Using Stakeholder Input to Guide Data Visualization and Reporting to Promote Evidence-based Practice Use in Public Schools

Jill Locke[1] · Catherine M. Corbin[1] · Clayton R. Cook[2] · Mark G. Ehrhart[3] · Chayna Davis[1] · Aaron R. Lyon[1]

Abstract

Background High-quality implementation is crucial for students to reap the benefits of school-based evidence-based practices (EBP). Despite data being routinely used to support EBP delivery to students, there is a dearth of data-driven decision-making related to school-wide implementation of universal EBPs like Multi-Tiered Systems of Support (MTSS). The lack of specialized knowledge (e.g., what data to collect and how to interpret it) and systems (e.g., data teams) required to support data collection, analysis, and presentation act as barriers to school teams' effective use of data to guide and be responsive to implementation efforts.

Methods Guided by principles of human-centered design, semi-structured interviews were conducted with 30 school-based implementation practitioners and stakeholders (principals, administrators, and expert intermediaries) to guide the development of data reports that provided information on a school or district's implementation context (leadership, climate, and citizenship behavior).

Results Four themes emerged from the interviews including: (1) interpretation and description of the report/feedback; (2) application of data to districts and/or schools; (3) reactions to the report; and (4) use of the report. The results of this qualitative case study point to the importance of incorporating stakeholder feedback as a methodology to ensure the end product (e.g., data report) is meaningful and applicable to the setting and has direct implications for how to incorporate stakeholder feedback to help shape and improve data visualization and interpretation for better use in schools' decision-making process to support MTSS and other EBP implementation.

Conclusions Practical implications related to report redesign and the utility of well-designed data products to support school-based implementation are discussed.

Keywords Schools · Data visualization · Evidence-based practices · Stakeholders

The research-to-practice gap is a pervasive reality for schools and limits the effectiveness of evidence-based practices (EBPs) for students' social, emotional, behavioral, and academic achievement (Sanetti & Collier-Meek, 2019). Multi-tiered Systems of Support (MTSS) is a widely used prevention framework that uses data to facilitate schools' selection of appropriate EBPs aligned with student need (e.g., Tier 1 is universal programming delivered to all students), monitor students' progress, and adjust service delivery based on that progress (Fuchs & Fuchs, 2006). However,

✉ Jill Locke
jjlocke@uw.edu

Catherine M. Corbin
cmcorbin@uw.edu

Clayton R. Cook
clay@characterstrong.com

Mark G. Ehrhart
mark.ehrhart@ucf.edu

Chayna Davis
chay.j.davis@kp.org

Aaron R. Lyon
lyona@uw.edu

[1] Department of Psychiatry and Behavioral Sciences, University of Washington, 6200 NE 74th Street, Suite 100, Seattle, WA 98115, USA

[2] Character Strong, Minneapolis, MN, USA

[3] Department of Psychology, University of Central Florida, Orlando, FL, USA

low-quality implementation attenuates students' gains making clear that schools need ways to be continuously responsive to faltering implementation. Access to data can be a powerful tool in this regard. School teams equipped with information about their school's implementation context can guide resources toward enhancing aspects of that context (leadership behaviors, climate, citizenship behaviors) known to undergird high-quality implementation (Williams et al., 2020). However, data-driven decision-making often is constrained because school teams lack the specialized knowledge required to effectively collect, analyze, and interpret data (Kippers et al., 2018). Human-centered design (HCD) is a useful approach to improve educators' data-driven decision-making. HCD leverages the strengths, limitations, and experiences of people who use a product (e.g., a data report) to inform redesign of that product to better meet the needs of the user (Lyon et al., 2020a, 2020b). Including key stakeholders in the process of developing tools like data reports is crucial to effective use of those tools to enhance implementation.

Data-Driven Decision-Making

MTSS offers a way to organize the selection and delivery of EBPs (Sugai & Horner, 2009) and includes four key elements: screening, progress monitoring, multiple tiers of support based on student need, and data-driven decision-making (Chard et al., 2008). Central to the success of MTSS is making timely, data-driven decisions to prevent and address social, emotional, behavioral, and academic struggles (McIntosh & Goodman, 2016). While student data are commonly used to monitor progress of and adapt service delivery (Buzhardt et al., 2020), data also can be used to inform and support implementers' use of EBPs within a MTSS framework (Sanetti & Collier-Meek, 2015). The Consolidated Framework for Implementation Research (CFIR; Damschroder et al., 2009) is widely applied to understand EBP implementation and outlines a wide range of implementation determinants, including those at the level of the inner setting (e.g., school implementation climate). Key implementation determinants can be measured to provide formative feedback that school-based implementation teams can use to guide, adapt, and monitor EBP implementation. However, there are several barriers that constrain the ability of schools to engage in the data-driven decision-making that would best support EBP implementation.

Barriers to Data-Driven Decision-Making

Many school teams make use of distributed leadership structures that gather and use data to inform efforts to improve aspects of the school organizational context in relation to implementation. There are currently no established

procedures that provide teams with data on implementation-specific organizational factors as part of a structured decision-making and action planning process. In general, school teams need to have the right members (e.g., formal and informal leaders) to create distributed leadership and establish key norms (e.g., adopt solution-focused conversations, limit side-bar conversations, engage in respectful communication) that enable effective collaboration and productivity (Salas et al., 2018). These alone, however, are insufficient to promote a system-level assessment-to-action process that drives improvements in implementation-specific organizational factors (Rosenfield et al., 2018). In addition, school teams also need feasible, low-cost yet effective supports that enable them to gather reliable and valid data on implementation-specific organizational factors and use these data to inform decisions to develop, deploy, and monitor the impact of action plans (i.e., detailed plans regarding what, for whom and by whom, when, and to achieve what goal) tailored to context-specific needs in their school. However, it is difficult to set up systems to collect, analyze, and present data in a user-friendly way. This limits the extent to which data can be used to support the high-quality implementation necessary for EBPs to have their intended effect. Fortunately, HCD can help to identify aspects of this process that can be optimized to best meet the needs of school teams.

