



Decision-Making in Mental Health Care: Measuring Provider and Supervisor Use of Evidence

Alayna L. Park¹ · Kimberly D. Becker² · Maya M. Boustani³ · Bruce F. Chorpita¹

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Use of evidence to inform clinical decisions has been shown to improve the quality and effectiveness of services. This study piloted an observational coding system for understanding providers and supervisors' use of evidence in their clinical decision-making. The Action Cycle and Use of Evidence Behavioral Observation Coding System (ACE-BOCS) is based on Graham et al. (Contin Educ Health Prof 26:13–24, 2006) conceptual framework for knowledge management, which articulates a sequence relevant to integrating evidence into decisions and actions, including identifying and selecting a problem and choosing, planning, and rehearsing a solution or action. Using the ACE-BOCS, two coders rated the extensiveness with which evidence was used to inform decisions made in clinical supervision sessions. In these clinical supervision sessions, supervisor-provider dyads discussed cases ($N=30$; age range 8–19 years; 80% Latino/a or Hispanic ethnicity) that were identified as potentially being at risk for low treatment engagement in school mental health services. Results indicated that the ACE-BOCS can reliably and validly measure use of evidence and distinguish between strategic and indiscriminate use of evidence. The ACE-BOCS has value and utility for studying use of evidence, as it incorporates multiple actions related to service delivery and has the potential to be adapted for other aspects of mental healthcare decision-making beyond clinical supervision, as well as decision making within fields outside of mental health.

Keywords Use of evidence · Measurement · Supervision · School mental health · Treatment engagement

The past 25 years have witnessed an increase in the role of science to guide practice (Chorpita and Regan 2009; Flaspohler et al. 2012; Shlonsky et al. 2011; Tseng and Nutley 2014). Findings from multiple service sectors, including education (e.g., Ronnes and Hoagwood 2000; Slavin et al. 2009), social welfare (e.g., Dawson and Berry 2002; Gershater-Molko et al. 2003), health (e.g., Darmstadt et al. 2005; Vale et al. 2002), and mental health (Weisz et al. 2006; 2013), have demonstrated consistently that some programs, interventions, or solutions are more effective than others. In the area of mental health care, this accumulation of positive findings has spurred policy that has increasingly prioritized mental health service delivery backed by evidence (e.g.,

Chambless and Hollon 1998; National Advisory Mental Health Council Workgroup on Child and Adolescent Mental Health Invention and Deployment 2001).

Despite these efforts, research evidence tends to be underutilized in usual care contexts (Gyani et al. 2014; Weiss et al. 2008). Attempts to bridge this science-practice gap have adopted a diverse set of strategies. One especially dominant paradigm has involved the implementation of evidence-based treatments (EBTs), whose ultimate strategy is to reproduce in a service context those therapeutic procedures that were demonstrated to be successful in research trials. This aim is typically pursued through fully-developed, manualized specification of those procedures, with implementation efforts ensuring that they are performed with fidelity to the manualized program, through various combinations of training, consultation, and practice measurement (Fixsen et al. 2005; Greenhalgh et al. 2004; Real and Poole 2005). For example, efforts to implement an EBT for improving child anxiety might involve: (a) developing a treatment manual that delineates a specific set of therapeutic procedures to be delivered in a specific sequence (e.g.,

✉ Bruce F. Chorpita
chorpita@ucla.edu

¹ Department of Psychology, University of California, Los Angeles, Box 951563, Los Angeles, CA 90095, USA

² University of South Carolina, Columbia, SC, USA

³ Loma Linda University, Loma Linda, CA, USA

assess readiness for change in the first session, provide psychoeducation about anxiety in the second session, introduce exposures for anxiety in the third session, and so forth); (b) testing the efficacy of the treatment manual in a randomized clinical trial (e.g., comparing providers' delivery of the treatment manual for child anxiety with usual care); (c) [if found to be efficacious] training providers to deliver the treatment manual (e.g., holding workshops where treatment experts instruct providers on how and when to deliver the specified therapeutic procedures with their clients); and (d) monitoring fidelity of the delivery of the treatment manual through consultation with treatment experts (e.g., assessing providers' fidelity to the treatment manual and having treatment experts offer corrective feedback on the providers' adherence to and competency delivering the treatment manual).

A related, long-standing paradigm seeking to connect science and practice involves the "use of evidence" (UE; Leslie et al. 2014; McDonnell and Weatherford 2013; Ness 2010; Weiss 1977), which, although a complex construct with multiple definitions, is commonly defined as the application of knowledge gained from research studies to an imminent decision or set of decisions (Gray et al. 2014; Weiss 1979). Relative to strategies aiming to integrate science and practice through EBT implementation, the UE perspective is less concerned with the procedural specifications for applying findings of science and is more concerned with the consideration of those scientific findings to inform decision-making. Notably, opportunities to use evidence within mental health services extend far beyond selecting and delivering an EBT. These opportunities include behaviors that come before the intervention (e.g., using evidence to determine what the problem is and whether it requires treatment) as well as after (e.g., using evidence to determine how well the intervention is working and what to do if progress is not occurring). Moreover, UE is important not only in clinical practice, but also to treatment planning, clinical supervision, and mental health administration. For example, efforts to increase UE in mental health service delivery might involve: referencing the literature to identify therapeutic procedures supported by research evidence (e.g., distilling common practice elements from EBTs for child anxiety; Chorpita and Daleiden 2009); using research evidence to inform the sequence in which the therapeutic procedures should be delivered (e.g., determining that psychoeducation for anxiety should be delivered before starting exposures, given the scientific literature on the role of psychoeducation in promoting treatment engagement); evaluating the success of the therapeutic procedures by measuring clinical outcomes following implementation (e.g., monitoring progress through routine administration of a child anxiety measure) and comparing observed clinical outcomes with clinical outcomes that would be expected based on research evidence (e.g., comparing mid-treatment scores on a child anxiety measure

with mid-treatment benchmarks for that measure), among other activities. Because the UE perspective focuses on the general application of evidence or research findings, it easily extends beyond a direct service realm, such that it is possible to examine the degree to which evidence is used in management, policy meetings, or workforce development, to name but a few example contexts (Florin et al. 2012; Palinkas et al. 2016; Triantafillou 2015). Across contexts, although UE is not as common as social scientists might hope, studies have elucidated facilitators and barriers to UE (e.g., mandates requiring research to inform decisions have been shown to increase UE, while inconclusive research findings and questionable external validity have been shown to inhibit UE; Weiss and Weiss 1981; Weiss et al. 2008); thus, more recent initiatives have focused on strategies to increase UE in various contexts.

