

Clinician Use of Standardized Assessments Following a Common Elements Psychotherapy Training and Consultation Program

Aaron R. Lyon · Shannon Dorsey · Michael Pullmann ·
Jessica Silbaugh-Cowdin · Lucy Berliner

Published online: 4 March 2014
© Springer Science+Business Media New York 2014

Abstract Despite increasing emphasis on the implementation of evidence-based treatments in community service settings, little attention has been paid to supporting the use of evidence-based assessment (EBA) methods and processes, a parallel component of evidence-based practice. Standardized assessment (SA) tools represent a key aspect of EBA and are central to data-driven clinical decision making. The current study evaluated the impact of a statewide training and consultation program in a common elements approach to psychotherapy. Practitioner attitudes toward, skill applying, and use of SA tools across four time points (pre-training, post-training, post-consultation, and follow-up) were assessed. Results indicated early increases in positive SA attitudes, with more gradual increases in self-reported SA skill and use. Implications for supporting the sustained use of SA tools are discussed, including the use of measurement feedback systems, reminders, and SA-supportive supervision practices.

Keywords Evidence-based assessment · Standardized assessment · Community mental health · Implementation · Training

Introduction

A growing body of literature has focused on improving the quality of mental health services available in community settings by increasing clinician use of evidence-based

practices (EBP) through targeted implementation efforts (Fixsen et al. 2005; McHugh and Barlow 2010). This literature has developed in response to numerous findings that research-supported interventions are used infrequently in “usual care” contexts where most children, families, and adults receive services (Garland et al. 2012), and has spawned multiple large-scale efforts to increase adoption and sustained use of EBP (e.g., Clark et al. 2009; Graham and Tetroe 2009) with varying degrees of success. Although EBP have generally outperformed usual care in clinical trials (Weisz et al. 2006), recent research has questioned the assertion that simply implementing EBP in public mental health settings is sufficient for decreasing symptoms or enhancing functioning beyond the effects of usual care (Barrington et al. 2005; Southam-Gerow et al. 2010; Spielmans et al. 2010; Weisz et al. 2009, 2013). These findings suggest that additional strategies may be necessary to improve the effectiveness of community services.

A complementary or alternative approach, as articulated by Chorpita et al. (2008), is to shift the focus of implementation from “using EBP” toward the super-ordinate objective of “getting positive outcomes.” This perspective acknowledges that the promise of improved outcomes is often most compelling to clinicians and policy makers. A body of literature is now developing which places increased emphasis on the explicit measurement of results rather than simply focusing on the use of EBP or fidelity to EBP models.

Evidence-Based Assessment (EBA)

EBP is comprised of both evidence-based treatment and EBA components. EBA, defined as assessment *methods* and *processes* that are based on empirical evidence (i.e.,

A. R. Lyon (✉) · S. Dorsey · M. Pullmann ·
J. Silbaugh-Cowdin · L. Berliner
University of Washington, 6200 NE 74th St., Suite 100, Seattle,
WA 98115, USA
e-mail: lyona@uw.edu

reliability and validity) and their clinical usefulness for prescribed populations and purposes (Mash and Hunsley 2005), is an element of nearly all EBP treatment protocols. Notably, this definition includes both assessment tools (methods) as well as mechanisms for effectively integrating assessment information into service delivery through strategies such as feedback and clinical decision making (processes). Use of standardized assessment (SA) tools is a central component of EBA (Jensen-Doss and Hawley 2010). SA serves a vital role in clinical practice because research has indicated that therapists generally are not effective in judging client progress or deterioration (Hannan et al. 2005). The use of SA tools for initial evaluation and ongoing progress monitoring also represents a core, evidence-based clinical competency (Sburlati et al. 2011). Additionally, evidence is accumulating to suggest that SA and progress monitoring, when paired with feedback to clinicians, have the ability to enhance communication between therapists and clients (Carlier et al. 2012) and may improve adult and youth outcomes independent of the specific treatment approach (Bickman et al. 2011; Brody et al. 2005; Lambert et al. 2003). For these reasons, use of assessment and monitoring protocols is increasingly being identified as an EBP in and of itself (Substance Abuse and Mental Health Services Administration 2012). In this way, EBA simultaneously satisfies the dual implementation aims of increasing the use of EBP (because the use of EBA can be considered an EBP) and monitoring and improving outcomes (Chorpita et al. 2008).

Although EBA methods and processes—and their role in diagnosis and outcome tracking—are increasingly a topic in the mental health literature (e.g., Poston and Hanson 2010; Youngstrom 2008), they are infrequently discussed in the context of the dissemination and implementation of EBP (Mash and Hunsley 2005). Indeed, the status of EBA has been likened to the state of evidence-based treatment a decade earlier, when compelling empirical support had begun to emerge about effectiveness, but few studies had explored how to support their uptake and sustained use (Jensen-Doss 2011; Mash and Hunsley 2005). Despite their advantages, components of EBA—such as SA tools—are used infrequently by community-based mental health providers regardless of their discipline (Hatfield and Ogles 2004; Gilbody et al. 2002). In a study of the assessment practices of child and adolescent clinicians, for example, Palmiter (2004) reported that only 40.3 % reported using any parent rating scales and only 29.2 % used child self-report scales. Garland et al. (2003) found that even practitioners who consistently received scored SA profiles for their clients rarely engaged in corresponding EBA processes, such as incorporating the assessment findings into treatment planning or progress monitoring.

There are multiple reasons for the low level of EBA penetration in community practice. Hunsley and Mash (2005) have observed that EBA is not a component of many training programs for mental health providers from a variety of backgrounds. Particularly problematic, SA tools have traditionally been the province of psychologists. In contrast, the vast majority of mental health providers in public mental health settings are from other disciplines (e.g., social work, counseling) (Ivey et al. 1998; Robiner 2006). Nevertheless, the advent of free, brief, reliable, and valid measures that can be easily scored and interpreted has significantly extended the utility and feasibility of SA of many mental health conditions. For example, the Patient Health Questionnaire (PHQ-9) is widely used internationally in medical contexts to identify depressed patients and monitor their progress (e.g., Clark et al. 2009; Ell et al. 2006).

Some have advocated for supplementing training in EBP with specific training in EBA methods and processes to increase clinician knowledge and skills (Jensen-Doss 2011). For instance, research examining attitudes toward SA and diagnostic tools suggests a link between attitudes and use (Jensen-Doss and Hawley 2010, 2011). Jensen-Doss and Hawley (2010) found that clinician attitudes about the “practicality” of SA tools (i.e., the feasibility of using the measures in practice) was an independent predictor of use in a large ($n = 1,442$) multidisciplinary sample of providers. Interestingly, the other attitude subscales, “psychometric quality” (i.e., clinician beliefs about whether assessment measures can help with accurate diagnosis and are psychometrically sound) and “benefit over clinical judgment” (i.e., clinician belief about whether assessment measures added value over clinical judgment), were not independently related to SA use. Targeted research studies, training initiatives, and implementation efforts that attend closely to attitudes and other predictors of use are needed to identify barriers to use and strategies to promote the utilization of SA in routine clinical practice.

