The TPOCS-self-reported Therapist Intervention Fidelity for Youth (TPOCS-SeRTIFY): A case study of pragmatic measure development

Emily M Becker-Haimes1,2, Melanie R Klein1, Bryce D McLeod3, Sonja K Schoenwald4,5,6, Shannon Dorsey5, Aaron Hogue6, Perrin B Fugo1, Mary L Phan1,5, Carlin Hoffacker1 and Rinad S Beidas1,7,8,9,10

Abstract

Background: Developing pragmatic assessment tools to measure clinician use of evidence-based practices is critical to advancing implementation of evidence-based practices in mental health. This case study details our community-partnered process of developing the Therapy Process Observation Coding Scale-Self-Reported Therapist Intervention Fidelity for Youth (TPOCS-SeRTIFY), a pragmatic, clinician-report instrument to measure cognitive behavioral therapy (CBT) delivery.

Approach: We describe a five-step community-partnered development process. Initial goals were to create a self-report instrument that paralleled an existing direct observation measure of clinician delivery of CBT use to facilitate later assessment of measure performance. Cognitive interviews with community clinicians (n = 6) and consultation with CBT experts (n = 6) were used to enhance interpretability and usability as part of an iterative refinement process. The instrument was administered to 247 community clinicians along with an established self-reported measure of clinician delivery of CBT and other treatments to assess preliminary psychometric performance. Preliminary psychometrics were promising.

Conclusion: Our community-partnered development process showed promising success and can guide future development of pragmatic implementation measures both to facilitate measurement of ongoing implementation efforts and future research aimed at building learning mental health systems.

Plain language summary

Developing brief, user-friendly, and accurate tools to measure how therapists deliver cognitive behavioral therapy (CBT) in routine practice is important for advancing the reach of CBT into community settings. To date, developing such “pragmatic” measures has been difficult. There is little known about how researchers can best develop these types of

1Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA
2Hall Mercer Community Mental Health, University of Pennsylvania Health System, Philadelphia, PA, USA
3Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA
4Oregon Social Learning Center, Eugene, OR, USA
5Department of Psychology, University of Washington, Seattle, WA, USA
6Partnership to End Addiction, New York, NY, USA
7Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA
8Department of Medical Ethics and Health Policy, University of Pennsylvania, Philadelphia, PA, USA
9Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA
10Penn Implementation Science Center (PISCE@LDI), Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA

Corresponding author:
Emily M Becker-Haimes, Department of Psychiatry, University of Pennsylvania, Perelman School of Medicine, 3535 Market Street, 3rd Floor, Philadelphia, PA 19104, USA.
Email: emily.haimes@pennmedicine.upenn.edu
assessment tools so that they (1) are easy for clinicians in practice to use and (2) provide valid and useful information about implementation outcomes. As a result, there are few well-validated measures in existence that measure therapist use of CBT that are feasible for use in community practice. This paper contributes to the literature by describing our community-partnered process for developing a measure of therapist use of CBT (Therapy Process Observation Coding Scale -Self-Reported Therapist Intervention Fidelity for Youth; TPOCS-SeRTIFY). This descriptive case study outlines the community-partnered approach we took to develop this measure. This case study will contribute to future research by serving as a guide to others aiming to develop pragmatic implementation measures. In addition, the TPOCS-SeRTIFY is a pragmatic measure of clinician use of CBT that holds promise for its use by both researchers and clinicians to measure the success of CBT implementation efforts.

Keywords
Assessment, implementation, pragmatic measurement, cognitive behavioral therapy

Implementation research is hampered by measurement limitations (Lewis et al., 2018). This is particularly so for measuring the delivery of evidence-based treatments by mental health clinicians in routine clinical care settings, which is a key implementation outcome (Bond & Drake, 2019; Proctor et al., 2011). Observational methods are the current gold-standard approach to ascertaining clinician use of evidence-based therapy techniques (Rodriguez-Quintana & Lewis, 2018; Schoenwald et al., 2011), but there are significant feasibility challenges for community settings in using these methods (Beidas et al., 2016). Despite requests for pragmatic measures that are feasible for ongoing use in the community to measure implementation efforts (Glasgow & Riley, 2013; Powell et al., 2017), there is little guidance for how to develop them. Retrofitting existing measures for use in clinician-report format may not be sufficient to create a pragmatic and clinically credible tool for everyday practice. This short report offers our collaborative experience designing an instrument with community partners as a case study of pragmatic measurement development in implementation science and behavioral health.

