



Improving Delivery Behaviors During Exposure for Pediatric OCD: A Multiple Baseline Training Trial With Community Therapists

Kristen G. Benito

Jennifer Herren

Jennifer B. Freeman

Abbe M. Garcia

Warren Alpert Medical School of Brown University

Paul Block

William James College

Cadence Consultants

Elizabeth Cantor

Cadence Consultants

Bruce F. Chorpita

University of California, Los Angeles

Brianna Wellen

University of Utah

Elyse Stewart

State University of New York at Binghamton

Christopher Georgiadis

Florida International University

Hannah Frank

Warren Alpert Medical School of Brown University

Temple University

Address correspondence to Kristen Benito, Warren Alpert Medical School of Brown University, 222 Richmond St., Providence, RI 02903; email: Kristen_Benito@brown.edu

0005-7894/© 2020 Association for Behavioral and Cognitive Therapies.
Published by Elsevier Ltd. All rights reserved.

Jason Machan

Lifespan Biostatistics Core, Lifespan Hospitals
Warren Alpert Medical School of Brown University
University of Rhode Island

This study tested whether a new training tool, the Exposure Guide (EG), improved in-session therapist behaviors (i.e., indicators of quality) that have been associated with youth outcomes in prior clinical trials of exposure therapy. Six therapists at a community mental health agency (CMHA) provided exposure therapy for 8 youth with obsessive-compulsive disorder (OCD). Using a nonconcurrent multiple baseline design with random assignment to baseline lengths of 6 to 16 weeks, therapists received gold-standard exposure therapy training with weekly consultation (baseline phase) followed by addition of EG training and feedback (intervention phase). The primary outcome was therapist behavior during in-session exposures, observed weekly using a validated coding system. Therapist behavior was evaluated in relation to *a priori* benchmarks derived from clinical trials. Additional outcomes included training feasibility/acceptability, therapist response to case vignettes and beliefs about exposure, and independent evaluator-rated clinical outcomes. Three therapists reached behavior benchmarks only during the EG (intervention) phase. Two therapists met benchmarks during the baseline phase; one of these subsequently moved away from benchmarks but met them again after starting the EG phase. Across all therapists, the percentage of weeks meeting benchmarks was significantly higher during the EG phase (86.4%) vs. the baseline phase (53.2%). Youth participants experienced significant improvement in OCD symptoms and global illness severity from pre- to posttreatment. Results provide initial evidence that adding the EG to gold-standard training can change in-session therapist behaviors in a CMHA setting.

Keywords: exposure; training; mechanism

ALTHOUGH EFFECTIVE TRAINING is critical for evidence-based practice (EBP) implementation, practicing therapists cite lack of training and uncertainty about what constitutes “sufficient” training among barriers to using EBPs (Gunter & Whittal, 2010; Herschell et al., 2010). Multiple studies have tested the effectiveness of different approaches to therapist training (Beidas & Kendall, 2010). Available evidence suggests that “gold-standard” approaches emphasize active learning via multiple components (e.g., didactics/-

workshops, tape review) and ongoing consultation (Herschell et al., 2010). Passive learning strategies such as manual reading or workshop alone have mostly been ineffective (Herschell et al., 2010). However, while gold-standard training has consistently improved therapist knowledge and may increase the use of EBPs, it has generally not optimized treatment delivery (i.e., the manner in which treatment is provided; Beidas & Kendall, 2010). This is true even at high training “doses” (e.g., specialist consultation for over 1 year; Sholomskas et al., 2005), suggesting a need for training implementation strategies that specifically target the quality of interventions being delivered.