Measurement

Given the increasing emphasis for science to inform services, it is important to have strategies at hand to measure the degree to which science or research is used. To date, service providers, researchers, and policymakers have heavily invested in developing and testing accountability tools related to EBT implementation. Measurement of UE as it relates to EBT implementation typically involves measuring practice fidelity, which is defined as the extent to which services were implemented as intended (Forgatch et al. 2005; Perepletchikova and Kazdin 2005; Perepletchikova et al. 2009) and emphasizes adherence to predetermined, specific, event-by-event procedures (e.g., Did a provider competently implement Practice A as the first event, according to its specified steps?). There are now dozens of self-report, provider-report, and observational fidelity measurement instruments (e.g., McLeod and Weisz 2010), such that many EBTs have their own unique measurement approach to assess the integrity of program implementation (Schoenwald et al. 2011).

Strategies to measure UE more generally (i.e., beyond the application of specific EBTs) are more limited than those strategies for evaluating EBT implementation (Gotimer and Crouse 2019). Recent efforts have in fact been made to measure UE within the context of mental health services, with promising results. For example, Palinkas et al. (2016) measured how state and county mental health, child welfare, and juvenile justice system leaders acquire, evaluate, and apply evidence through the Standard Interview for Evidence Use (SIEU), a 60-item, self-report measure that was developed based on a review of the literature on UE and interview and focus groups with mental health department directors, child welfare directors, and probation officers. As another example, Wulczyn et al. (2015) assessed how child welfare agency staff used evidence to articulate

the presence and causes of a perceived problem, identify interventions for solving the problem, and assess the effectiveness of the implemented solution. In addition, Brennan et al. (2017) evaluated how health policymakers engage with and use research through Seeking, Engaging with and Evaluating Research (SEER), a self-report measure with 15 items informed by existing measures of organizational-level intent to use research and by consultation with researchers, policymakers, and knowledge exchange specialists. Notably, existing strategies for measuring UE typically rely on self-report of the provider or another professional, and there is currently no general method for measuring UE from direct observation of practice, planning, or clinical supervision. Given the potential for UE measurement to extend beyond the application of specific EBTs programs, to be backwards compatible (i.e., to be compatible with older versions of an innovation) with EBT fidelity measures (i.e., the UE metric could serve as a proxy for many standalone fidelity measurement systems), and to generalize to contexts beyond direct service, further attempts to develop UE measurement strategies appear warranted.

To address this need, we therefore set out develop a generalized system that (a) could measure use of evidence through direct observation, (b) could span direct service activities as well as other activities such as supervision or management, and (c) was not specific to a particular treatment or program, but rather referenced the relevant research literature more generally. Graham et al. (2006) have written extensively about the application of knowledge in the context of the information management literature, and their model offered a highly relevant conceptual framework that guided our measurement development process.

One particular advantage to Graham's model relative to many existing measurement strategies is that it positions UE or the use of knowledge (i.e., in this context, information based on research evidence that is relevant to decision making; cf. Rowley 2007) within a stream of behaviors that Graham et al. (2006) refer to as an *action cycle*. Graham's action cycle involves: (a) identifying a problem that needs addressing (e.g., identifying a client's presenting problem); (b) identifying, reviewing, and selecting evidence relevant to the problem (e.g., administering standardized measures for assessing the presenting problem); (c) adapting the evidence to the local context (e.g., determining appropriate informants for completing standardized measures); (d) assessing barriers to using the evidence (e.g., considering provider competencies); (e) selecting, tailoring, and implementing interventions to promote the use of evidence (e.g., selecting and delivering therapeutic procedures for treating the presenting problem); (f) monitoring use of evidence (e.g., monitoring fidelity with which therapeutic procedures are implemented); (g) evaluating the outcomes of using evidence (e.g., re-administering standardized measures and

assessing progress); and (h) sustaining ongoing use of evidence (e.g., planning for continued use or logical adaptation of therapeutic procedures). We contend that this *action cycle* context is especially important for measuring UE, because it can put a lens on the decision-making and behavioral antecedents and consequences of UE. In other words, measurement of what comes before and after, and not just UE itself, has the potential to illuminate not only when that evidence is used, but possibly also why it is used (or why not).

Although Graham's action cycle is relevant to many contexts, using this framework to study UE within the context of clinical supervision may be particularly illuminating. Specifically, a variety of clinical decisions (e.g., how to conceptualize the case, which practices to use, how to prepare for the upcoming session) are often discussed in supervision (Bailin et al. 2018; Dorsey et al. 2018), making it a prime context for elucidating how mental health providers and their supervisors use evidence to inform their decision-making. Additionally, supervisors are a critical influence on providers' knowledge, attitudes, and actions (Greenhalgh et al. 2004) and can thereby be valuable resources for promoting the UE in mental health treatment delivery (e.g., Dorsey et al. 2013; Herschell et al. 2010).

Study Aims

This study involved a preliminary evaluation of a new observational coding system for evaluating UE, incorporating constructs relevant to the UE literature (Nutley et al. 2007), as well as selected constructs from Graham et al.'s (2006) action cycle. Specifically, our design objectives were for this system to: (1) reliably assess the UE and action cycle phases related to clinical decision making; (2) validly measure the UE and action cycle phases; and (3) detect strategic (i.e., using evidence that fits the problem and context; e.g., implementing procedures supported by research for treating child anxiety with an 8-year-old presenting with anxiety) versus indiscriminate (i.e., using evidence without considering its applicability; e.g., implementing procedures supported by research for treating child anxiety with a 16-year-old presenting with anxiety or an 8-year-old presenting with depression; Park et al. 2018) UE. This study focused specifically on clinical supervision (i.e., dyadic discussion between mental health service providers and their supervisors about individual youth cases), given the potential for supervisors to enhance the UE in mental health service delivery and given that one of our aims was to develop a generalized system that could encompass a wide variety of service activities. This system was designed to be potentially useful for assessing the quality of implementation across a wide variety of services and contexts and to help promote the integration of science and service in multiple service sectors.

Method

Data were collected from a pilot study testing an intervention for improving treatment engagement in school mental health services (Becker et al. 2019). Therefore, in Graham's action cycle, the target problem was low treatment engagement, the interventions related to practices for enhancing treatment engagement, and the evidence encompassed client-level data (i.e., youth and caregiver responses on a survey about treatment engagement) as well as research evidence (i.e., materials representing clinical procedures from the evidence base designed to enhance client treatment engagement). All study procedures were approved by the institutional review board of the University of California, Los Angeles as well as by the institutional review boards of participating organizations that requested independent reviews.

