



HHS Public Access

Author manuscript

Adm Policy Ment Health. Author manuscript; available in PMC 2022 May 01.

Published in final edited form as:

Adm Policy Ment Health. 2021 May ; 48(3): 464–481. doi:10.1007/s10488-020-01082-7.

Exploration, preparation, and implementation of standardized assessment in a multi-agency school behavioral health network

Elizabeth Connors, PhD¹, Gwendolyn Lawson, PhD², Denise Wheatley-Rowe, MSW, RN³, Sharon Hoover, PhD⁴

¹Department of Psychiatry, Yale University, 389 Whitney Avenue, Office 106, New Haven, CT 06511, USA

²Children's Hospital of Philadelphia, Roberts Center for Pediatric Research, 2716 South Street, 8th Floor, Philadelphia, PA 19146, USA

³Behavioral Health System Baltimore, 100 South Charles Street, Tower 2, 8th Floor, Baltimore, MD, 20201, USA

⁴National Center for School Mental Health, Division of Child and Adolescent Psychiatry, University of Maryland School of Medicine, 737 West Lombard Street, 4th Floor, Baltimore, MD, 21201, USA

Abstract

School mental health treatment services offer broad public health impact and could benefit from more widespread implementation and sustainment of standardized assessments (SA). This demonstration study describes one approach to increase SA use in a large school behavioral health network using the Exploration, Preparation, Implementation and Sustainment (EPIS) framework. Mental health administrator interviews with four participating agencies and a multidisciplinary planning team informed SA measure selection and implementation supports. The SA initiative was implemented during one school year, including system-wide training and ongoing implementation supports for 95 clinicians. Linear mixed effect models revealed improvements in clinician attitudes about the SA for clinical utility and treatment planning immediately following the half-day training ($N=95$, $p < .001$). Clinicians self-reported a significant increase in use of SA for new clients during intakes ($p < .001$) over time and 71.4% of expected SA data were submitted. Qualitative feedback, including recommendations to offer more SA choices and beginning new SA data collection earlier in the school year, was integrated to inform quality improvements and future sustainment efforts.

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <https://www.springer.com/aam-terms-v1>

Corresponding Author: elizabeth.connors@yale.edu phone: 1-203-815-0685.

Publisher's Disclaimer: This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

Compliance with Ethical Standards: None of the authors have any potential conflicts of interest. This study was reviewed and approved by the University of Maryland Human Research Protections Office as exempt from Institutional Review Board review. Clinicians were informed that their survey responses are confidential and would be used in aggregate to inform continuous quality improvement of the implementation initiative.

Keywords

school behavioral health services; standardized assessment; evidence-based practice; adoption; sustainment; implementation context

To improve and standardize the quality of mental health services, a large body of health services and implementation research has been dedicated to promoting mental health clinicians' use of evidence-based interventions. This work is complemented by a concurrent and growing focus on promoting clinicians' use of evidence-based approaches to assessment and measurement in clinical practice, often referred to as evidence-based assessment, feedback-informed treatment or measurement-based care (Arora et al., 2016; Bickman, Lyon & Wolpert, 2016; Fortney et al., 2017; Purbeck et al., 2019).

Reliable and valid standardized assessment (SA) tools improve the accuracy of clinical judgment and are considered a core component of an evidence-based approach to treatment (Jensen-Doss & Hawley, 2010; Lyon et al., 2017; Scott & Lewis, 2015). SA tools can be administered at all stages of clinical practice, including screening, treatment planning, treatment monitoring, and outcome evaluation. Using SA tools to track progress and inform treatment planning has been found to improve client outcomes. However, effect sizes range widely, from .28 to .70; larger effects are observed when a feedback component is included, clinical support tools are provided, and for cases "not on track" (Lambert, Whipple & Kleinstauber, 2018; Fortney et al., 2017; Krageloh, Czuba, Billington, Kersten, & Siegert, 2015; Shimohawa, Lambert & Smart, 2010). A recent Cochrane review found little to no benefit of using SA (Kendrick et al., 2016). Yet, studies in which clinicians used the SA data to adjust the treatment regimen based on patient-reported progress and feedback were excluded. This practice is regarded as a primary reason to collect SA to improve outcomes, making the results from this methodology difficult to interpret (Resnick & Hoff, 2019; Scott & Lewis, 2015; Lewis et al., 2019). A second Cochrane review intended to synthesize the evidence on use of client feedback tools in child and adolescent psychotherapy was inconclusive; only six randomized controlled trials that compared feedback to no feedback for youth psychotherapy were located and five of those studies had incomplete outcome data due to attrition, precluding the authors' ability to pool results across studies (Bergman et al., 2018). The attrition bias limitations among the few most rigorous trials of feedback in youth psychotherapy underscores the need for additional research on how to successfully implement SA practices in youth psychotherapy. Also, the preponderance of meta-analyses and systematic reviews of SA implementation is based on a narrow set of adult SA outcome measures, and there is limited consensus on which patient-reported outcome measures are ideal for usability, sensitivity to change over time and and/or patients of diverse cultural and linguistic backgrounds, for adults but especially for youth (Becker-Haimes, et al., 2020; Kendrick et al., 2016). Despite the need for ongoing research about which SA tools and practices optimize outcomes for patients across settings, presenting concerns and age ranges, SA implementation in usual care practice is important to ensure a systematic, reliable, patient-informed and patient-centered process for assessing and treating mental health disorders (Fortney et al., 2017; Valenstein et al., 2009).

Mental health treatment provided in school settings could especially benefit from increased adoption and implementation of SA due to its substantial reach and potential for public health impact (Bohnenkamp, Glascoe, Gracey, Epstein, & Benningfield, 2015; Lyon et al., 2017). Schools are sometimes underrecognized as a critical location for delivery of mental health treatment services. Yet, 45-80% of children who receive treatment do so at school (Burns et al., 1995; Green et al., 2013).

Use of SA in School Mental Health Treatment

Use of SA among *school* mental health clinicians is not well known. However, school mental health clinicians are often asked to collect and report academic indicators of success for students they serve and many schools emphasize response to intervention practices to multi-tiered systems of academic and behavioral supports that in theory require ongoing assessment of student progress in interventions (Connors, Arora, Curtis & Stephan, 2015; Mellard, McKnight & Woods, 2009). Thus, use of SA is a logical fit for mental health treatment delivered in schools. Unfortunately, research documenting practices of mental health clinicians working in non-school settings consistently indicate that routine collection and use of assessment data is not a widespread practice. A recent study found that 61.5% of a diverse sample of mental health practitioners reported never using standardized progress measures, and only 13.9% use them at least monthly (Jensen-Doss et al., 2016). In other studies, approximately 20% of surveyed community mental health clinicians report collecting data routinely in treatment (Bickman et al., 2000; Gilbody, House, & Sheldon, 2001). It is likely that rates of SA use by community-partnered school behavioral health clinicians, who are employed by mental health agencies, have similarly low rates of SA use (for more information on the community-partnered school behavioral health model, see Connors et al., 2019; Lever et al., 2015).

Barriers and Facilitators to Implementing SA in School Mental Health

Barriers and facilitators to implementing SA in schools are not well known, but likely resemble those in community-based treatment settings (e.g., demands on limited time and concern about response burden on patients, as outlined in Lewis et al., 2019). Unique factors affecting evidence-based practice implementation in schools that would apply to SA include school organizational factors (e.g., principal leadership, technological resources, mandates for professional development), treatment delivered by mental health professionals from a diversity of disciplines and implications of the nine-month school calendar (Owens et al., 2014).

Implementation Strategies to Increase SA in School Mental Health

The impact of implementation strategies such as training and consultation to address barriers to implementation and increase SA use among school mental health clinicians is not well-established. One study with 15 mental health clinicians in school-based health centers indicates that use of a digital measurement-feedback system in addition to consultation may be an effective strategy to increase use of assessment tools in schools (Lyon et al., 2017). However, SA implementation research with non-school clinicians provides plausible directions for selecting specific implementation strategies relevant to school clinicians. For

instance, Lyon and colleagues found that clinicians who participated in an intensive statewide training and consultation program for an evidence-based intervention that emphasized SA showed early increases in positive SA attitudes and gradual increases in self-reported SA skill and use (Lyon, Dorsey, Pullmann, Silbaugh-Cowdin, & Berliner, 2015). Moreover, dynamic, multicomponent training plus ongoing consultation and coaching has been shown to increase practitioner's implementation of evidence-based practices generally (Herschell, Kolko, Baumann, & Davis, 2010). Organizational variables such as implementation climate and leadership support have also been found to influence clinicians' practice change in response to training and supports (Aarons, Ehrhart, Farahnak, & Sklar, 2014; Beidas & Kendall, 2010). Successful implementation of other evidence-based practices in schools has been attributed to not only high-quality training and ongoing support but also contextual and organizational factors including administrative support and peer support among clinicians implementing the practice (Forman, Olin, Hoagwood, Crowe, & Saka, 2009; Langley, Nadeem, Kataoka, Stein & Jaycox, 2010).

School districts are well positioned to select and implement standardized assessments on a systemic level for their schools in an effort to improve quality of mental health care offered to students (Bohnenkamp et al., 2015; Sander, Everts, & Johnson, 2011). Guidance documents have been released by national technical assistance centers in response to mounting requests from schools and districts on how to track student progress and outcomes of mental health interventions (see Connors, Wigand, Moffa, Hoover & Lever, 2019 and Wright, with Center for Applied Research Solutions, 2018). However, no guidance exists in the peer-reviewed literature for school district¹ administrators or their community partners providing school-based mental health treatment to adopt and implement SA district-wide. Overall, literature detailing effective, pragmatic efforts to increase school mental health clinician attitudes toward and use of SA – especially for larger school behavioral health systems - is underdeveloped.

Current Study

The current study describes one approach to implementing SA in a school behavioral health network serving a large, urban school district. The methods of this demonstration study are grounded in the EPIS model and pay explicit attention to outer and inner context factors during the Exploration and Preparation phases that influence implementation plans. Implementation phase outcomes are clinician attitudes toward SA and self-reported SA practices to explore effectiveness of training and support on clinician attitudes, knowledge and practices.

The overall goals of this study were to 1) demonstrate one approach to explore, plan and implement SA in a large school behavioral health network, guided by the EPIS framework and 2) explore the impact of training and support on clinician attitudes, practices and experience.