Human-Centered Design

Problematic design can undermine otherwise appealing and effective products (Karsh, 2004). HCD is dedicated to the development of products that are compelling, intuitive, and effective (International Standards Organization, 1999), and includes methods that ground the development process in information about the needs and desires of people who will ultimately use the product (Courage & Baxter, 2005; Norman & Draper, 1986). HCD processes typically involve: understanding users, their needs, and their contexts/constraints; iteratively developing design solutions; and evaluating the extent to which those solutions meet identified needs (ISO, 1999; Lyon et al., 2020a, 2020b). These steps are intended to result in parsimonious and accessible designs that are usable for the broadest possible range of stakeholders. There has been a recent surge in the application of HCD principles and methods to the development and implementation of psychosocial interventions for use in applied clinical and community settings (Altman et al., 2018; Lyon et al., 2021, 2020a, 2020b). To ensure a product's appropriateness, development should incorporate key processes including iterative prototyping and frequent feedback from users in the context of ongoing usability testing (Rubin et al., 2008; Zhou, 2007).

Although they bring novel methods, HCD processes are highly compatible with existing implementation frameworks. In particular, multilevel implementation determinants such as those identified in the CFIR are particularly useful for supporting early steps of the HCD process and surfacing information about users, needs, and contexts that can drive subsequent redesign efforts (Lyon et al., 2019). While innovation-level determinants (e.g., design quality) might lead directly to high-priority usability issues (Munson et al., 2022), individual, inner, and outer setting factors can yield important design constraints that design solutions should be able to address.

Appropriateness of products is crucial to high-quality implementation in schools. Once evaluations end, permanent products are left in the implementation setting (e.g., schools, classrooms) without research-related resources to support ongoing use. EBPs often are implemented without the collaboration of stakeholders in the setting in which the EBP will be used. Stakeholders are infrequently invited to partner in the implementation effort to offer input and feedback as to whether an EBP is usable, feasible, or appropriate for schools, which leads to low use and sustainment (Triplett et al., 2021). Local stakeholders are uniquely situated to provide feedback about aspects of the implementation context that facilitate or inhibit successful EBP implementation; they can offer insight to the leadership practices (e.g., proactively removing barriers to implementing an EBP), climate (e.g., recognition of staff expertise related to EBP use), and citizenship behaviors (e.g., helping others to ensure proper implementation of an EBP) known to support implementation (Lyon et al., 2018; Williams et al., 2020). As frontline implementers, local stakeholders also are uniquely situated to provide feedback on whether a product intended to support implementation efforts (e.g., implementation context feedback report) is usable in the context of active EBP implementation.

Present Study

This study occurred as part of a larger federally-funded project that examined the iterative adaptation and validation of measures capturing key constructs of the school organizational implementation context conducted in the Northwestern and Midwestern United States of America (Lyon et al., under review). Teachers were asked to complete a suite of measures that capture key organizational factors in a school that influence educators' EBP adoption, delivery, and sustainment. The measures assessed implementation leadership, implementation climate, and implementation citizenship behavior (Lyon et al., 2022; Thayer et al., 2022). Data from the original study were aggregated at the school and district levels, and a feedback report was generated. The purpose of this qualitative case study was to gather stakeholder feedback from multiple perspectives (expert intermediaries, district administrators, and school principals) about the feedback report to inform the best ways to use data to drive MTSS implementation in public schools. This study serves as an example for how implementation scientists can gather formative feedback from implementation practitioners and other key stakeholders to help them strengthen various aspects of implementation processes, including data-driven implementation strategies, which is a gap in the literature. We addressed the following research question: How does including school-based implementation practitioners' feedback support the development of a meaningful and useable end product (e.g., data report)?

Methods

Setting and Participants

The University of Washington institutional review board approved the study (Study No. 52311). Participants were purposively recruited in two ways. First, principals ($n = 52$) and district administrators ($n = 11$) of schools that participated in the larger study were eligible to participate. Second, principals and district administrators nominated other principals and district administrators. Eight expert intermediaries, defined as organizations that disseminate information about and/or provide support for the adoption of specific EBPs (Proctor et al., 2019) were recruited from two prominent purveyor organizations for the selected EBPs (i.e., Positive Behavior and Intervention Supports (PBIS) and Promoting Alternative Thinking Strategies (PATHS)). PBIS (Sugai & Horner, 2009) and PATHS (Kusché & Greenberg, 2005) were selected in the parent grant to have two different evidence-based practices to be able to show the measures that were developed—School Implementation Climate Scale (Thayer et al., 2022), School Implementation Leadership Scale (Lyon et al., 2022), School Implementation Citizenship Behavior Scale (Ehrhart et al., 2015; Lyon et al., 2018)—were implementation agnostic processes. Expert intermediaries were prioritized because they work to implement interventions in schools and districts across the United States, which gives them a unique perspective of the barriers to educators' use of data to support effective EBP implementation. The data report generated for this study is intended for use by school leaders and district administrators (for similar example, see Skar et al., 2022), which is why other school and district personnel were not recruited. Potential participants were contacted via email or phone, invited to participate in a 60–90-min interview, and informed of the $100 gift card incentive. This recruitment process resulted in a total of 30 individuals ($n = 4$ expert intermediaries, $n = 8$

district administrators, and $n = 18$ principals). All principals and four district administrators were from the parent project and represented four relatively large school districts in the Northwestern United States and two large school districts in the Midwestern United States. Four district administrators from three additional Northwestern districts were successfully recruited via nomination from other district administrators. Thirty-seven principal and district administrator participants from the parent study did not respond to this study's invitation. See Table 1 for recruitment rates by participant type and district. All participants provided informed consent. Qualitative research prioritizes saturation of information over a specific number of respondents. However, there is growing consensus that 20–30 semi-structured interviews are needed to reach saturation, though the number could be as low as 9 given a homogenous target population (Boddy, 2016; Hennick & Kaiser, 2022).