To optimize implementation and effectiveness outcomes, training should ensure that key aspects of delivery are clear and reproducible for learning therapists. However, conceptualization and measurement of optimal therapist delivery vary considerably across studies in both research and practice settings and may reflect one or more overlapping constructs. These include adherence (whether or not prescribed manual components are delivered), fidelity (the degree to which components are delivered as intended by treatment developers), competence (overall skill in delivery), and integrity (a blend of adherence and competence; McLeod et al., 2013; McLeod et al., 2017; Proctor et al., 2011). There is a presumed connection between these constructs and patient outcomes, yet literature does not consistently support this, and measurement is fraught with challenges (McLeod et al., 2017; Proctor et al., 2011). Moreover, it could be that optimal delivery differs from that intended by treatment developers or used in clinical trials. Ultimately, it may be most important that therapists learn aspects of delivery that contribute to better patient outcomes—a concept that has been termed *delivery quality* (McLeod et al., 2013).

Indicators of delivery quality in the form of specific therapist *behaviors*—rather than more global ratings—might have advantages for training. Behaviors offer a concrete basis for teaching what quality “looks like,” which could improve replicability and communication between trainers and learners (who might use the same technique labels but have different ideas about what they entail). It may also match the needs of practice settings,

where attention to psychotherapy process is valued (Castonguay et al., 2010). To the degree that specific behaviors are relevant for mechanisms of treatment change, targeting these during training might maximize treatment potency (Kazdin, 2009). In behavioral treatments, theorized mechanisms relate to a core learning process (e.g., reinforcement, extinction) and therapist behaviors that influence learning are best defined using function rather than topography. For example, a therapist could provide a child with a small toy (a topographically defined therapist behavior), but it may be more relevant to determine whether this functions as positive reinforcement (e.g., by evaluating whether the toy is desirable and its provision contingent upon a goal behavior). Moreover, when patient learning can be detected in practice, it could be a valuable source of proximal feedback to aid training and treatment tailoring. This idea is central to treatments such as applied behavior analysis, where within-session change in target behaviors is used as a proximal outcome (i.e., an early sign that treatment is working or needs adjustment; Hanley et al., 2003).

TRAINING AND QUALITY IN EXPOSURE THERAPY

Exposure-based cognitive behavioral therapy (CBT) is a striking example of the science-to-service delivery gap. Despite robust efficacy (Higa-McMillan et al., 2016), it has not consistently outperformed usual care in practice settings (e.g., Weisz et al., 2013). Exposure as a single treatment element is particularly underutilized (Becker-Haimes et al., 2017) and associated with the biggest research versus practice differences in adherence and competence (McLeod et al., 2017; Smith et al., 2017). When exposure is used in practice, therapists endorse techniques that differ from specialist recommendations, including more fear-decreasing or “cautious” behaviors and fewer fear-increasing or “intense” behaviors (Deacon et al., 2013; Whiteside et al., 2016). These differences are more pronounced among therapists with negative beliefs about exposure (Whiteside et al., 2016). Even after intensive training, therapists identify exposure as difficult to implement, cite concerns about its effectiveness over time, and report it as the least sustained of all CBT elements (Chu et al., 2015). This raises the possibility that exposure quality influences sustainability (i.e., if therapists try exposure, do not see improvement, and then discontinue). Together, findings suggest that difficulty reproducing exposure quality could contribute to limited effectiveness and underutilization in practice settings.

Targeting mechanism-relevant therapist behaviors may be a particularly good match for exposure training. Exposure is underpinned by a clear theory of behavioral mechanism that centers on fear extinction, and specific variants of this basic theory emphasize different patterns of fear change that might signal exposure learning (Craske et al., 2014). Several studies have examined multiple types of within-exposure fear change in the context of treatment. Results have been mixed—even across different studies using the same sample (e.g., Kircanski & Peris, 2015; and Chu et al., 2015)—suggesting that findings may be influenced by measurement. Aiming to address several measurement limitations (e.g., nonindependent reporters, inconsistent sampling frames), Benito et al. (2012) used continuous observational coding in a sample of youth receiving exposure in three clinical trials. Greater *habituation*—fear decreases occurring “on their own”—predicted better clinical outcomes (Benito et al., 2018). In contrast, fear decreases that were better explained by observed antecedents (e.g., avoidance, exposure stimulus change) did not predict outcomes, nor did other types of fear changes (fear activation, fear variability). The distinction between habituation and other kinds of fear decreases is consistent with behavioral theory, which suggests that fear-reducing antecedents maintain symptoms over time. Although habituation is commonly thought of as specific to one variant of mechanistic theory (i.e., emotional processing theory; Craske et al., 2014), when measured this way it can be conceptualized in a manner consistent with multiple variants (e.g., as reduction occurring through fear toleration, as a way to disconfirm beliefs/violate expectancy about the permanence of fear, or simply as ensuring that exposures do not end with escape). This form of habituation also has implications for therapist behavior during exposure. Behaviors that reduce short-term fear could impede habituation, whereas those that maintain or increase it could provide greater opportunity for habituation and facilitate clinical improvement. This idea matches the recommendations of exposure specialists, who generally advise that therapists use a fear-increasing or “intense” approach rather than a fear-decreasing or “cautious” approach that might interfere with exposure learning (e.g., Abramowitz et al., 2019).

