3	39.7	40.4
Years in role (average)	5.3	12.46	11.0
Degree			
High school	0.0	0.0	7.7
Associate's	0.0	0.0	5.6
Bachelor's	0.0	36.4	32.0
Master's	100.0	63.1	53.0
Doctoral	0.0	0.5	0.3
Prefer not to disclose	0.0		1.5
EBP implementation experience			
Yes	100.0	94.0	88.8
PGD implementation experience			
Yes	94.4	67.6	62.4

Note. Three school personnel declined to provide demographic information. EBP = evidence-based practice; PGD = Positive Greetings at the Door.

1% had a doctorate. Teachers reported an average of 12 years of experience in their current role (*SD* = 9.0, range = 1–30) with 94% reporting experience implementing EBPs and 68% reporting experience implementing Positive Greetings at the Door (PGD; see Procedure). The number of teachers in each school ranged from 10 to 31. Demographic information was not collected from district leaders.

Procedure

We obtained informed consent and collected pre-training measures approximately two weeks before the school year began via REDCap, an online data collection system. Next, pre-training data were used for stratified block randomization at the school level to ensure equal numbers of schools in each condition and a balance of important school-level covariates (i.e., school-average implementation climate, implementation leadership, % free/reduced lunch, and school-level racial composition) using K nearest neighbor matched median splits. Teachers were masked to their school's condition and the primary study purpose and hypotheses. There was no school or leader attrition post-randomization. Following collection of the baseline measures and random assignment process, all schools were trained in PGD (see below), and schools randomized to HELM (*n* = 5) received a 2-day workshop following PGD training, whereas schools randomized to implementation as usual (*n* = 5) received the first of a series of informational webinars on general leadership skills. Each participating school regardless of randomization received a $500 stipend and free PGD resources, training, and follow-up support for serving as a partner in this study. Coaching for principals randomized to HELM and organizational strategy development with district administrators was conducted via Zoom and/or in person. Participants in both conditions completed the same measures via REDCap mid-year and at the end of the school year. Individual participants received $40 at each time point. PGD fidelity assessments were conducted monthly.

HELM. HELM is a 9-month, data-driven organizational and leadership implementation strategy that entails eight core components, which largely map onto the original components of the LOCI strategy (Aarons, Farahnak, et al., 2014): (a) *Assessment and Feedback*. 360° surveys that gather feedback from different sources (i.e., self-ratings and educator/staff ratings of the distributed leadership team) measuring implementation leadership and climate were administered at three times during the school year. These data were synthesized into a detailed feedback report, which was shared with the distributed leadership team and used to create a personal leadership development plan to support implementation coaching throughout the year. (b) *Initial Training*. Two, 2-hr didactic and interactive trainings were provided to principals, their distributed leadership teams, and district-level leaders about a week apart that covered developing strategic implementation leadership behavior and building a positive EBP implementation climate in their schools. (c) *Leadership Development Plan*. During the initial training, principals and their distributed leadership teams worked individually with their HELM coach to review their 360° assessment data and develop goals for improving implementation leadership and climate around EBP implementation. (d) *Individual Coaching*. HELM coaches provided monthly 1-hr coaching sessions in person or via Zoom to review progress and update the leadership development plan. The coaching structure included reflective questions about (i) broader school updates (10 min); (ii) EBP

implementation (10 min); (iii) Leadership Development Plan progress (20 min); (iv) barriers to implementation and solution generation (10 min); (v) next steps (5 min); and (vi) "what other support is needed" from the district and/or HELM coaches (5 min). (e) *Group Coaching*. Coaches offered *optional* monthly 1-hr group coaching calls with all HELM principals and distributed leadership teams to review progress and share strategies across schools for idea generation and implementation support. (f) *Organizational Strategy Development*. Two 1-hr meetings with district-level leaders were held, one in the Fall semester and one in the Spring semester to develop and update an organizational implementation strategy plan (Climate Development Plan). This meeting provided a structured discussion of alignment between school- and district-level efforts to support EBP implementation. (g) *Professional Learning Collaboratives*. Two professional learning collaboratives were held with principals and their distributed leadership teams (in January and May of the school year) to review content (align HELM strategies with principles from the National Educational Leadership Standards (NELS) and EBP sustainment for the following school year) and share strategies across participants. The NELS is the professional standards that define the nature and quality of the principal profession and guide how principals are prepared, hired, developed, supervised, and evaluated in their profession in the United States. The NELS was used in the development of HELM to ensure alignment between the duties and responsibilities of the principalship in U.S. public schools. (h) *Graduation*. During graduation, principals' and their distributed leadership teams' final feedback were supposed to be reviewed, and progress for the past year was supposed to be celebrated. However, due to unexpected circumstances within the school district, Graduation was canceled (see Limitations).

IAC. Schools assigned to the IAC condition received an online, self-paced, independent studies program broadly focused on leadership and management. The program did not specifically discuss implementation leadership or climate. The program included four content modules hosted on a web-based learning portal: (a) Motivating and Engaging Employees, (b) Authentic Leadership, (c) Managing Diverse Teams, and (d) Fostering an Idea Culture. Each module consisted of one content-specific 60-min webinar.

PGD. PGD is the EBP that was used to pilot test HELM. PGD is a universal, classroom-based, prevention EBP that has evidence for its impact on student social, emotional, and behavioral well-being; is appropriate for use with a diverse range of schools and students; has an established fidelity assessment rubric; and has high variability in its implementation (thus necessitating targeted implementation supports) (Cook et al., 2018). PGD is a multi-component proactive strategy that involves: (a) authentic greetings and welcoming of each student as they enter the classroom, (b) pre-corrective statements to remind students of expected behavior as they transition into the classroom, (c) high-interest activity that students engage in after they enter the classroom, and (d) restorative conversations with students for whom they had negative interactions with during the previous period or day.