We describe our five-stage process for developing a brief, user-friendly, self-reported instrument for indexing clinician cognitive behavioral therapy (CBT) delivery in publicly funded mental health settings: the TPOCS-SeRTIFY (Beidas et al., 2016). The TPOCS-SeRTIFY is a self-report instrument to measure clinician CBT delivery for youth that parallels the CBT intervention items on the TPOCS-Revised Strategies (TPOCS-RS), a gold-standard observational coding system for clinician behavior (McLeod et al., 2015) that is often used to index community CBT use (e.g., Smith et al., 2017). Development goals were to design items to parallel the TPOCS-RS to allow for construct validation with direct observation, while creating a user-friendly instrument that would be comprehensible, appropriate, brief, and easy to use within a community mental health setting (i.e., is responsive to both an implementation gap and a stakeholder need). We use the exemplar of CBT for youth mental health as it has strong empirical support (Dorsey et al., 2017; Higa-McMillan et al., 2016; Hofmann et al., 2012; McCart & Sheidow, 2016) and because there have been nationwide implementation efforts designed to increase CBT use for youth (e.g., Bruns et al., 2008; Dorsey et al., 2016; Powell et al., 2016). Widespread CBT implementation will not be possible without companion self-report instruments with the capability to monitor and facilitate the adjustment of CBT delivery integrity to maximize outcomes across clinical settings.

Our process was informed by a conceptual model for conducting community-partnered research that places emphasis on bidirectional communication with community stakeholders throughout the research process (Pellecchia et al., 2018). Our pragmatic development steps include (1) develop an initial instrument draft based on prior discussions with community clinicians identifying the need for a brief, self-reported instrument of clinician CBT use in collaboration with content experts; (2) refine the instrument with community stakeholder and independent expert input; (3) collect preliminary feasibility and psychometric data; (4) conduct formal psychometric evaluation; and (5) identify next steps for instrument refinement, dissemination, and implementation, in collaboration with stakeholders. As the processes for conducting Steps 4 and 5 are better articulated in the literature (Lewis et al., 2018; Martin & Savage-McGlynn, 2013), we focus on Steps 1 through 3.

Approach

Step 1: initial measure development

Based on community partner feedback that existing clinician-report instruments measuring clinician CBT use are lengthy and contain too much jargon, the initial draft emphasized brevity and non-technical language and contained two versions: one to index CBT use at a session level and one across sessions representing overall treatment with a client. Twelve cognitive behavioral intervention items derived from the TPOCS-RS (e.g., cognitive distortion, parenting skills) were drafted. Items were reworded to remove technical language (e.g., changing "respondent strategies" to "exposure strategies") and provide
brief definitions for each item’s meaning. An exemplar was written for each intervention to illustrate what a clinician delivering this intervention might say (e.g., Psychoeducation: “It seems like you’ve been lashing out at the people you care about a lot lately. That is a common thing we see with depression, where it can make you more irritable.”). This initial draft was sent to three experts (authors SKS, SD, and AH) in CBT and measurement development, and a TPOCS-RS expert (author BDM) for review. Based on feedback, we refined item wording to further reduce technical language and make instructions more user-friendly (e.g., bolding relevant aspects of the directions). One expert recommended trying to reduce the instrument to one page; a second version was drafted that contained briefer definitions to fit onto a single page.

**Step 2: instrument refinement with stakeholders and independent experts**

We conducted cognitive interviews with six racially diverse, female clinicians who worked across four community mental health agencies. Given the intended goal of the TPOCS-SeRTIFY to monitor CBT use in the context of implementation efforts in publicly funded settings, we purposively interviewed stakeholders who held a primary clinician or supervisor role. Four clinicians were members of a Community Advisory Board formed by our team that provides consultation and feedback on the relevance, feasibility, and public health impact of research projects, consistent with our conceptual model for a community-partnered approach (Pellecchia et al., 2018). Cognitive interviewing is an interview procedure used in survey development to evaluate potential sources of response error (Willis, 2004). We used a combination of a “think aloud” and verbal probing approach to cognitive interviewing. Clinicians completed the instrument in the presence of a trained interviewer while talking out loud their thought process for arriving at each answer. The interviewer verbally probed for additional information to understand the interviewees’ comprehension and interpretation of each item. The instrument was iteratively revised following each interview, prioritizing items that were confusing or interpreted differently by participants. Example changes included refining instructions to clarify that each item represented a distinct CBT intervention, clarify item descriptions, change the item layout, so that conceptually similar items were grouped together to facilitate respondent’s ability to differentiate between them, change item names to make them more interpretable to the target user population (e.g., change “functional analysis of behavior” to “Antecedents, Behavior, and Consequences [ABC] model”), and refine examples to make them more realistic. The research team came to consensus on changes needed through weekly team meetings; in the rare instance of lack of consensus, the interview guide was amended to include specific queries of our community partners to resolve dispute. Saturation was reached after the fifth interview; no changes were made after the sixth interview.