When discussing strategies to increase SA use, Mash and Hunsley (2005) have warned against simply suggesting or mandating that clinicians adopt SA tools: “Blanket recommendations to use reliable and valid measures when evaluating treatments are tantamount to writing a recipe for baking hippopotamus cookies that begins with the instruction ‘use one hippopotamus,’ without directions for securing the main ingredient” (p. 364). Indeed, some authors have suggested that inadequate preparation of clinicians in the proper use of SA measures in their clinical interactions (e.g., via mandates without adequate accompanying supports) carries potential iatrogenic consequences for the recipients of mental health services, including the possibility that assessment questions could be viewed as irrelevant or that results could be used to limit service

access (Wolpert 2014). Instead, SA implementation efforts should provide specific instruction and ongoing consultation related to the selection, administration, and use of SA tools in a manner consistent with the implementation science literature. Recommendations include active initial training and continued contact and support during a follow-up period, both of which appear essential to achieving lasting professional behavior change (Beidas and Kendall 2010; Fixsen et al. 2005). Such efforts should include information about how to select appropriate measures at different phases of the intervention process (e.g., longer screening tools at intake and shorter, problem-specific assessments at regular intervals thereafter) to maximize their relevance.

Although authors have called for improved training in the use of SA (Hatfield and Ogles 2007) and some recent evidence suggests that SA use may represent a particularly malleable behavior change target in response to training in a larger intervention approach (e.g., Lyon et al. 2011a), there remains very little research focused on the uptake of SA following training.

SA and the “Common Elements” Approach

SA and progress monitoring have particular relevance to emerging *common elements* approaches to the dissemination and implementation of EBP. The common elements approach is predicated on the notion that most evidence-based treatment protocols can be subdivided into meaningful practice components (Chorpita et al. 2005a). Furthermore, recent common elements approaches make explicit use of *modularized design* (e.g., Weisz et al. 2012). In modular interventions, individual components can be implemented independently or in complement with one another to bring about specific treatment outcomes (Chorpita et al. 2005b). Findings in support of the feasibility and effectiveness of such approaches are now emerging. For instance, Borntrager et al. (2009) found that a modularized, common elements approach focused on anxiety, depression, and conduct problems among youth was more acceptable to community-based practitioners than more typical, standard-arranged evidence-based treatment manuals. Furthermore, the results of a recent randomized controlled trial (RCT) found that the same intervention outperformed standard manuals and usual care (Weisz et al. 2012).

SA is the “glue” that holds modular interventions together because assessment results are used to guide decisions about the application or discontinuation of treatment components (Daleiden and Chorpita 2005). SA tools are used initially to identify presenting problems and formulate a treatment approach. Over time, assessments provide important indicators of client progress and can help

to determine if shifts in the treatment approach are indicated (Daleiden and Chorpita 2005). For this reason, training in SA is among the most essential components to consider when implementing a modularized, common elements approach. As Huffman et al. (2004) have stated, “the provision of education and training experiences to practitioners regarding the use of empirical methods of practice is likely to improve their attitudes about outcomes measurement and to increase its implementation across all phases of treatment” (p. 185). Despite this, no studies have examined the specific impact of training on the use of SA tools.

The Washington State CBT+ Initiative

Starting in 2009, clinicians from public mental health agencies have received training in a version of a modularized, common elements approach to the delivery of cognitive behavioral therapy (CBT) for depression, anxiety, behavior problems, and posttraumatic stress disorder through the Washington State CBT+ Initiative (see Dorsey et al., under review). Due to the centrality of SA in modular approaches, the Initiative includes a strong focus on the use of SA tools (see “CBT+ Training” section below) and associated EBA processes. CBT+ is funded by federal block grant dollars administered by the state mental health agency (Department of Social and Health Services, Division of Behavioral Health Recovery). The CBT+ Initiative developed from an earlier statewide trauma-focused cognitive behavioral therapy (TF-CBT) (Cohen et al. 2006) initiative designed to expand EBP reach to the broader range of presenting problems among children seeking care in public mental health. Funding for CBT+ and the preceding TF-CBT Initiative has been relatively modest, ranging from \$60 to 120 thousand a year, depending on the specific activities. Between 2009 and 2013, five cohorts of clinicians ($n = 498$), from 53 agencies statewide, have participated.

Current Aims

In light of the documented impact of EBA on client outcomes and the central role of SA—a key EBA component—in effectively delivering a modularized, common elements approach to psychotherapy, the current study was designed to evaluate the specific impact of the CBT+ training and consultation program on knowledge about, attitudes toward, and use of SA tools. Longitudinal information about SA was collected from a subset of trainees from each of the cohorts who participated in the CBT+ Initiative to address the following research questions: (a) to what extent do community providers’ SA attitudes, skill, and use change over the course of training and

consultation, and (b) which factors predict change in the use of SA tools?

Method

Participants

Participants in the current study included 71 clinicians and supervisors from the larger group of CBT+ training participants in three training cohorts (fall 2010, spring 2011, and fall 2011) who agreed to participate in a longitudinal evaluation of the CBT+ Initiative. CBT+ training participation was available to clinicians in community public mental health agencies serving children and families in Washington State. To be included in this study, participants in the larger training initiative (~400 registered participants) had to agree to participate in the longitudinal evaluation, complete the baseline evaluation, and complete the research measures again at one or more of three follow-up time points. Participants were predominantly female (82 %), Caucasian (89 %), masters-level providers (87 %) in their late twenties and thirties (68 %; see Table 1).

Table 1 Participant demographics ($N = 71$)

Variables	<i>n</i>	%
Current role ^a		
Clinician	60	84.5
Supervisor	13	18.3
Female	58	81.7
Ethnicity or race		
Caucasian	63	88.7
African American	3	4.2
Hispanic or Latino	2	2.8
Asian	2	2.8
Other	1	1.4
Educational background		
High school degree	1	1.4
Degree from 4-year college	6	8.5
Master's in social work	21	29.6
(Other) Master's degree	41	57.7
(Other) Doctoral degree	2	2.8
Age		
Under 25	2	2.8
25–29	18	25.4
30–39	30	42.3
40–49	10	14.1
Over 50	11	15.5

^a Current role adds to more than 100 % because participants could choose both options

CBT+ Training

The CBT+ Initiative training model includes a three-day active, skills-based training and 6-months of biweekly, expert-provided phone consultation (provided by the second and last authors, and other Initiative faculty). These training procedures were designed to be consistent with best practices identified in the implementation literature (Beidas and Kendall 2010; Fixsen et al. 2005; Herschell et al. 2010; Lyon et al. 2011b). Participating agencies are required to send one supervisor and two or three clinicians to the training to ensure that each has a clinician cohort and a participating supervisor. Additionally, participating supervisors are asked to attend an annual 1-day supervisor training and a monthly CBT+ supervisor call. Training incorporates experiential learning activities (e.g., cognitive restructuring activity for a situation in the clinicians' own life), trainer modeling and video demonstration of skills, trainee behavioral rehearsal of practice elements with both peer and trainer feedback and coaching, and small and large group work. Training is tailored to be applicable to the circumstances of public mental health settings in that it includes an explicit emphasis on engagement strategies and skills, focuses on triage-driven decision making when faced with comorbidity (e.g., pick a clinical target for initial focus), and teaches how to deliver interventions to accommodate the typical 1-h visit (for a full description of the practice element training, see Dorsey et al., under review).

In addition to training on the common treatment elements, the EBA component of CBT+ includes a specific focus on brief SA tools to (a) identify the primary target condition and (b) measure treatment improvement. Training includes active practice (e.g., scoring standardized measures) and behavioral rehearsal of key EBA processes (e.g., role plays in which trainees provide feedback to children and caregivers in small groups using a case vignette description and scored measures). Assessment measures were selected based on the following characteristics: (a) limited items, (b) ability to score quickly by hand, (c) available in the public domain (i.e., no cost to agencies), and (d) strong psychometric properties. Measures selected based on these criteria included the Pediatric Symptom Checklist (Gardner et al. 1999), the Mood and Feelings Questionnaire (Angold and Costello 1987), the Screen for Child Anxiety-Emotional Related Disorders (SCARED)—5-item version and traumatic stress subscales (Birmaher et al. 1997; Muris et al. 2000), and the Child Posttraumatic Stress Symptoms Checklist (Foa et al. 2001).