Although a considerable body of translational work focuses on ways to optimize exposure task setup (e.g., ordering, number/type of stimuli, setting; Craske et al., 2014), few studies have examined therapist behaviors that occur *during*

exposure tasks and only a handful have done so in clinical treatment. Therapists' encouraging patient engagement has been linked to higher short-term fear during exposure (Chu et al., 2015). Evidence also supports a link between therapist behaviors, habituation, and clinical outcomes. Benito et al. (2020) found that a lower proportion of therapist fear decreasing behaviors and a higher proportion of fear increasing behaviors were associated with greater subsequent habituation within the same exposure task. These behaviors also predicted improved clinical outcomes via increased habituation across treatment. These findings support the relevance of specific therapist behaviors for proximal (i.e., habituation during exposure) and clinical outcomes—suggesting that these behaviors may be key quality indicators and are promising targets for therapist training.

SUMMARY AND CURRENT STUDY

Difficulty replicating high-quality delivery during the translation from clinical trial into practice may contribute to inconsistent effectiveness of EBPs, yet gold-standard training approaches have not closed this gap. Exposure is a clear example of this delivery gap and is guided by a clear theory of mechanism as well as evidence that links specific therapist behaviors to outcomes. To enhance current gold-standard training efforts, the Exposure Guide (EG) was developed as an implementation strategy that identifies specific therapist behaviors that could improve exposure quality and match the needs of a practice setting. Just as knowledge about treatment mechanism has many advantages (Kazdin, 2009), knowledge about *training* mechanisms may also have advantages that include improved training potency and fit within practice settings. The current study is part of a series that was designed together to test therapist behaviors as a mechanism—or *target*—through which a training and quality monitoring tool (the EG) might improve exposure effectiveness and implementation outcomes. Figure 1 presents the guiding conceptual model for this study series and work

completed to date. Although there are many ways to test mechanisms (see Kazdin, 2009, for a review), this series uses an experimental therapeutics framework, which emphasizes a specific set of prescribed steps. Methodological strengths of this framework include experimental manipulation of the hypothesized mechanism and replication of effects across samples. The present study aims to test *initial target engagement* (i.e., whether there is an initial signal that the EG produces a change in targeted behaviors). Evidence of initial target engagement is helpful for determining whether an intervention warrants continued study, and is most compelling when intervention-related change is observed (a) at a threshold that has relevance for clinical functioning and is set *a priori*, and (b) in the absence of concurrent change in other domains that could also be plausible mechanisms (i.e., demonstration of *target specificity*).

This study used a nonconcurrent multiple baseline design to test whether adding the EG to gold-standard training changes targeted behaviors during exposure in a community mental health agency (CMHA). Specifically, we aimed to determine whether introduction of the EG during consultation is associated with observed changes in therapist behavior (exposure quality) relative to clinical trial benchmarks, and to explore whether it is associated with changes in the level or trend of therapist behaviors (*initial target engagement*, Aim 1). We also examined changes in therapist responses to a case vignette and beliefs about exposure across study phases (*target specificity*, Aim 2). Finally, we aimed to explore youth treatment outcomes (*effectiveness*, Aim 3) and characterize *feasibility and acceptability* of EG training for therapists (Aim 4).