We explicitly asked stakeholders for feedback about their preference of the short versus long form of the instrument by showing both and asking them directly which version they preferred. We counterbalanced which instrument they received first. As feedback was equivocal, we turned to our measurement experts for guidance, who recommended proceeding with the long form to enhance item interpretability.

Finally, to ensure scientific rigor and maximize concordance between the self-report and direct observation versions of the TPOCS-RS to allow for later comparison, the instrument was sent to two independent CBT experts alongside the original item descriptions on the TPOCS-RS. Experts rated how similar each item on the TPOCS-SeRTIFY were to the items on the TPOCS-RS on a seven-point scale (1 = Not At All Similar to 7 = Extremely Similar). Mean similarity was 5.69, with reviewers within 1 point of each other on 68% of intervention items, suggesting overall comparability between TPOCS-RS items and TPOCS-SeRTIFY items. However, two items (Coping Skills, Exposure, or Trauma Narrative) were rated 4 out of 7 by one rater. These items were discussed with the measurement and independent CBT experts and who had previously reviewed the instrument, and with the TPOCS-RS developer. These items were further revised to more explicitly map on to the TPOCS-RS constructs (i.e., changing coping skills to cognitive coping skills, more directly linking use of the coping skill to coping with strong feelings, revising the wording on the exposure or trauma narrative item to more explicitly state that exposures refer to structured practice). Revised items were sent back to the CBT experts who initially provided similarity ratings; revised items each received average similarity ratings of 6.5. Table 1 shows final items.

**Step 3: collect preliminary psychometric data**

We administered the final TPOCS-SeRTIFY to 247 mental health clinicians in a publicly funded mental health system to examine item performance and preliminary construct validity as compared with a lengthier, but commonly used, self-report measure: the Therapy Procedures Checklist-Family Revised (TPC-FR; Weersing et al., 2002). The TPC-FR is a well-validated self-report instrument (Kolko et al., 2009) that produces three subscales: use of CBT, family therapy, and psychodynamic techniques. All three subscales showed excellent internal consistency in this sample (α range = .89–.92). Clinicians were asked to select a representative client on their caseload and indicate the extent to which they used various techniques with this client on the TPC-FR (see Beidas et al., 2013; Supplemental File 1). The TPOCS-SeRTIFY asked respondents to report on the extent to which they used the 12 listed CBT interventions with the same client.
Table 1. Final TPOCS-RS SeRTIFY items and item descriptions.

| TPOCS-SeRTIFY intervention item | Final measure item descriptors |
|---------------------------------|--------------------------------|
| **Psychoeducation about the problem** | I provided information to the client to teach him or her about the nature of the problem for which he or she is seeking treatment |
| **Brief definition** | You may have provided education to the client about topics such as child development, parent-child relationships, or, symptoms, causes or treatment of the problem for which the client is seeking treatment |
| **Expanded definition** | Example statement “It seems like you’ve been lashing out at the people you care about a lot lately. That is a pretty common thing we see with depression, where it can make you more irritable.” |
| **Cognitive education** | I discussed with the client that his or her thoughts can impact how she or he acts and feels |
| **Brief definition** | You may have taught how thoughts influence body feelings and behavior; pointed out examples of the link between thoughts, body, feelings, and behavior from the client’s own life |
| **Expanded definition** | Example statement “Getting butterflies in your stomach when you are thinking about taking a big test is an example of how your thoughts relate to how you feel in your body.” |
| **Cognitive distortion** | I discussed with the client how she or he can identify unhelpful ways of thinking that influence how he or she feels and behaves and learn other ways of thinking that may be more helpful |
| **Brief definition** | You may have helped the client identify thoughts that may not be accurate or helpful; taught the client to become more aware of his or her thoughts; encouraged the client to challenge his or her thoughts and develop helpful ways of thinking |
| **Expanded definition** | Example statement “I heard you say that you feel like you’re stupid because you got a bad grade on your assignment. That sounds like an unhelpful thought. What’s another way you could think about it?” |
| **Antecedents, behaviors, and consequences model** | I discussed with the client how understanding what happens before and after a specific behavior (e.g., school refusal, a temper tantrum) is important for learning how to change that behavior |
| **Brief definition** | You may have identified things that happen before a specific behavior (antecedents); identified things that happen after a specific behavior (consequences); taught the client about how and why to identify antecedents and consequences |
| **Expanded definition** | Example statement “You’ve been skipping school a lot lately. What’s been going on in the morning on those days you are skipping school?” |
| **Relaxation** | I discussed with the client strategies she or he can use to relax to cope with strong feelings |
| **Brief definition** | You may have taught or encouraged the client to use deep breathing, muscle relaxation strategies, meditation, pleasant mental (i.e., guided) imagery, or other relaxation strategies |
| **Expanded definition** | Example statement “I can tell that you are upset by what your mom said to you earlier. Why don’t you try taking some deep breaths to relax.” |
| **Exposure or trauma narrative** | I planned, conducted, or reviewed the client’s attempt to complete structured activities designed to help them face their fears |
| **Brief definition** | You may have provided education about exposures; had the client make a list of uncomfortable/fearred situations and rank them from easy to hard; prepared for an exposure; helped the client gradually face uncomfortable feelings or situations (e.g., anxiety, reminders of trauma) in a supported way; had the client construct or share a trauma narrative |
| **Expanded definition** | Example statement “You’ve been working so hard to write down the story of the trauma that you went through. Today I was thinking we could work on the next chapter of your narrative.” “Are you ready to face your fear of spiders? Let’s start by looking at some pictures of a spider that I have here.” |