In the CBT+ model, clinicians begin group consultation calls within 3 weeks after training, focused on implementing assessment and treatment with clients on their caseloads. Each call group included 3–4 agency teams with

~10–15 trainees. Calls are led by experts in CBT, use a clinical case presentation format, and involve reviewing assessment data to determine clinical focus, applying CBT+ components to cases, and problem-solving challenges with child and caregiver engagement. Trainees are expected to present at least one case during the consultation period and attend 9 of 12 calls to receive a certificate of participation.

Data Collection

All CBT+ attendees participated in a basic program evaluation in which they were asked to complete brief questionnaires at two time points (i.e., before the training and following the 6-month consultation period) as a general check on training quality. This paper reports on a subset of the CBT+ attendees (i.e., the “longitudinal sample”) who participated in a more intensive evaluation that included completion of additional measures pre-training (T1), immediately post-training (T2), immediately after consultation was completed (T3), and at a follow-up time point 3 months after the conclusion of consultation activities (T4). All attendees were invited to participate in the longitudinal evaluation. Attendees who agreed to participate in the longitudinal evaluation received a \$10 gift card for completion of assessment measures at each assessment point. Research activities were submitted to the Washington State IRB and determined to be exempt from review.

Measures

Demographics

All Washington State CBT+ participants completed a questionnaire collecting demographic information (e.g., age, gender, ethnicity), agency, role in the agency, and years of experience.

Consultation Dose

For each consultation call, clinician attendance was reported by the expert consultants who led the calls.

Attitudes Toward Standardized Assessment Scales (ASA)

The ASA (Jensen-Doss and Hawley 2010) is a 22-item measure of clinician attitudes about using SA in practice. It includes three subscales: Benefit over Clinical Judgment, Psychometric Quality, and Practicality (described previously). All subscales have been found to demonstrate good psychometrics. Higher ratings on all subscales have been

associated with a greater likelihood of SA use (Jensen-Doss and Hawley 2010). The ASA was administered at each of the four time points.

Clinician-Rated Assessment Skill

As part of a larger self-assessment of understanding and skill in delivering components of treatment for anxiety, depression, behavioral problems, and PTSD/trauma, study clinicians reported on their understanding and skill administering assessment measures and giving feedback for each of the four clinical targets. Items were rated on a 5-point Likert scale from “Do Not Use” to “Advanced.” The four-item scale had high internal consistency (Cronbach’s $\alpha = .90$), and items were averaged to create a total assessment skill score. Assessment skill was administered at each of the four time points.

Current Assessment Practice Evaluation (CAPE)

The CAPE (Lyon and Dorsey 2010) is a four-item measure of clinician ratings of SA use across different phases of intervention (e.g., at intake, ongoing during treatment, at termination). Items capture the use of SA tools and associated EBA processes (e.g., incorporation of assessment results into treatment planning, provision of SA-based feedback to children/families) and are scored on a 4-point scale [None, Some (1–39 %), Half (40–60 %), Most (61–100 %)]. Total score internal reliability in the current sample was acceptable ($\alpha = .72$). Because there was no opportunity for clinicians to change their actual use of SA tools immediately following training, the CAPE was collected at three time points: pre-training, post-consultation, and at the 3-month post-consultation follow-up.

Results

Crosstabulations with χ^2 tests and t-tests were run to examine differences between participants in the longitudinal sample ($n = 71$) and all other participants for whom data were available ($n = 314$). At baseline, there were no statistically significant differences between the groups by reported role (supervisor or therapist), gender, race/ethnicity, age, level of education, whether they had ever provided therapy, number of years providing therapy, whether they received supervision, whether they ever provided supervision, frequency of use of CBT, and the four clinician-rated assessment skill questions. Those in the longitudinal sample were more likely to report receiving supervision in a specific evidence-based treatment model (50.7 vs. 35.7 %, $\chi^2_{(1)} = 5.03$, $p = .025$) and attended

Table 2 Means and standard deviations of attitudes toward standardized assessments, assessment skill, and current assessment practice for total sample size at each time point

	Possible range	Time point <i>M</i> (<i>SD</i>)			
		Pre-training	Post-training	Post-consultation	Follow-up
<i>n</i> with complete data	–	67	71	52	47
Benefit over clinical judgment	1–5	3.13 (0.30)	3.18 (0.29)	3.21 (0.31)	3.25 (0.29)
Psychometric quality	1–5	3.79 (0.41)	4.08 (0.39)	4.15 (0.33)	4.17 (0.41)
Practicality	1–5	3.25 (0.53)	3.70 (0.43)	3.76 (0.37)	3.78 (0.39)
Standardized assessment skill	1–5	2.49 (1.06)	3.26 (1.12)	3.91 (0.59)	3.91 (0.47)
Standardized assessment total use (CAPE)	1–4	2.08 (0.73)	–	2.62 (0.72)	2.46 (0.79)
CAPE item 1: Percentage of intake clients administered SA in last month	1–4	2.79 (1.23)	–	3.51 (0.78)	3.28 (0.99)
CAPE item 2: Percentage of caseload administered SA in last week	1–4	1.88 (0.98)	–	2.16 (0.95)	2.32 (0.96)
CAPE item 3: Percentage of clients given feedback from SA during last week	1–4	2.20 (1.06)	–	2.63 (1.07)	2.57 (1.07)
CAPE item 4: Percentage of clients changed treatment plan based on SA scores in last week	1–4	1.48 (0.61)	–	2.24 (0.95)	1.68 (0.84)

slightly more consultation calls ($M = 9.2$ vs. 8.4 , $t_{(138.2)} = -2.37$, $p = .019$).

Table 2 depicts the number of participants with complete data at each time point. Of the 71 participants, there were 68 (95.7 %) with data at baseline, 71 (100 %) with data post-training, 52 (73 %) with complete data at post-consultation, 47 (66 %) with complete data at 3-month post-consultation follow-up, and 56 (79 %) with complete data for at least one of these time points. To determine if missingness was related to pre-training values of outcome variables or to the descriptive characteristics of the population, we ran a series of t-tests and χ^2 tests using three dichotomous independent variables reflecting missing status (i.e., at post-consultation, at follow-up, and at both post-consultation and follow-up) and the same series of variables described above. There were no significant differences between participants with missing data and participants with complete data.

Additionally, we obtained data on the number of consultation calls attended for 64 of the 71 participants. These participants attended an average of 8.6 calls ($SD = 2.6$). Forty-four (68.8 %) attended 9 or more calls, which was the number required to obtain a certificate of completion.