Measures

Primary Outcome

PGD Fidelity. A PGD expert outside of the research team was contracted to train assistant principals in collecting live PGD fidelity. Assistant principals conducted live observations ($n = 15$) of a randomly selected subset of three classrooms across 10 observation windows from October 2022 to May 2023. The participating district provided the study team with a list of classroom teachers in each school. Three teachers were randomly selected for observation in each window, blocking by grade span (K-2, 3–5) to ensure observations captured the full developmental range of students. The process for randomly selecting the third teacher changed every other window, either randomly selecting across grades K-5 or among specialist teachers (e.g., music). Assistant principals observed students' transition into the classroom and scored whether teachers practiced each of the four PGD components: greeting, pre-corrective statements, high-interest activity, and restorative conversations. Teachers also self-reported their fidelity to the same four PGD components at mid- and end-of-year.

Secondary Outcomes (Collected at All Three Time Points)

School Implementation Leadership Scale (SILS; Lyon, Corbin, et al., 2022): Adapted from the original *Implementation Leadership Scale* (Aarons, Ehrhart, et al., 2014), the SILS has 21 items loading onto seven subscales: Proactive, Knowledgeable, Supportive, Perseverant, Communication, Vision, and Available. Subscale internal consistencies range from 0.98 to 0.99 across time points.

School Implementation Climate Scale (SICS; Thayer et al., 2022): Adapted from the original Implementation Climate Scale (Ehrhart et al., 2014), the SICS includes 21 items loading onto seven 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: 0.95–0.97).

School Implementation Citizenship Behavior Scale (SICBS; Corbin, Ehrhart, et al., 2024; Ehrhart et al., 2015): This 12-item scale has four subscales: Helping Others, Keeping Informed, Taking Initiative, and Advocacy. Internal consistencies ranged from .96 to .97 across time points. Recent work has supported the use of the SICBS in schools (Corbin et al., 2024).

Commitment to Organizational Change Scale (COCS; Herscovitch & Meyer, 2002): A modified version of the

COCS was used to measure school staff perceptions of implementation initiative stability. Two subscales were used: affective commitment ($n = 5$ items; e.g., "I believe in the value of this change"), and normative commitment ($n = 4$ items; e.g., "I want to be a team player and continue to support this change"). Responses were made using a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Internal consistency for initiative stability ranged from .91 to .93 across time points.

For all measures above (SILS, SICS, SICBS, COCS), the referent at baseline was "EBP" as participants had no prior exposure to PGD; however, at Time 2 and Time 3, the referent was changed to "PGD."

Data Analysis

Baseline equivalency was examined between matched pairs of schools and by assessing differences between HELM and IAC schools on demographic information not available pre-randomization (i.e., teacher race, gender, and years of experience). Variables that demonstrated differences at baseline were included as covariates in all models. A series of two-level growth curve models were used to examine the effect of HELM on each of the teacher-reported fidelity (primary outcome) and mechanisms of change (secondary outcomes; level 1 = time, level 2 = teacher) (Little, 2013, pp. 247–260). We also examined the impact of HELM participation on subscales for each mechanism of change (see Supplementary Appendix A). This approach was exploratory and pursued to facilitate a more detailed understanding of potential effects of this pilot study. Adopting the notation of Little (2013), models for all secondary outcomes took the form,

$$\text{Level 1:} y_{io} = \pi_{0i} + \pi_{1i}(time) + \varepsilon_{io} \quad (1)$$

$$\text{Level 2:} \pi_{oi} = \gamma_{00} + \gamma_{10}(HELM) + \gamma_{1n}(covars) + \zeta_{0i}$$

$$\pi_{1i} = \gamma_{01} + \gamma_{11}(HELM) + \gamma_{1m}(covars) + \zeta_{1i}$$

where y_{io} are implementation leadership, climate, citizenship, and initiative stability scores for individual i at time point o, π_{0i} is the intercept for individual i, π_{1i} is the slope for individual i across time points, and ε_{io} is the residual variance for individual i at time point o. The level 1 intercept (π_{0i}) and slope (π_{1i}) were each predicted at level 2 (i.e., teacher) as a function of the intercept (γ_{00}, γ_{01}), treatment indicator (γ_{10}, γ_{11}), and a vector of covariates (γ_{1n}, γ_{1m}). The residual variance for the level 1 intercept (ζ_{0i}) and slope (ζ_{1i}) were allowed to randomly vary across individuals.

Models with no independent variables (i.e., unconditional models) were first examined to determine (a) the intraclass correlation coefficient (ICC) or amount of between teacher variance in dependent variables, (b) the average rate of change across time, and (c) whether there was significant variance around the slope—a necessary condition to assess between teacher treatment impacts. There were too few schools ($n = 10$) to estimate and model a third level (i.e., between-school variance), which means that between-school variance is conflated with the variance between teachers (i.e., level 2). To control for between-school variation, we included fixed effects for matched school pairs at level 2, excluding the largest pair as the referent. Emergent work demonstrates that in the absence of intervention, providers have shared and individual perceptions of implementation leadership and climate (Corbin et al., 2022; Zhang et al., 2023), both of which lead to changes in hypothesized implementer outcomes like implementation citizenship behavior (Corbin, Zhang, et al., 2024).

Three classrooms per school were observed at each time point (i.e., 15 classrooms for each condition at each time point), resulting in too few classrooms observed to quantitatively assess treatment effects on observed PGD fidelity. Instead, mean scores were calculated for each PGD fidelity component across observation windows to roughly align with administration of the three survey assessments, resulting in mean scores for three points. Multiple linear regression was used to assess the impact of HELM on teacher-reported fidelity, which was reported by all participating teachers at Times 2 and 3, to each of the four PGD components.