¹The organizing body for a group of schools is called a school district in many states. However, the organizing body of a group of schools might also be referred to as the local education authority, town, county, region, school administrative unit, charter organization or private school company.

Study Context

This study was conducted with a multi-agency school behavioral health network in Baltimore City, Maryland. Behavioral Health System Baltimore (BHSB) is the local behavioral health authority for Baltimore City. BHSB oversees a network of predominantly private, non-profit, behavioral health agencies that deliver services to over 68,000 Baltimore City residents. As part of its portfolio, BHSB also oversees an Expanded School Mental Health Network (ESMH) Network of seven behavioral health agencies (four of which specialize in mental health and participated in the current study) authorized through a competitive request for proposals process to deliver school-based mental health promotion, prevention, early intervention and treatment services and supports in Baltimore City Public Schools. This study was exempt from continuing review by the University of Maryland Human Research Protections Office as implementation outcome data were collected voluntarily and anonymously from agency leaders and clinicians as a component of this quality improvement initiative.

EPIS Framework

Recognizing the need for a phased approach and the importance of multi-level factors influencing implementation, this study used the Exploration, Preparation, Implementation and Sustainment (EPIS) framework to guide its methods (Aarons, Hurlburt, & Horwitz, 2011). EPIS focuses on effective implementation of evidence-based practices in public, child-serving systems with explicit attention to the influence of service contexts on each of the four phases. **Exploration** refers to the awareness that an issue needs attention or improvement (where neither profit nor investigator-initiated research is driving this). **Preparation** refers to the adoption decision and early experimentation with and/or planning for implementation. **Implementation** refers to actual integration or addition of the innovation into the service system. **Sustainment** refers to the continued use of an innovation or practice. At every phase, outer context factors such as the service environment, inter-agency environment, and consumer support or advocacy, as well as inner context factors such as intra-agency characteristics and individual adopter (i.e., clinician) characteristics are considered for their relative influence on the process. In the current study, the entire process of Exploration, Preparation and Implementation lasted about 18 months.

Phase I: Exploration

The ESMH Network leadership wanted to implement a consistent approach to SA among its participating behavior health agencies to improve the quality of school-based mental health services. There were several factors in the outer context that influenced interest in adopting SA. First, BHSB needed to communicate the value of ESMH services to school district partners to justify school mental health services. Second, ESMH Network members participated in an evidence-based assessment workgroup; this provided a catalyst for BHSB's decision to adopt and implement SA. Inner context factors, including the agency's mission to utilize evidence-based care and dynamic and motivated leadership, also inspired SA interest. The Exploration phase helped ESMH leaders develop a network of expertise to support their vision for the initiative, including individuals within their own organization (i.e., BHSB), university partners and other colleagues exploring this topic nationally in

school mental health. During this time, the ESMH Network increased its readiness to make an adoption decision in the Preparation phase.

Phase II: Preparation

At the start of the Preparation phase, ESMH Network leaders made the decision to adopt SA to achieve system-level and clinician-level goals related to care quality. System-level goals were to 1) build capacity to monitor and aggregate student progress and outcomes and 2) demonstrate impact of services on student mental health symptoms. Clinician-level goals were to 1) promote SA practices among the 100+ clinicians serving 126 public schools in Baltimore City and 2) incorporate systematic data collection and feedback to families at quarterly intervals into routine care. However, there were numerous considerations about the intervention-setting fit of SA considering the size of the ESMH Network, workforce capacity within each agency, clinician and leadership knowledge and skills, and agency values. The Preparation phase involved consideration of these factors to optimize the fit of SA tools and practices with each agency. The first year was viewed as a “pilot” year to explore feasibility, the effectiveness of training and implementation supports, and build capacity for this effort at the clinician and system levels. Therefore, the Preparation phase involved consideration of which SA tools to select, data systems and reporting procedures, as well as processes for managing data, communicating with each agency, and supporting agency leaders and individual clinicians during implementation.

Phase II Method

Procedures—Preparation phase procedures consisted of two main activities: 1) phone interviews with each of the four agency administrators and 2) multidisciplinary leadership planning. Phone interviews were conducted to assess how outer context (e.g., accountability requirements from other sources related to demonstrating student outcomes) and inner context (e.g., agency capacity, leadership needs and interests in adopting SA, agency structures such as electronic health records and requirements for their clinicians’ use of SA throughout treatment) factors would influence implementation procedures and outcomes. Interviews lasted 15-20 minutes and were conducted by an invited academic partner (first author). An interview summary was sent to the administrator to check for accuracy before it was finalized and shared with the leadership team. Interview summaries informed multidisciplinary leadership team decisions about selection of the SA(s) for this initiative and plans for making the implementation pilot compatible with all agency goals and structures.

Phase II Results

All agency administrators indicated that routine use of one SA across agencies would be a major shift in clinical and administrative workflow. However, each administrator expressed unsolicited enthusiasm about the initiative and a genuine interest in increasing capacity to monitor student progress and outcomes at the student, clinician, agency, and ESMH Network level. All four agencies had some version of a standardized approach to initial assessment but no ongoing SA use other than what is required for insurance reauthorization every six months. Agency 1 used a SA for their partial hospitalization program but not

school-based mental health providers. Agency 2 recently adopted a new electronic health record that could be customized to add a SA but due to length and training required, hadn't decided whether to use that feature. Agency 3 was planning to ask clinicians to submit a few key indicators of academic success (e.g., attendance) to agency leadership via excel but hadn't started this process. Agency 4 arguably had the most experience with providing professional development and support for SA use to their clinicians; all schools were equipped with an "Assessment Toolkit" of free assessment measures to facilitate their individualized selection of and access to assessment tools and all clinicians were required to have at least 1 diagnosis-specific assessment in the medical record aligned with the primary diagnosis that was checked during annual chart reviews. Otherwise, the clinical teams, patient population, insurance providers, and school settings were equivalent across agencies.

Next, the multidisciplinary leadership team, which included ESMH Network leaders, data team members and academic partners, reviewed findings from administrator interviews and evaluated Network capacity for data systems and technical assistance to support implementation. The leadership team also conducted a comprehensive review of SA tools to inform selection of specific tools to use in the initiative. Decisions about which SA tools to select and the frequency of data collection were informed by agency leader interviews, the service system context, and findings about barriers to the use of SA tools (Jensen-Doss & Hawley, 2010). Finally, prioritized criteria for psychometrically strong and pragmatic SA were developed by the leadership team (see Figure 1) and used for evaluating reviewed measures. After close review of leading options, the team selected the Pediatric Symptom Checklist (17-item version; Jellinek et al., 1988) supplemented by a substance use screening tool, the CRAFFT (Knight et al., 1999). These tools were the only public domain measures that met all criteria shown in Figure 1, specifically covering a broad age range (the PSC assesses for mental health concerns for students ages 4-18 and the CRAFFT assesses for substance use concerns for children over the age of 14). There are very few free, brief, valid measures available in the public domain for children (Becker-Haimes, Tabachnick, Last, Stewart, Hasan-Granier, & Beidas, 2020).

The team decided to pilot assessments on small scale (i.e., four of the most recently-enrolled cases at each school, only two data collection intervals) to test feasibility. Implementation strategies (see Table 1) for the Implementation phase were selected based on evidence that effective training in evidence-based practices must include active initial training, as well as continued implementation supports (Beidas & Kendall, 2010).

Phase III: Implementation

Phase III Method

Participants: Clinicians from four mental health and three substance use agencies participated in the implementation phase. Due to very low numbers of clinicians in the substance use agencies (N=3), we opted to use only mental health clinician data for the current analyses. Ninety-six clinicians completed the survey at baseline and attended the training. Of those, 83 (86%) completed the post-training survey and 76 (79%) completed the exit survey at the end of the year. Fifty-nine (61%) had complete survey data at all time periods. See Table 2 for clinician characteristics. There were seven clinicians in the sample

who either didn't attend the initial training or didn't complete the pre survey beforehand. In those cases, clinicians would have had to rely on their agency leaders' support, presentation and educational materials distributed and ongoing implementation supports to learn about the initiative and expectations.

Procedures: One half-day educational training was provided in the beginning of the school year that included information on the specific SA tools including development, evidence base, scoring and use, as well as background information about the use of systematically-collected patient data throughout the course of treatment to monitor progress, provide feedback to the student and family and inform treatment decisions. Clinicians rehearsed introduction, administration, scoring and feedback in the training and received feedback based on observed skills. Clinicians were required by their agencies to attend this training.

Implementation supports following training included agency supports (e.g., consultation with the ESMH Network about data collection and feedback preferences, e-mail reminders and clarifications to agency administrators before data collection started), two virtual office hours for clinicians and administrators to discuss implementation experiences, ongoing email communication with updated "Frequently Asked Questions," and implementation support materials (e.g., visual aid of response options). A full list of discrete implementation strategies used for clarity in our reporting is included in Table 1, some of which reflect strategies that were part of the Exploration (i.e., develop academic partnerships) and/or Preparation phase (i.e., conduct local consensus discussions, use advisory boards and workgroups).

Measures—Attitudes toward SA, self-reported practice change related to SA, and experiences with implementation were assessed via clinician self-report surveys at three time points: up to one week prior to the half-day training ("pre-training"), immediately following the half-day training ("post-training") and at six-month follow-up ("follow-up").

Attitudes Toward SA.—The Attitudes toward Standardized Assessment Scale – Monitoring and Feedback (ASA-MF; Jensen-Doss et al., 2016) is an 18-item measure of clinicians' attitudes toward SA. It includes three subscales: Clinical Utility (8 items; e.g., "standardized progress measures provide more useful information than other assessments like informal interviews or observations"), Treatment Planning (5 items; e.g., "information from standardized progress measures can help me plan for sessions"), and Practicality (5 items; e.g., "standardized progress measures can efficiently gather information"). Subscales have shown acceptable internal consistency (alphas between .81 and .85). In the current sample, the subscales also show acceptable internal consistency at the three time points (Treatment Planning alphas between .78 and .82; Clinical Utility alphas between .73 and .74; Practicality alphas between .70 and .75).

Clinician-Reported Practice.—The Current Assessment Practice Evaluation (CAPE; Lyon et al., 2017) is a four-item measure of clinician ratings of standardized assessment practices (see items listed in Table 3). Items are scored on a 4-point scale: None, Some (1-39%), Half (40-60%), Most (60-100%). It has shown acceptable internal consistency (alpha = .72). In the current sample, the CAPE shows relatively acceptable internal

consistency ($\alpha = .73, .74, \text{ and } .57$) at all time points. The CAPE was administered at all three time points, but scores (i.e., use of SA practices) would not be expected to increase immediately after the initial training because clinicians did not have an opportunity to apply practices yet.