(Continued)
### Behavioral activation

**Brief definition**
I discussed with the client how participating in pleasant or fun activities can lead to improvements in mood.

**Expanded definition**
You may have taught the client about the relationship between pleasurable activities and mood; engaged in pleasurable activities in session to demonstrate the impact of these activities on mood; assigned the client to participate in pleasurable activities to improve mood.

**Example statement**
“Let’s make a list of fun activities that you can do this week. I would like you to pick a few activities to do and track what happens to your mood before and after you do those activities.”

### Cognitive coping skills

**Brief definition**
I taught or reviewed how to use cognitive coping skills (e.g., problem-solving) with the client to help him or her effectively deal with difficult situations, such as strong feelings (e.g., losing temper, feeling nervous) or social situations with peers or family.

**Expanded definition**
You may have taught or encouraged the client to use thinking strategies to improve social functioning, improve his or her ability to solve problems, manage anger, communicate more effectively, or generally improve daily functioning in response to stress using a cognitively oriented coping strategy.

**Example statement**
“It seems like it has been hard for you to spend the time you want with your friends lately. Can you use your problem-solving steps to think through ways that we can make this more possible for you?”

### Social or communication skill building

**Brief definition**
I taught or reviewed how to use behavioral coping skills (e.g., social skills, communication, assertiveness) to help the client effectively deal with difficult social situations with peers or family.

**Expanded definition**
You may have taught or encouraged the client to use behavioral skills to help improve his or her life, such as assertiveness skills and social skills; identified situations in which the client could use a specific skill.

**Example statement**
“I would like us to practice meeting new people. What are some ways we should act when we approach a new person?”

### Reinforcement strategies—child

**Brief definition**
I discussed with the child how rewards and consequences can be used to change his or her behavior or directly used rewards in session with the child to shape his or her behavior.

**Expanded definition**
You may have taught the child information about rewards or consequences; taught the child how to set up rewards for himself or herself; reviewed how previously developed reward systems are working; used rewards with a child directly (e.g., provided a reward or verbal praise for homework completion).

**Example statement**
“I know that making that phone call is going to be pretty hard. What can you do to reward yourself after you are successful at making the call?”

### Reinforcement strategies—parent

**Brief definition**
I discussed with a caregiver how rewards and consequences can be used to change his or her child’s behavior.

**Expanded definition**
You may have taught a caregiver strategies such as how to use positive attention, rewards, time-out, ignoring, or give consequences to change his or her child’s behavior; encouraged a caregiver to use one or more of these strategies; reviewed one or more of these strategies that were discussed in previous sessions (e.g., checked on how a reward system has been working).

**Example statement**
“It seems like your child is getting a lot of attention from you when she throws a temper tantrum when she doesn’t want to do her homework. Let’s talk about why ignoring this behavior might be helpful.”

### Other parenting skills

**Brief definition**
I discussed with a caregiver how to improve parenting skills to manage his or her child’s behavior or improve communication with his or her child.

**Expanded definition**
You may have taught a caregiver how to set limits, give instructions to his or her child in an effective way, provide appropriate supervision, monitor his or her child’s behavior, or communicate effectively.