Primary SA Outcome Analyses

Table 2 also depicts the mean scores and standard deviations at each time point for the five outcome variables. Because of missing data and shared variance in observations due to repeated measures and nesting of clinicians within agencies, we used 3-level longitudinal multilevel modeling (agency, clinician, time point) with full maximum likelihood estimation to examine changes over time for the five outcome

variables, as well as item-level analysis of the four items on the SA use measure (CAPE) in order to determine longitudinal changes by specific type of assessment-related behavior. For the attitudes toward assessment and assessment skill outcome measures, independent models estimated the piecewise rate of change for three slopes: pre-training to post-training, post-training to post-consultation, and post-consultation to follow-up. Because we did not have post-training data on SA use, we estimated two slopes (pre-training to post-consultation, and post-consultation to 3-month follow-up for these outcomes. Fully random effects models failed to converge due to limited sample size, therefore, only the intercept term was permitted to randomly vary. Intraclass Correlations (ICCs), a measure of within-group similarity, indicated that clinician responses at Level 3 (agency) were dissimilar for all attitude DVs (ICCs ranged from .0001 to .0004). Clinicians within agency were more similar for SA skill (ICC = .20) and SA total use (ICC = .10). Most of the similarity for SA total use was due to the item measuring SA use at intake (ICC = .27). The ICCs for items about the percentage of clients administered a SA for the total caseload, who were given feedback based on SA, and who had a change in their treatment plan were all relatively small (ICC = .08, .002, .0001, respectively). Model results are shown in Table 3 and graphically depicted in Fig. 1 to simplify interpretation.

SA Attitudes

Practicality and *psychometric quality* both exhibited a statistically significant improvement after training (T1–T2 $\beta = .45$ and $.29$, respectively, $p < .001$), but did not change during consultation and did not have any significant

Table 3 Mixed model regression coefficient estimates of fully piecewise models for rate of change over time on attitudes toward standardized assessments, assessment skill, and current assessment practice

	Pre-training estimate (SE)	T1–T2 slope (SE)	<i>p</i>	T2–T3 slope (SE)	<i>p</i>	T3–T4 slope (SE)	<i>p</i>
Benefit over clinical judgment	3.13 (.04)	.05 (.05)	.38	.04 (.05)	.35	.03 (.04)	.45
Practicality	3.24 (.08)	.45 (.09)	<.001	.06 (.05)	.22	-.01 (.05)	.87
Psychometric quality	3.79 (.06)	.29 (.05)	<.001	.06 (.03)	.09	.01 (.06)	.86
Skill at assessments	2.41 (.17)	.76 (.10)	<.001	.62 (.10)	<.001	-.05 (.08)	.56
	Pre-training estimate (SE)	T1–T3 slope (SE)	<i>p</i>			T3–T4 slope (SE)	<i>p</i>
Standardized assessments total use (CAPE)	2.03 (.11)	.27 (.04)	<.001			-.17 (.08)	.05
CAPE item 1: Percentage of intake clients administered SA in last month	2.67 (.22)	.35 (.10)	<.001			-.23 (.10)	.03
CAPE item 2: Percentage of caseload administered SA in last week	1.83 (.13)	.14 (.04)	<.001			.15 (.12)	.21
CAPE item 3: Percentage of clients given feedback from SA during last week	2.17 (.11)	.20 (.05)	<.001			-.07 (.17)	.69
CAPE item 4: Percentage of clients changed treatment plan based on SA scores in last week	1.47 (.08)	.38 (.06)	<.001			-.56 (.10)	<.001

Standardized assessment use (CAPE) was only administered at T1, T3, and T4; therefore, we do not include a slope for T1–T2. Fixed effects include robust standard errors

T1 time 1, pre-training, *T2* time 2, post-training, *T3* time 3, post-consultation, *T4* time 4, 3-month follow-up, *SE* standard error

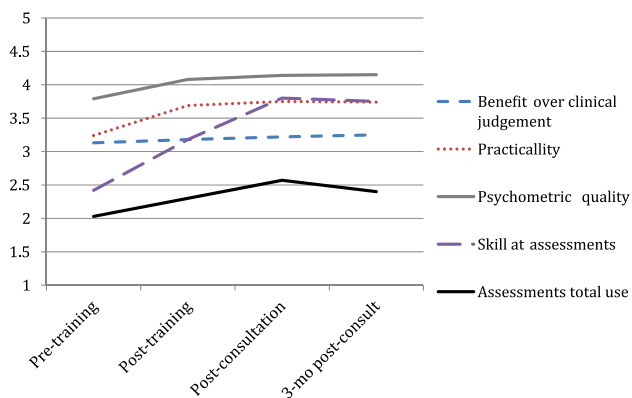


Fig. 1 Mixed-model longitudinal point estimates of self-reported ratings of the value of assessments, skill at using assessments, and use of assessments

post-consultation change. The rates of change on *benefit over clinical judgment* were not statistically significant.

SA Skill

Skill at SA significantly increased through both training and consultation ($T1-T2 \beta = .76, p < .001$; $T2-T3 \beta = .62, p < .001$), and leveled off after consultation.

SA Use

SA total use significantly improved from pre-training to post-consultation ($T1-T3 \beta = .27, p < .001$), and when consultation ended, it demonstrated a small, but significant, decrease ($T3-T4 \beta = -.17, p = .05$).

We also examined individual items from the CAPE to determine whether specific assessment items were driving the observed change. Two of the four items increased from pre-training to post-consultation, and then plateaued from post-consultation to follow-up. These included therapist report of the percentage of clients who received an assessment in the prior week ($T1-T3 \beta = .14, p = .049$) and percentage of clients who were given feedback based on SA in the prior week ($T1-T3 \beta = .20, p = .007$). Therapist report for two other items significantly increased from pre-training to post-consultation and then significantly decreased to near pre-training levels at follow-up: percentage of clients administered an assessment at intake in the prior month ($T1-T3 \beta = .35, p < .001$; $T3-T4 \beta = -.23, p = .03$), and percentage of clients who had their treatment plan changed based on SA scores in the prior week, ($T1-T3 \beta = .38, p < .001$; $T3-T4 \beta = -.56, p < .001$).

Table 4 Final longitudinal multilevel models predicting assessment use from demographics and attitudes towards standardized assessments

Fixed effects		Coefficient	SE	<i>t</i> -ratio	<i>df</i>	<i>p</i>
Intercept						
Intercept		2.07	.10	21.81	23	<.001
Pre-training skill at standardized assessments		.27	.05	5.80	41	<.001
Pre-training to post-consultation						
Slope		.27	.04	6.97	171	<.001
Post-consultation to 3-month follow-up						
Slope		-.18	.11	-1.70	174	.092
Pre-training psychometric quality		-.63	.24	-2.67	174	.008
Random effects	<i>SD</i>	Variance component		<i>df</i>	χ^2	<i>p</i>
Intercept	.43	.18		42	115.83	<.001
Level-1 e	.50	.25				
Level 3 intercept	.13	.02		23	31.90	.102

Fixed effects include robust standard errors. All non-time trend variables are grand centered

Exploratory Analyses

Exploratory analyses examined the relations between pre-training variables and assessment use across time points. Individual-predictor longitudinal models were fit to identify which variables were most promising for inclusion in a complete model. A full model was run using those variables with statistically significant *t*-ratios at $p < .05$. For the individual-predictor longitudinal models, a series of multilevel models tested the main effects of an array of pre-training variables on the intercept (i.e., score on SA total use at pre-training), the slope from T1 to T3, and the slope from T3 to T4. No slope was tested for T2 because SA use was not collected immediately post-training. Fully random effects models failed to converge due to limited sample size, therefore, only the intercept term was permitted to randomly vary. No significant relationships were found between intercept, either slope, and any of the following variables: sex, age, primary role (therapist or supervisor), having a Master's degree in Social Work, providing supervision in EBP, the use of CBT, or baseline ratings of *benefit over clinical judgment*. A final combined model was constructed using the remaining variables that were significant at $p < .10$, in line with parsimonious model-building guidelines (Tabachnick and Fidell 2007). All non-time trend and non-dichotomous variables were grand mean centered to aid interpretation.