Sample

Supervisor participants ($N=4$) were employed by the Los Angeles Unified School District (LAUSD) School Mental Health Clinic and Wellness Center program. All supervisors were female and averaged 43.25 ($SD=4.50$) years of age. Two supervisors identified as White or Caucasian, one supervisor identified as Spanish, Hispanic, or Latino, and one supervisor identified as Black or African American. All supervisors were licensed by the state of California and had a master's degree in social work. They had an average of 13.63 ($SD=4.71$) years of full-time clinical experience since obtaining their highest degree, and 15.00 ($SD=3.37$) years of experience working for the LAUSD.

Provider participants ($N=17$) were also employed by the LAUSD. Of these providers, 59% were Spanish, Hispanic, or Latino, 24% were White or Caucasian, 12% were Asian, and 6% were Black or African American. Providers were predominantly female (94%) and averaged 36.56 ($SD=6.83$) years of age. All providers had obtained their master's degree in social work; 35% were licensed by the state of California, and the remainder were working toward their licensure. Providers had an average of 9.35 ($SD=6.35$) years of clinical experience since earning their highest degree, and 9.36 ($SD=8.51$) years of experience working for the LAUSD.

For this study, three supervisors provided clinical supervision to four providers each, and one supervisor provided clinical supervision to five providers. Providers received clinical supervision from their usual supervisor with whom they met regularly outside of the study. Each supervisor was based out of a different school mental health clinic or wellness center in the LAUSD; providers

and their supervisors were located at the same school mental health clinic or wellness center.

Action Cycle and UE Behavioral Observation Coding System (ACE-BOCS)

The ACE-BOCS (Chorpita et al. 2018) is an observational coding system designed to evaluate the UE in service activities.

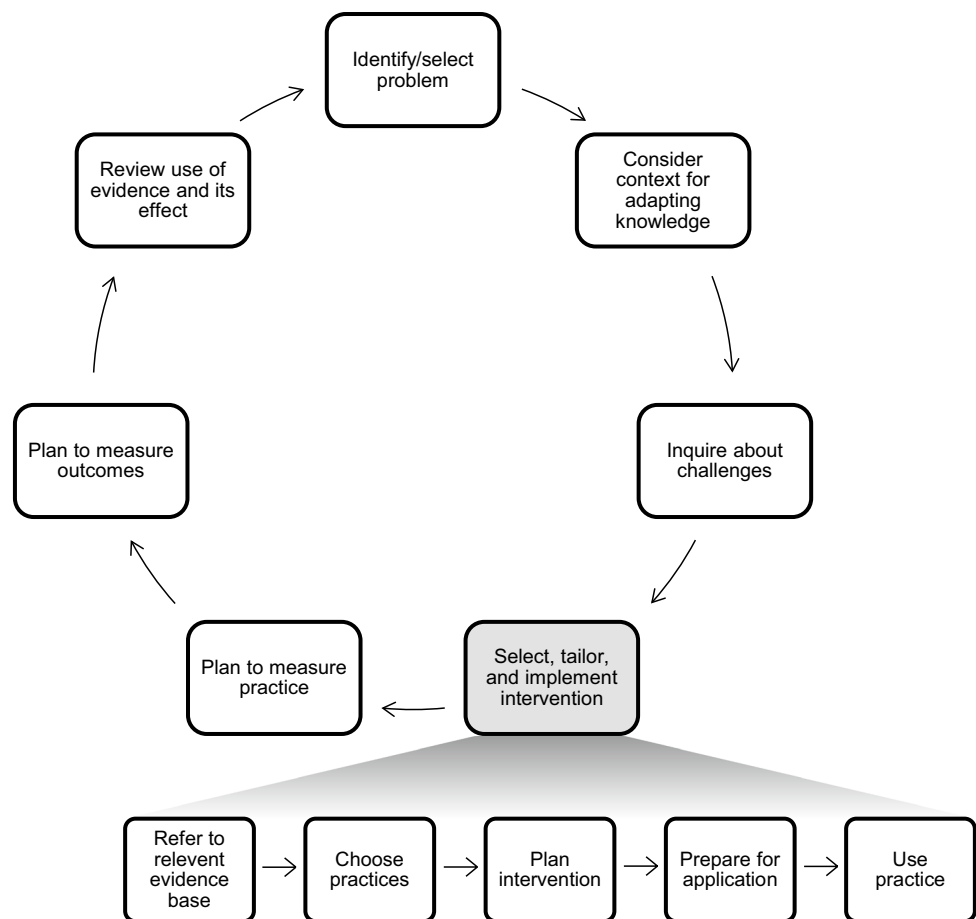
Step 1: Code Development The process of developing codes for the ACE-BOCS began with referencing Graham et al. (2006) conceptual framework for translating knowledge into action. Phases of Graham et al.'s (2006) action cycle that were relevant to UE were identified. Some action cycle phases were elaborated upon in the ACE-BOCS for increased applicability to service activities. For example, Graham et al. (2006) phase of "selecting, tailoring, and implementing an intervention" was divided into referencing the evidence base on interventions, choosing an intervention, planning the intervention, preparing for application, and using the intervention. The resulting ACE-BOCS codes were: (1) identify problem(s) that need addressing, and select problem to address; (2) consider context for adapting knowledge; (3) inquire about challenges for applying research evidence to case; (4) refer to relevant evidence base; (5) choose practice(s) for addressing selected problem(s); (6) plan intervention; (7) prepare for application by reviewing or rehearsing practice(s); (8) use practice(s); (9) plan to measure if practice was used; (10) plan to measure outcomes following practice use; and (11) revisit practice use and outcomes over time. The basic structure of the ACE-BOCS is displayed in Fig. 1.

Step 2: Scoring Strategy The ACE-BOCS scoring strategy involves extensiveness ratings designed to measure the degree to which evidence was used to inform each phase of the action cycle. Extensiveness ratings range from 0 (not present) to 5 (present with extensive UE). For example, consideration of multiple informants' scores on standardized assessment measures to identify a problem that needs addressing would be captured under the identify/select problem code with an extensiveness rating of 5. Table 1 provides the name and definition of each code, as well as anchors for low and high extensiveness ratings.

Step 3: Manual Development Next, a coding manual was developed, and one graduate student and one postdoctoral scholar independently coded 10 digital recordings of clinical supervision. Coders met weekly to review, discuss, and refine the ACE-BOCS manual. Following completion of this pilot coding phase, a final version of the ACE-BOCS manual (Chorpita et al. 2018) was produced.