All models were run in *Mplus* v8 (Muthén & Muthén, 1998–2017) using full information maximum likelihood to recover missing data (9% across all study constructs) (Larsen, 2011). Continuous covariates in multilevel models were grand mean-centered. Effect sizes were calculated in alignment with recommendations for growth models ($d = \boldsymbol{\gamma}_{11}(time) / \boldsymbol{SD_{raw}}$) and are interpreted such that .20, .50, and .80 represent small, medium, and large effects (Feingold, 2009). A post hoc power analysis indicated that given the parameters of this study (randomization, nested data structure, ICC of approximately .35, eight covariates), we can detect with 80% power an effect size of $d = 0.94$ or higher.

Results

Baseline Equivalency

Although efforts were made to match school pairs on all pre-selected demographic characteristics, each of the variables were inequivalent within at least one school pair at baseline. There also were significant differences at baseline between HELM and IAC teachers on all variables used to randomize schools to treatment condition. No baseline differences emerged between HELM and IAC teachers on demographic variables collected as part of the first assessment. Therefore, all models included free reduced-price lunch, percent white students, and school size as covariates. Table 3 presents descriptive statistics for all mechanisms of change by treatment condition.

Observed PGD Fidelity

For both conditions, there was a ceiling effect for greetings (T1 = 91%, T2 = 95%, T3 = 97%). Pre-corrective statements also were high at each time point (e.g., T1 = 90%, T2 = 81%, T3 = 88%), and high-interest activities were lower on average (T1 = 78%, T2 = 84%, T3 = 85%). On average, teachers engaged in less than one restorative conversation at each time point (T1 = 0.70, T2 = 0.71, T3 = 0.56). Visual inspection of average PGD fidelity component scores by treatment condition (Figure 1) revealed that HELM and IAC teachers greeted students at the door most of the time (Panel A). HELM teachers increased and then decreased implementation of a high-interest activity, whereas IAC teachers were stable and then increased use of the practice from Time 2 to Time 3 (Panel B). HELM teachers had lower fidelity to pre-corrective statements (Panel C) and engaged in fewer restorative conversations than IAC teachers (Panel D).

Teacher-Reported PGD Fidelity

HELM did not positively impact any of the four teacher-reported PGD fidelity components (i.e., greetings, pre-corrective statements, high-interest activity, and restorative conversations) relative to IAC (Table 4).

Mechanisms of Change

Preliminary Analyses. Table 5 presents results from unconditional models (i.e., no covariates) for each mechanism of change. ICCs ranged from .39 to .47, indicating substantial between-teacher variation for all outcomes. On average, implementation leadership ($b = -0.07$, $p = .07$) and climate ($b = -0.15$, $p \leq .001$) significantly decreased across the academic year. There were significant differences between teachers' rate of change (i.e., variance of the slope) for each. There was no average change in implementation citizenship behavior across time, though there was significant variability in how teachers changed over time. Initiative stability did not demonstrate significant change across time or variability around its slope, failing to meet the conditions to assess impacts.

Treatment Effects. Participation in HELM did not significantly impact the level of implementation leadership at the start of the academic year (i.e., random intercept) or the rate of change in overall implementation leadership across time (i.e., random slope; Table 5). HELM positively impacted the rate of change (i.e., random slope) for two implementation leadership subscales such that the average decline was slowed relative to IAC teachers (Supplementary Appendix A): the extent to which leaders persevere through implementation challenges ($b = 0.20$, $p = .10$, $d = 0.62$) and communicate about the importance and prioritization of EBP use ($b = 0.20$, $p = .10$, $d = 0.61$). No other impacts on implementation leadership subscales emerged.

Average decreases in implementation climate over the academic year were significantly improved (i.e., positive impact) for HELM teachers relative to IAC teachers ($b = 0.23$, $p \leq .05$, $d = 0.79$). There was no effect of HELM on teachers' implementation climate scores at the start of the academic year (i.e., random intercept). HELM positively impacted the rate of change (i.e., random slope) for three implementation climate subscales, again showing slower decreases over time (Supplementary Appendix B): use of existing supports to deliver EBP ($b = 0.29$, $p \leq .05$, $d = 0.76$), recognition for use of EBP ($b = 0.23$, $p \leq .10$, $d = 0.65$), and rewards for use of EBP ($b = 0.35$, $p \leq .01$, $d = 0.95$). There were no other significant impacts.