Table 4 depicts the final model. Each estimated one-point increase in pre-training assessment skill was associated with a .27 increase in SA use pre-training. Participants increased their use of SA by .27 per time point on average from pre-training to post-consultation. With borderline significance ($p = .09$), there was a trend toward decreased use of assessments by .18 from post-consultation to 3-month follow-up.

However, participants who rated psychometric quality as important for assessment at baseline were more likely to experience decreases in their self-reported use of assessments after consultation ended. For every point higher that participants had rated importance of pre-training psychometric quality, they decreased their use of assessments from post-consultation to 6-month follow-up by an additional .63 points.

Discussion

This study was conducted to evaluate how trainee use of SA tools changed following participation in a common elements psychotherapy training and consultation program which placed heavy emphasis on EBA methods and processes. Although findings indicated significant main effect increases over the four time points (pre-training, post-training, post-consultation, and 3-month post-consultation follow-up) for each of the five outcomes, assessment of baseline change predictors and piecewise evaluation of the changes across time revealed a more complex picture. Below, these findings are discussed and recommendations made around future EBA training and consultation.

Attitudes and Skill

Across all time periods, two SA attitude subscales—practicality and psychometric quality—appeared to increase immediately following training and then remained at high levels for the duration of the consultation and follow-up periods. In comparison, SA skill and SA use followed a more gradual path (SA use is discussed further below). Attitudes toward new practices are commonly referenced in implementation models as a precondition for initial or

sustained practitioner behavior change (e.g., Aarons et al. 2011). Furthermore, attitudinal changes have previously been documented following training in EBP (Borntrager et al. 2009). It is therefore somewhat intuitive to expect that attitudes may have been influenced more easily or quickly in CBT+ and might have been most responsive to the initial training, given the in-training practice and experience with SA. In contrast, clinician self-rated skill improved continuously over the course of training and consultation—the time in which clinicians received ongoing support—but then leveled off with no significant change during the follow-up period once those supports were removed. This suggests that both the training and consultation phases were important in supporting the development of provider competence surrounding the use of SA for depression, anxiety, conduct problems, and trauma.

Standardized Assessment Use

Interestingly, SA use improved over the training and consultation period in a manner similar to SA skill, but a trend just reaching significance was observed in which participating clinicians may have lost some of those gains at the 3-month post-consultation follow-up. The trajectories of clinician assessment *skill* (i.e., the ability to use a new practice), which plateaued from consultation to post-consultation, and the decrease in two assessment *behaviors* (i.e., clients administered SA at intake and treatment plan alteration on the CAPE) over the same time period, may reflect other findings that knowledge or skill improvements may not necessarily translate into long-term practice change (Fixsen et al. 2005; Joyce and Showers 2002). Instead, partial sustainment of new practices is typical (Stirman et al. 2012). Even in training programs that require trainees to be able to demonstrate a certain level of competence at an initial training post-test often report much lower levels of actual implementation (Beidas and Kendall 2010).

In the current study, however, the observed increase in overall use of SA during consultation, and then maintenance, with only a small (although significant) decrease, holds promise. Prior research in mental health and other areas (e.g., pain assessment) has documented declines in clinicians' use of SA tools at follow-up (de Rond et al. 2001; Koerner 2013; Pearsons et al. 2012). Following an assessment initiative, Close-Goedjen and Saunders (2002) documented sustained increases in clinician attitudes about SA relevance and ease of use (i.e., practicality), but actual use of SA returned to baseline levels 1 month after the removal of supports. Although it is unclear why provider use of assessments remained above baseline in CBT+, one possible explanation involves the additional supports that are

available (e.g., on-line resources, listserv, ongoing supervisor support). For instance, SA is stressed in the yearly 1-day supervisor training and is often a topic of discussion on the CBT+ supervisor listserv and on the monthly CBT+ supervisor calls. In addition, there has been a concomitant CBT+ Initiative effort to promote organizational change in support of EBP adoption (Berliner et al. 2013). In this, routine use of SA is explicitly identified as a characteristic of an "EBP organization" and a number of participating agencies have integrated SA tools into their intake and treatment planning procedures (e.g., screening for trauma, using reductions in SA scores as a treatment goal). It is encouraging that many of the observed gains in SA use were maintained at the 3-month follow up, a finding that compares favorably to the complete loss of gains over a 1-month period in the study by Close-Goedjen and Saunders (2002).

Item-level analyses revealed that the decreasing trend observed was driven by providers' reports that they were less likely to change their treatment plans based on the results of assessments and, to a lesser extent, that they were administering SA tools to fewer of their clients at intake. In contrast, more routine administration of SA tools to youth already on their caseloads and provision of assessment-driven feedback maintained their gains, exhibiting little change during the follow-up period. One explanation for the decrease in SA-driven treatment plan changes may be that, prior to their involvement in CBT+, providers had little exposure to SA tools; a finding confirmed by anecdotal consultant reports. As provider administration of SA became more common, new information may have come to light about their current clients' presenting problems, resulting in treatment plan changes. Over time, as providers became more proficient in the collection and interpretation of SA data, it likely became *less necessary* to alter treatment plans in response to assessments. Furthermore, many providers, particularly those from agencies new to CBT+, may not have been subject to an organizational expectation that they would continue to use SA tools at intake after the end of CBT+. Therefore, following the conclusion of CBT+ consultation, the observed decrease in use of SA at intake may have resulted from the removal of some accountability. This interpretation is supported by the large ICC observed at the agency level for SA administration at intake, suggesting that some sort of organizational policy or norm may have systematically influenced responses to that item. While providers appear to have continued using SA tools with youth already on their caseloads, new cases may have been less likely to receive assessments for these reasons. Future research should further explore the patterns through which different assessment-related behaviors change as a result of training.

Analyses indicated that few baseline variables were found to predict change in SA use in the final model. This

general lack of individual-level predictors of outcome in the current study is consistent with the larger training literature in which the identification of basic clinician characteristics (e.g., age, degree) that predict training outcomes has proved elusive (Herschell et al. 2010). Even though it limits the extent to which trainings may be targeted to specific groups of professionals, this finding is nonetheless positive, as it suggests that a wide variety of clinicians may respond well to training in EBA.

The result that higher clinician ratings of psychometric quality—an attitude subscale—at baseline were associated with a post-consultation drop in SA use was counterintuitive, and there is relatively little information currently available to shed light on this finding. The psychometric quality subscale is comprised of items that address the importance of reliability and validity as well as items discussing the role of SA in accurate clinical diagnosis. Notably, psychometric quality began higher than the other attitude subscales and changed little over the course of training. Training specific to the psychometrics of the measures used was also limited in the CBT+ Initiative, potentially reducing its ability to affect this variable. This is appropriate, however, as psychometric quality is arguably less relevant than benefit over clinical judgment or practicality to actual clinical practice. Because the measures provided to participating clinicians as a component of the training (see “[CBT+ Training](#)” section) had already met minimum standards of psychometric quality, there was less need to review those concepts. Related to the role of SA in diagnosis, it may be the case that providers who entered the training with positive attitudes toward the use of SA for initial diagnostic purposes were less invested in or prepared to use the tools to routinely monitor outcomes *over time*.