This coding manual supplemented an engagement codebook that included 5 broad categories of engagement challenges (i.e., relationship, expectancy, attendance, clarity,

Fig. 1 Basic structure of the action cycle and use of evidence behavioral coding system (ACE-BOCS)



homework/participation) and 12 engagement procedures from the empirical literature that were well-suited to address these concerns (e.g., positive expectation setting, psychoeducation; Becker et al. 2018). We mention this supplemental codebook here because UE related to problem identification and intervention selection and application requires a specified target or focus, yet a detailed understanding of these supplemental codes is not required for interpretation of the ACE-BOCS. (We direct the interested reader to (Becker et al. 2019) for a more detailed description of the engagement materials).

Coding Sample

Thirty digital recordings of clinical supervision were coded using the ACE-BOCS. In each clinical supervision recording, a supervisor-provider dyad discussed a case that was identified by their provider to be potentially at risk for low treatment engagement. Thus, all of the ACE-BOCS codes were anchored to supervisors and providers' UE and actions related to detecting engagement problems and creating a plan to improve engagement. The clinical supervision session immediately following identification of these cases was recorded. Each supervisor-provider

dyad recorded treatment planning discussions about either one case ($n = 4$ dyads) or two cases ($n = 13$ dyads), for a total of 30 youth cases. Youth ranged in age from 8 to 19 years ($Mean = 13.71$, $SD = 3.06$), and were predominantly of Latino/a or Hispanic ethnicity (80%); 17% identified as White or Caucasian, and 3% identified as Black or African American. Half of the youth participants were males. Most youth reported that English was their primary language (67%); 30% of youth reported that Spanish was their primary language. Clinical supervision session recordings ranged in time from 2:36 to 25:33 ($M = 11:45$, $SD = 5:52$) minutes.

Coding Procedure

Two coders (one graduate student, one postdoctoral scholar) received training on the ACE-BOCS. Coder training consisted of reading the coding manual, reviewing segments of digital recordings, and practicing code application. Coders then independently coded two digital recordings of clinical supervision per week and met weekly to discuss any issues and to prevent coder drift, until all 30 digital recordings were double-coded.

Table 1 Action cycle and use of evidence codes

Code	Definition	Extensiveness rating (1)	Extensiveness rating (5)	ICC
Identify/select problem	The extent to which a problem was identified/selected	Identified/selected problem (or lack thereof) without referencing any supporting and objective evidence	Identified/selected problem (or lack thereof) by referencing evidence with established validity and reliability	.97
Consider context for adapting knowledge	The extent to which the value, usefulness, and appropriateness of evidence was considered for application with a particular client	Inquired about the generalizability of evidence to particular client (e.g., "I wonder if this practice is appropriate for the youth's age")	Discussed the generalizability of evidence to a particular client, and planned implementation with or without adaptation based on thorough reflection of fit between evidence and client's background and characteristics	–
Inquire about challenges	The extent to which barriers or facilitators that may influence the use of evidence are assessed	Identified barrier or facilitator to applying evidence to case (e.g., provider competencies; organizational barriers; client attitudes and experiences)	Assessed for barriers or facilitators to using evidence and made a plan to address barriers or capitalize on facilitators	.94
Refer to relevant evidence base	The extent to which evidence for facilitating the selection and implementation of an intervention was discussed	Identified source of research evidence (e.g., literature, theory) or local/client-specific evidence (e.g., provider expertise, clinic data, client data)	Discussed research evidence or local/client-specific evidence from more than one source (e.g., discussed relevant treatment protocol <i>and</i> identified provider with relevant expertise)	.92
Choose practice	The extent to which an intervention was identified	Identified intervention for addressing problem without referencing any relevant evidence	Identified an intervention using multiple sources of relevant research evidence or local/client-specific evidence (e.g., discussed recommendations outlined in treatment protocol <i>and</i> offered by another provider with pertinent expertise)	.97
Plan intervention	The extent to which an intervention was selected and planned	Selected intervention without providing any rationale	Made an evidence-informed plan to implement an intervention (e.g., determined who will be involved, the duration of the intervention, and so forth based on recommendations outlined in treatment protocol and those offered by another provider with pertinent expertise)	.96
Prepare for application	The extent to which an intervention (or its related information or materials) was reviewed or rehearsed	Identified intervention content to be reviewed or rehearsed (e.g., recommended review of treatment protocol without discussing in detail)	Role played or modeled intervention content	.92
Use practice	The extent to which identified intervention was implemented	Implemented part of identified intervention	Implemented identified intervention with high quality	–
Plan to measure practice	The extent to which a plan to measure intervention delivery was established	Discussed plan to measure if intervention was implemented without providing any details	Made detailed plan to measure if intervention was implemented and provided rationale	–
Plan to measure outcomes	The extent to which a plan to measure outcomes was established	Discussed plan to measure outcomes following implementation of intervention without providing any details	Made detailed plan to measure outcomes following implementation of intervention and provided rationale	1.00

Table 1 (continued)

Code	Definition	Extensiveness rating (1)	Extensiveness rating (5)	ICC
Review use of evidence and its effect	The extent to which intervention delivery and outcomes were reviewed	Reviewed previous measures of practice use or outcomes (e.g., reviewed trajectory of recent scores on outcome measure, without considering recent practice delivery)	Discussed relationship between practice use and outcomes (e.g., reviewed how trajectory of recent scores on outcome measure maps onto recent practice delivery)	–

ICC intraclass correlation coefficient. ICCs could not be calculated for Consider Context for Adapting Knowledge, Plan to Measure Practice, Review Use of Evidence and its Effect, and Use Practice codes because extensiveness ratings were zero across all digital recordings

Data Analyses

To determine whether the ACE-BOCS could reliably assess evidence use and action cycle phases, two-way mixed, consistency intraclass correlation coefficients (ICCs) were used to evaluate inter-rater reliability of extensiveness ratings for each code. Based on standards set by Koo and Li (2016), ICCs above .90 would be reflective of “excellent” agreement between coders.

Preliminary validity of the ACE-BOCS was evaluated through bivariate Pearson correlations examining the association between extensiveness ratings for each code. Given that the UE in one direct service task is conceptualized to be associated with the UE in another direct service task (e.g., Graham et al. 2006), it was expected that ACE-BOCS codes would have positive correlations. In addition, it was expected that these correlations would be moderate in strength, as the ACE-BOCS codes were developed to measure distinct aspects of evidence use and action cycle phases. Preliminary validity was also assessed using descriptive statistics, such that varying extensiveness ratings across codes would support the design objective of measuring distinct aspects of evidence use and action cycle phases. Specifically, given previous research indicating that mental health providers infrequently use evidence in their clinical decision-making (Gyani et al. 2014), it was expected that extensiveness ratings across codes would be low. However, it was expected that extensiveness ratings for certain codes would be higher than others. For example, extensiveness ratings for the identify/select problem code were anticipated to be higher than extensiveness ratings for the plan to measure outcomes code, as studies suggest that mental health providers occasionally use standardized assessment measures at intake (Jensen-Doss and Hawley 2011) but rarely use such measures to monitor treatment progress (Jensen-Doss et al. 2018).