**Example statement**
“When you are giving your son a task to do, it is helpful to make sure that he is paying attention to you and to make sure to keep the instructions simple and clear.”
Both measures were administered during a one-time, 2-hr meeting as part of a larger study (Beidas et al., 2013; six clinicians did not complete the TPOCS-SeRTIFY, yielding a final sample of 241. All procedures were approved by the (University of Pennsylvania and City of Philadelphia) Institutional Review Boards and all clinicians completed informed consent before participating. Participant characteristics were representative of the mental health workforce and were largely female (n=189, 78.4%) and master’s level (n=184, 76.3%). Clinicians identified as 43.6% (n=105) White, 25.7% (n=62) as Black or African American, 7.9% (n=19) as Asian, and 18.7% as “Other” (n=45); 4.1% did not disclose their race; 18.7% (n=45) of participants identified as Hispanic or Latinx. Clinicians averaged 37.5 years of age (SD=11.6) and 8.4 years of full-time clinical experience (SD=8.7).

Given that this instrument is designed to assess CBT use broadly across youth presenting concerns and not of prescribed CBT interventions for a particular disorder or problem area, we did not derive a scale score through calculating an “average” CBT score or item reduction analysis (e.g., factor analysis). Instead, we examined general item performance and preliminary score validity by examining the associations of each CBT item on the TPOCS-SeRTIFY to subscale scores on the TPC-FR (CBT, family, and psychodynamic strategies), using mixed-effects models with random intercepts to account for the nested nature of the data. Given that all three subscales of the TPC-FR are highly correlated with one another (rs range from .58 to .68), consistent with the broad range of techniques used in community-based treatment (Garland et al., 2010), we hypothesized that TPOCS-SeRTIFY item scores would be positively associated with all three TPC-FR subscales. However, we expected that TPOCS-SeRTIFY items would be more strongly associated with the TPC-FR CBT Strategies subscale, relative to the family and psychodynamic technique use subscales.

**TPOCS-SeRTIFY item performance.** Little’s (1988) MCAR test indicated that TPOCS-SeRTIFY data were missing completely at random ($\chi^2=181.5$, df=248, $p>.05$), and no item was missing more than 6% of its values, supporting measure feasibility. Skewness and kurtosis values indicated that all TPOCS-SeRTIFY items were normally distributed (absolute skewness <1, absolute kurtosis <1.5, well-within normal range per recommended parameters; West et al., 1995). Table 2 shows the descriptive statistics for TPOCS-SeRTIFY items for the sample. Clinicians reported using a moderate level of CBT with their representative clients across all 12 interventions.

**Preliminary construct validity.** Table 3 shows that all TPOCS-SeRTIFY items positively related to the CBT subscale of the

| Table 2. TPOCS-SeRTIFY descriptive statistics. |
|-----------------|---------|--------|
| **Adherence**   |         |        |
| Psychoeducation about the problem | 5.35 (1.30) | 1–7   |
| Cognitive education | 5.39 (1.36) | 1–7   |
| Cognitive distortion | 5.03 (1.57) | 1–7   |
| ABC model | 5.05 (1.53) | 1–7   |
| Relaxation | 5.39 (1.51) | 1–7   |
| Exposure or trauma narrative | 3.89 (1.88) | 1–7   |
| Behavioral activation | 4.76 (1.79) | 1–7   |
| Cognitive coping skills | 5.35 (1.35) | 1–7   |
| Social or communication skill building | 5.03 (1.51) | 1–7   |
| Reinforcement strategies: child | 4.47 (1.69) | 1–7   |
| Reinforcement strategies: parent | 4.70 (1.77) | 1–7   |
| Other parenting skills | 4.92 (1.68) | 1–7   |

| Table 3. Predictors of TPOCS-SeRTIFY item scores: TPC subscale scores (N=241). |
|-----------------|---------|--------|
| **TPOCS-SeRTIFY item** | **TPC-FR: CBT strategies** | **TPC-FR: psychodynamic strategies** | **TPC-FR: family therapy strategies** |
| | b (SE) | p | b (SE) | p | b (SE) | p |
| Psychoeducation about the problem | .52 (.13) | <.001 | .30 (.12) | .01 | .27 (.10) | .01 |
| Cognitive education | .87 (.13) | <.001 | .53 (.12) | <.001 | .15 (.10) | .15 |
| Cognitive distortion | 1.10 (.14) | <.001 | .68 (.14) | <.001 | .33 (.12) | .01 |
| ABC model | 1.25 (.13) | <.001 | .87 (.13) | <.001 | .46 (.11) | <.001 |
| Relaxation | .98 (.14) | <.001 | .58 (.14) | <.001 | .21 (.11) | .07 |
| Exposure or trauma narrative | 1.38 (.16) | <.001 | 1.04 (.16) | <.001 | .52 (.14) | <.001 |
| Behavioral activation | 1.28 (.16) | <.001 | .93 (.16) | <.001 | .42 (.13) | .002 |
| Cognitive coping skills | 1.03 (.12) | <.001 | .57 (.12) | <.001 | .43 (.10) | <.001 |
| Social or communication skill building | 1.09 (.13) | <.001 | .58 (.14) | <.001 | .37 (.11) | .001 |
| Reinforcement strategies: child | 1.65 (.13) | <.001 | .98 (.14) | <.001 | .73 (.11) | <.001 |
| Reinforcement strategies: parent | 1.07 (.17) | <.001 | .60 (.16) | <.001 | .92 (.12) | <.001 |
| Other parenting skills | .92 (.16) | <.001 | .67 (.15) | <.001 | .95 (.11) | <.001 |