Clinician Data Submission.—In addition to subjective ratings of SA practices, we also include SA data submission rates by school, within agency. Data were collected electronically or via paper and pencil and then submitted electronically by clinicians via a secure, web-based survey system directly to the ESMH Network data team. These data were collected throughout implementation to provide weekly feedback on data submission progress during data collection intervals. The expectation was for clinicians to submit SA data twice, for four cases per school. This measure is specific to the four expected cases with two administrations during the implementation period.

Implementation Experiences and Recommendations.—Clinicians answered several open-ended questions on the follow-up survey about their experiences, both positive and negative, and recommendations for future implementation and sustainment. Clinicians were also presented a list of issues that might have occurred during their data collection from students and families (e.g., “collecting data had a negative impact on rapport with the caregiver”) and were asked to indicate whether they experienced each issue with “No cases,” “Some cases,” “Most cases,” or “All cases.” These items were developed based on specific concerns clinicians or agency leaders reported anticipating at the time of initial training and/or experiencing during implementation. They were queried on the survey specifically to assess how widespread these barriers were throughout the entire sample of clinicians. Concerns were balanced with positive experiences clinicians could rate (e.g., “collecting data had a positive impact on rapport with the caregiver”).

Analyses

Quantitative Analyses. Data were summarized using descriptive statistics, such as means and standard deviations, counts and percentages. Linear mixed effect models were used to examine clinicians’ attitudes toward standardized assessment (per the ASA) and self-reported standardized assessment practices (per the CAPE, hereafter, “practices”) over the three assessment intervals. Linear mixed modeling was selected to 1) include cases with missing data and 2) examine the contribution of random effects of individual clinicians, above and beyond fixed effects, on outcomes. For each model, the following continuous and categorical fixed effects were considered: age, years of experience, agency and time. Random intercept, indicating individual clinician’s deviation from the average outcome value, was the first random effect used in models. Likelihood ratio tests were used to examine the benefit of adding random slopes, indicating deviations from the effect of time. Intraclass correlation coefficients (ICCs) were estimated from the random-intercept models, quantifying the proportion of variability in the outcome explained by underlying clinician characteristics not captured by the fixed effects. Nine models were estimated, one for each outcome of interest including overall attitude score, three attitude subscale scores, overall practice score, and four individual practice items. Agency was included as a fixed effect to explore its role as an independent variable *and* adjust for any potential effect of clustering

because the number of agencies was too small to include agency as a random effect. Initial models examining differences between all agency comparisons revealed numerous agency differences compared to Agency 4, which had the highest clinician attitude and practice scores. Thus, Agency 4 was thus used as the referent agency for final models presented in results. As not all pairwise comparisons of agency differences could be thoroughly examined using this method, pairwise comparisons with Agency 1 as an opposite referent group were requested in final model statements and significant differences are reported. Correlation in the residuals was modeled using different covariance patterns (e.g., compound symmetry, AR(1) and unstructured). Nested models were compared using the likelihood ratio test, and the best model was chosen that incorporated a random intercept and AR(1) correlation structure. Significance was established at alpha of 0.025 for the effect of time because we were interested in two pair-wise comparisons of: 1) pre-training (Time 1) to post-training (Time 2) and 2) post-training to follow-up (Time 3). Alpha of 0.05 was used for other effects. Results were summarized as parameter estimates with standard errors in the tables, and least square means in the figures. Confidence intervals (95%) were examined and interval ranges are reported in results. SPSS Version 26 was used to implement the analyses.

Clinician experiences captured by quantitative ratings of pre-populated experiences with SA were organized by examining the distribution of clinicians who reported each experience with “none,” “some,” “most” or “all” cases.

Qualitative Analyses. Qualitative feedback to open-ended questions was the primary source for understanding clinician experiences and recommendations; this feedback on the survey was compiled by question and independently analyzed by two coders (first and second author). Modeled after a grounded theory approach (Charmaz, 2014), coders first reviewed the data and applied and discussed open codes. Next, focus codes were iteratively applied based on consensus conversations about the data and memos were made throughout the coding process to inform the identification of themes. Themes were then developed based on focus code content and the saturation of each theme in the sample was quantified by counting the number of clinicians who commented on that theme.

Mixed Methods. Although SA training and implementation supports were primarily evaluated based on quantitative survey data, these data were connected to qualitative clinician feedback at follow-up which expanded on quantitative findings. Consistent with recommendations to intentionally integrate and report on mixing of quantitative and qualitative data in mixed methods studies, our approach can be described as “QUAN→qual” for the function of *expansion* by the process of *connecting* (Palinkas et al., 2011).

Phase III Results

The ICCs were large for attitudes (range = .44 to .59) and practices (range = .28 to .40), indicating substantial variability in individual responses on all outcomes. Consistent with the high ICC values, the random intercept was highly significant ($p < .001$) for all models, indicating substantial individual deviations from the sample average. Final models did not include random slope of time as likelihood ratio tests were nonsignificant when slope was added.

Attitudes—Self-reported attitudes toward standardized assessment increased between pre-training and post-training for overall attitude scores ($\beta = 0.12$, $t = 2.81$, $p = .006$), clinical utility ($\beta = 0.19$, $t = 3.87$, $p < .001$) and treatment planning ($\beta = 0.19$, $t = 3.42$, $p < .001$) but returned to pre-training levels at the time of follow-up. Attitudes about SA practicality did not change over time. Pairwise comparisons also showed that clinicians in Agency 4 had significantly better attitudes toward SA with respect to overall attitude scores, clinical utility and treatment planning attitude subscales as compared to Agency 1 and 2, respectively (see Table 4). Agency 4 clinicians had better attitudes about SA practicality than Agency 2. The few clinicians in the 51 and older age group reported highest attitude scores but clinician age was not normally distributed in this sample so pairwise differences for age should be interpreted with caution. See Table 4 for attitude model estimates and Figure 2 for estimated marginal (or least square) means at all three time points examined. Confidence intervals (95%) ranged from 0.16 to 0.22 on either side of marginal means at all time points.

Qualitative feedback from clinicians provide additional insight about quantitative reports of attitudes toward SA, particularly with respect to mixed opinions about use of SA for clinical utility and treatment planning. For example, despite the initial increase of clinical utility ratings that was not maintained at the end of the school year, some clinicians felt the measures were clinically useful and, in some cases, improved parent engagement and communication. Other clinicians expressed concerns about the clinical relevance of the CRAFFT for students without substance use risk. Some clinicians also felt the PSC had limited utility after the initial assessment; they perceived a global measure as clinically useful to identify diagnostic considerations and general concerns at the beginning of treatment, but believed that a targeted measure would be more useful later to assess progress on specific treatment goals. These concerns could explain why attitudes about SA did not stay high through the follow-up interval.

Qualitative comments also offer potential explanations about perceived practicality. First, some clinicians wanted a larger time frame between baseline and follow-up data collection in order for the Network to detect aggregate student outcomes. This concern about practicality of the follow-up time frame indicates that they may have felt pressure that their students didn't have time to "improve" before the follow-up assessment. It also indicates that clinicians might have viewed the primary purpose of the pilot mostly as a way for the Network to track student treatment outcomes. Perhaps if clinicians felt the pilot was more focused on the clinician-level goals (i.e., to improve care quality through regular SA data collection and feedback with families), clinicians would have viewed SA as more practical. Practicality of SA might also be particularly difficult to affect in this service context, where some clinicians reported difficulty with parent involvement to obtain parent-reported data.

Clinician-Reported Practices—Descriptive statistics about clinician scores on practice items at baseline are shown in Table 3. Clinicians were more likely to report administering SA than providing feedback or altering the treatment plan based on SA. 61% of clinicians who reported taking on clients in the past month reported administering SA during at least some of those intakes. However, although 52% reported administering SA to at least some of their cases in the past week, only 40% reportedly gave feedback about SA and 21% changed the treatment plan based on the SA in that same time frame. There were no significant

effects of time on most practice scores. One exception was the significant improvement in clinicians' self-reported increase in the percentage of new cases with whom they administered SA between pre-training (Time 1) and follow-up (Time 3, $\beta = 0.67$, $t = 3.83$, $p < .001$, see Figure 3). At baseline, 31% of clinicians who took on cases in the past month reported using a SA for "most" of their clients compared to 54% at follow-up. However, we recommend interpreting this finding with some caution as more clinicians reported not having new intakes at follow-up than at pre-training (which makes sense given full caseloads at the end of the school year) and the 95% CIs are 0.42 to 0.46 on either side of the estimated means which are the widest CIs observed among all attitude and practice outcomes. However, the increase is substantial enough to indicate that clinicians' SA data collection for new cases did improve to some extent. There were also some pairwise differences among agencies for various SA practices (see Table 5). In general, clinicians from Agency 4 reported significantly higher assessment practices overall and with new cases as compared to clinicians from all other agencies. Agency 4 clinicians also reported significantly higher practices than Agency 2 clinicians on percent of total cases they collected SA from in the past week ($\beta = -0.51$, $t = -1.77$, $p = .007$), percent of cases they gave feedback about SA in the last week ($\beta = -0.66$, $t = -4.49$, $p < .001$), and percent of cases for which they changed the treatment plan based on SA in the past week ($\beta = -0.23$, $t = -2.09$, $p = .039$). Additional pairwise comparisons revealed that Agency 1 clinicians reported significantly higher practices than Agency 2 clinicians overall (CAPE Total, mean difference = .27, $p < .05$), on the percentage of new cases they collected SA from (CAPE item 1, mean difference = 0.55, $p < .05$), and on the percentage of cases for which they changed the treatment based on SA in the past week (CAPE item 6, mean difference = .25, $p < .05$). As was noted in the clinician-reported attitudes results, the few clinicians in the 51 and older age group reported highest practice scores but clinician age was not normally distributed in this sample so pairwise differences should be interpreted with caution; age was also entered as a continuous variable to each model to further understand its role as a fixed effect and was not significant for any outcomes in that format.

Clinician Data Submission—Objective data submission results indicate that 71.4% of expected SA data were submitted. This was calculated based on comparing the number of cases with data submitted ($N=277$) compared to the number of cases with data projected (i.e., 4 cases per 97 schools; $N=388$). The mean data completion rate by agency was 73.4% (Range = 55.6% for Agency 2 to 83.8% for Agency 4).