Finally, HELM teachers reported significantly slower declines in implementation citizenship behavior across
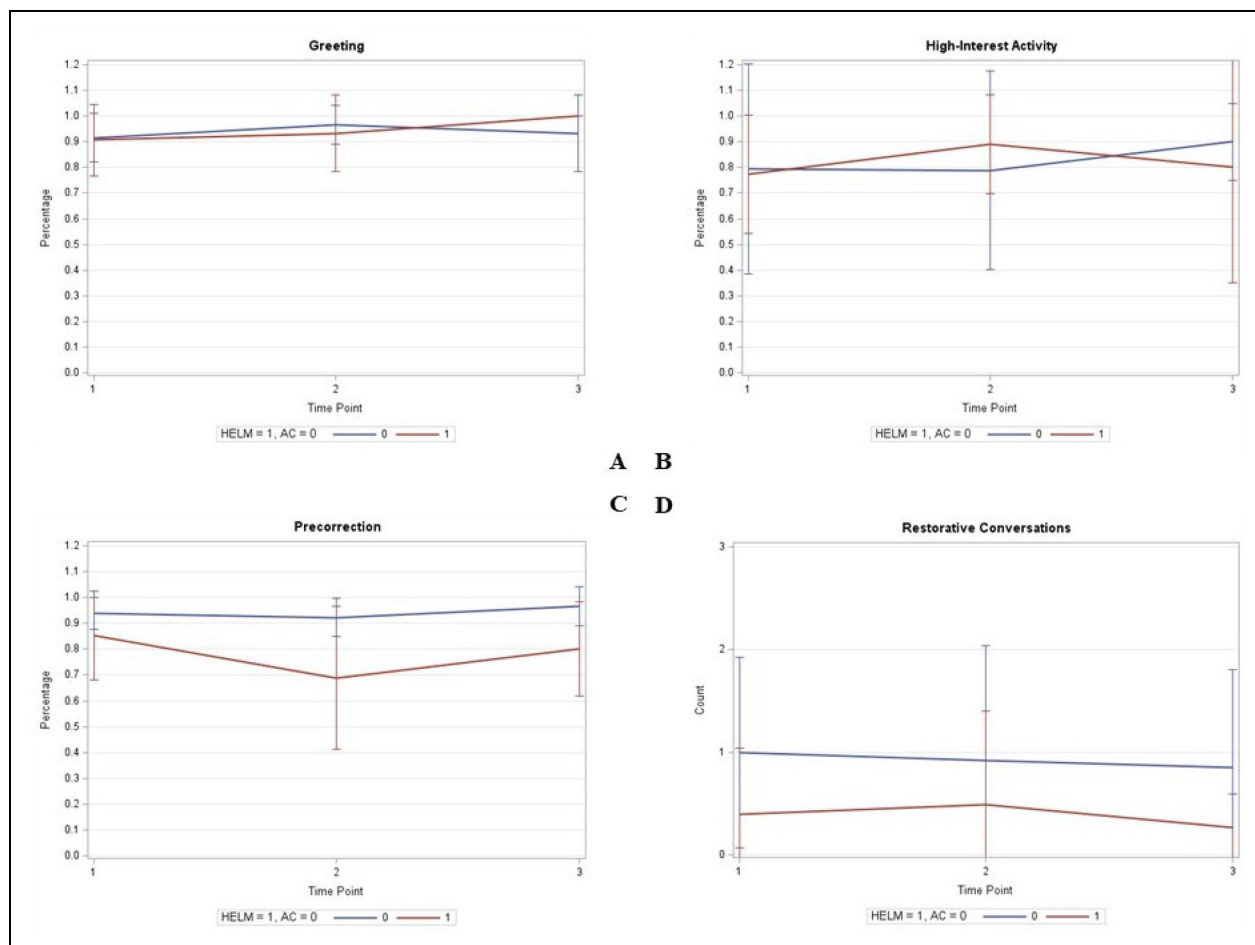
Table 3

Univariate Statistics for All Mechanisms of Change

	Implementation leadership		Implementation climate		Implementation citizenship		Initiative stability	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
HELM								
Time 1	2.90	0.78	2.19	0.82	2.20	0.85	5.16	0.62
Time 2	3.12	0.75	2.14	0.75	2.22	0.91	5.17	0.59
Time 3	2.88	0.96	2.06	0.91	2.37	0.90	5.13	0.60
IAC								
Time 1	3.00	0.68	2.31	0.76	2.31	0.87	5.20	0.59
Time 2	2.88	0.88	1.75	0.89	1.97	1.02	5.14	0.52
Time 3	2.73	1.02	1.77	0.99	2.06	1.09	5.12	0.48

Note. HELM = Helping Educational Leadership Mobilize Evidence; IAC = implementation attention control.

Figure 1
Observed Positive Greetings at the Door Fidelity Scores Across Time by Treatment Condition

Note. HELM = Helping Educational Leaders Mobilize Evidence; AC = active control. Plotted means with standard deviations at each time point.

Table 4
Results From Multiple Linear Regression Models Assessing the Impact of HELM on Teacher-Reported Fidelity to the Four PGD Components

	Greetings b (SE)	Pre-correction b (SE)	High-interest activity b (SE)	Restorative conversations b (SE)
Intercept	1.54* (0.65)	2.00 (1.12)	1.53 (0.89)	0.57 (1.79)
Time 1 greeting	0.434*** (0.16)	0.22*** (0.08)	0.37*** (0.08)	0.32*** (0.09)
HELM	0.12 (0.09)	0.01 (0.12)	−0.10 (0.11)	0.08 (0.24)
r^2	.30	.08	.19	.13

Note. $n = 210$ teachers. Point estimates are presented with standard errors in parentheses. Covariates not presented. HELM = Helping Educational Leaders Mobilize Evidence; PGD = Positive Greetings at the Door. *$p \leq .05$; ***$p \leq .001$.

the school year compared to IAC teachers ($b = 0.18$, $p \leq .10$, $d = 0.56$). Treatment condition had no impact on teachers' implementation citizenship behavior scores at the start of the academic year (i.e., random intercept). Results indicated significant differences between teachers' rate of change (i.e., variance of the slope) for all four SICBS subscales. Average decreases in the extent to which teacher kept informed about PGD (Supplementary Appendix C) were significantly slowed for HELM compared to IAC teachers ($b = 0.35$, $p \leq .01$, $d = 0.95$). No other significant effects emerged. Participants did not report any harms or

Table 5
Results From Random Intercept and Slope Models Assessing the Impact of HELM on Mechanisms of Change