Principals and district administrators worked in districts that were socioeconomically (approximately 55.2% identify as low income) and racially/ethnically diverse (34.8% White; 23.7% Hispanic/Latinx; 18.9% Black /African American; 11.2% Asian; 8.1% Multiethnic; 1.4% Native Hawaiian/Pacific Islander; 1.0% American Indian or Alaskan Native). The sample was predominantly female ($n = 15$, 51.7%) with an average age range of 35–44 years. The ethnic backgrounds were: 82.8% white, 6.9% Asian, 3.4% African American, and 3.4.% multiracial/multiethnic. Their highest educational attainment was: 31.0% had a doctoral degree, 62.1% had a master's degree, and 3.4% had a bachelor's degree. See Table 2.

Procedures

Data from the original study were aggregated at the school and district levels, and a report was generated. During the week prior to the interview, principals and district administrators were emailed a report with aggregated data for their individual school or district, respectively; expert intermediaries were sent a district-level report as an example. The report included: (1) mean ratings for each construct (school implementation leadership [SILS], school implementation climate [SICS], and school implementation citizenship behavior [S-ICBS]) for that school/district; (2) mean ratings for each construct for that school/district; and (3) benchmark ratings. Organizational theory (e.g., theory of organizational change for readiness; theory of strategic implementation leadership and implementation climate) suggests that these aspects of the implementation context should be similar among implementers and consistently enacted to meaningfully influence EBP implementation (Aarons et al., 2011; Weiner et al., 2011; Williams et al., 2020). When aspects of the implementation context are similarly shared among implementers, there is alignment with regard to perceptions of the organization and suggest a more positive and conducive implementation context. Empirical evidence shows organizational averages of these constructs range from 1.93 (Ehrhart et al., 2014) to 2.42 (Aarons et al., 2014), suggesting a benchmark would need to be set higher to ensure room for improvement was possible (scale range for both constructs is 0–4). As such, benchmarks were set to "3-To a Very Great Extent" for implementation leadership and climate and "3-Fairly Often" for S-ICBS (see Figs. 1–4 in the supplementary materials for an excerpt from a sample data report).

It was infeasible to have all respondents provide feedback on each of the 18 subscales included across the SILS ($n = 7$), SICS ($n = 7$), and S-ICBS ($n = 4$) (see Table 3 for description of each subscale). Therefore, to ensure adequate information was gathered for each subscale while not compromising data integrity due to interviewee fatigue, participants were randomly assigned subscales to review for each construct. Participants also were allowed to select subscales that seemed important to them based on the data. Selections were well distributed with at least two participants opting to focus their

Table 1 Recruitment and participation rates by district and participant type

	Principals			District administrators		
	Recruited	Consented	Percent (%)	Recruited	Consented	Percent (%)
Parent study						
District 1	12	4	33	2	0	0
District 2	10	4	40	2	1	50
District 3	4	2	50	2	2	100
District 4	13	5	38	1	1	100
District 5	4	1	25	0	0	n/a
District 6	9	2	22	0	0	n/a
Nominations						
District 7				1	1	100
District 8				1	1	100
District 9				2	2	100

Table 2 Demographics

Variable	Whole Sample (%)	Central Admin (%)	Principals (%)	Experts (%)
Age				
18 to 24 years old	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
25 to 34 years old	5 (34.50)	1 (14.30)	4 (23.50)	0 (0.00)
35 to 44 years old	10 (26.70)	4 (57.10)	5 (29.40)	1 (25.00)
45 to 54 years old	8 (27.60)	1 (14.30)	5 (29.40)	2 (50.00)
55 to 64 years old	4 (13.80)	0 (0.00)	3 (17.60)	1 (25.00)
65 to 74 years old	1 (3.40)	1 (14.30)	0 (0.00)	0 (0.00)
75 years or older	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Gender				
Male	13 (44.80)	4 (51.70)	8 (47.10)	1 (25.00)
Female	15 (51.70)	3 (32.90)	9 (52.90)	3 (75.00)
Race/Ethnicity				
Am. Indian or Alaskan Native	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Asian	2 (6.90)	1 (14.30)	1 (5.90)	0 (0.00)
Black or African Am.	1 (3.40)	0 (0.00)	1 (5.90)	0 (0.00)
Native Hawaiian or Other Pacific Islander	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
White or Caucasian	24 (82.80)	5 (71.40)	15 (88.20)	4 (100.00)
Multiethnic	1 (3.40)	1 (14.30)	0 (0.00)	0 (0.00)
Advanced degree				
Bachelor's degree	1 (3.40)	0 (0.00)	1 (5.90)	0 (0.00)
Master's degree	18 (62.10)	5 (71.40)	13 (76.50)	0 (0.00)
Doctoral degree	9 (31.00%)	2 (28.60)	3 (17.60)	4 (100.00)

$N = 1$ principal and 1 central administrator did not provide demographic information. Percentages exclude these datapoints to sum to 100%

Table 3 Description of School Implementation Leadership, Climate, and Implementation Citizenship Behavior subscales