TPC-FR: Therapy Procedures Checklist-Family Revised; CBT: cognitive behavioral therapy.
TPC-FR (unstandardized coefficients ranged from .52 to 1.65, all ps < .001). Consistent with hypotheses, associations between TPOCS-SeRTIFY items and the TPC-FR psychodynamic and family strategies subscales were of a smaller magnitude than the TPC-CBT strategies subscale. Consistent with established patterns of community practice, whereby clinicians employ techniques from multiple treatment families, all TPOCS-SeRTIFY items also correlated with the TPC-FR psychodynamic and family strategies scales, except for relaxation and cognitive education, which were unrelated to the TPC-FR family strategies subscale.

Discussion

Pragmatic measurement development addresses implementation science instrumentation challenges (Glasgow & Riley, 2013). This case study describes our five-step approach to developing a pragmatic, clinician-report instrument to measure CBT delivery, which offers a model for others hoping to develop pragmatic implementation measures. We highlight two critical aspects. First, our emphasis on involving community partners at all stages of research allowed us to make early measure changes to enhance its feasibility and perceived utility. Second, collecting preliminary feasibility and psychometric data prior to full-scale evaluation allowed us to ensure the measure was feasible to implement and had concordance with another, similar instrument before formal evaluation. Based on conversations with our community partners, the instrument has since been adapted to be a primarily session-level instrument. Work is ongoing to validate the TPOCS-SeRTIFY against direct observation of client sessions (Step 4) and will be shared with community partners to inform future research (Step 5).

Our specific measure, the TPOCS-SeRTIFY, also may fill an important youth mental health measurement gap. While primarily designed as a tool to monitor CBT, such measures can also contribute to practice-based learning efforts and future research aimed at building learning mental health systems, which requires feasible ways to index implementation outcomes (Beidas & Wiltsey Stirman, 2020). Developing the TPOCS-SeRTIFY is among the first efforts (see also, Hogue et al., 2014, 2015) to enhance self-report measures to increase their accuracy and clinical utility that (1) can be applied transdiagnostically, regardless of a specific manual or protocol a clinician may be using and (2) parallels an existing direct observational measure. This latter point will facilitate future work examining strategies for enhancing the accuracy of clinician’s self-rating skills to optimize the TPOCS-SeRTIFY’s utility. Related, examining how systematic clinician biases inform their self-report will be important for refining the TPOCS-SeRTIFY to enhance clinicians’ accuracy to self-rate.

This study had several strengths, including the rigorous, stepped, community-partnered development process. While data provide only preliminary information regarding the TPOCS-SeRTIFY’s psychometric performance, results are promising and suggest the benefit of additional work in this area. Measures designed collaboratively with community stakeholders may also afford benefits over researcher-developed gold-standard implementation instruments, which are often developed on samples without historically excluded groups. The primary limitation is the lack of direct observation against which to compare results of clinician ratings on the TPOCS-SeRTIFY and our comparisons to an existing self-report measure that also has not been tested in comparison with direct observation. An additional limitation is lack of gender diversity in our stakeholder sample.

A self-report instrument that maps directly on to a gold-standard coding system, provides clearly operationalized clinician behaviors of interest, and is designed to facilitate clinician self-ratings, fills an important measurement gap for youth mental health service implementation research and practice. This is a critical step in a growing body of research (Hogue et al., 2015) to enhance accurate and feasible self-reported measurement of clinician practice use. Our process can serve as a guide for the development of pragmatic measures across implementation science.

Authors’ note

This study was registered with ClinicalTrials.gov (Identifier NCT02820623).