Implications for Training and Consultation

As studies continue to document the effectiveness of EBA tools and processes for improving client outcomes, it is likely that an increased focus on EBA implementation will follow. This study provides support for the importance of specifically attending to assessment in EBP training. Well-designed implementation approaches are essential, given the potential for negative consequences of poorly thought out or inadequately supported initiatives (Wolpert 2014). As the primary mechanisms for supporting the implementation of new practices, training and consultation will be central to these new efforts, and training-related EBA recommendations are beginning to surface. For instance, in a recent review of methods for improving community-based mental health services for youth, Garland et al. (2013) suggested prioritizing clinician training in the utility of outcome monitoring. Related to SA tools, Jensen-Doss

(2011) suggested that reviews of specific measures could determine the trainability of each tool to aid in the selection of those that can be more easily introduced, thus enhancing clinician uptake. Although this study was focused on SA training within the context of a common elements initiative, it may be that comprehensive training dedicated to EBA tools and processes could be valuable in isolation. Future research is necessary to determine the value of well-designed trainings in EBA alone to enhance usual care.

Additionally, although the current study emphasized SA tools, future clinical trainings also could extend beyond SA to include idiographic monitoring targets (Garland et al. 2013). Idiographic monitoring can be defined as the selection of quantitative variables that have been individually selected or tailored to maximize their relevance to a specific client (Haynes et al. 2009). Such variables may include frequencies of positive or negative behaviors that may or may not map onto psychiatric symptoms (e.g., self-injury, school attendance, prosocial interactions with peers) or scaled (e.g., 1–10) ratings of experiences provided at specific times of day. In trainings, targets with particular relevance to individual clients (e.g., Weisz et al. 2012) or common service settings (e.g., school mental health; Lyon et al. 2013) can both be prioritized for idiographic monitoring and evaluated alongside the results from SA of mental health problems. In other words, important emphasis should be placed on the value of measurement processes, both initially and at reoccurring intervals, in addition to specific tools. Indeed, it is possible that both clinicians and clients might resonate with the routine measurement of an individualized, tailored target in addition to a general condition (e.g., anxiety) targeted by SA.

EBA training will also be enhanced when supportive technologies, such as measurement feedback systems (MFS) (Bickman 2008) are more widely available in public mental health. These systems allow for the automated collection and display of client-level outcome data to inform clinical decision-making and provider communications. The CBT+ Initiative invested in a MFS, the “CBT+ Toolkit,” as a means of documenting competence in EBP and the use of SA. The CBT+ Initiative has also increased the requirements for receiving a certificate of completion to include entry of SA at baseline and at another point during treatment, as well as the selection of a primary treatment target (e.g., depression, behavior problems, etc.) and documentation of the elements (e.g., exposure, behavioral activation) provided in at least six treatment sessions. Freely available to CBT+ participants during and after consultation, the CBT+ Toolkit allows participating supervisors and clinicians to score a range of SA tools, track progress over time, and observe the relationship between outcomes and the delivery of specific treatment content on a dashboard. In addition, the Toolkit

includes a pathway to CBT+ provider “Rostering” for clinicians who have demonstrated (albeit via self-report) SA use, treatment target determination, and delivery of appropriate common elements for the treatment target based on review by an expert consultant.

These results show that a specific emphasis on SA during EBP training and consultation increases use, but that there may be a risk of “leveling off” or declining use of some advantageous assessment behaviors (e.g., SA at intake) once active outside consultation ends. This suggests the need for strategies at the organizational level that incentivize, support, or require baseline and ongoing SA use as part of routine practice. One approach is the incorporation of point-of-care (POC) reminder systems into MFS to prompt clinicians to engage in basic EBA practices (e.g., SA measure administration, feedback to clients, etc.). POC reminders are prompts given to practitioners to engage in specific clinical behaviors under predefined circumstances and represent a well-researched approach to promoting lasting behavioral change among healthcare professionals (Lyon et al. 2011b; Shea et al. 1996; Shojania et al. 2010). Due to their proven ability to effect concrete and specific practitioner behavior changes, POC reminders may be well suited to the promotion of assessment use and feedback.

A second approach may be to require completion of SA as well as evidence-based interventions for service reimbursement. A number of public mental health contexts are now implementing centralized web-based data entry MFS that require clinicians to enter data on treatment content and to complete outcome assessments. The Hawaii mental health delivery system has the only state-wide system that requires routine completion of a SA measure (HigamcMillan et al. 2011; Nakamura et al. 2013). A large-scale Los Angeles County EBP initiative is also requiring SA (Southam-Gerow et al. in press). Although it is not entirely possible to disentangle the impact of the interventions from the use of SA in these systems (when both are required), the projects demonstrate both the feasibility of SA administration and their associations with client-level improvements.

A third approach, to some degree already in place in the CBT+ Initiative, is to improve organizational-level supports for SA use. Current findings indicating that some assessment behaviors may vary by agency lend additional credence to this approach. It is likely that the high within-agency ICC for percentage of clients who were administered a SA at intake is due to agency policies, procedures and supports. This may also be related to the moderately high within-agency ICC for self-rated skill using assessments. As mentioned, the CBT+ Initiative includes some supervisor-level supports (e.g., yearly supervisor training, listserv, monthly supervisor call), and SA use is one of the covered topics, among many, related to supervising

clinicians in EBP. An ongoing RCT in Washington State (see Dorsey et al. 2013) is testing the use of routine symptom and fidelity monitoring during supervision against symptom and fidelity monitoring plus behavioral rehearsal of upcoming EBP elements. Outcomes from this study may yield information surrounding the role of supervisors in encouraging EBA processes, specifically SA administration and feedback (one of the behavioral rehearsals involves practicing SA feedback). Additionally, as part of CBT+ activities in 2013, an organizational coaching guide was created, one section of which focuses on institutionalizing SA at the organizational level. If organizational leaders and administrators better prioritize, support, and encourage SA, there may be less potential for leveling or decreasing SA use following the end of consultation.

Limitations

The current project had a number of limitations. First, CBT+ was not only a training in EBA tools and processes, but also included exposure to common elements of evidence-based treatments (findings from the more general components of the training can be found elsewhere; Dorsey et al., under review). Part of the uptake observed could have been due to the fact that modularized, common elements approaches are generally optimized to integrate assessment results into data-driven decision making. Second, only a subset of the CBT+ trainees participated in the current study. Although our analyses indicate that there were few baseline differences between participants and nonparticipants, the findings should be generalized with caution. Third, all dependent variables were based on clinician self report. Fourth, although provider changes in attitudes, skill, and behavior were consistent with what would be expected in response to training, there was no control group for the purpose of drawing stronger causal inferences. Finally, the effects of training and consultation implementation are confounded by possible time effects.

Conclusion

Our findings indicate that providing training and consultation in an approach to modularized, common elements psychotherapy that emphasizes EBA tools and processes may positively impact clinicians’ self-reported attitudes, skill, and use of SA, and that changes persist over time. Findings also indicate that EBA implementation may need to emphasize methods of sustaining practice through supervision or other means. Focusing on EBA implementation while many important questions remain about evidence-based treatments and their implementation has been

likened to “taking on a 900-pound gorilla while still wrestling with a very large alligator” (Mash and Hunsley 2005, p. 375), but we believe it is a worthwhile fight. As data continue to accrue to support EBA as a bona fide EBP, training and consultation in EBA tools and processes will likely continue to be an area of research that most explicitly satisfies the two primary foci of implementation science, using EBP and getting positive outcomes.