Relation Between Attitudes and Practices—Post-hoc analyses were also conducted to assess attitudes as a possible fixed effect predicting practices. However, although most bivariate correlations between attitudes (total scores and subscales) and practices (overall and individual items) were consistently significant at time 3 (range of $r = .20$ to $.40$; $p < .05$) attitude total scores and subscales were never significant when added to the LME models displayed in Table 5. Patterns of significance didn't change with the addition of attitudes variables, except that the intercept of CAPE item 1 (i.e., percentage of new cases administered a SA) was no longer significant ($\beta = 1.64$, $t = 1.41$, $p = 0.16$). This could mean that the individual clinician variability is somewhat due to clinician attitudes for this item only, but as a fixed effect, attitudes are not significant predictors of this item in this sample.

Of note, some clinicians reported in qualitative feedback that measures were clinically useful but they still wanted more flexibility to select different or more specific SA measures for their clients.

Implementation Experiences and Recommendations—Quantitatively-ranked experiences are displayed in Figure 4. Despite initial concerns raised by clinicians that collecting data might have a negative impact on rapport with students or caregivers, these experiences were reported relatively infrequently. In fact, data collection having a positive impact on rapport with the caregiver was the most frequent experience reported after implementation.

Qualitative clinician feedback about positive experiences with the SA implementation was organized into five themes (see Table 6). Clinician feedback about barriers to implementation and clinician recommendations for future implementation support and sustainment were organized into six primary themes (see Table 7). In addition to themes in Table 7, clinicians also commented on data collection being time consuming (N=4), data collection negatively affecting rapport with parents (N=3), technical difficulties related to data entry (N=3), and that the item wording was difficult for parents or students to understand (N=2). Also, the ESMH Network provided reports to agency administrators with the aggregate student data submitted, and some clinicians wrote that these reports were not consistently shared with clinicians. For example, three respondents reported that the ability to see their individual students' SA results would be rewarding for clinicians.

Discussion

The use of evidence-based, standardized assessments (SA) at the beginning of and throughout treatment is imperative to measurement-based care and an evidence-based practice orientation to mental health service delivery. The implementation of SA is particularly critical in school settings, where a large portion of children's mental health treatment is provided. However, school settings pose unique implementation considerations for any evidence-based innovation (Domitrovich et al., 2008; Owens et al., 2014). This study demonstrates a phased approach to piloting SA in a network of four behavioral health agencies delivering treatment in schools, with the Exploration, Preparation, and Implementation Phases informing the Sustainment Phase (Aarons et al., 2011).

The Exploration phase of this initiative started with the Network's interest in communicating the value of its services and agency commitment to evidence-based care. Outer and inner context factors contributed to their forward momentum toward Preparation phase and clarification of multilevel (i.e., agency and clinician) goals. The Preparation phase involved interviews with agency administrators to consider factors that may impact implementation outcomes as well as inform multidisciplinary planning team. Throughout the Preparation phase, the organizing role of the ESMH Network was a critical outer context factor for SA adoption and preparation. The inter-agency ties among partnering agencies were long-standing and powerful to foster experimentation, collaborative learning and adoption of best practices. These agencies had shared missions, values, outcome goals, patient populations,

reimbursement structures for public insurance payors, and service contexts in the same school system.

The Implementation phase primarily involved initial training and basic implementation support via virtual learning, facilitation, and audit and feedback, as well as a number of other implementation strategies informed by the literature and the project planning team. Quantitative results revealed that attitudes increased initially after training but gains were not maintained. This was the case for overall attitudes toward SA, clinical utility and treatment planning, but not practicality. Because the coefficients are unstandardized, they are in the metric of the outcome and function as a parameter to interpret the size of the effect. That is, the magnitude of the change for the ASA total was 0.11 between time 1 and 2 on a five-point Likert scale, which is fairly small and could reflect a possible ceiling effect. This quadratic pattern of initial improvement in implementation outcomes not sustained by year-end has been found in other studies (Aarons & Palinkas, 2007; Nakamura, Higa-McMillan, & Chorpita, 2012).

In terms of self-reported practices, administering SA to new clients increased significantly between pre-training and year-end follow-up. Changing the treatment plan based on SA approached significant increases over time, but overall SA practices did not appear to improve throughout the implementation support period. Importantly, the fixed effects tested (i.e., agency, age, clinicians' years of experience and time) did not sufficiently describe the variation in the outcome variables (i.e., attitudes and practices) above and beyond individual response variations. These data consistently underscore the salience of individual clinician differences with respect to SA attitudes and practice which we were unable to explain by age, years of experience, or agency affiliation.

Although the clinician self-reported practice data indicated that a sizeable proportion of clinicians did not collect SA from new or active cases as frequently as predicted from their entire caseload, the objective submission data on the four expected cases paint a more positive picture of implementation. Although our implementation efforts were intended to influence overall practice change, clinicians were only expected to submit SA data from four cases. Thus, taking practice change results from the CAPE and data submission together, we can presume that many clinicians chose not to use SA with any cases other than the four required.

One possible explanation of limited effects among quantitative measures of overall attitudes and caseload-wide practice change is that the participatory process with clinicians should have been more intentional or explicitly balanced with system-level goals to demonstrate feasibility of data collection for future outcomes reporting. When approaching an implementation project with system-level goals for monitoring aggregate outcomes of patients served and clinician-level goals for having data-driven signals of patient progress, the measures and methods of data collection and feedback must reflect both levels' needs (Connors et al., 2020). This has been referred to as the "golden thread" of data-informed decision making (Douglas, Button, & Casey, 2016). Another possible explanation of the limited effects among quantitative measures is that the implementation supports should have been more intensive in frequency or duration. The implementation supports in this study

were less intensive than ongoing consultation or coaching referenced as the “gold standard” (Herschell et al., 2010). However, this degree of implementation support is difficult to achieve in practicality due to resources and costs involved.

Indeed, more implementation supports may not be the best answer. Tailored implementation supports based on determinants of practice are emerging as an optimal approach to improve implementation outcomes. In fact, given the results of our study that unmeasured clinician variations had a substantial effect on implementation outcomes, detailed information about determinants at the individual clinician level could be very useful to inform tailored implementation strategies. There is growing evidence that using multilevel implementation strategies tailored to context-specific barriers to change is optimal to improve implementation outcomes (Powell et al., 2017), so understanding specific barriers at the clinician, organizational, and service delivery context level is imperative for future implementation practice and research in school behavioral health. Finally, although Agency 4 provided more emphasis on SA use prior to implementation than other agencies, the four agencies were quite similar in their SA operations queried so additional inquiry into inner setting factors will be important for future research in multi-site implementation studies of SA.

Phase IV. Sustainment

Sustainment of this implementation effort was beyond the time frame of this study but may be informed by initial implementation results. First, training was provided at the beginning of the school year and data collection began in the middle of the school year. However, in school mental health, many students are enrolled at the beginning of the school year which is a more natural “baseline” for SA collection. Thus, future-year implementation included training before the school year started, so that the first SA data collection interval could begin in the fall. Finally, there are ongoing conversations about the eligibility criteria (e.g., only new intakes) and timeline (e.g., aligning with the Medicaid reimbursement interval) of data collection as well as strategies for scaling this initiative up to a larger proportion of students served.

Limitations and Future Directions—There are some primary limitations to be aware of when considering the results of this study. First, in terms of the implementation methods, the dosage (i.e., frequency and intensity) of implementation supports was relatively limited due to large number of participants and resource constraints. Related, the “bundled” implementation support approach limits ability to detect effects of each strategy. Parsing the effects of distinct implementation strategies is an important future direction in implementation science (Lewis et al., 2017). Also, the mandatory nature of the initiative may have negatively influenced clinician attitudes toward the implementation, despite early efforts to engage in participatory decision-making with agency leadership.

Also, the substantial between-subject variability estimates in our models after predictors were added suggest there may be other predictors of clinicians’ attitude and practice scores which were not captured by our study. Possible unmeasured variables could be training background emphasizing assessment as a part of evidence-based practice, access to

supervision or consultation to support use of standardized assessments, and/or school context factors making the logistics of collecting and using data easier or more difficult. Of note, only about one third of clinicians who completed the exit survey provided codable qualitative results, which we hypothesize was related to perceived effort and time to write in detailed feedback. It's possible the qualitative data underrepresent the full breadth of experiences of participating clinicians, and that with more qualitative data we could begin to hypothesize about additional unmeasured variables. However, future research should continue to explore factors related to SA implementation in school mental health using additional mixed methods designs. Nonetheless, we expect this model and its results to be a useful approximation to the true process of learning about and implementing SA in schools with the set of implementation supports provided in this study. Standard cautions related to self-reported attitudes and behavior also apply.

Implications—Our findings underscore several considerations for multi-agency behavioral health networks intending to embark on SA implementation. First, although clinicians may come to view SA more positively for their clinical utility and value in treatment planning following training and implementation support, SA may still not be perceived as practical to implement. Practical barriers to SA implementation have been widely documented (Lewis et al., 2019), which has implications not only for improving the practicality of the measures themselves but also how clinicians can reasonably collect, score, submit, view and use the data in their everyday workflow. Participatory action research with practicing clinicians may be one approach to identifying solutions and strategies to make SA more practical. Also, peer support networks have been one strategy suggested for private practice clinicians and would likely be useful for school behavioral health clinicians who also do not often have clinical colleagues on site at the school (Jensen-Doss et al., 2016; Koerner, & Castonguay, 2015).

Despite clinicians' improved attitudes toward SA for the possibility of treatment planning, and data collection from new cases increased, clinicians didn't seem to use SA data for treatment planning. Feedback to the client and use of measures to guide treatment is a hallmark of measurement-based care and clinicians likely need additional supports to integrate SA into routine practice (Bickman et al., 2016). This is consistent with previous research showing that with implementation supports, SA data collection improves, but using those data to give feedback and change the treatment plan occurs less often (Lyon et al., 2015).