Lastly, descriptive statistics were used to depict how the ACE-BOCS can differentiate strategic (i.e., using evidence that fits the problem and context) versus indiscriminate (i.e., using evidence without necessarily considering its applicability) UE. Specifically, we examined the number of cases where the selected engagement practice fit the selected engagement problem (i.e., strategic UE; e.g., a practice supported by research for improving therapy attendance was chosen for a case with attendance problems) and the number of cases where the selected engagement practice did not fit the selected engagement problem (i.e., indiscriminate UE; e.g., a practice supported by research for improving therapy attendance was chosen for a case with regular attendance but inconsistent homework completion or for a case with no engagement problems).

Table 2 Correlations between extensiveness ratings for action cycle and use of evidence codes

	Action cycle codes			Use of evidence codes			
	Identify/ select prob- lem	Inquire about chal- lenges	Plan to meas- ure outcomes	Refer to evidence	Choose practice	Plan intervention	Prepare for application
Identify/select problem	1.00	.47**	.18	.68**	.76**	.47**	.46**
Inquire about challenges	–	1.00	.24	.35	.37*	.48**	.38*
Plan to measure outcomes	–	–	1.00	.00	.16	.12	-.13
Refer to evidence	–	–	–	1.00	.75**	.38*	.65**
Choose practice	–	–	–	–	1.00	.54**	.52**
Plan intervention	–	–	–	–	–	1.00	.73**
<i>n</i> Cases Code Observed	24	17	1	12	18	16	15
Code Observed <i>Mean (SD)</i>	2.75 (.63)	3.00 (1.41)	1.00	2.50 (.80)	1.89 (.72)	3.04 (1.05)	1.43 (.60)

** $p < .01$, * $p < .05$

Results

Inter-rater Reliability

As seen in Table 1, inter-rater reliability for 7 of 11 codes was in the excellent range. ICCs could not be calculated for four codes (i.e., consider context for adapting knowledge, plan to measure practice, review UE and its effect, and use practice) because extensiveness ratings for these codes were zero across all digital recordings. Of note, some of these codes (e.g., review UE and its effect; use practice) were not relevant to a single supervision discussion from the early stages of a youth's treatment, so the absence of these action cycle phases was consistent with expectations. It was also expected that some of these codes might pertain to multiple supervision discussions (e.g., reviewing UE and its effect) or to activities outside of clinical supervision (e.g., practice use would be observed in treatment sessions as opposed to supervision).

Preliminary Validity

Results revealed several moderate, positive correlations between extensiveness ratings of codes in ways that were conceptually sound, which indicate that ACE-BOCS codes measure related but distinct constructs (Table 2). For example, identifying and selecting a problem that needs addressing was moderately, positively correlated with inquiring about challenges for applying research evidence to a case ($r = .47$, $p < .01$), referring to the relevant evidence base for a case ($r = .68$, $p < .01$), planning the intervention ($r = .47$, $p < .01$), and preparing for the application of that intervention ($r = .46$, $p < .01$). In other words, each of the action cycle and UE indicators are part

of a behavioral sequence that would be expected to follow after a supervisor and provider identify and select an engagement problem to target. Most of the other codes also demonstrated multiple positive correlations with other codes. The one exception was that planning to measure outcomes following practice use was not associated with any other action cycle or UE codes, which might suggest that this behavior is not part of the regular repertoire of clinical behaviors for supervisors and providers.

Results showed that extensiveness ratings ranged from 1.00 (plan to measure outcomes) to 3.04 (plan intervention) (Table 2). The range of these extensiveness ratings are consistent with the literature on the limited UE in usual care settings (e.g., Gyani et al. 2014) and provide further support that the ACE-BOCS codes measure distinct constructs. Additionally, codes representing behaviors that other studies have found to be infrequently observed in usual care settings (e.g., plan to measure outcomes, prepare for application; Bailin et al. 2018; Jensen-Doss et al. 2018) were shown to have the lowest extensiveness ratings.

Strategic Versus Indiscriminate UE

Results showed that supervisor-provider dyads discussed a practice from the research evidence base (e.g., Psychoeducation, Goal Setting) for 18 (60%) cases, discussed an engagement problem but no practice for 6 (20%) cases, and discussed neither an engagement problem nor practice for 6 (20%) cases. Although supervisors and providers used evidence to inform their practice identification and selection for 18 cases, they only used evidence strategically for 15 of those cases (i.e., discussed an evidence-based

practice that was indicated for the selected problem). For the remaining three cases, supervisors and providers engaged in indiscriminate UE. Specifically, for two cases, supervisor-provider dyads discussed a practice that was not supported by research evidence for the selected problem¹ (e.g., providing Psychoeducation about mental health services for a client who is attending therapy infrequently; cf. Becker et al. 2018); for the last case, the supervisor-provider dyad discussed a practice for a problem that was not found to be clinically-impairing (e.g., considering Goal Setting to address problems with in-session participation for youth who is actively engaged in treatment). Such results indicate that the ACE-BOCS can differentiate between strategic and indiscriminate UE and lend insights into how service providers may be using evidence indiscriminately.

Discussion

This paper presents findings related to the development of an initial observational coding system for measuring the UE in clinical supervision. The ACE-BOCS is based on relevant constructs from the UE literature (Nutley et al. 2007) and from Graham et al.'s (2006) action cycle and was developed to be the first generalized system for measuring UE and action cycle phases in direct service and related activities using behavioral observation.

Results from our inter-rater reliability analyses indicate that the ACE-BOCS is a reliable resource for evaluating the UE and action cycle phases in clinical supervision. Such results indicate that UE can be quantified through direct observation, which extends the possible options for evaluating UE beyond self-reports of providers or other professionals (e.g., Palinkas et al. 2016). Given the advantages of observational coding (McLeod et al. 2013), this system may serve as a valuable tool for objectively assessing UE and may illuminate additional opportunities for enhancing the quality of mental health services (e.g., evaluating how providers actually use evidence to inform their clinical decisions to supplement existing measures on perceived UE). Although extensiveness ratings across action cycle and UE codes were typically low, they varied across our sample

of 30 digital recordings of clinical supervision—suggesting that the ACE-BOCS is sensitive to the content of these activities. These findings illustrate the potential for the ACE-BOCS to evaluate UE in a variety of activities related to direct service delivery. This information can contribute to existing knowledge about the implementation of evidence-based assessment and treatment procedures by elucidating how evidence is used to inform these clinical activities. Through better understanding how evidence informs service delivery, we can develop supports for encouraging the UE across many clinical decisions to enhance the quality of mental health services.