	Implementation leadership		Implementation climate		Implementation citizenship behavior		Initiative stability
	Empty	Impact	Empty	Impact	Empty	Impact	Empty
Level-1 (within)							
Residual variance (ε_{io})	0.30*** (0.10)	0.30*** (0.40)	0.28*** (0.03)	0.28*** (0.03)	0.35*** (0.04)	0.35*** (0.04)	0.18*** (0.03)
Level-2 (between)							
Random intercept (π_{0i})							
Intercept (γ_{00})	3.00*** (0.04)	3.27*** (0.13)	2.36*** (0.07)	2.23*** (0.13)	2.12*** (0.05)	2.62*** (0.16)	5.18*** (0.04)
HELM (γ_{10})		−0.08 (0.12)		−0.06 (0.13)		−0.11 (0.14)	
Residual variance (ζ_{0i})	0.22*** (0.06)	0.17*** (0.05)	0.44*** (0.14)	0.24*** (0.05)	0.39*** (0.07)	0.33*** (0.06)	0.16*** (0.04)
Random slope (π_{1i})							
Intercept (γ_{01})	−0.07† (0.04)	−0.23* (0.10)	−0.15*** (0.04)	−0.55*** (0.10)	−0.01 (0.04)	−0.25* (0.11)	−0.03 (0.02)
HELM (γ_{11})		0.12 (0.10)		0.23* (0.10)		0.18† (0.11)	
Residual variance (ζ_{1i})	0.14*** (0.04)	0.11** (0.04)	0.11*** (0.03)	0.06** (0.03)	0.12*** (0.03)	0.08* (0.03)	0.01 (0.02)
Covariance (π_{0i}, π_{1i})	−0.01 (0.04)	−0.02 (0.03)	−0.12* (0.06)	0.01 (0.10)	−0.03 (0.04)	−0.03 (0.04)	−0.02 (0.02)

Note. $n_{teachers} = 241$. Point estimates are presented with standard errors in parentheses. Notation aligns with equation (I). HELM = Helping Educational Leaders Mobilize Evidence.
Empty: models with no covariates.
†$p \leq .10$. *$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.

unintended consequences as a result of HELM activities during exit interviews, which were administered to gather feedback on further refinements of HELM components, nor did the study team observe any harms during sessions. However, we learned of emergent crises unrelated to HELM implementation (e.g., student death, legal case) during coaching sessions and meetings with the district administrators as part of the HELM activities.

Discussion

This study examined the impact of an adapted organizational implementation strategy, HELM, on theorized mechanisms of change (i.e., implementation leadership, implementation climate), educator-level factors (i.e., teacher implementation citizenship), and implementation outcomes (i.e., fidelity, initiative stability). Evidence from implementation trials in outpatient mental health clinics indicates implementation strategies that address implementation leadership and climate can improve fidelity to, and effectiveness of, mental and behavioral health EBPs (Aarons, Farahnak, et al., 2015; Aarons, Moullin, et al., 2017; Brookman-Frazee & Stahmer, 2018; Skar et al., 2022; Williams et al., 2023). However, leadership-focused implementation strategies have not been systematically developed for and evaluated in educational settings. In this study, HELM positively impacted two dimensions of implementation leadership (perseverant leadership and communication), three dimensions of implementation climate (recognition, rewards, and existing supports) and total implementation climate, and one dimension of implementation citizenship (keeping informed). No significant effects were found with regard to implementation outcomes (i.e., fidelity, initiative stability). These results provide evidence that leadership-focused implementation strategies may be impactful in the education sector. Implications are discussed below.

The data indicated that HELM may buffer the decline in EBP implementation efforts over the school year. While HELM was not designed or developed to be an implementation strategy to slow the pace of decline in EBP use, this phenomenon warrants further exploration given that a pattern of gradually lower EBP implementation is common in schools as the academic year progresses; including across EBPs that address different domains (e.g., socioemotional learning, academics, etc.) and intended recipients (e.g., students in general vs. special education; Lyon, Liu, et al., 2022). It may be that the HELM components that target sustained structures and supports (e.g., recognition and rewards system, feedback loops around PGD implementation, PGD fidelity checks) in schools help slow the decline of EBP use as the year progresses, but that consistent and sustained organizational supports in the local implementation context may still be needed. Our results were not as robust as we hypothesized based on previous LOCI trials, where LOCI had a more

immediate impact on implementation leadership followed by climate further in the LOCI process (Aarons, Farahnak, et al., 2015; Aarons, Moullin, et al., 2017; Aarons, Sklar, et al., 2024; Skar et al., 2022; Williams et al., 2024).

In this pilot trial, HELM had a more direct and immediate impact on implementation climate. Although this school district had a number of emergent and significant crises during the year of study participation, the observed outcomes align with the efforts the schools and district administrators contributed during the year and have implications as to what can be immediately addressed in a relatively short amount of time. Most of the school leadership development plans as well as the district climate development plan focused on short-term goals that align with the changes in observed leadership, climate, and citizenship dimensions. For example, distributed leadership teams focused on providing opportunities for individualized recognition for PGD implementation efforts, noting that some individuals prefer large group shout-outs and others prefer private celebratory emails/notes, or some individuals prefer tangible rewards (e.g., coffee gift cards) and others prefer other opportunities (e.g., an additional prep period during the week) in recognition of their implementation efforts. Furthermore, district administrators and school leadership teams also had a strong communication plan focused on PGD implementation that was regularly shared with all educators. Such strengthened communication systems explicitly focused on PGD implementation and allowed all educators to have access to various PGD resources and specific examples of strategies that were successful in different classrooms. These may be more visible and accessible aspects of implementation leadership, climate, and citizenship in schools, where a little bit of change is noticed as opposed to other dimensions which may require more time than a school year to change.