Phased implementation, including adequate preparation, was critical to ensure the implementation approach was appropriately tailored to outer and inner context factors. Often related to changes in funding or external demands, health system leaders may not have the benefit of robust exploratory and preparatory processes. However, omission of these phases may result in significant investment with little gain or durability. Also, requiring clinicians to only collect data from four cases per school allowed for a system-wide pilot to test feasibility and durability of system changes before going to scale. Ideally, future projects may choose to test SA on an even smaller scale before testing a pilot implementation throughout the whole system; this rapid-cycle approach is often recommended in the quality improvement literature (American Diabetes Association, 2004; Taylor et al., 2014). An

implementation planning guide for SA within the overall practice of MBC, such as the one developed by Dollar and colleagues, also provides a helpful blueprint for agencies looking to take on this work (Dollar et al., 2019).

Conclusion

There is mounting emphasis on the importance of implementing SA, particularly as a component of measurement-based care, to improve quality of care. However, studies demonstrating feasibility and implementation of measures across multiple usual care sites within one larger system are scant. Moreover, multi-site demonstrations in usual care systems are typically of adult patient samples (Resnick and Hoff, 2019; Trivedi et al., 2006; Unützer et al., 2012) and there are very few studies with children (Bickman et al., 2011; Kotte et al., 2016). This is the first multi-agency demonstration piloting a uniform approach to SA for school-based mental health treatment in a large district. Given the large proportion of usual care mental health services provided to children in the education sector, this study provides evidence that SA adoption, pilot implementation and capacity building is possible for school districts working in partnership with several care agencies. Specifically, results indicated that the multi-stage implementation approach with a set of strategies selected for and tailored to the project yielded a significant increase in collection of SA from new cases over one school year and 74% of projected case data were submitted by clinicians. Two critical areas for future research are to explore additional clinician characteristics and inner context factors that explain currently unmeasured variance in clinician SA practices.

Acknowledgements:

This study was funded by Behavioral Health System Baltimore (Pis: Connors and Hoover). We are deeply appreciative of participating behavioral health agencies and their clinicians' time and collaboration, as well as the contributions of Dr. Julia Goolsby and Ms. Sabrina Ereshefsky for their support reviewing public domain measures during the Preparation Phase of this project. We also thank Dr. Jose Arbelaez for his support of this project on the multidisciplinary planning team and review of this manuscript. Finally, we are very grateful for Dr. Veronika Shabanova's guidance and feedback on our statistical analyses.

References

- Aarons GA, Ehrhart MG, Farahnak LR, & 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*, 255–74. doi: 10.1146/annurev-publhealth-032013-182447
- Aarons GA, Hurlburt M, & Horwitz SM (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*(1), 4–23. doi: 10.1007/s10488-010-0327-7
- Aarons GA, & Palinkas LA (2007). Implementation of evidence-based practice in child welfare: Service provider perspectives. *Administration and Policy in Mental Health and Mental Health Services Research, 34*(4), 411–419. doi: 10.1007/s10488-007-0121-3
- American Diabetes Association. (2004). The breakthrough series: IHEs collaborative model for achieving breakthrough improvement. *Diabetes Spectrum, 17*(2), 97–101. doi: 10.2337/diaspect.17.2.97.
- Arora PG, Connors EH, George MW, Lyon AR, Wolk CB, & Weist MD (2016). Advancing evidence-based assessment in school mental health: Key priorities for an applied research agenda. *Clinical Child and Family Psychology Review, 19*(4), 271–284. doi:10.1007/s10567-016-0217-y
- Becker-Haimes EM, Tabachnick AR, Last BS, Stewart RE, Hasan-Granier A, & Beidas RS (2020). Evidence base update for brief, free, and accessible youth mental health measures. *Journal of*

- Clinical Child & Adolescent Psychology, 49(1), 1–17. doi: 10.1080/15374416.2019.1689824 [PubMed: 31825683]
- Beidas RS, & Kendall PC (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, 17(1), 1–30. doi: 10.1111/j.1468-2850.2009.01187.x [PubMed: 20877441]
- Bergman H, Kornør H, Nikolakopoulou A, Hanssen-Bauer K, Soares-Weiser K, Tollefsen TK, & Bjørndal A (2018). Client feedback in psychological therapy for children and adolescents with mental health problems. *Cochrane Database of Systematic Reviews*, (8), Art. No.: CD011729.
- Bickman L, Douglas SR, De Andrade Ana Regina Vides, Tomlinson M, Gleacher A, Olin S, & Hoagwood K. (2016). Implementing a measurement feedback system: A tale of two sites. *Administration and Policy in Mental Health and Mental Health Services Research*, 43(3), 410–425. doi: 10.1007/s10488-015-0647-8 [PubMed: 25876736]
- Bickman L, Kelley SD, Breda C, de Andrade AR, & Riemer M (2011). Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services*, 62(12), 1423–1429. doi: 10.1176/appi.ps.002052011 [PubMed: 22193788]
- Bickman L, Lyon AR, & Wolpert M (2016). Achieving precision mental health through effective assessment, monitoring, and feedback processes. *Administration and Policy in Mental Health and Mental Health*, 73(3), 271–276. doi: 10.1007/s10488-016-0718-5
- Bickman L, Rosof-Williams T, Salzer MS, Summerfelt WT, Noser K, Wilson ST, & Karver MS (2000). What information do clinicians value for monitoring adolescent client progress and outcomes? *Professional Psychology: Research and Practice*, 31(1), 70. doi: 10.1037/0735-7028.31.1.70
- Bohnenkamp JH, Glascoe T, Gracey KA, Epstein RA, & Benningfield MM (2015). Implementing clinical outcomes assessment in everyday school mental health practice. *Child and Adolescent Psychiatric Climes*, 24(2), 399–413. doi: 10.1016/j.chc.2014.11.006
- Borntrager C, & Lyon AR (2015). Monitoring client progress and feedback in school-based mental health. *Cognitive and Behavioral Practice*, 22, 74–86. doi: 10.1016/j.cbpra.2014.03.007 [PubMed: 26257508]
- Burns BJ, Costello EJ, Angold A, Tweed D, Stangl D, Farmer EM, & Erkanli A (1995). Children’s mental health service use across service sectors. *Health Affairs*, 14(3), 147–159. doi: 10.1377/hlthaff.14.3.147 [PubMed: 7498888]
- Charmaz K (2014). *Constructing grounded theory*, 2nd Edition. Sage Publications.
- Connors EH, Arora P, Curtis L, & Stephan SH (2015). Evidence-based assessment in school mental health. *Cognitive and Behavioral Practice*, 22(1), 60–73. doi: 10.1016/j.cbpra.2014.03.008
- Connors EH, Douglas S, Jensen-Doss A, Landes SJ, Lewis CC, McLeod BD, ... & Lyon AR. (2020). What gets measured gets done: How mental health agencies can leverage measurement-based care for better patient care, clinician supports, and organizational goals. *Administration and Policy in Mental Health and Mental Health Services Research*, 1–16. doi: 10.1007/s10488-020-01063-w [PubMed: 31549276]
- Connors EH, Schiffman J, Stein K, LeDoux S, Landsverk J, & Hoover S (2019). Factors Associated with Community-Partnered School Behavioral Health Clinicians’ Adoption and Implementation of Evidence-Based Practices. *Administration and Policy in Mental Health and Mental Health Services Research*, 46(1), 91–104. doi: 10.1007/s10488-018-0897-3
- Connors E, Wigand K, Moffa K, Hoover S & Lever N (2019). Student Information Systems. National Center for School Mental Health, Baltimore, MD. <http://bit.ly/SISbrief>
- Dollar KM, Kirchner JE, DePhilippis D, Ritchie MJ, McGee-Vincent P, Burden JL, & Resnick SG (2019). Steps for implementing measurement-based care: Implementation planning guide development and use in quality improvement. *Psychological Services*. Advance online publication, doi: 10.1037/ser0000368
- Domitrovich CE, Bradshaw CP, Poduska JM, Hoagwood K, Buckley JA, Olin S, ... & Ialongo NS (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, 1(3), 6–28. doi: 10.1080/1754730X.2008.9715730 [PubMed: 27182282]

- Douglas S, Button S, & Casey SE (2016). Implementing for sustainability: Promoting use of a measurement feedback system for innovation and quality improvement. *Administration and Policy in Mental Health and Mental Health Services Research*, 43(3), 286–291. doi: 10.1007/s10488-014-0607-8 [PubMed: 25344848]
- Forman SG, Olin SS, Hoagwood KE, Crowe M, & Saka N (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health*, 1(1), 26. doi: 10.1007/s12310-008-9002-5
- Fortney JC, Uniitzer J, Wrenn G, Pyne JM, Smith GR, Schoenbaum M, & Harbin HT (2017). A tipping point for measurement-based care. *Psychiatric Services*, 68, 179–188. doi: 10.1176/appi.ps201500439 [PubMed: 27582237]
- Gilbody SM, House AO, & Sheldon TA (2001). Routinely administered questionnaires for depression and anxiety: Systematic review British Medical Association. Retrieved from <http://survey.hshsl.umaryland.edu/?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsjsr&AN=edsjsr.25466218&site=eds-live>
- Green JG, McLaughlin KA, Alegría M, Costello EJ, Gruber MJ, Hoagwood K, ... & Kessler RC (2013). School mental health resources and adolescent mental health service use. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(5), 501–510. doi: 10.1016/j.jaac.2013.03.002
- Herschell AD, Kolko DJ, Baumann BL, & Davis AC (2010). The role of therapist training in the implementation of psychosocial treatments: A review and critique with recommendations. *Clinical Psychology Review*, 30(4), 448–466. doi: 10.1016/j.cpr.2010.02.005 [PubMed: 20304542]
- Jellinek MS, Murphy JM, Robinson J, Feins A, Lamb S, & Fenton T (1988). Pediatric symptom checklist: Screening school-age children for psychosocial dysfunction. *The Journal of Pediatrics*, 112(2), 201–209. doi: 10.1016/s0022-3476(88)80056-8 [PubMed: 3339501]
- Jensen-Doss A, Haimes EMB, Smith AM, Lyon AR, Lewis CC, Stanick CF, & Hawley KM (2016). Monitoring treatment progress and providing feedback is viewed favorably but rarely used in practice. *Administration and Policy in Mental Health and Mental Health Services Research*, 45(1), 48–61. doi:10.1007/s10488-016-0763-0
- Jensen-Doss A, & Hawley KM (2010). Understanding barriers to evidence-based assessment: Clinician attitudes toward standardized assessment tools. *Journal of Clinical Child and Adolescent Psychology*, 39(6), 885–896. doi:10.1080/15374416.2010.517169 [PubMed: 21058134]
- Kendrick T, El-Gohary M, Stuart B, Gilbody S, Churchill R, Aiken L, ... & Moore M (2016). Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. *Cochrane Database of Systematic Reviews*, (7).
- Knight J, Shrier L, Bravender T, Farrell M, Vander Bilt J, Shaffer H A new brief screen for adolescent substance abuse. *Archives of Pediatrics & Adolescent Medicine*, 1999, 153(6), 591–6. doi: 10.1001/archpedi.153.6.591 [PubMed: 10357299]
- Koerner K, & Castonguay LG (2015). Practice-oriented research: What it takes to do collaborative research in private practice. *Psychotherapy Research*, 25(1), 67–83. doi: 10.1080/10503307 [PubMed: 25087919]
- Kotte A, Hill KA, Mah AC, Korathu-Larson PA, Au JR, Izmirian S, ... & Higa-McMillan CK (2016). Facilitators and barriers of implementing a measurement feedback system in public youth mental health. *Administration and Policy in Mental Health and Mental Health Services Research*, 43(6), 861–878. doi: 10.1007/s10488-016-0729-2
- Krageloh C, Czuba K, Billington R, Kersten P, & Siegert R (2015). Using feedback from patient-reported outcome measures in mental health services: A scoping study and typology. *Psychiatric Services*, 66(3), 563–570. doi: 10.1176/appi.ps.201400141
- Lambert MJ, Whipple JL, Hawkins EJ, Vermeersch DA, Nielsen SL, & Smart DW (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice*, 10(3), 288–301. doi: 10.1093/clipsy/bpg025
- Lambert MJ, Whipple JL, & Kleinstaub M (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy*, 55(4), 520–537. doi 10.1037/pst0000167 [PubMed: 30335463]