Description	
School implementation leadership	
Knowledgeable	Conveys accurate information about and answers questions about the EBP
Supportive	Supports and appreciates staffs' efforts to use and successfully implement the EBP
Proactive	Develops plans to facilitate and identifies/removes obstacles to EBP implementation
Perseverant	Successfully navigates the challenges that arise from EBP implementation
Available	Is available to staff to discuss and/or troubleshoot EBP implementation
Communication	Establishes and encourages clear and open lines of communication about EBP implementation
Vision/Mission	Links EBP implementation with the mission of the school and improved student outcomes
School implementation climate	
Focus on EBP	EBP implementation is a top school priority that staff thinks is important
Educational support for EBP	Resources (e.g., trainings, conferences) are made available to support EBP implementation
Recognition for EBP	Staff are rewarded/acknowledged for their successful implementation of EBP
Rewards for EBP	Small perks or incentives (e.g., coffee cards) are provided to staff who use EBP
Use of data for EBP	Data are collected/used to identify EBP implementation barriers, progress, and improvements
Existing supports for EBP	EBP implementation is supported by protecting staff time and use of structured meetings (PLCs)
Integration of EBP	EBP implementation is integrated with existing EBPs and systems (performance evaluations)
School implementation citizenship behaviors	
Helping others	Staff help one another to implement EBP
Keeping informed	Staff remain abreast of the schools' policies and practices surrounding EBP implementation
Taking initiative	Staff voluntarily take on additional tasks or responsibilities to support EBP implementation
Advocacy/Boosterism	Staff are open advocates for an proponents of EBP implementation

planning efforts on each available subscale, which ensured adequate representation that each subscale was reviewed and discussed. Individual semi-structured interviews were held for each stakeholder type to allow for comparisons across groups.

A systematic and comprehensive interview guide to capture feedback on the data report was used. The interview guide reviewed definitions of the organizational factors, checked for comprehension, and followed a think-aloud protocol (Benbunan-Fich, 2001) where participants told the interviewer what they were thinking as they processed the report. Interview questions elicited open-ended responses about the: (1) interpretation of data (e.g., "On the report, please review one graph under implementation climate by describing what you are looking at and how you are interpreting what you are seeing."); (2) accuracy of the report (e.g., "How accurately does this report reflect implementation leadership in your school?" and (3) use of report (e.g., "Please describe some strategies you might use to address this issue."). Three female research study coordinators with BA or higher degrees who had at least one year of qualitative research experience conducted all interviews via videoconference. Interviewers had no prior interactions with participants. Only the interviewer and the interviewee were present. At the start of the interview, interviewers provided the purpose of the research study. All interviews were audio recorded and lasted approximately 60 min.

Data Analysis

Interviews were transcribed and uploaded to NVivo QSR 12 for data management. The coding scheme was developed using a rigorous, systematic, transparent, and iterative approach using the following steps. First, four members of the research team independently coded two initial transcripts to identify recurring codes. Second, they met as a group to discuss recurring codes and developed a codebook using an integrated approach as certain codes were conceptualized during the interview guide development (i.e., deductive approach), and other codes were developed through a close reading of the two transcripts (i.e., inductive approach; Bradley et al., 2007; Neale, 2016). Next, the research team met to discuss common codes interpreted from the transcripts to include in the final codebook. Then, operational definitions of each code were documented as well as examples of when to use and not use the code. See Table 4.

The coding scheme was applied to the data to produce a descriptive analysis of each code and refined throughout the data analytic process (Bradley et al., 2007). Two members of the research team coded all data and overlapped on 20% of randomly selected transcripts to determine interrater reliability. They met together on a weekly basis to discuss, clarify, verify, and compare emerging codes to ensure

consensus. Agreement between raters was excellent (percent agreement, which relies on the proportion of agreement of coded units between two independent judges = 92.1% for expert intermediaries, 94.7% for district administrators, and 96.3% for principals). Data saturation was reached at the point at which no new insights were obtained, and no new themes were identified when the codebook was applied across the text segments (Guest et al., 2006; Saunders et al., 2018).

Results

Interpretation/Description of the Report/Feedback

Many participants felt the data report was useful, well organized, straightforward, easily digestible with a nice, visually appealing layout. The majority preferred the consistent use of color and formatting (spacing, font size) on all graphs as well as consistent presentation of data (e.g., scales, number of respondents, median vs. mean reporting). One principal noted, "the line graphs clearly show where the median is and where my school compares to others—which is very easy to decipher." Participants also noted legends were helpful. Highlights of the data reports included the discussion/interpretation blurbs of the data. An expert participant commented, "…above the bar graph there's the reminder, "Implementation citizen behavior scale assesses…" So, [I know] what am I looking at. I think that the keys and the summaries under the item graph is helpful." And a district administrator noted:

> The plot shows how the example school's leader is perceived to engage in these behaviors.' So that's cool, I like that. I felt like I wanted that above some of my tables before. I'm actually looking at those to see if I missed it, because I really like that, it reminds me about what the components were. I like that there's no acronyms, but everything's completely written out. It's really helpful.

Participants highlighted the visual representation of individual schools in relation to other schools in their district and the benchmark ratings in terms of "areas of growth for the school" as indicated by "handy little messages down there that tell me—'zero percent of the respondents answered favorably.' So, I need to pay attention to this." One expert participant noted that the visual representation allowed for easy comparisons between the school and the benchmark ratings:

> [It's] easy to compare the schools to one another within the district, I'd be really interested in what this looks like over time because this is a single snapshot…it's

Table 4 Definitions of Codes

Code	Definition and examples
Interpretation/ description of the report/feedback	How the participant interprets, describes, and processes the data for the report. This is their unemotional description of the report; emotional interpretations will be coded in "Reactions to Report." Do not code if they are quoting the report for the primary purposes of the other codes (i.e., Application, Use of Report). Do not code just because they are reading aloud a certain sentence of the report I. Participant description of the report II. General understanding of the report III. What they think the report means Suggestions for things to change/clarify within existing aspects of the report or constructs (scales and subscales). What they liked and things that worked well about the report. I. Visual improvements (e.g., different colors) II. The usefulness of the report (e.g., understandable; "Nothing was bad about the report")
Application of data to district/ school in district	How the participant applies and evaluates the organizational implementation context constructs to their school. This code is more of the background information on the school and is not action-plan oriented I. Justification of score II. Accuracy of score based on example III. Examples of constructs in school
Reactions to report	How the participant emotionally reacts to, judges, self-reflects, and internalizes the results of the report. This entails both positive and negative feelings about the report and data I. How the data "feels" to the participant II. Participant emotions III. Self-reflection of school's performance IV. Participant bias and perception
Use of reports	How the participant plans to use the data from the report to action-plan or make implementation changes in their school positively or negatively. In what ways they will use the report (how they use the report. [Note]: It is okay to apply this code if responding via a hypothetical situation (e.g., "if additions/modifications were made, then I could… [insert action plan]." I. Areas of improvement to focus on or strength to leverage II. Action planning

really easy to look at that school and say oh, my gosh, we got to do something for them. And it's easy to look and say overall, we're doing pretty well in this area, and maybe room to grow in other ones.