Acknowledgements

The authors gratefully acknowledge Judy Shea, David Mandell, Steven Marcus, Jessica Fishman, Catherine Maclean, and Michael French for their contributions to the larger clinical trial and Courtney Gregor, Kelly Zentgraf, and Adina Lieberman for their coordination support. They are grateful for the support that the Department of Behavioral Health and Intellectual Disability Services has provided us to conduct this work within their system, for the Evidence-Based Practice and Innovation (EPIC) group, and for the partnership provided to us by participating agencies, therapists, and clients.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Dr Beidas discloses that she is an Associate Editor at Implementation Research and Practice. She also receives royalties from Oxford University Press and has consulted for Camden Coalition of Healthcare Providers. Dr. Beidas provides consultation to United Behavioral Health. Dr. Beidas serves on the Clinical and Scientific Advisory Board for Optum Behavioral Health. Dr Schoenwald is co-founding editor of Implementation Research and Practice. All other authors have no conflict of interest to disclose.
Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by Grant R01 MH108551 (PI: Beidas) and Grant K23 MH099179 (PI: Beidas) from NIMH.

ORCID iDs
Sonja K Schoenwald https://orcid.org/0000-0002-0560-8440
Mary L Phan https://orcid.org/0000-0003-4382-2867

Supplemental material
Supplemental material for this article is available online.

References
Beidas, R. S., Aarons, G., Barg, F., Evans, A., Hadley, T., Hoagwood, K., & Mandell, D. S. (2013). Policy to implementation: evidence-based practice in community mental health–study protocol. Implementation Science, 8(1), 1–9. https://doi.org/10.1186/1748-5908-8-38

Beidas, R. S., Maclean, J. C., Fishman, J., Dorsey, S., Schoenwald, S. K., Mandell, D. S., Shea, J. A., McLeod, B. D., French, M. T., Hogue, A., Adams, D. R., Lieberman, A., Becker-Haimes, E. M., & Marcus, S. C. (2016). A randomized trial to identify accurate and cost-effective fidelity measurement methods for cognitive-behavioral therapy: Project FACTS study protocol. BMC Psychiatry, 16, 323. https://doi.org/10.1186/s12888-016-1034-z

Beidas, R. S., & Willsey Stirman, S. (2020). Realizing the promise of learning organizations to transform mental health care: Telespsychiatry care as an exemplar. Psychiatric Services, 72, 86–88. https://doi.org/10.1176/appi.ps.202000257

Bond, G. R., & Drake, R. E. (2019). Assessing the fidelity of evidence-based practices: History and current status of a standardized measurement methodology. Administration and Policy in Mental Health and Mental Health Services Research, 47, 874–884. https://doi.org/10.1007/s10488-019-00991-6

Bruns, E. J., Hoagwood, K. E., & Hamilton, J. D. (2008). State implementation of evidence-based practice for youths, part I: Responses to the state of the evidence. Journal of the American Academy of Child & Adolescent Psychiatry, 47(4), 369–373. https://doi.org/10.1097/CHI.0b013e31816485f4

Dorsey, S., Berliner, L., Lyon, A. R., Pullmann, M. D., & Murray, L. K. (2016). A statewide common elements initiative for children’s mental health. Journal of Behavioral Health Services and Research, 43(2), 246. https://doi.org/10.1007/s11414-014-9430-y

Dorsey, S., McLaughlin, K. A., Kerns, S. E., Harrison, J. P., Lambert, H. K., Briggs, E. C., Revillion Cox, J., & Amaya-Jackson, L. (2017). Evidence base update for psychosocial treatments for children and adolescents exposed to traumatic events. Journal of Clinical Child and Adolescent Psychology, 46(3), 303–330. https://doi.org/10.1080/15374 416.2016.1220309

Garland, A. F., Brookman-Frazee, L., Hurlburt, M. S., Accurso, E. C., Zoffness, R. J., Haine-Schlagel, R., & Ganger, W. (2010). Mental health care for children with disruptive behavior problems: A view inside therapists’ offices. Psychiatric Services, 61(8), 788–795. https://doi.org/10.1176/ps.2010.61.8.788

Glasgow, R. E., & Riley, W. T. (2013). Pragmatic measures: What they are and why we need them. American Journal of Preventive Medicine, 45(2), 237–243. https://doi.org/10.1016/j.amepre.2013.03.010

Higa-McMillan, C. K., Francis, S. E., Rith-Najarian, L., & Chorpita, B. F. (2016). Evidence base update: 50 years of research on treatment for child and adolescent anxiety. Journal of Clinical Child and Adolescent Psychology, 45(2), 91–113. https://doi.org/10.1080/15374416.2015.1046177

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. Cognitive Therapy and Research, 36(5), 427–440. https://doi.org/10.1007/ s10608-012-9476-1

Hogue, A., Dauber, S., & Henderson, C. E. (2014). Therapist self-report of evidence-based practices in usual care for adolescent behavior problems: Factor and construct validity. Administration and Policy in Mental Health and Mental Health Services Research, 41(1), 126–139. https://doi.org/10.1007/s10488-012-0442-8