- Langley AK, Nadeem E, Kataoka SH, Stein BD, & Jaycox LH (2010). Evidence-based mental health programs in schools: Barriers and facilitators of successful implementation. *School Mental Health*, 2(3), 105–113. doi: 10.1007/s12310-010-9038-1 [PubMed: 20694034]
- Lewis CC, Boyd M, Puspitasari A, Navarro E, Howard J, Kassab H, ... & Simon G (2019). Implementing Measurement-Based Care in Behavioral Health: A Review. *JAMA Psychiatry*, 76(3), 324–335. doi: 10.1001/jamapsychiatry.2018.3329
- Lewis CC, Stanick C, Lyon A, Darnell D, Locke J, Puspitasari A., ... Landes S,J. (2017, 9 7-9). Proceedings of the fourth biennial conference of the society for implementation research collaboration (SIRC) 2017: Implementation mechanisms: What makes implementation work and why? Paper presented at the Seattle, WA. doi: 10.1186/s13012-018-0714-0
- Lever N, Stephan S, Castle M, Bernstein L, Connors E, Sharma R, et al. (2015). Community-partnered school behavioral health: State of the field in Maryland. Baltimore: Center for School Mental Health.
- Lyon AR, Dorsey S, Pullmann M, Silbaugh-Cowdin J, & Berliner L (2015). Clinician use of standardized assessments following a common elements psychotherapy training and consultation program. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(1), 47–60. doi: 10.1007/s10488-014-0543-7 [PubMed: 24590606]
- Lyon AR, Pullmann MD, Whitaker K, Ludwig K, Wasse JK, & McCauley E (2017). A digital feedback system to support implementation of measurement-based care by school-based mental health clinicians. *Journal of Clinical Child & Adolescent Psychology*, 1–12. doi: 10.1080/15374416.2017.1280808
- Mellard DF, McKnight M, & Woods K (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice*, 24(4), 186–195. doi: 10.1111/j.1540-5826.2009.00292.x
- Moullin JC, Dickson KS, Stadnick NA, Rabin B, & Aarons GA (2019). Systematic review of the exploration, preparation, implementation, sustainment (EPIS) framework. *Implementation Science*, 14(1), 1. doi: 10.1186/s13012-018-0842-6 [PubMed: 30611302]
- Nakamura BJ, Higa-McMillan C, & Chorpita BF (2012). Sustaining Hawaii's evidence-based service system in children's mental health. *Dissemination and Implementation of Evidence-Based Psychological Interventions*, 166–186.
- Owens JS, Lyon AR, Brandt NE, Warner CM, Nadeem E, Spiel C, & Wagner M (2014). Implementation science in school mental health: Key constructs in a developing research agenda. *School Mental Health*, 6(2), 99–111. doi: 10.1007/s12310-013-9115-3 [PubMed: 26413173]
- Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, & Landsverk J (2011). Mixed method designs in implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(1), 44–53. doi: 10.1007/s10488-010-0314-z
- Powell BJ, Beidas RS, Lewis CC, Aarons GA, McMillen JC, Proctor EK, & Mandell DS (2017). Methods to improve the selection and tailoring of implementation strategies. *The Journal of Behavioral Health Services & Research*, 44(2), 177–194. doi: 10.1007/s11414-015-9475-6 [PubMed: 26289563]
- Purbeck CA, Briggs EC, Tunno AM, Richardson LM, Pynoos RS, & Fairbank JA (2019). Trauma-informed measurement-based care for children: Implementation in diverse treatment settings. *Psychological Services*. Advance online publication, doi: 10.1037/ser0000383
- Resnick SG, & Hoff RA (2019). Observations from the national implementation of Measurement Based Care in Mental Health in the Department of Veterans Affairs. *Psychological Services*. Advance online publication, doi: 10.1037/ser0000351
- Sander MA, Everts J, & Johnson J (2011). Using data to inform program design and implementation and make the case for school mental health. *Advances in School Mental Health Promotion*, 4(4), 13–21. doi: 10.1080/1754730X.2011.9715639
- Scott K, & Lewis CC (2015). Using measurement-based care to enhance any treatment. *Cognitive and Behavioral Practice*, 22, 49–59. doi:10.1016/j.cbpra.2014.01.010 [PubMed: 27330267]
- Shimokawa K, Lambert MJ, & Smart DW (2010). Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance

system. *Journal of Consulting and Clinical Psychology*, 78(3), 298. doi: 10.1037/a0019247 [PubMed: 20515206]

- Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, & Reed JE (2014). Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ Quality & Safety*, 23(4), 290–298. doi: 10.1136/bmjqs-2013-001862
- Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, ... & Shores-Wilson K (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *American Journal of Psychiatry*, 163(1), 28–40. doi: 10.1176/appi.ajp.163.1.28
- Unützer J, Chan YF, Hafer E, Knaster J, Shields A, Powers D, & Veith RC (2012). Quality improvement with pay-for-performance incentives in integrated behavioral health care. *American Journal of Public Health*, 102(6), e41–e45. doi: 10.1176/appi.ajp.163.1.28 [PubMed: 22515849]
- Wright AJ with Center for Applied Research Solutions (2018, March). Using data to improve student mental health. Now Is The Time Technical Assistance Center. Washington, DC: Substance Abuse and Mental Health Services Administration.
- Valenstein M, Adler DA, Berlant J, Dixon LB, Dulit RA, Goldman B, ... & Sonis WA. (2009). Implementing standardized assessments in clinical care: now's the time. *Psychiatric Services*, 60(10), 1372–1375. doi: 10.1176/ps.2009.60.10.1372

The selected measure should meet the following criteria:

- Reliable*
- Valid**
- Brief (fewer than 50 items)
- Free or very low cost
- Sensitive to change (for progress monitoring)
- Used in schools
- Parent and student self-report available
- Measure internalizing, externalizing and substance use progress and outcomes for students ages 4-18 years old

*Reliability: Internal consistency and test-retest reliability was coded as Adequate, Good, or Excellent based on the criteria set forth by Hunsley & Mash (2008)

**Validity: High, Medium and Low categories were developed based on degree of published concurrent, predictive, convergent and discriminant validity with specific attention to whether the measure is predictive of academic success and/or validated with low-income, urban, and/or minority youth.