Participants, especially principals, also explicitly noted the readability of each graph as the layout and distribution allows for "very quick skim [ming] across to draw some conclusions from, whereas a lot of times when you have a lot of data that's compiled into one report it can be a little bit overwhelming." Other participants noted that it was important to receive "legitimate staff feedback" on EBP implementation and school practices. One principal said, "It's a great way to kind of assess where we're at."

Interestingly, participants noted a number of points of confusion related to each of the constructs and recommended clear definitions of constructs be included and repeated in all relevant sections. One principal suggested including specific school examples underneath each item. Many participants, despite role, requested modifications that pertained to stylistic visualization of the data including color choice (avoid yellow and colors with a traditionally negative connotation such as "red" which can be associated with "incorrect/wrong") and font size. District participants also noted that data points on individual graphs should correspond to specific school sites to aid in data interpretation and application. One district administrator commented, "I'm struggling because I would like to know what schools they are, so if you could ideally put schools, that would give me a little bit more context…" Many participants also preferred more detailed information with specific recommendations to address areas of improvement. One expert participant commented,

> …I would love to see the district averages for those because for this school that I'm looking at, it looks like 'Rewards' is incredibly low…I'd want to know if that's consistent across the district or is this a particular area that we want to focus on.

Participants recommended that: (1) important information be highlighted to draw the reader's attention (e.g., strengths and areas of improvement); (2) the number of graphs be limited to maintain the reader's attention; (3) information at the school and district level be presented and readily digested; and (4) provide specific, actionable strategies to address constructs moving forward. With regard to the latter, participants noted specific strategies that may have worked for other schools that achieved higher marks were of interest.

Application of Data to District/School in District

Participants' description of the application of data to the overall district and/or individual schools varied. Some participants noted that the constructs represented may be inappropriate in some domains. For example, the data highlighted the disconnect between staff EBP implementation and rewards (e.g., promotion and other incentives, etc.). For rewards, a number of participants commented on how schools were not able to provide time off as a reward or other traditional incentives such as bonuses and raises. A district administrator commented,

> I definitely would throw out, after looking at "Rewards," and going back to the definition of what rewards is, the use of financial incentives, bonus and raises. This isn't even applicable really to public schools because we don't do bonuses and raises associated with the use of anything. There could be some comp time, potentially associated with some of the work, but that varies so much by building. So, I'm not going to focus as much [on Rewards], even though that's the lowest [score].

Principals also commented on the reason for low "Rewards" scores:

> I look at "Rewards" as being so low, and I'd want to look into what's the definition of "Rewards" to make sure, are we really that bad? 'Our school provides perks, incentives, coffee cards [to teachers/staff] who use evidence-based practices.' I would refuse to do that ever. We're not allowed to gift public funds, and we are a taxpayer institution, so the only time they get perks or incentives are if our community organizations donate.

Another expert participant highlighted that schools may include personnel who use the data and others who do not use collected data, which may be a challenge for applicability. Several participants noted that the data were positive as noted here:

> As I was going through the list here I see that "Advocacy" is an area that I think there's some good positives across the board-- moderate to great and higher. And so that shows that we can build from the fact that people here do want kids to be successful, that there's a strong sense of advocacy which is why they're working in a Title I school with high mobility and high free and reduced lunch. So, that's a thing that brings people to work in this particular environment.

Other participants illuminated areas that they were not aware were an issue (e.g., schedules need to be amended so teachers have more time for collaboration and training,

communication between the district and school buildings with regard to EBP implementation). One principal remarked,

> So, 'this school uses professional development time to support staff to use evidence-based practices over time.' My first thought is I've done nothing but use evidence-based practices in both academics and behavior, so those numbers make me go, "what!?" Yeah. Every single thing that we do, it comes out of research, or I don't do it. So that means that the respondent lens— I'm not communicating effectively enough that it's all research-based. That's what that tells me.

Across all levels, participants provided explanations for low scores. These included a heavy workload, school and district size, and communication barriers between district administrators and principals as well as principals and staff. Principals reflected on data accuracy and how the data reports met their expectations. There were a few principals who were surprised by their data citing existing systems that would apply to items on each scale that they thought were in place. For example, one principal commented:

> The school connects implementation of EBP to teacher's school performance evaluations, 30 percent say "not at all". Only 40 percent say to a "great extent". And then another 30 are in the yellow. So, on this graph, there's red now for "not at all", which I'm actually kind of surprised at those numbers, based on what I know our evaluation tool is. So, I'm really surprised at the red and the "slight extent". I mean, we're evaluating teachers, and knowing that best practices are part of their evaluation.