Of note, there were no significant differences in implementation outcomes between HELM and non-HELM schools, whereas LOCI trials have seen an immediate effect on EBP use (Aarons, Farahnak, et al., 2015; Aarons, Moullin, et al., 2017; Skar et al., 2022; Williams et al., 2023). One explanation may be the complexity of PGD—it is a relatively easy to implement EBP which may explain the high-fidelity use of PGD across schools immediately following training. The schools were highly engaged and motivated to use PGD, perhaps not leaving room for HELM to make improvements on implementation outcomes. It could be that HELM may be more impactful for more complex EBPs that require more focus across time. These data highlight the importance of sequencing HELM activities in future trials where HELM training precedes EBP training, allowing HELM to have an influence during preparation as well as implementation, instead of after implementation has already begun.

Limitations and Future Directions

Several limitations are noted. First, recruitment was geographically limited to one school district in Washington State. The inclusion of other geographically diverse school districts may provide insights into how HELM may fit within different public school structures across multiple levels (e.g., district, school, educator). Future studies should test HELM in different school districts across the United States to ensure HELM content and logistical considerations (e.g., timing of 360° assessments, frequency of coaching, etc.) are applicable and generalizable to other settings. Second, the pilot test occurred during the COVID-19 pandemic, and most HELM components (e.g., 360° assessments, initial training, coaching, organizational strategy meetings, etc.) with the exception of a few in-person coaching visits per the school's request were conducted remotely. While we had a set of agreements during the conduct of each activity (e.g., one person speaking at a time), virtual platforms do not allow natural conversations to emerge in the same way as in-person sessions. Third, the district in which this pilot study was conducted experienced a number of unexpected and emergent significant crises during the course of the study year that prevented Graduation from being held as planned and could have more broadly impacted the outcomes of the study. Fourth, while previous studies indicate using a generic referent (i.e., "EBP") versus a specific referent (i.e., "PGD") measures the same underlying construct (e.g., implementation leadership) (Lyon, Corbin, et al., 2022), there were changes to the referent mid-data collection, which could have impacted the outcomes of the study (e.g., participants reporting higher or lower levels of the underlying construct). Fifth, although HELM was designed with a focus on activating distributed leadership teams in schools, distributed leadership was not explicitly measured or tested in this pilot study. Distributed leadership is an important construct to measure in future studies as it could be a potential condition under which HELM performs differently (e.g., schools with more collaborative distributed leadership teams may have stronger teamwork to build more conducive implementation contexts). Schools with existing distributed leadership teams may have systems and processes in place that allow them to immediately function more effectively compared to schools that are newly developing a distributed leadership team. Finally, we were unable to assess treatment impacts at the level of randomization, and the study was only powered to detect large effects between teachers. An appropriately powered trial is needed to determine the efficacy of HELM with confidence.

Conclusions

The results of this study support the effects of HELM on implementation leadership and climate for EBP

implementation. Given the deterioration of implementation outcomes across the school year observed in public schools, these results have important implications for the consideration of organizationally focused implementation strategies to support EBP use in educational settings.

Acknowledgments

Author Contributions

JL and ARL are the co-principal investigators for the study, generated the idea and designed the study. JL was the primary writer of the manuscript and approved all changes. CMC conducted all study analyses and drafted portions of the manuscript. RG, VKC, MGE, KH, and CE are core contributors to the research study in terms of HELM development and coaching 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.

Availability of Data and Material

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

Declaration of Competing Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: ARL is an Associate Editor for *Implementation Research and Practice*. As such, he had no part in the peer review process. All other Authors declare that there is no conflict of interest.

Funding

Ethical Approval

All procedures were approved by the University of Washington IRB (Study No. STUDY00010282).

Informed Consent

This study was determined to be IRB exempt.

ORCID iDs

Jill Locke https://orcid.org/0000-0003-1445-8509
Catherine M. Corbin https://orcid.org/0000-0002-2674-4425
Vaughan K. Collins https://orcid.org/0000-0002-5303-7291
Aaron R. Lyon https://orcid.org/0000-0003-3657-5060

Supplemental Material

Supplemental material for this article is available online.

References

Aarons, G. A., Ehrhart, M. G., & Farahnak, L. R. (2014). The Implementation Leadership Scale (ILS): Development of a brief measure of unit level implementation leadership. *Implementation Science*, *9*(1), Article 45. https://doi.org/10.1186/1748-5908-9-45

Aarons, G. A., Ehrhart, M. G., Farahnak, L. R., & Hurlburt, M. S. (2015). 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. *Implementation Science*, *10*, Article 11. https://doi.org/10.1186/s13012-014-0192-y

Aarons, G. A., Ehrhart, M. G., Farahnak, L. R., & Sklar, M. (2014). Aligning leadership across systems and organizations to develop a strategic climate for evidence-based practice implementation. *Annual Review of Public Health*, *35*, 255–274. https://doi.org/10.1146/annurev-publhealth-032013-182447

Aarons, G. A., Ehrhart, M. G., Moullin, J. C., Torres, E. M., & Green, A. E. (2017). Testing the leadership and organizational change for implementation (LOCI) intervention in substance abuse treatment: A cluster randomized trial study protocol. *Implementation Science*, *12*(1), Article 29. https://doi.org/10.1186/s13012-017-0562-3

Aarons, G. A., Farahnak, L. R., & Ehrhart, M. G. (2014). Leadership and strategic organizational climate to support evidence-based practice implementation. In R. S. Beidas & P. C. Kendall (Eds.), *Dissemination and implementation of evidence-based practices in child and adolescent mental health* (pp. 82–97). Oxford University Press.