Figure 1.
Criteria for Candidate SA Measures

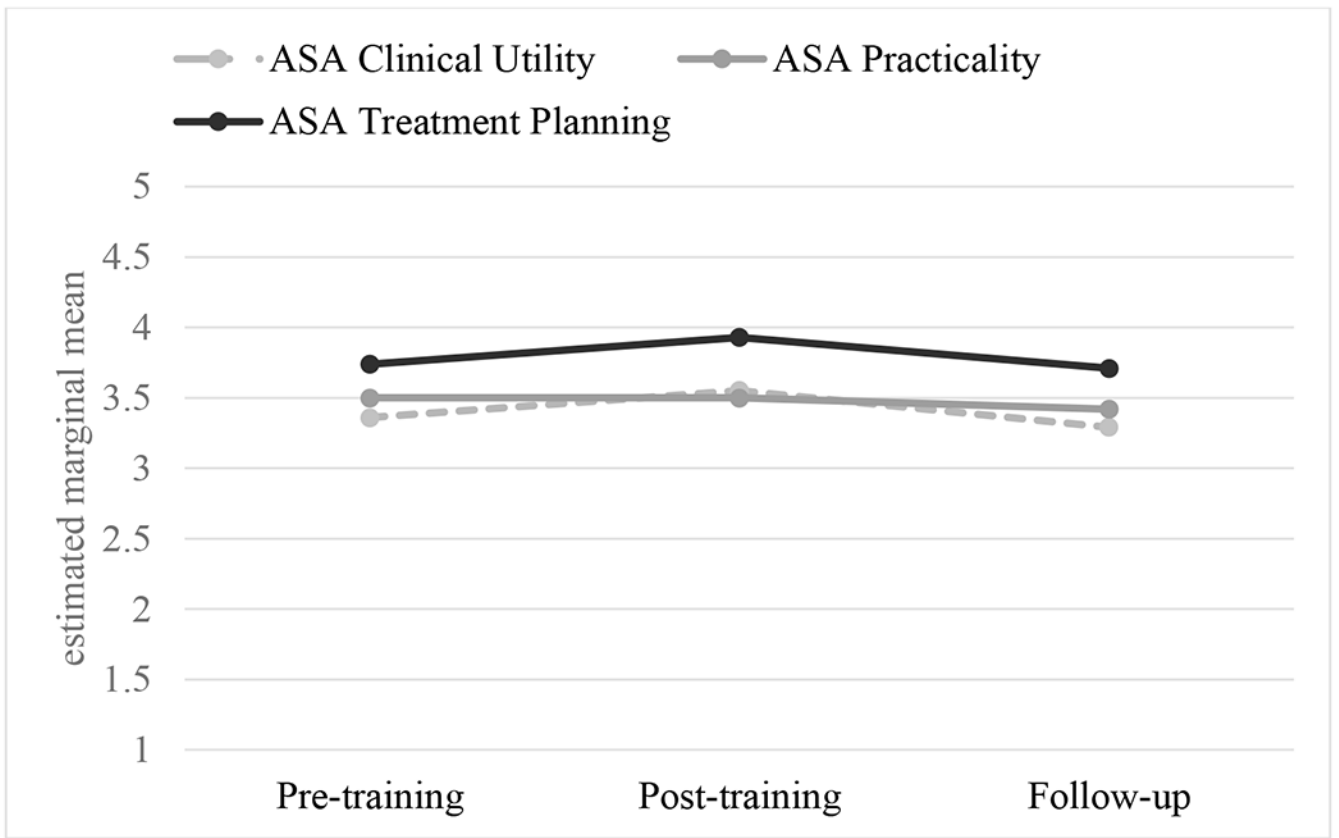


Figure 2.
Clinician Attitudes Toward Standardized Assessment

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

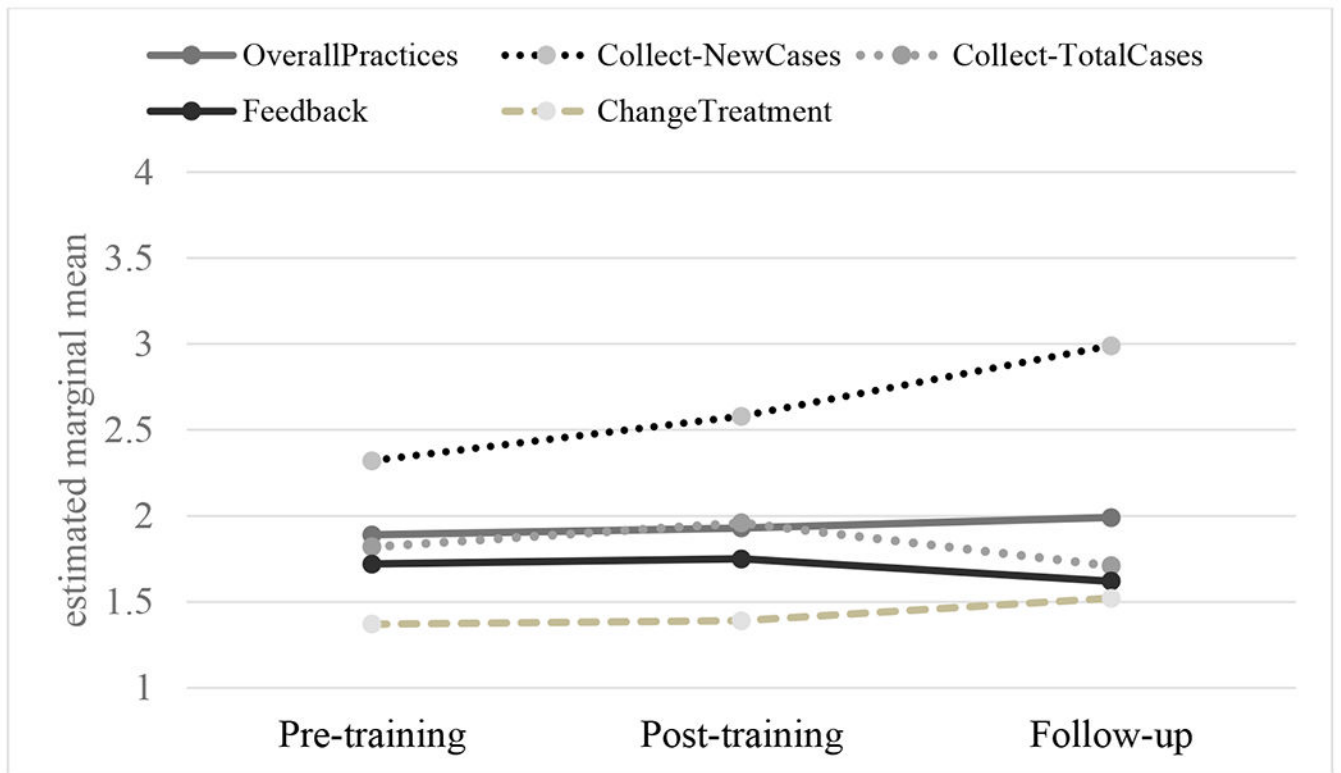


Figure 3.
Clinician Self-reported Standardized Assessment Practices

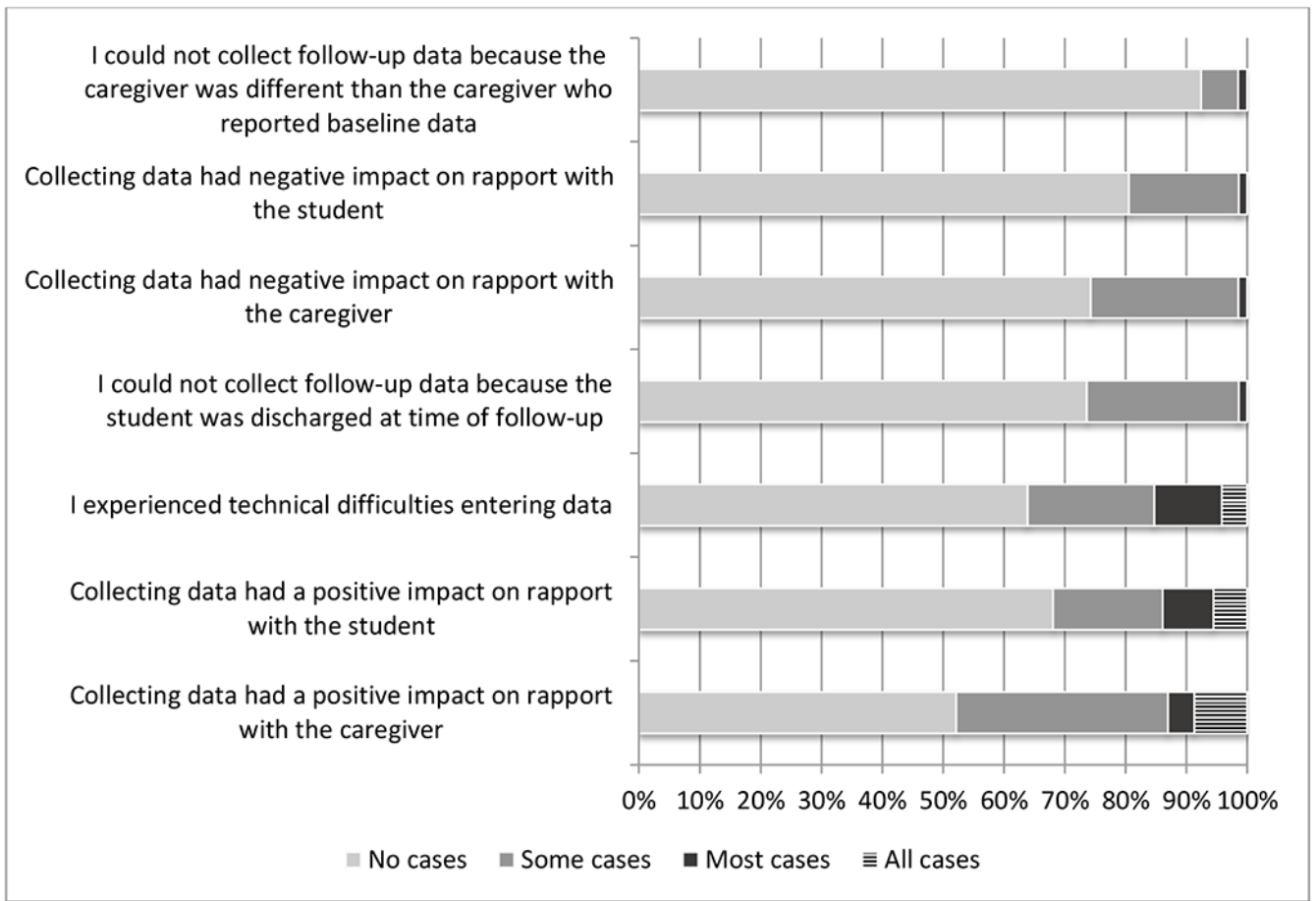


Figure 4.
Clinician Experiences with Data Collection (N=82)

Table 1

Discrete Implementation Strategies Used

Strategy*	EPIS Phase	Description
Develop academic partnerships	Exploration	ESMH Network partnered with colleagues at the University of Maryland for the purposes of training and implementation support
Conduct local consensus discussions	Preparation	Include agency administrators in discussions (interviews) that address whether tracking student outcomes is important and whether a consistent approach to SA across agencies is appropriate
Develop educational materials	Preparation	Develop and format training materials, guidance documents and materials to use with students and families to make it easier for stakeholders to learn about SA and for clinicians to learn how to deliver SA
Use advisory boards and workgroups	Preparation	Create and engage a formal group of multiple kinds of stakeholders (ESMH Network Regional Implementation Meeting and School Mental Health Evidence Based Assessment Workgroup) to provide input and advice on implementation efforts
Mandate change	Preparation, Implementation	Have ESMH Network leadership and agency leaders declare the priority of the SA initiative and their determination to have it implemented
Distribute educational materials	Implementation	Distribute educational materials in-person (at initial training) and electronically. Ensure clinicians have hard copy materials at their school sites as needed.
Audit and provide feedback	Implementation	Collect and summarize SA data collection performance during data collection intervals and give it to clinicians and administrators to monitor, evaluate and modify provider behavior
Make training dynamic	Implementation	Initial training content was designed for mental health and substance use clinicians, tailored to the school mental health context, and included didactic content, behavioral rehearsal, feedback and discussion.
Provide local technical assistance/facilitation**	Implementation	Develop and use a system to deliver responsive, problem-solving oriented technical assistance focused on implementation issues using supportive, local personnel (i.e., clinicians and agency leaders reported implementation progress to university partners, ESMH Network data team and leadership who worked as a team to provide support as needed)
Remind clinicians	Implementation	Develop reminder systems (i.e., emails and some personal communications) to clinicians and agency leaders to help clinicians to recall information and prompt them to use SA

* Strategy terms are from the ERIC project, and descriptions are adjusted for the current project (Powell et al., 2015)

** Provide local technical assistance and facilitation were combined for this project

Table 2

Clinician Participant Characteristics (N=103)