Methods

PROCEDURES

All procedures were approved by the Institutional Review Board. Therapist and youth participants provided informed consent (minor children pro-

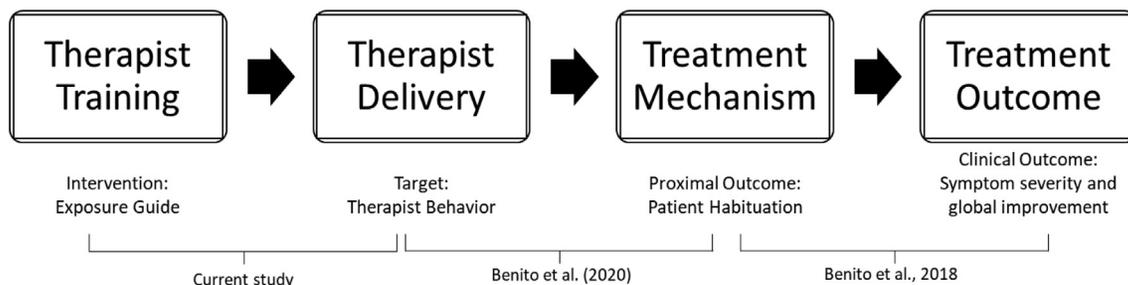


FIGURE 1. Conceptual model guiding the current study and past studies in this experimental therapeutics series.

vided assent). To avoid identifying individual participants in this small sample (particularly therapists, for whom colleagues were aware of study participation and characteristics of assigned youth), we report demographic and other relevant characteristics by group rather than individual.

EXPOSURE THERAPY

In both study phases, therapist participants provided exposure therapy for youth participants using a treatment manual designed for flexible use while retaining essential components of exposure and treatment elements that support exposure (manual is available from corresponding author upon request). As in usual practice, length of treatment was variable; therapists worked with families to determine the appropriate end for each person's treatment. The manual provided content for five core procedures: psychoeducation, using a reward system, hierarchy building, exposure, and relapse prevention. Therapists were encouraged to begin treatment with psychoeducation and hierarchy building and then to incorporate exposure during most treatment sessions. Otherwise, therapists could combine, order, or duplicate core procedures as needed throughout treatment (e.g., revisiting hierarchy development when a new treatment target has been identified, using rewards as-needed for addressing adherence).

PARTICIPANTS

Therapist participants ($n = 6$) were licensed clinicians (3 licensed clinical social workers, 2 psychologists, 1 mental health counselor); all were Caucasian (100%) and most were female (83%). Therapists were employed at a moderately sized CMHA serving approximately 750 youth per year and accepting most major insurance types (public and private). Therapists were recruited on a volunteer basis; we enrolled the first six who expressed interest in study participation (no therapists declined or were excluded, consort diagram in [Figure 1 of Supplementary Materials](#)). To ensure a sufficient number of baseline and intervention training weeks, we replaced youth participants that ended treatment before their therapists had completed both phases. Therefore, two therapists treated two youth each and four therapists treated one youth each. Therapists were compensated as usual by the agency for sessions; the study provided compensation for study activities (i.e., completing measures, training, and consultation).

Youth participants were recruited through community referrals of families seeking treatment at an anxiety specialty clinic in a hospital setting. Interested families were assessed for eligibility, then

referred to our partner CMHA where they were assigned to a participating therapist using usual agency procedures. Sessions were reimbursed using usual agency procedures (i.e., through insurer). Families were reimbursed for time completing study assessments. Inclusion criteria were (a) age 5–17, (b) primary or co-primary diagnosis of OCD, and (c) insurance accepted by CMHA; exclusion criteria were (a) concurrent psychotherapy for OCD, (b) psychosis, autism spectrum disorder, or moderate to severe intellectual disability, and/or (c) change in psychotropic medication in the previous 6 weeks. The resulting sample ($n = 8$, consort diagram in [Supplementary Materials Figure 2](#)) was 37.5% female with a mean age of 10 years ($M = 10.12$ years, $SD = 2.53$) and moderate OCD severity (Children's Yale-Brown Obsessive Compulsive Scale $M = 18.50$, $SD = 4.50$). A majority (62.5%) had at least one comorbidity (specific phobia $n = 3$; ADHD $n = 1$; generalized anxiety disorder $n = 1$; tic disorder $n = 1$). One participant was taking psychotropic medication (stable > 6 weeks and no changes during study).