Overall, principals felt that the data report provided evidence that current internal programs, changes, and growth plans were working and that these should continue forward in subsequent years. Some principals said that the report gave information on how to change for the next year, including highlighting lower scoring areas that need to be improved and building on higher scoring items. For example, one principal commented:

> It makes me smile inwardly, the difference on the last graph—'teachers/school staff here are proponents of evidence-based practices' and 'very great' slanted all the way over from 'moderate' all the way to 'very great extent' to the implementations on the far left side— 'the teachers and school staff advocate for EBP implementations in their interactions with other staff.' So, I'm doing it, but I'm not sure if my colleagues are doing it, and we don't talk about it, so we can't tell if we're doing it or not. Just kind of calls to the work that we're doing as this being my second year as principal

at [school] around the culture of collaboration and just general culture of—culture and climate of the staff.

Reactions to the Report

Participants reacted to the data report in both positive and negative ways. Some participants were glad that schools were at or around the benchmark for each construct and encouraged by strong displays of leadership in some schools. One principal commented, "Regardless of whatever the baseline was or who set it, I notice, oh, good, my school is above and below the mean of the other schools." A number of principals reported that they were motivated to "jump off items for improvements (e.g., team effort, growth, and communication) moving forward." Others noted the strong ratings of principals and attributed those ratings to principals' approachability/collaborative problem solving and buy-in of EBPs and various school initiatives.

Some participants had negative reactions to the data report that primarily surrounded concerns around low scores, interpretation of the data, and the need for growth in areas that were scored less favorably. One expert participant noted the disconnect between the data collection and data dissemination to teachers as seen here:

> This one is better than we see in a lot of schools, because a lot of schools collect data, but a lot of them don't know how to use it. So, the fact that a high percentage already are seen as 'moderate or above,' I think is pretty good. We need to know, are there tools that they need to do better decision-making with the data? And then, 'this school collects data about how well the EBP is being implemented.' 12.5 percent say, 'very great.' So, my guess is they're doing it, but most people don't know about it, so how is it they can let people know about it. And then, 'this school provides data-driven feedback to all staff about their delivery.' So, with this one, it's more a negative than a positive. But data-driven feedback—it could be that staff don't know if the feedback they're getting is data-driven. So maybe when they're presenting more of their results, they need to talk about how those are come up with, because some people say to a "great extent."

Some district administrators felt concerned and disappointed about certain schools falling below the benchmark, that teachers/staff reported low confidence and support from leadership, and low EBP use in some schools. One district administrator commented,

> Wow. This is bad. This means that it's the perception of the people who are supposed to be implementing the behavior-based practices—and no one believes that they're getting support or recognition for it. They're

not using data to support it. And it doesn't look like it's being integrated with any kind of fidelity, because it's well-below the benchmark.

Many principals were shocked, hurt, disappointed, and concerned because of low scores on the report and/or the mismatch between their ratings and those of comparable schools in their district. One principal noted,

> The school continues to improve in effort… I am pleased that 63 percent say 'to a great extent' that we have continuous improvement efforts. So, that's good. I'm glad that they recognize that. Concerning though, that there's quite a larger red section in the next one, 'school connects implementation to teacher school staff performance evaluation.'

Use of Reports

Participants noted that the data reports were helpful in identifying strengths and areas of improvement for individual schools. One principal commented it was helpful to see the implementation climate scores as they "were a little more on the lower side" and suggests a more significant area of focus for future work. The data reports also spurred thinking around initiatives that could be implemented district-wide. For example, one district administrator commented:

> I think in order for adults to change behavior, they need time and support to do that. And I think a little bit of recognition goes a long way. So, I would start and focus on those because that might also bring up the delivery of evidence-based practices, which could improve data or the use of data. That people see the benefit of what they're doing and people notice and recognize the hard work it takes.

Other participants said the data reports would help generate action plans for further data collection, monitoring, and analysis. For example, one principal noted,

> But I think "Use of Data" would probably be the area I would do specific work. I mean, all of it needs to be improved, to be honest. But I think the data piece is central to some of the other things—like the recognition, the rewards and even communicating things that come up in other areas. I think the data piece is central for that to be effective.

Participants also recommended that the data reports could be used to improve communication among school staff to increase buy-in around EBP use and between staff and the community as a means to disseminate information as well as inform schools about the status of other schools in the district, which could be an opportunity to learn and grow from higher performing schools. In addition, the data reports

could be a useful tool to discuss among individual school teams. One principal said, "I will share [these data] with the PBIS leadership team…and then ask them to tell me what we're going to do." Lastly, participants said the data reports would be helpful to identify what individual school staff want and need for successful EBP implementation. The data reports also generated strategies for areas of improvement including communication, proactiveness, and availability. Some strategies included coordinating changes with the district office, sharing data, and emphasizing outcomes with all staff, involving the greater community (families, students), and utilizing existing supports. One district administrator noted where they would start based on these data:

> If I go back up to the top, looking at other districts surveyed and where we're at, even with the highest, I do think it starts with leadership. So again, if conditions aren't right—even though again that's the highest, I don't think we can just go to the teachers and say okay, we need to work on this without actually doing some very strategic work. Leadership, both at the school level and also at the district level, and how we're aligning our supports and services –that's probably where I would want to start.

Discussion

Participants provided a rich description of the ways in which educators interpret and process research data as well as concrete recommendations to make data reports more useable and meaningful to support EBP implementation in public schools. The results of this qualitative case study: (1) point to the importance of incorporating stakeholder feedback as a methodology to ensure the end product (e.g., data report) is meaningful and applicable to the setting; and (2) has direct implications for how to incorporate

stakeholder feedback to help shape and improve data visualization and interpretation for better use in schools' decision-making process to support MTSS and other EBP implementation. The most practical implication of these results is the well-defined list of recommendations that educators offered to support the utility and use of data in schools (see Table 5).