Demographic Characteristics	N (%)
Gender	
Female	96 (93.2%)
Male	7 (6.8%)
Race	
White/Caucasian	71 (68.9%)
Black/African-American	20 (19.4%)
Asian	6 (5.8%)
Mixed Race	5 (4.9%)
American Indian	1 (1.0%)
Ethnicity	
Non-Hispanic/Latino	97 (94.2%)
Hispanic/Latino	6 (5.8%)
Age	
21-30	64 (62.1%)
31-40	27 (26.2%)
41-50	9 (8.7%)
51 and over	3 (2.9%)
Professional Characteristics	
Years of Experience	
2 years or less	17 (16.5%)
3-5 years	44 (42.7%)
6-10 years	28 (27.2%)
11-15 years	10 (9.7%)
16 years or more	4 (3.9%)
Field	
Social Work	59 (57.3%)
Clinical/Counseling	20 (19.4%)
Psychology	16 (15.5%)
Professional Counseling	6 (5.8%)
Marriage and Family Therapy	2 (1.9%)
School Psychology	
Degree	
Associate's degree	1 (1.0%)
Master's degree	100 (97.1%)
Doctoral degree	2(1.9%)

Table 3

Standardized Assessment Practices as Self-Reported on the CAPE at Baseline (N=95)

CAPE item	None N (%)	Some N (%)	Half N (%)	Most N (%)
% of new cases collected SA in last month [*]	33(32.0%)	18(17.5%)	7(6.8%)	26(25.2%)
% of total cases collect SA in last week	46(44.7%)	36(35.0%)	8(7.8%)	5(4.9%)
% of cases gave feedback about SA in last week	57(55.3%)	28(27.2%)	7(6.8%)	3(2.9%)
% of cases altered treatment plan based on SA in the last week ^{**}	74(71.8%)	17(16.5%)	2(1.9%)	1(1.0%)

Note. SA = standardized assessment

^{*} 11 (10.7%) of participants reported they did not take on any new clients in the past month.

^{**} N=94

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Linear Mixed Effects Model Summaries: Attitudes toward Standardized Assessment (ASA)

	ASA Total				ASA Clinical Utility				ASA Practicality				ASA Treatment Planning			
	Coeff.	SE	t	p	Coeff.	SE	t	p	Coeff.	SE	t	p	Coeff.	SE	t	p
Intercept	4.19	0.32	12.99	<.001	4.11	0.33	12.35	<.001	3.99	0.42	9.43	<.001	4.54	0.39	11.59	<.001
Agency																
1 vs 4	-0.34	0.12	-2.78	.006	-0.32	0.13	-2.56	.012	-0.27	0.16	-1.67	.099	-0.43	0.15	-2.91	.004
2 vs 4	-0.32	0.11	-3.05	.003	-0.30	0.11	-2.76	.007	-0.37	0.14	-2.67	.009	-0.31	0.13	-2.38	.019
3 vs 4	0.15	0.13	1.16	.247	0.18	0.13	1.35	.180	0.14	0.17	0.83	.410	0.15	0.15	0.98	.331
Age ⁺																
21-30	-0.48	0.24	-1.99	.049	-0.58	0.25	-2.31	.023	-0.40	0.32	-1.27	.208	-0.54	0.29	-1.85	.068
31-40	-0.59	0.24	-2.41	.018	-0.73	0.25	-2.92	.004	-0.52	0.32	-1.62	.110	-0.58	0.30	-1.95	.054
41-50	-0.43	0.25	-1.67	.087	-0.55	0.26	-2.14	.035	-0.41	0.33	-1.24	.220	-0.41	0.30	-1.36	.178
Experience	-0.04	0.05	-0.79	0.43	-0.04	0.05	-0.83	.407	0.02	0.06	0.25	.800	-0.08	0.06	-1.48	.143
Time																
2 vs 1	0.12	0.04	2.81	.006	0.19	0.05	3.87	<.001	-0.00	0.06	-0.02	.981	0.19	0.06	3.42	.001
3 vs 1	-0.05	0.04	-1.09	.277	-0.07	0.06	-1.22	.223	-0.08	0.07	-1.19	.237	-0.03	0.05	-0.55	.582

⁺Reference category is 51 or more years old

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Linear Mixed Effects Model Summaries: Current Assessment Practices (CAPE)

	CAPE Total				% New Cases Administered SA (CAPE1)				% All Cases Administered SA (CAPE2)				% Cases provided Feedback (CAPE3)				% Cases changed treatment plan (CAPE6)			
	Coeff	SE	t	p	Coeff	SE	t	p	Coeff	SE	t	p	Coeff	SE	t	p	Coeff	SE	t	p
Intercept	2.66	0.42	6.33	<.001	3.51	0.83	4.20	<.001	2.66	0.55	4.82	<.001	2.65	0.44	6.07	<.001	1.85	0.32	5.70	<.001
Agency																				
1 vs 4	-0.37	0.16	-2.26	.026	-0.74	0.32	-2.28	.025	-0.34	0.21	-1.62	.108	-0.45	0.17	-2.71	.008	0.02	0.12	0.16	.870
2 vs 4	-0.64	0.15	-4.38	<.001	-1.29	0.29	-4.44	<.001	-0.51	0.19	-1.77	.007	-0.66	.015	-4.49	<.001	-0.23	0.11	-2.09	.039
3 vs 4	0.38	0.17	2.11	.030	0.85	0.34	2.51	.014	0.39	0.22	1.77	.080	0.29	0.17	1.67	.098	0.16	0.13	1.22	.224
Age*																				
21-30	-0.57	0.31	-1.82	.072	-0.50	0.61	-0.81	.420	-0.68	0.41	-1.64	.106	-0.74	0.32	-2.29	.025	-0.50	0.24	-2.01	.042
31-40	-0.63	0.32	-2.00	.049	-0.95	0.62	-1.52	.132	-0.85	0.42	-2.03	.045	-0.73	0.33	-2.21	.030	-0.48	0.24	-1.99	.050
41-50	-0.80	0.32	-2.46	.016	-0.91	0.64	-1.43	.157	-0.95	0.43	-2.23	.029	-0.90	0.34	-2.67	.009	-0.52	0.25	-2.10	.040
Exper.	0.03	0.06	0.50	.620	0.05	0.12	.040	.693	0.04	0.08	0.45	.656	0.01	0.06	0.10	.925	-0.00	0.05	-0.11	.917
Time																				
2 vs 1	0.04	0.07	0.49	.629	0.26	0.15	1.78	.077	0.15	0.10	1.52	.132	0.03	0.09	0.34	.733	0.02	0.08	0.21	.837
3 vs 1	0.10	0.08	1.24	.219	0.67	0.18	3.83	<.001	-0.11	0.10	-1.10	.279	-0.10	0.08	-1.29	.202	0.15	0.07	1.97	.051

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 6

Positive Feedback about SA Implementation (N=22 clinicians with codable comments)

Theme	N Clinicians	Illustrative Comment
Measures were clinically-useful	9	<i>"[The SAs] provided a different perspective to clinical treatment."</i>
The process was feasible (i.e., follow-up interval was appropriate, materials were accessible)	6	<i>"The materials were easy to access; the goals were reasonable (two test dates four clients)."</i>
The measures were feasible (i.e., to administer and discuss with families)	5	<i>"The PSC was relatively easy and fast to administer and discuss with patients families."</i>
SA improved parent engagement and communication	4	<i>"[The PSC-17] was useful in assessing treatment and parents could see the progress which increased their participation."</i>
Initiative offered an organized, standardized system to monitor outcomes	4	<i>"I appreciate the attempt to consistently monitor student progress across sites. It was also well organized and easy to support."</i>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 7

Barriers and Recommendations for Sustainment (N=69 clinicians with codable comments)

Theme	N Clinicians	Illustrative Comment
Data collection timelines should be different (Baseline should begin sooner/at beginning of school year (N=9) and follow-up window should be longer (N=5) to align with insurance authorization schedules and provide more time in treatment for improved outcomes)	14	<i>"If we are continuing this initiative next year, I would recommend starting at the beginning of the year with only new intakes and doing follow ups at 6 month intervals that coincide with the authorization periods so we are not redoing the same assessment multiple times when it may not be clinically indicated." "I do not believe the pre post-test were spaced far enough apart to measure significant clinical changes."</i>
Parental involvement made data collection difficult (Clinicians requested student report only, teacher report, or to choose cases for data collection to select for the most engaged parents)	7	<i>"It is difficult to get in touch with parents to get the response for the screening. " "Maybe have the option of giving scales to teachers instead of parents, as parents can be hard to reach sometimes."</i>
Allow more clinician choice (Clinicians wanted more flexibility in measures used, follow-up date flexibility, select more or different clients to assess which would optimize clinician buy-in)	7	<i>"Perhaps we can submit data reflected in assessments we are already using in our clinical work with the child. For example, if I have a child for whom I normally collect Vanderbilts from teachers, maybe I can enter in the data from the Vanderbilts into an online system, instead of having to collect yet another survey from teachers parents. "</i>
SAs should be aligned with current requirements (Collecting these SAs felt redundant with the reimbursement-required SA and created additional time burden for data collection)	6	<i>"Maybe having the new questionnaires either take the place of the OMS (I don't know if that's possible for authorization purposes). It is really difficult to keep track of all the deadlines for the different surveys and hunt down parents, especially for those with high caseloads. There is so much going on already on a daily basis at the schools that needs to be prioritized to be there for the clients and school. "</i>
Provide additional implementation supports (Suggestions include providing more handouts or hard copies of measures, additional in-person training, greater clarity of expectations for clinicians and how to handle follow-up data collection if a patient is discharged, leadership support to make data collection more routine)	6	<i>"A spring meeting like we had in the fall so we can all get together to learn and share ideas. " "It will be helpful if supervisors can take the lead in insuring that administering the PSC-17 becomes part of the intake packet, and is automatically completed with each treatment plan review (with results noted in the treatment plan review, and treatment plan amended accordingly). "</i>
Global SAs not clinically relevant (Clinicians commented that the PSC-17 was not helpful after diagnosis, CRAFFT was irrelevant for many clients, PSC age range was too wide, and a preference for individualized progress monitoring targets)	5	<i>"The CRAFFT was irrelevant to a lot of clients, and the PSC covers too wide of an age range. They were also very general scales, which feels useless when you're already on track to client diagnosis. "</i>