Acknowledgments This publication was made possible in part by funding from the Washington State Department of Social and Health Services, Division of Behavioral Health Recovery and from Grant Nos. F32 MH086978, K08 MH095939, and R01 MH095749, awarded from the National Institute of Mental Health (NIMH). Drs. Lyon and Dorsey are investigators with the Implementation Research Institute (IRI), at the George Warren Brown School of Social Work, Washington University in St. Louis; through an award from the National Institute of Mental Health (R25 MH080916) and the Department of Veterans Affairs, Health Services Research & Development Service, Quality Enhancement Research Initiative (QUERI).

References

- Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research*, 38, 4–23.
- Angold, A., & Costello, E. J. (1987). *Mood and feelings questionnaire (MFQ)*. Durham, NC: Developmental Epidemiology Program, Duke University.
- Barrington, J., Prior, M., Richardson, M., & Allen, K. (2005). Effectiveness of CBT versus standard treatment for childhood anxiety disorders in a community clinic setting. *Behaviour Change*, 22, 29–43.
- Beidas, R. S., & Kendall, P. C. (2010). Training providers in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, 17, 1–30.
- Berliner, L., Dorsey, S., Merchant, L., Jungbluth, N., Sedlar, G. (2013). *Practical Guide for EBP Implementation in Public Health*. Washington State Division of Behavior Health and Recovery, Harborview Center for Sexual Assault and Traumatic Stress and University of Washington, School of Medicine, Public Behavior Health and Justice Policy.
- Bickman, L. (2008). A measurement feedback system (MFS) is necessary to improve mental health outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47, 1114–1119.
- Bickman, L., Kelley, S. D., Breda, C., de Andrade, A. R., & Riemer, M. (2011). Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services*, 62, 1423–1429.
- Birmaher, B., Khetarpal, S., Brent, D., & Cully, M. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(4), 545–553.
- Borntrager, C. F., Chorpita, B. F., Higa-McMillan, C., & Weisz, J. R. (2009). Provider attitudes toward evidence-based practices: Are the concerns with the evidence or with the manuals. *Psychiatric Services*, 60, 677–681.
- Brody, B. B., Cuffel, B., McColloch, J., Tani, S., Maruish, M., & Unützer, J. (2005). The acceptability and effectiveness of patient-reported assessments and feedback in a managed behavioral healthcare setting. *The American Journal of Managed Care*, 11(12), 774–780.
- Carlier, I. V., Meuldijk, D., Van Vleit, I. M., Van Fenema, E., Van der Wee, N. J., & Zitman, F. G. (2012). Routine outcome monitoring and feedback on physical or mental health status: Evidence and theory. *Journal of Evaluating in Clinical Practice*, 18(1), 104–110.
- Chorpita, B. F., Bernstein, A., Daleiden, E. L., & The Research Network on Youth Mental Health. (2008). Driving with roadmaps and dashboards: Using information resources to structure the decision models in service organizations. *Administration and Policy in Mental Health and Mental Health Services Research*, 35, 114–123.
- Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005a). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research*, 7, 5–20.
- Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005b). Modularity in the design and application of therapeutic interventions. *Applied and Preventive Psychology*, 11(3), 141–156.
- Clark, D. M., Layard, R., Smithies, R., Richard, D. A., Suckling, R., & Wright, B. (2009). Improving access to psychological therapy: Initial evaluation of two UK demonstration sites. *Behaviour Research and Therapy*, 47, 910–920.
- Close-Goedjen, J., & Saunders, S. M. (2002). The effect of technical support on clinician attitudes toward an outcome assessment instrument. *Journal of Behavioral Health Services and Research*, 29, 99–108.
- Cohen, J. A., Mannarino, A. P., & Deblinger, E. (2006). *Treating trauma and traumatic grief in children and adolescents*. New York: Guilford Press.
- Daleiden, E., & Chorpita, B. F. (2005). From data to wisdom: Quality improvement strategies supporting large-scale implementation of evidence based services. *Child and Adolescent Psychiatric Clinics of North America*, 14, 329–349.
- de Rond, M., de Wit, R., & van Dam, F. (2001). The implementation of a pain monitoring programme for nurses in daily clinical practice: Results of a follow-up study in five hospitals. *Journal of Advanced Nursing*, 35, 590–598.
- Dorsey, S., Berliner, L., Lyon, A. R., Pullmann, M., & Murray, L. K. (under review). A state-wide common elements initiative for children’s mental health.
- Dorsey, S., Pullmann, M. D., Deblinger, E., Berliner, L., Kerns, S. E., Thompson, K., et al. (2013). Improving practice in community-based settings: a randomized trial of supervision–study protocol. *Implementation Science*, 8, 89.
- Ell, K., Unützer, J., Aranda, M., Sanchez, K., & Lee, P. J. (2006). Routine PHQ-9 depression screening in home health care: Depression prevalence, clinical treatment characteristics, and screening implementation. *Home Health Care Services Quarterly*, 24(4), 1–19.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication No. 231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.
- Foa, E. B., Johnson, K. M., Feeny, N. C., & Treadwell, K. H. (2001). The Child PTSD Symptom Scale: A preliminary examination of its psychometric properties. *Journal of Clinical Child Psychology*, 30(3), 376–384.
- Gardner, W., Murphy, M., Childs, G., Kelleher, K., Pagano, M., Jellinek, M., et al. (1999). The PSC-17: A brief pediatric symptom checklist with psychosocial problem subscales. A