Results also showed moderate to strong positive correlations between many of the ACE-BOCS codes. For example, the extent to which supervisors and providers identified and selected a problem that needed addressing was associated with nearly all other action cycle and UE codes. These findings highlight the potential utility of the ACE-BOCS to expand current knowledge about UE. That is, based in part on Graham et al.'s (2006) action cycle, the ACE-BOCS features a set of actions related to direct service delivery, which can be studied and elucidated through use of this system. For instance, the ACE-BOCS can be used to help determine the occurrence of behaviors related to direct service delivery (e.g., How often do providers identify a problem that needs addressing? How often do providers review or rehearse an intervention before implementation?), understand patterns of behaviors (e.g., Do providers tend to make plans for measuring both practice delivery and outcomes in a single activity?), and illuminate barriers or facilitators to using evidence (e.g., How likely are providers to use an evidence-informed practice if no problem was identified?).

Additionally, results indicate that the ACE-BOCS has the potential to differentiate between strategic and indiscriminate UE. Specifically, the ACE-BOCS was designed to be paired with a supplemental codebook of problems and corresponding solutions, in order to facilitate assessment of whether evidence is being used strategically (e.g., whether a chosen solution is supported by research to address an identified problem, such as treatment engagement in the current study)—particularly in instances when the problem that needs addressing is not fixed, such as in usual care or in the delivery of dynamic or principle-based interventions. For instance, as shown in our results, providers chose an evidence-informed solution for nearly half of the cases in the sample. However, for a portion of these cases, the chosen evidence-informed solution was not indicated for the identified problem. In other words, consideration of the empirical linking of problems and solutions yielded a different determination about the UE (i.e., indiscriminate versus strategic UE, respectively). Accordingly, it is important for any supplemental codebooks being used with the ACE-BOCS to contain (1) a

¹ The ACE-BOCS is intended to measure, among other actions, identification of a problem that needs addressing, identification of a practice, and match between the identified problem and practice as well as several other “use of evidence” behaviors. Match between the identified problem and practice should be determined based on the best available evidence (e.g., randomized clinical trials, theory, provider expertise, etc.). However, mismatch between the identified problem and practice does not necessarily mean that the identified practice is not effective for treating the identified problem—only that existing evidence does not support that practice for addressing that problem.

list of problems, (2) a list of solutions, and (3) empirically linking of problems and solutions (e.g., research supports Solution A to address Problem A)—as a focus on only problems or only solutions may result an obscured index of service quality. The identification of evidence-informed solutions for problems of interest and empirical linking of problems and solutions may be accomplished through approaches such as the distillation and matching model, which distills common features of successful solutions (e.g., practices from EBTs) and matches those features to specific problems (e.g., the distillation and matching model indicates that exposure is a common feature of EBTs for anxiety) (Chorpita and Daleiden 2009; Chorpita et al. 2005).

By designing the ACE-BOCS to be paired with a supplemental codebook of problems and corresponding solutions—rather than embedding problem and solution codes into the coding system—the ACE-BOCS can be easily adapted for use in a variety of contexts. For example, the ACE-BOCS could be paired with a supplemental codebook of mental health problems and corresponding solutions to illuminate opportunities for encouraging UE in the delivery of psychotherapy in usual care contexts and for better integrating research and practice. For instance, the code of identify/select problem could evaluate use of evidence-based assessment techniques for various mental health problems, with use of standardized assessment measures and data from multiple informants contributing to higher extensiveness ratings. For further detail, separate extensiveness ratings could be assigned to each mental health problem under investigation. As an example, a provider might use a validated self-report measure and a structured interview to assess for depression, yielding a high extensiveness rating for identifying depressive problems, but might rely on information from the phone screen to assess for mania, yielding a low extensiveness rating for identifying manic symptoms. The ACE-BOCS could also be used in educational settings to plan curricula for the upcoming academic year, if paired with a supplemental codebook of academic problems and corresponding solutions. For instance, the code *refers to evidence* could evaluate the extent to which material for planning successful curricula was referenced, with consideration of information from multiple sources (e.g., journal articles, chapters, past school curricula) contributing to higher extensiveness ratings. Although this study and these examples represent only a snapshot of the ACE-BOCS, the potential scope of this coding system is promising. We encourage professionals in other fields to test whether the ACE-BOCS may be a feasible and reliable option for measuring UE and action cycle phases in their contexts.

Limitations

This study has several strengths, including its evaluation of a new strategy for measuring the UE that (a) adds to the presently limited measurement options for assessing UE in direct service activities (Gotimer and Crouse 2019); (b) uses direct observation, which is the gold standard for measuring behaviors (McLeod et al. 2013); and (c) is not specific to a particular treatment or program; however, some limitations should be noted. For instance, although this coding system was found to be reliable for evaluating UE using digital recordings of clinical supervision, it is currently unknown whether this coding system would demonstrate similar reliability using different sources of information (e.g., direct interactions, notes) or different sets of problems and corresponding solutions. Additionally, because four codes (i.e., considers context for adapting knowledge; plans to measure practices; reviews UE and its effect; uses practice) were never observed in our sampling of 30 supervision sessions, inter-rater reliability could not be calculated and, accordingly, the reliability of these codes remains unknown. Future research should therefore apply the ACE-BOCS to a larger sample of behaviors, such as other direct service or supervision activities, management, policy meetings, or workforce development, to not only test this coding system's ability to measure UE in other contexts but to also increase the likelihood of observing all of the ACE-BOCS codes. Another limitation is that this study's assessment of validity did not involve a comparison between ACE-BOCS extensiveness ratings and scores from other measures of UE. To date, few measures of UE have been developed, and no supplementary UE measurement strategies were employed in the current study. It would thus be worthwhile for future research to apply the ACE-BOCS alongside other strategies for measuring UE to assess the construct validity of this coding system. Additionally, the UE in clinical supervision described in this study should be interpreted with caution, as the coding sample consisted of only one to two supervision session recordings per supervisor-provider dyad, with discussions that could be as short as a few minutes in duration. It is unknown whether these supervision sessions are reflective of supervisors and providers' typical supervision; however, the purpose of this study was to evaluate the ACE-BOCS as an observational coding strategy for measuring UE rather than to describe UE in clinical supervision. Given results indicating that the ACE-BOCS may be a feasible and reliable strategy for measuring UE, future research should apply the ACE-BOCS to a larger sample of direct service activities, including clinical supervision, to explore UE in those contexts. As we await those studies, however, this preliminary investigation of the ACE-BOCS indicates that this coding system can reliably detect

UE in recorded interactions, presenting a new option for measuring and studying UE in a potentially wide variety of contexts.