Hogue, A., Dauber, S., Lichvar, E., Bobek, M., & Henderson, C. E. (2015). Validity of therapist self-report ratings of fidelity to evidence-based practices for adolescent behavior problems: Correspondence between therapists and observers. Administration and Policy in Mental Health and Mental Health Services Research, 42(2), 229–243. https://doi.org/10.1007/s10488-014-0548-2

Kolko, D. J., Cohen, J. A., Mannarino, A. P., Baumann, B. L., & Knudsen, K. (2009). Community treatment of child sexual abuse: A survey of practitioners in the national child traumatic stress network. Administration and Policy in Mental Health and Mental Health Services Research, 36(1), 37–49. https://doi.org/10.1007/s10488-008-0180-0

Lewis, C. C., Puspitasari, A., Boyd, M. R., Scott, K., Marriott, B. R., Hoffman, M., Navarro, E., & Kassab, H. (2018). Implementing measurement based care in community mental health: A description of tailored and standardized methods. BMC Research Notes, 11(1), 76. https://doi.org/10.1186/s13104-018-3193-0

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. Journal of the American Statistical Association, 83(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Martin, C. R., & Savage-McGlynn, E. (2013). A “good practice” guide for the reporting of design and analysis for psychometric evaluation. Journal of Reproductive and Infant Psychology, 31(5), 449–455. https://doi.org/10.1080/0264 838.2013.835036

McCurt, M. R., & Sheidow, A. J. (2016). Evidence-based psychosocial treatments for adolescents with disruptive behavior. Journal of Clinical Child and Adolescent Psychology, 45(5), 529–563. https://doi.org/10.1080/15374416.2016.11 46990

McLeod, B. D., Smith, M. M., Southam-Gerow, M. A., Weisz, J. R., & Kendall, P. C. (2015). Measuring treatment differentiation for implementation research: The Therapy Process Observational Coding System for Child Psychotherapy.
Revised Strategies Scale. *Psychological Assessment*, 27(1), 314. https://doi.org/10.1037/pas0000037

Pellecchia, M., Azad, G., Maddox, B., Nuske, H. J., Skriner, L. C., Stewart, R., Wolk, C. B., Mandell, D. S., & Beidas, R. S. (2018). Community-academic partnerships in implementation research. *Journal of Community Psychology*, 46, 941–952. https://doi.org/10.1002/jcop.21981

Powell, B. J., Beidas, R. S., Rubin, R. M., Stewart, R. E., Wolk, C. B., Matlin, S. L., Weaver, S., Hurford, M. O., Evans, A. C., & Hadley, T. R. (2016). Applying the policy ecology framework to Philadelphia’s behavioral health transformation efforts. *Administration and Policy in Mental Health and Mental Health Services Research*, 43(6), 909–926. https://doi.org/10.1007/s10488-016-0733-6

Powell, B. J., Stanick, C. F., Halko, H. M., Dorsey, C. N., Weiner, B. J., Barwick, M. A., Damschroder, L. J., Wensing, M., Wolfenden, L., & Lewis, C. C. (2017). Toward criteria for pragmatic measurement in implementation research and practice: A stakeholder-driven approach using concept mapping. *Implementation Science*, 12(1), 118. https://doi.org/10.1186/s13012-017-0649-x

Proctor, E. K., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G. A., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(2), 65–76. https://doi.org/10.1007/s10488-010-0319-7

Rodriguez-Quintana, N., & Lewis, C. C. (2018). Observational coding training methods for CBT treatment fidelity: A systematic review. *Cognitive Therapy and Research*, 42(4), 358–368. https://doi.org/10.1007/s10608-018-9898-5

Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(1), 32–43. https://doi.org/10.1007/s10488-010-0321-0

Smith, M. M., McLeod, B. D., Southam-Gerow, M. A., Jensen-Doss, A., Kendall, P. C., & Weisz, J. R. (2017). Does the delivery of CBT for youth anxiety differ across research and practice settings? *Behavior Therapy*, 48(4), 501–516. https://doi.org/10.1016/j.beth.2016.07.004

Weersing, V. R., Weisz, J. R., & Donenberg, G. R. (2002). Development of the therapy procedures checklist: A therapist-report measure of technique use in child and adolescent treatment. *Journal of Clinical Child and Adolescent Psychology*, 31(2), 168–180. https://doi.org/10.1207/S15374424JCCP3102_03

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). SAGE.

Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. SAGE. https://doi.org/10.4135/9781412983655