The US Department of Education (2008) reports how confident educators feel about their knowledge and skills in data analysis and interpretation affects data collection and the prevalence of data-informed decision-making in public schools. Stakeholders highlighted many elements that could potentially ease the burden of data interpretation and usage. For example, stylistic elements like font choice and size and color of graphical representation helped ease educators' review and processing of data. In addition, stakeholders recommended that data reports need to be simple with limited information (e.g., a few visualizations or graphs/images) that draw the reader to what to focus on (e.g., strengths/areas of improvement), and be clear and interpretable (e.g., all graphical representation should include defined scales, legends, and a brief description of how to disentangle school-specific vs. district level data). Stakeholders also highlighted the importance for data reports to clearly provide definitions of constructs (e.g., implementation leadership, implementation climate, and citizenship) and school-specific examples to illustrate those constructs and recommendations for how to bolster areas of improvement. These stakeholder-generated suggestions for redesign also reflect the advantages of including stakeholders in the development of tools they will be expected to use. Guided by principles of HCD (Lyon et al., 2021), we now have information to improve upon the design of these reports in ways that are likely to promote principals', administrators', and districts' use of them to guide data-driven decision-making in support of high-quality EBP implementation.

Table 5 Stakeholder recommendations by qualitative code

Qualitative codes	Recommendation for improvement
Interpretation/description of the report/feedback	Important information should be highlighted to draw the reader's attention (strengths and improvements).
	The number of graphs should be limited to maintain the reader's attention.
	Information at the school and district level should be presented and digestible.
	Include specific, actionable strategies to address constructs moving forward.
Application of data to district/school in district	Ensure data collected is appropriate to the context.
	Collect data from implementing staff to promote applicability of the results.
Reactions to the report	N/A
Use of reports	Use the report to identify strengths and areas for improvement.
	Use the report to brainstorm district-wide initiatives and/or school level action plans for further data collection, monitoring, and analysis.
	Reports can be used to improve communication among staff to increase buy-in around EBP and/or to support school teams.
	Use reports to identify what staff want and need for successful EBP implementation.

Limitations and Future Directions

First, participants were from six school districts in the Northwestern and Midwestern United States, which limits the geographic generalizability of the findings. Future studies should include a more representative sample given school systems may vary in how they use data to support EBP implementation in their local contexts. The sample represented a subset of principals and administrators that participated in the original study. It is possible that those who agreed to participate in the present study differed from those that did not in ways that artificially constrained the information provided (e.g., those with more familiarity with data may have been more motivated to participate). Moreover, while this study included multiple stakeholders involved with EBP implementation, it did not examine the perspectives of teachers who also play an influential role. Teachers often are asked to lead EBP implementation efforts in their classroom and their perspective on the use of data would be helpful in ensuring accurate interpretation and application of data to drive implementation and sustainment. Lastly, while we did incorporate stakeholder feedback, there was no member checking of results, which can strengthen the accuracy and credibility of qualitative data.

Future research should examine the utility and use of data in schools to inform implementation efforts. There may be a practical and empirical need for a school-level assessment-to-action organizational implementation strategy or system that informs educators' data-driven decisions and actions within a continuous improvement process. A functional and scalable process that capitalizes on school-wide collaboration and commitment to support implementation of EBP may foster a more conducive school organizational context to promote better implementation and outcomes for students. An organizational implementation strategy that uses a quality improvement cycle that when repeated over time generates site-specific action plans that include implementation strategies tailored to cultivate and maintain a supportive organizational implementation context seems worthy of future exploration and research given the importance and underutilization of data-driven decision-making in schools.

Conclusion

This study illustrates the importance of structured reports that capture and organize stakeholder input to support the interpretation and application of data in driving MTSS implementation efforts in public schools. Results have important implications for the design and application of data reports to increase the use of research data in supporting educators' delivery of MTSS, and potentially other EBPs, in public school settings. These findings demonstrate that frontline implementers can interpret and use data about their school or district's implementation context to identify ways to improve upon it. While promising, these findings also demonstrate improvements that can be made to make the data more visually appealing and/or easier to understand, increasing the likelihood that the reports will be helpful for implementers. Such insights underscore the importance of inviting feedback from those who will need to digest and act upon data reports to support MTSS and other EBP implementation in schools.

Data Availability Data are available upon request of the senior author (AL).

Declarations

Conflict of interest All Authors declare that there is no conflict of interest.

Ethical Approval All procedures were approved by the University of Washington IRB (Study No. 52311).

Informed Consent Informed consent was obtained from all participants included in the study.

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*. https://doi.org/10.1186/1748-5908-9-45

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research, 38*, 4–23. https://doi.org/10.1007/s10488-010-0327-7

Altman, M., Huang, T. T. K., & Breland, J. Y. (2018). Design thinking in health care. *Preventing Chronic Disease, 15*, 180128. https://doi.org/10.5888/pcd15.180128

Benbunan-Fich, R. (2001). Using protocol analysis to evaluate the usability of a commercial website. *Information & Management, 39*, 151–163. https://doi.org/10.1016/S0378-7206(01)00085-4

Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research, 19*, 426–432. https://doi.org/10.1108/QMR-06-2016-0053

Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: Developing taxonomy, themes, and theory. *Health Services Research, 42*, 1758–1772. https://doi.org/10.1111/j.1475-6773.2006.00684.x

Buzhardt, J., Greenwood, C. R., Jia, F., Walker, D., Schneider, N., Larson, A. L., Valdovinos, M., & McConnell, S. R. (2020). Technology to guide data-driven intervention decisions: Effects on language growth of young children at risk for language delay. *Exceptional Children, 87*, 74–91. https://doi.org/10.1177/0014402920938003

Chard, D. J., Harn, B. A., Sugai, G., Horner, R. H., Simmons, D. C., & Kame'enui, E. J. (2008). Core features of multi-tiered systems of reading and behavior support. In C. R. Greenwood, T. R. Kratochwill, & M. Clements (Eds.), *Schoolwide prevention models: Lessons learned in elementary schools* (pp. 31–58). The Guilford Press.