Conclusion

UE holds great promise for connecting science to practice. However, few instruments for measuring UE currently exist, and all of those rely on providers or other professionals' self-report. The ACE-BOCS was designed as an observational coding system to feasibly and reliably evaluate UE across direct service activities. As discussed, such a coding system has the potential to serve as a proxy for many standalone fidelity measurement systems, yield an index of service quality for usual care and the delivery of dynamic or principle-based interventions, and encompass many service-relevant activities (e.g., direct intervention, supervision). Within the context of mental health service delivery, the ACE-BOCS has the potential to evaluate the use of evidence in clinical activities ranging from assessment to EBT delivery and sustainment and to identify opportunities for encouraging the integration of research and practice. Given the promise of this system, we hope that the ACE-BOCS will help to extend current knowledge about UE in mental health and beyond.

Compliance with Ethical Standards

Conflict of interest All Authors declare that they have no conflict of interest.

Ethical Approval Approval by the institutional review board of the University of California, Los Angeles as well as by those institutional review boards of participating service agencies that requested independent reviews was obtained before the study was conducted. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

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

References

- Bailin, A., Bearman, S. K., & Sale, R. (2018). Clinical supervision of mental health professionals serving youth: Format and microskills. *Administration and Policy in Mental Health and Mental Health Services Research*. <https://doi.org/10.1007/s10488-018-0865-y>.
- Becker, K. D., Boustani, M., Gellatly, R., & Chorpita, B. F. (2018). Forty years of engagement research in children's mental health services: Multidimensional measurement and practice elements. *Journal of Clinical Child and Adolescent Psychology*, 47(1), 1–23. <https://doi.org/10.1080/15374416.2017.1326121>.
- Becker, K. D., Park, A. L., Boustani, M. M., & Chorpita, B. F. (2019). A pilot study to examine the feasibility and acceptability of a coordinated intervention design to address treatment engagement challenges in school mental health services. *Journal of School Psychology*, 76, 78–88. <https://doi.org/10.1016/j.jsp.2019.07.013>.
- Brennan, S. E., McKenzie, J. E., Turner, T., Redman, S., Makkari, S., Williamson, A., et al. (2017). Development and validation of SEER (Seeking, Engaging with and Evaluating Research): A measure of policymakers' capacity to engage with and use research. *Health Research Policy and Systems*, 15(1), 1–19. <https://doi.org/10.1186/s12961-016-0162-8>.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66(1), 7–18.
- Chorpita, B. F., Becker, K. D., & Park, A. L. (2018). Action cycle and use of evidence behavioral observation coding system (ACE-BOCS) [Measurement instrument]. Unpublished instrument.
- Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology*, 77(3), 566–579. <https://doi.org/10.1037/a0014565>.
- Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research*, 7(1), 5–20. <https://doi.org/10.1007/s11020-005-1962-6>.
- Chorpita, B. F., & Regan, J. (2009). Dissemination of effective mental health treatment procedures: Maximizing the return on a significant investment. *Behaviour Research and Therapy*, 47(11), 990–993. <https://doi.org/10.1016/j.brat.2009.07.002>.
- Darmstadt, G. L., Bhutta, Z. A., Cousens, S., Adam, T., Walker, N., De Bernis, L., et al. (2005). Evidence-based, cost-effective interventions: How many newborn babies can we save? *The Lancet*, 365(9463), 977–988. [https://doi.org/10.1016/S0140-6736\(05\)71088-6](https://doi.org/10.1016/S0140-6736(05)71088-6).
- Dawson, K., & Berry, M. (2002). Engaging families in child welfare services: An evidence-based approach to best practice. *Child Welfare: Journal of Policy, Practice, and Program*, 81(2), 293–317.
- Dorsey, S., Kerns, S. E. U., Lucid, L., Pullmann, M. D., Harrison, J. P., Berliner, L., et al. (2018). Objective coding of content and techniques in workplace-based supervision of an EBT in public mental health. *Implementation Science*, 13(1), 19. <https://doi.org/10.1186/s13012-017-0708-3>.
- 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, 1–11. <https://doi.org/10.1186/1748-5908-8-89>.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). Implementation research: A synthesis of the literature. Tampa, FL: University of South Florida: Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network. (FMHI Publication #231).
- Flaspohler, P., Lesesne, C. A., Puddy, R. W., Smith, E., & Wandersman, A. (2012). Advances in bridging research and practice: Introduction to the second special issue on the interactive system framework for dissemination and implementation. *American Journal of Community Psychology*, 50(3–4), 271–281. <https://doi.org/10.1007/s10464-012-9545-3>.
- Florin, J., Ehrenberg, A., Wallin, L., & Gustavsson, P. (2012). Educational support for research utilization and capability beliefs regarding evidence-based practice skills: A national survey of senior nursing students. *Journal of Advanced Nursing*, 68(4), 888–897. <https://doi.org/10.1111/j.1365-2648.2011.05792.x>.
- Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent

- adherence to the Oregon model of parent management training. *Behavior Therapy*, 36(1), 3–13.
- Gershater-Molko, R. M., Lutzker, J. R., & Wesch, D. (2003). Project SafeCare: Improving health, safety, and parenting skills in families reported for, and at-risk for child maltreatment. *Journal of Family Violence*, 18(6), 377–386. <https://doi.org/10.1023/A:1026219920902>.
- Gotimer, D. H., & Crouse, K. (2019). *Studying the use of research evidence: A review of methods*. New York, NY: William T. Grant Foundation.
- Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J., Caswell, W., et al. (2006). Lost in knowledge translation: Time for a map? *Journal of Continuing Education in the Health Professions*, 26, 13–24. <https://doi.org/10.1002/chp.47>.
- Gray, M., Sharland, E., Heinsch, M., & Schubert, L. (2014). Connecting research to action: Perspectives on research utilisation. *The British Journal of Social Work*, 45(7), 1952–1967. <https://doi.org/10.1093/bjsw/bcu089>.
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *The Milbank Quarterly*, 82, 581–629. <https://doi.org/10.1111/j.0887-378X.2004.00325.x>.
- Gyani, A., Shafraan, R., & Myles, P. (2014). The gap between science and practice: How therapists make their clinical decisions. *Behavior Therapy*, 45(2), 199–211. <https://doi.org/10.1016/j.beth.2013.10.004>.
- 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. <https://doi.org/10.1016/j.cpr.2010.02.005>.
- Jensen-Doss, A., Haines, E. M. B., Smith, A. M., Lyon, A. R., Lewis, C. C., Stanick, C. F., et al. (2018). Monitoring treatment progress and providing feedback is viewed favorably but rarely used in practice. *Administration and Policy in Mental Health and Mental Health Services Research*, 45(1), 48–61. <https://doi.org/10.1007/s10488-016-0763-0>.
- 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. <https://doi.org/10.1007/s10488-011-0334-3>.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>.
- Leslie, L. K., Maciolek, S., Biebel, K., Debordes-Jackson, G., & Nicholson, J. (2014). Exploring knowledge exchange at the research–policy–practice interface in children's behavioral health services. *Administration and Policy in Mental Health and Mental Health Services Research*, 41(6), 822–834. <https://doi.org/10.1007/s10488-014-0535-7>.
- McDonnell, L. M., & Weatherford, M. S. (2013). Evidence use and the Common Core State Standards movement: From problem definition to policy adoption. *American Journal of Education*, 120(1), 1–25. <https://doi.org/10.1086/673163>.
- McLeod, B. D., Southam-Gerow, M. A., Tully, C. B., Rodríguez, A., & Smith, M. M. (2013). Making a case for treatment integrity as a psychosocial treatment quality indicator for youth mental health care. *Clinical Psychology: Science and Practice*, 20(1), 14–32. <https://doi.org/10.1111/cpsp.12020>.
- McLeod, B. D., & Weisz, J. R. (2010). The therapy process observational coding system for child psychotherapy strategies scale. *Journal of Clinical Child & Adolescent Psychology*, 39(3), 436–443. <https://doi.org/10.1080/15374411003691750>.
- National Advisory Mental Health Council Workgroup on Child and Adolescent Mental Health Intervention Development and Deployment. 2001. *Blueprint for Change: Research on Child and Adolescent Mental Health. A Report by the National Advisory Mental Health Council's Workgroup on Child and Adolescent Mental Health Intervention Development and Deployment*. Bethesda, MD: National Institutes of Health, National Institute of Mental Health.
- Ness, E. C. (2010). The role of information in the policy process: Implications for the examination of research utilization in higher education policy. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 1–49). Dordrecht: Springer.
- Nutley, S. M., Walter, I., & Davies, H. T. O. (2007). *Using evidence: How research can inform public services*. Bristol: The Policy Press.
- Palinkas, L., Garcia, A., Aarons, G., Finno-Velasquez, M., Holloway, I., Mackie, T., et al. (2016). Measuring use of research evidence: The Structured Interview for Evidence Use. *Research on Social Work Practice*, 26, 550–564. <https://doi.org/10.1177/1049731514560413>.
- Park, A. L., Tsai, K. H., Guan, K., & Chorpita, B. F. (2018). Unintended consequences of evidence-based treatment policy reform: Is implementation the goal or the strategy for higher quality care? *Administration and Policy in Mental Health and Mental Health Services Research*. <https://doi.org/10.1007/s10488-018-0853-2>.
- Perepletchikova, F., Hilt, L. M., Chereji, E., & Kazdin, A. E. (2009). Barriers to implementing treatment integrity procedures: Survey of treatment outcome researchers. *Journal of Consulting and Clinical Psychology*, 77(2), 212–218. <https://doi.org/10.1037/a0015232>.
- Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*, 12(4), 365–383. <https://doi.org/10.1093/clipsy.bpi045>.
- Real, K., & Poole, M. S. (2005). Innovation implementation: Conceptualization and measurement in organizational research. In D. A. Noumair & A. B. Shani (Eds.), *Research in organizational change and development* (pp. 63–134). Bingley: Emerald Group Publishing Limited.
- Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review*, 3(4), 223–241. <https://doi.org/10.1023/A:1026425104386>.
- Rowley, J. (2007). The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of Information Science*, 33, 163–180. <https://doi.org/10.1177/0165551506070706>.
- Schoenwald, S. K., Garland, A. F., Southam-Gerow, M. A., Chorpita, B. F., & Chapman, J. E. (2011). Adherence measurement in treatments for disruptive behavior disorders: Pursuing clear vision through varied lenses. *Clinical Psychology: Science and Practice*, 18, 331–341. <https://doi.org/10.1111/j.1468-2850.2011.01264.x>.
- Shlonsky, A., Noonan, E., Littell, J. H., & Montgomery, P. (2011). The role of systematic reviews and the Campbell Collaboration in the realization of evidence-informed practice. *Clinical Social Work Journal*, 39(4), 362–368. <https://doi.org/10.1007/s10615-010-0307-0>.
- Slavin, R. E., Lake, C., & Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research*, 79(2), 839–911. <https://doi.org/10.3102/0034654308330968>.
- Triantafyllou, P. (2015). The political implications of performance management and evidence-based policymaking. *The American Review of Public Administration*, 45(2), 167–181. <https://doi.org/10.1177/0275074013483872>.
- Tseng, V., & Nutley, S. (2014). Building the infrastructure to improve the use and usefulness of research in education. In K. S. Finnigan

- & A. J. Daly (Eds.), *Using research evidence in education* (pp. 163–175). Cham: Springer.
- Vale, M. J., Jelinek, M. V., Best, J. D., & Santamaria, J. D. (2002). Coaching patients with coronary heart disease to achieve the target cholesterol: A method to bridge the gap between evidence-based medicine and the “real world”—randomized controlled trial. *Journal of Clinical Epidemiology*, 55(3), 245–252. [https://doi.org/10.1016/S0895-4356\(01\)00460-7](https://doi.org/10.1016/S0895-4356(01)00460-7).
- Weiss, C. H. (1977). *Using social research in public policy making*. Lanham: Lexington Books.
- Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, 39(5), 426–431.
- Weiss, C. H., Murphy-Graham, E., Petrosino, A., & Gandhi, A. G. (2008). The fairy godmother—And her warts: Making the dream of evidence-based policy come true. *American Journal of Evaluation*, 29(1), 29–47. <https://doi.org/10.1177/1098214007313742>.
- Weiss, J. A., & Weiss, C. H. (1981). Social scientists and decision makers look at the usefulness of mental health research. *American Psychologist*, 36(8), 837–847. <https://doi.org/10.1037/0003-066X.36.8.837>.
- Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *The American Psychologist*, 61(7), 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. <https://doi.org/10.1001/jamapsychiatry.2013.1176>.
- Wulczyn, F., Alpert, L., Monahan-Price, K., Huhr, S., Palinkas, L. A., & Pinsoneault, L. (2015). Research evidence use in the child welfare system. *Child Welfare*, 94(2), 141–165.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.