Aarons, G. A., Sklar, M., Ehrhart, M. G., Roesch, S., Moullin, J. C., & Carandang, K. (2024). Randomized trial of the leadership and organizational change for implementation (LOCI) strategy in substance use treatment clinics. *Journal of Substance Use and Addiction Treatment*, *165*, Article 209437. https://doi.org/10.1016/j.josat.2024.209437

Bolden, R. (2011). Distributed leadership in organizations: A review of theory and research. *International Journal of Management Reviews*, *13*(3), 251–269. https://doi.org/10.1111/j.1468-2370.2011.00306.x

Brookman-Frazee, L., & Stahmer, A. C. (2018). Effectiveness of a multi-level implementation strategy for ASD interventions: Study protocol for two linked cluster randomized trials. *Implementation Science*, *13*(1), 66. https://doi.org/10.1186/s13012-018-0757-2

Camburn, E., Rowan, B., & Taylor, J. E. (2003). Distributed leadership in schools: The case of elementary schools adopting comprehensive school reform models. *Educational Evaluation and Policy Analysis*, *25*(4), 347–373. https://doi.org/10.3102/01623737025004347

Collins, V. K., Corbin, C. M., Locke, J. J., Cook, C. R., Ehrhart, M. G., Hatch, K. D., & Lyon, A. R. (2024). Usability evaluation of a leadership-focused implementation strategy to support Tier 1 programs in schools. *School Mental Health*, *16*(3), 667–680. https://doi.org/10.1007/s12310-024-09635-z

Cook, C. R., Fiat, A., Larson, M., Daikos, C., Slemrod, T., Holland, E. A., Thayer, A. J., & Renshaw, T. (2018). Positive greetings at the door: Evaluation of a low-cost, high-yield proactive classroom management strategy. *Journal of Positive Behavior Interventions*, 20(3), 149–159. https://doi.org/10.1177/1098300717753831

Corbin, C. M., Ehrhart, M. G., Brown, E. C., Locke, J., Cook, C. R., Aarons, G. A., & Lyon, A. R. (2024). "Going above and beyond" in the education sector: Extension and validation of the School Implementation Citizenship Behavior Scale (SICBS). Research Square. https://doi.org/10.21203/rs.3.rs-4682900/v1

Corbin, C. M., Hugh, M. L., Ehrhart, M. G., Locke, J., Davis, C., Brown, E. C., Cook, C. R., & Lyon, A. R. (2022). Teacher perceptions of implementation climate related to feasibility of implementing Schoolwide Positive Behavior Supports and Interventions. *School Mental Health*, 14, 951–966. https://doi.org/10.1007/s12310-022-09528-z

Corbin, C. M., Zhang, Y., Ehrhart, M. G., Locke, J., & Lyon, A. R. (2024). Testing an organizational implementation process model related to teachers' implementation-related attitudes and behaviors: A multilevel mediation analysis. *Prevention Science*, 25(7), 1053–1064. https://doi.org/10.1007/s11121-024-01722-6

Diamond, J. B., & Spillane, J. P. (2016). School leadership and management from a distributed perspective: A 2016 retrospective and prospective. *Management in Education*, 30(4), 147–154. https://doi.org/10.1177/0892020616665938

Ehrhart, M. G., Aarons, G. A., & Farahnak, L. R. (2014). Assessing the organizational context for EBP implementation: The development and validity testing of the Implementation Climate Scale (ICS). *Implementation Science*, 9(1), Article 157. https://doi.org/10.1186/s13012-014-0157-1

Ehrhart, M. G., Aarons, G. A., & Farahnak, L. R. (2015). Going above and beyond for implementation: The development and validity testing of the Implementation Citizenship Behavior Scale (ICBS). *Implementation Science*, 10(1), Article 65. https://doi.org/10.1186/s13012-015-0255-8

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods*, 14(1), 43–53. https://doi.org/10.1037/a0014699

Heck, R. H., & Hallinger, P. (2009). Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *American Educational Research Journal*, 46(3), 659–689. https://doi.org/10.3102/0002831209340042

Herscovitch, L., & Meyer, J. (2002). Commitment to organizational change: Extension of a three-component model. *Journal of Applied Psychology*, 87(3), 474–487. https://doi.org/10.1037/0021-9010.87.3.474

Hickey, N., Flaherty, A., & Mannix McNamara, P. (2022). Distributed leadership: A scoping review mapping current empirical research. *Societies*, 12(1), 15–20. https://doi.org/10.3390/soc12010015

Jantzi, D., & Leithwood, K. (2000). The effects of transformational leadership on organizational conditions and student engagement with school. *Journal of Educational Administration*, 38(2), 112–129. https://doi.org/10.1108/09578230010320064

Klein, K. J., Conn, A. B., & Sorra, J. S. (2001). Implementing computerized technology: An organizational analysis.

Journal of Applied Psychology, 86(5), 811–824. https://doi.org/10.1037/0021-9010.86.5.811

Larsen, R. (2011). Missing data imputation versus full information maximum likelihood with second-level dependencies. *Structural Equation Modeling: A Multidisciplinary Journal*, 18(4), 649–662. https://doi.org/10.1080/10705511.2011.607721

Lewis, C. C., Frank, H. E., Cruden, G., Kim, B., Stahmer, A. C., Lyon, A. R., Albers, B., Aarons, G. A., Beidas, R. S., Mittman, B. S., Weiner, B. J., Williams, N. J., Powell, B. J., Fernandez, M. E., Masyuko, S., Pullmann, M., Saldana, L., Wagner, A. D., Walsh-Bailey, C., & Stirman, S. W. (2024). A research agenda to advance the study of implementation mechanisms. *Implementation Science Communications*, 5(1), Article 98. https://doi.org/10.1186/s43058-024-00633-5

Lewis, C. C., Klasnja, P., Powell, B. J., Lyon, A. R., Tuzzio, L., Jones, S., Walsh-Bailey, C., & Weiner, B. (2018). From classification to causality: Advancing understanding of mechanisms of change in implementation science. *Frontiers in Public Health*, 6, Article 136. https://doi.org/10.3389/fpubh.2018.00136

Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.