Courage, C., & Baxter, K. (2005). *Understanding your users: A practical guide to user requirements methods, tools, and techniques*. Elsevier.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research finings into practice: A consolidated framework for advancing implementation science. *Implementation Science, 4*, 50. https://doi.org/10.1186/1748-5908-4-50

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*. 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*, 65. https://doi.org/10.1186/s13012-015-0255-8

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*(1), 93–99. https://doi.org/10.1598/RRQ.41.1.4

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*, 59–82.

Hennick, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine, 292*, 114523. https://doi.org/10.1016/j.socscimed.2021.114523

International Standards Organization. (1999). *Ergonomics of human system interaction-Part: Human-centered design for interactive systems*. Retrieved from https://www.iso.org/standard/77520.html

Karsh, B. T. (2004). Beyond usability: Designing effective technology implementation systems to promote patient safety. *BMJ Quality & Safety, 13*, 388–394. https://doi.org/10.1136/qshc.2004.010322

Kippers, W. B., Poortman, C. L., Schildkamp, K., & Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention? *Studies in Educational Evaluation, 56*, 21–31. https://doi.org/10.1016/j.stueduc.2017.11.001

Kusché, C. A., & Greenberg, M. T. (2005). *Promoting alternative thinking strategies*. Channing Bete Company.

Lyon, A. R., Brewer, S. K., & Areán, P. A. (2020a). Leveraging human-centered design to implement modern psychological science: Return on an early investment. *American Psychologist, 75*, 1067–1079. https://doi.org/10.1037/amp0000652

Lyon, A. R., Coifman, J., Cook, H., McRee, E., Liu, F. F., Lidwig, K., Dorsey, S., Koerner, K., Munson, S. A., & McCauley, E. (2021). The Cognitive Walkthrough for Implementation Strategies (CWIS): A pragmatic method for assessing implementation strategy usability. *Implementation Science Communications*. https://doi.org/10.1186/s43058-021-00183-0

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*. 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*, 48. https://doi.org/10.1186/s13012-022-01222-7

Lyon, A. R., Dopp, A. R., Brewer, S. K., Kientz, J. A., & Munson, S. A. (2020b). Designing the future of children's mental health services. *Administration and Policy in Mental Health and Mental Health Services Research, 47*, 735–751.

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), e14990.

McIntosh, K., & Goodman, S. (2016). *Integrated multi-tiered systems of support: Blending RTI and PBIS*. Guilford.

Munson, S. A., Friedman, E. C., Osterhage, K., Allred, R., Pullmann, M. D., Areán, P. A., Lyon, A. R., UW ALACRITY Center Researchers. (2022). Usability issues in evidence-based psychosocial interventions and implementation strategies: Cross-project analysis. *Journal of Medical Internet Research, 24*(6), e37585.

Neale, J. (2016). Iterative categorization (IC): A systematic technique for analysing qualitative data. *Addiction, 111*, 1096–1106. https://doi.org/10.1111/add.13314

Norman, D. A., & Draper, S. W. (Eds.). (1986). *User centered system design: New perspectives on human-computer interaction* (1st ed.). CRC Press.

Proctor, E., Hooley, C., Morse, A., McCrary, S., Kim, H., & Kohl, P. L. (2019). Intermediary/purveyor organizations for evidence-based interventions in the US child mental health: Characteristics and implementation strategies. *Implementation Science*. https://doi.org/10.1186/s13012-018-0845-3

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

Rubin, J., Chisnell, D., & Spool, J. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests* (2nd ed.). Wiley.

Salas, E., Reyes, D. L., & McDaniel, S. H. (2018). The science of teamwork: Progress, reflections, and the road ahead. *American Psychologist, 73*, 593–600. https://doi.org/10.1037/amp0000334

Sanetti, L. M. H., & Collier-Meek, M. A. (2015). Data-driven delivery of implementation supports in a multi-tiered framework: A pilot study. *Psychology in the Schools, 52*, 815–828. https://doi.org/10.1002/pits.21861

Sanetti, L. M. H., & Collier-Meek, M. A. (2019). Increasing implementation science literacy to address the research-to-practice gap in school psychology. *Journal of School Psychology, 76*, 33–47. https://doi.org/10.1016/j.jsp.2019.07.008

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Heather Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity, 52*(4), 1893–1907.

Skar, A. 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 Leadership and Organizational Change for Implementation (LOCI) intervention on implementation and transformational leadership, and implementation climate. *BMC Health Services Research, 22*, 298. https://doi.org/10.1186/s12913-022-07539-9

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

Thayer, A. J., Brown, E. C., Cook, C. R., Ehrhart, M. G., Locke, J., Davis, C., Picozzi, E., Aarons, G. A., & Lyon, A. R. (2022). Construct validity of the School Implementation Climate Scale (SICS). *Implementation Research and Practice, 3*, 1–14. https://doi.org/10.1177/26334895221116065

Triplett, N. S., Woodard, G. S., Johnson, C., Nguyen, J. K., AlRasheed, R., Song, F., Stoddard, S., Mugisha, J. C., Sievert, K., & Dorsey, S. (2021). How engaged are stakeholders in evidence-based treatment implementation projects? Results from a scoping review of children's mental health treatment projects. *PsyArXiv*. https://doi.org/10.3134/osf.io/pkr8d

U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2008). *Teachers' use of student data systems to improve instruction: 2005 to 2007*. Author.

Weiner, B. J., Belden, C. M., Bergmire, D. M., & Johnston, M. (2011). The meaning and measurement of implementation climate. *Implementation Science*. https://doi.org/10.1186/1748-5908-6-78

Williams, N. J., Wolk, C. B., Becker-Haimes, E., & 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*. https://doi.org/10.1186/s13012-020-0970-7

Zhou, R. (2007). How to quantify user experience: Fuzzy comprehensive evaluation model based on summative usability testing. In N. Aykin (Ed.), *Usability and Internationalization. Global and local user interfaces* (pp. 564–573). Springer.