- report from PROS and ASPN. *Ambulatory Child Health*, 5, 225–236.
- Garland, A. F., Haine-Schlagel, R., Accurso, E. C., Baker-Ericzen, M. J., & Brookman-Frazee, L. (2012). Exploring the effect of therapists' treatment practices on client attendance in community-based care for children. *Psychological Services*, 9, 74–88.
- Garland, A. F., Haine-Schlagel, R., Brookman-Frazee, L., Baker-Ericzen, M., Trask, E., & Fawley-King, K. (2013). Improving community-based mental health care for children: Translating knowledge into action. *Administration and Policy in Mental Health and Mental Health Services Research*, 40, 6–22.
- Garland, A. F., Kruse, M., & Aarons, G. A. (2003). Clinicians and outcome measurement: What's the use? *Journal of Behavioral Health Services and Research*, 30, 393–405.
- Gilbody, S. M., House, A. O., & Sheldon, T. A. (2002). Psychiatrists in the UK do not use outcome measures. *The British Journal of Psychiatry*, 180, 101–103.
- Graham, I. D., & Tetroe, J. (2009). Learning from the US Department of Veterans Affairs Quality Enhancement Research Initiative: QUERI Series. *Implementation Science*, 4(1), 13.
- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., et al. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, 61(2), 155–163.
- Hatfield, D. R., & Ogles, B. M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice*, 35, 485–491.
- Hatfield, D. R., & Ogles, B. M. (2007). Why some clinicians use outcome measures and others do not. *Administration and Policy in Mental Health and Mental Health Services Research*, 34, 283–291.
- Haynes, A. B., Weiser, T. G., Berry, W. R., Lipsitz, S. R., Breizat, A. H., Dellinger, E. P., et al. (2009). A surgical safety checklist to reduce morbidity and mortality in a global population. *New England Journal of Medicine*, 360, 491–499.
- Herschell, A. D., Kolko, D. J., Baumann, B. L., & Davis, A. C. (2010). The role of therapist training in the implementation of psychosocial treatments: A review and critique with recommendations. *Clinical Psychology Review*, 30, 448–466.
- Higa-McMillan, C. K., Powell, C. K., Daleiden, E. L., & Mueller, C. W. (2011). Pursuing an evidence-based culture through contextualized feedback: Aligning youth outcomes and provider practices. *Professional Psychology: Research and Practice*, 42, 137–144.
- Huffman, L. C., Martin, J., Botcheva, L., Williams, S. E., & Dyer-Friedman, J. (2004). Practitioners' attitudes toward the use of treatment progress and outcomes data in child mental health services. *Evaluation and the Health Professions*, 27(2), 165–188.
- Hunsley, J., & Mash, E. J. (2005). Introduction to the special section on developing guidelines for the evidence-based assessment (EBA) of adult disorders. *Psychological Assessment*, 17, 251–255.
- Ivey, S. L., Scheffler, R., & Zazzali, J. L. (1998). Supply dynamics of the mental health workforce: Implications for health policy. *The Milbank Quarterly*, 76, 25–58.
- Jensen-Doss, A. (2011). Practice involves more than treatment: How can evidence-based assessment catch up to evidence-based treatment? *Clinical Psychology: Science and Practice*, 18(2), 173–177.
- Jensen-Doss, A., & Hawley, K. M. (2010). Understanding barriers to evidence-based assessment: Clinician attitudes toward standardized assessment tools. *Journal of Clinical Child and Adolescent Psychology*, 39, 885–896.
- Jensen-Doss, A., & Hawley, K. M. (2011). Understanding clinicians' diagnostic practices: Attitudes toward the utility of diagnosis and standardized diagnostic tools. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(6), 476–485.
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Koerner, K. (2013, May). *PracticeGround: An online platform to help therapists learn, implement, and measure the impact of EBPs*. Paper presented at the Seattle Implementation Research Collaborative National Conference, Seattle, WA.
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice*, 10, 288–301.
- Lyon, A. R., Borntrager, C., Nakamura, B., & Higa-McMillan, C. (2013). From distal to proximal: Routine educational data monitoring in school-based mental health. *Advances in School Mental Health Promotion*, 6(4), 263–279.
- Lyon, A. R., Charlesworth-Attie, S., Vander Stoep, A., & McCauley, E. (2011a). Modular psychotherapy for youth with internalizing problems: Implementation with therapists in school-based health centers. *School Psychology Review*, 40, 569–581.
- Lyon, A. R., & Dorsey, S. (2010). The Current Assessment Practice Evaluation. Unpublished measure, University of Washington.
- Lyon, A. R., Stirman, S. W., Kerns, S. E. U., & Bruns, E. J. (2011b). Developing the mental health workforce: Review and application of training strategies from multiple disciplines. *Administration and Policy in Mental Health and Mental Health Services Research*, 38, 238–253.
- Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology*, 34, 362–379.
- McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments. A review of current efforts. *American Psychologist*, 65, 73–84.
- Muris, P., Merckelbach, H., Korver, P., & Meesters, C. (2000). Screening for trauma in children and adolescents: The validity of the Traumatic Stress Disorder scale of the Screen for Child Anxiety Related Emotional Disorders. *Journal of Clinical Child Psychology*, 29, 406–413.
- Nakamura, B. J., Mueller, C. W., Higa-McMillan, C., Okamura, K. H., Chang, J. P., Slavin, L., et al. (2013). Engineering youth service system infrastructure: Hawaii's continued efforts at large-scale implementation through knowledge management strategies. *Journal of Clinical Child and Adolescent Psychology*. doi:10.1080/15374416.2013.812039.
- Palmiter, D. J. (2004). A survey of the assessment practices of child and adolescent clinicians. *American Journal of Orthopsychiatry*, 74, 122–128.
- Pearsons, J. B., Thomas, C., & Liu, H. (2012, November). *Can we train psychotherapists to monitor their clients' progress at every therapy session? Initial findings*. Paper presented at the Association for Behavioral and Cognitive Therapies Annual Convention, National Harbor, MD.
- Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment*, 22(2), 203–212.
- Robiner, W. N. (2006). The mental health professions: Workforce supply and demand, issues, and challenges. *Clinical Psychology Review*, 26(5), 600–625.
- Sburlati, E. S., Schniering, C. A., Lyneham, H. J., & Rapee, R. M. (2011). A model of therapist competencies for the empirically supported cognitive behavioral treatment of child and adolescent anxiety and depressive disorders. *Clinical Child and Family Psychology Review*, 14(1), 89–109.

- Shea, S., DuMouchel, W., & Bahamonde, L. (1996). A meta-analysis of 16 randomized controlled trials to evaluate computer-based clinical reminder systems for preventive care in the ambulatory setting. *Journal of the American Medical Informatics Association*, 3, 399–409.
- Shojania, K. G., Jennings, A., Mayhew, A., Ramsay, C., Eccles, M., & Grimshaw, J. (2010). Effect of point-of-care computer reminders on physician behavior: A systematic review. *Canadian Medical Association Journal*, 182, E216–E225.
- Southam-Gerow, M. A., Daleiden, E. L., Chorpita, B. F., Bae, C., Mitchell, C., Faye, M., et al. (in press). MAPping Los Angeles County: Taking an evidence-informed model of mental health care to scale. *Journal of Clinical Child and Adolescent Psychology*. doi:10.1080/15374416.2013.833098.
- Southam-Gerow, M. A., Weisz, J. R., Chu, B. C., McLeod, B. D., Gordis, E. B., & Connor-Smith, J. K. (2010). Does CBT for youth anxiety outperform usual care in community clinics? An initial effectiveness test. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49, 1043–1052.
- Spielmanns, G., Gatlin, E., & McFall, E. (2010). The efficacy of evidence-based psychotherapies versus usual care for youths: Controlling confounds in a meta-reanalysis. *Psychotherapy Research*, 20, 234–246.
- Stirman, S. W., Kimberly, J., Cook, N., Calloway, A., Castro, F., & Charns, M. (2012). The sustainability of new programs and innovations: A review of the empirical literature and recommendations for future research. *Implementation Science*, 7(17), 1–19.
- Substance Abuse and Mental Health Services Administration. (2012, January). Partners for Change Outcome Management System (PCOMS): International Center for Clinical Excellence. Retrieved from the National Registry of Evidence-based Programs and Practices Web site, <http://www.nrepp.samhsa.gov/ViewIntervention.aspx?id=249>. Accessed 27 Feb 2014.
- Tabachnick, B. C., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education, Inc.
- Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., et al. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archives of General Psychiatry*, 69, 274–282.
- Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *American Psychologist*, 61, 671–689.
- Weisz, J. R., Kuppens, S., Eckshtain, D., Ugueto, A. M., Hawley, K. M., & Jensen-Doss, A. (2013). Performance of evidence-based youth psychotherapies compared with usual clinical care: A multilevel meta-analysis. *JAMA Psychiatry*, 70(7), 750–761.
- Weisz, J. R., Southam-Gerow, M. A., Gordis, E. B., Connor-Smith, J. K., Chu, B. C., Langer, D. A., et al. (2009). Cognitive-behavioral therapy versus usual clinical care for youth depression: An initial test of transportability to community clinics and clinicians. *Journal of Consulting and Clinical Psychology*, 77, 383–396.
- Wolpert, M. (2014). Uses and abuses of patient reported outcome measures (PROMs): Potential iatrogenic impact of PROMs implementation and how it can be mitigated. *Administration and Policy in Mental Health and Mental Health Services Research*, 41(2), 141–145.
- Youngstrom, E. A. (2008). Evidence based assessment is not evidence based medicine: Commentary on evidence-based assessment of cognitive functioning in pediatric psychology. *Journal of Pediatric Psychology*, 33, 1015–1020.