Locke, J., Corbin, C. M., Collins, V., Ehrhart, M., Goosey, R., Hatch, K., Espeland, C., Cook, C. R., & Lyon, A. (2024). Helping Educational Leadership Mobilize Evidence (HELM): The iterative redesign of the Leadership for Organizational Change for Implementation (LOCI) intervention for use in schools. *Implementation Research and Practice*, 5, Article 26334895241242523. https://doi.org/10.1177/26334895241242523

Locke, J., Lawson, G. M., Beidas, R. S., Aarons, G. A., Xie, M., Lyon, A. R., Stahmer, A., Seidman, M., Frederick, L. K., Oh, C., Spaulding, C., Dorsey, S., & Mandell, D. S. (2019). Individual and organizational factors that affect implementation of evidence-based practices for children with autism in public schools: A cross-sectional analysis. *Implementation Science*, 14(39). https://doi.org/10.1186/s13012-019-0877-3

Lyon, A. R., Cook, C. R., Brown, E. C., Locke, J., Davis, C., Ehrhart, M., & Aarons, G. A. (2018). Assessing organizational implementation context in the education sector: Confirmatory factor analysis of measures of implementation leadership, climate, and citizenship. *Implementation Science*, 13(1), Article 5. https://doi.org/10.1186/s13012-017-0705-6

Lyon, A. R., Corbin, C. M., Brown, E. C., Ehrhart, M. G., Locke, J., Davis, C., Picozzi, E., Aarons, G. A., & Cook, C. R. (2022). Leading the charge in the education sector: Development and validation of the School Implementation Leadership Scale (SILS). *Implementation Science*, 17, Article 48. https://doi.org/10.1186/s13012-022-01222-7

Lyon, A. R., Liu, F. F., Connors, E. H., King, K. M., Coifman, J. I., Cook, H., McRee, E., Ludwig, K., Law, A., Dorsey, S., & McCauley, E. (2022). How low can you go? Examining the effects of brief online training and post-training consultation dose on implementation mechanisms and outcomes for measurement-based care. *Implementation Science Communications*, 3(1), Article 79. https://doi.org/10.1186/s43058-022-00325-y

Lyon, A. R., Munson, S. A., Renn, B. N., Atkins, D. C., Pullmann, M. D., Friedman, E., & Areán, P. A. (2019). 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*, 8(10), Article e14990. https://doi.org/10.2196/14990

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Skar, A. M. S., Braathu, N., Peters, N., Bækkelund, H., Endsjø, M., Babaii, A., Borge, R. H., Wentzel-Larsen, T., Ehrhart, M. G., Sklar, M., Brown, C. H., Aarons, G. A., & Egeland, K. M. (2022). 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 Services Research*, 22(1), Article 298. https://doi.org/10.1186/s12913-022-07539-9

Thayer, A. J., Cook, C. R., Davis, C., Brown, E. C., Locke, J., Ehrhart, M. G., Aarons, G. A., Picozzi, E., & Lyon, A. R. (2022). Construct validity of the school-implementation climate scale. *Implementation Research and Practice*, 3, Article 26334895221116065. https://doi.org/10.1177/26334895221116065

Tian, M., Risku, M., & Collin, K. (2016). A meta-analysis of distributed leadership from 2002-2013: Theory development, empirical evidence and future research focus. *Educational Management Administration & Leadership*, 44(1), 146–164. https://doi.org/10.1177/1741143214558576

Williams, N. J. (2016). Multilevel mechanisms of implementation strategies in mental health: Integrating theory, research, and practice. *Administration and Policy in Mental Health and Mental Health Services Research*, 43, 783–798.

Williams, N. J., Benjamin-Wolk, C., Becker-Haimes, E. M., & Beidas, R. S. (2020). Testing a theory of strategic implementation leadership, implementation climate, and clinicians' use of evidence-based practice: A 5-year panel analysis. *Implementation Science*, 15(1), 1–15. https://doi.org/10.1186/s13012-020-0970-7

Williams, N. J., Ehrhart, M. G., Aarons, G. A., Esp, S., Sklar, M., Carandang, K., Vega, N. R., Brookman-Frazee, L., & Marcus, S. C. (2024). Improving measurement-based care implementation in youth mental health through organizational leadership and climate: A mechanistic analysis within a randomized trial. *Implementation Science*, 19, Article 29. https://doi.org/10.1186/s13012-024-01356-w

Williams, N. J., Hugh, M. L., Cooney, D. J., Worley, J., & Locke, J. (2022). Testing a theory of implementation leadership and climate across autism evidence-based behavioral health intervention of varying complexity. *Behavior Therapy*, 53(5), 900–912. https://doi.org/10.1016/j.beth.2022.03.001

Williams, N. J., Marcus, S. M., Ehrhart, M. G., Sklar, M., Esp, S. M., Carandang, K., Vega, N., Gomes, A. E., Brookman-Frazee, L., & Aarons, G. A. (2023). Randomized trial of an organizational implementation strategy to improve measurement-based care fidelity and youth outcomes in community mental health. *Journal of the American Academy of Child and Adolescent Psychiatry*, 63(10), 991–1004. https://doi.org/10.1016/j.jaac.2023.11.010

Zhang, Y., Cook, C., Fallon, L., Corbin, C., Ehrhart, M., Brown, E., Locke, J., & Lyon, A. (2023). 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 cross-sectional study. *Administration and Policy in Mental Health*, 50(3), 427–449. https://doi.org/10.1007/s10488-022-01248-5