STUDY DESIGN

We selected a nonconcurrent multiple baseline design for feasibility in practice, where treatment is initiated at different times. Therapists were randomly assigned to baseline lengths (6, 8, 10, 12, 14, or 16 weeks); assignment was masked for all study staff except the statistician and the consultants providing training. For logistical reasons, therapists were aware of their assigned baseline length. Therapist behavior, the primary outcome, was observed weekly on video by independent coders using a validated microanalytic coding system (see "Measures," below). Prior to randomization, therapists completed a 1-day workshop on exposure therapy for OCD. The workshop included didactics, experiential exercises, video examples, modeling, and role-plays. Workshop content focused on principles of exposure and components in the flexible treatment manual. Throughout baseline and intervention phases, therapists participated in weekly consultation provided by the first and second authors. Therapists worked with the same consultant across both phases; a protocol was used to ensure consistency across consultants and adherence to the training model during each study phase. Consultants were highly adherent to the protocol (>90%, based on consultant-rated activities during each meeting).

Baseline Phase: Standard Consultation

To ensure an adequate number of observations in the baseline phase, baseline periods started at the

first week with an exposure session (i.e., session where exposure occurred) and at least three baseline observation points were required to enter the intervention phase. During this phase, therapists participated in 1-hour weekly group consultation with a study consultant. Consultation techniques included case discussion, modeling, role-plays, and group review of session videos. During the baseline phase, consultants provided support for basic principles of exposure, and flexible use of the treatment manual, as well as feedback about exposure procedures based on video review. Consultants did not provide direct feedback about prescribed or proscribed therapist behaviors during exposure tasks.

Intervention Phase: Exposure Guide

After entering the intervention phase, therapists transitioned to a different consultation group, where training remained identical except that therapists began using the EG. The EG was developed to support quality monitoring and improvement in exposure therapy. Initial EG development included (1) drafting items based on the Exposure Process Coding System (EPCS), (2) iterative revisions to incorporate stakeholder feedback about optimal design for utility in community sites, (3) iterative pilot testing with four study raters, and (4) review by three exposure therapy researchers for theoretical fidelity. The resulting EG is a four-page written guide that takes 2–5 minutes for therapists to complete after each session. EG sections prompt therapist reflection about preparation (e.g., collaborative selection of an exposure), therapist behaviors during exposure, postprocessing with youth/family, and the exposure overall (e.g., was the exposure too hard or too easy). The EG incorporates these areas of reflection with the principles of exposure and common exposure decision points, using an interactive format. The first page uses a “skip out” procedure to determine whether an exposure took place. If it did not, therapists are asked to consider which barriers were present. If an exposure did occur, therapists proceed to checklist items for preparation, therapist behaviors (detailed in [Table 1](#)), and postexposure reflection. The final page provides brief troubleshooting related to potential problem areas. For example, if an exposure was “too hard,” suggestions are provided to address this in the next session (e.g., revisiting the hierarchy to choose an easier task, using intensifying behaviors judiciously). The organizing framework of the EG allows for scientific input (e.g., items, structure) in conjunction with real time decisions in treatment based on a therapist’s observations, which follows recommendations for coor-

dated strategic action ([Chorpita & Daleiden, 2018](#)). The EG is available from the corresponding author upon request.

At the start of the intervention phase, therapists were oriented to the EG with 1 hour of individual training. In each subsequent week, therapists completed an EG after each session (completion rate: 91.7%). Consultation included focus on EG self-evaluation, and consultants encouraged therapists to use behaviors that increase/maintain short-term fear, such as discussion of the exposure stimulus (encourage approach). They encouraged therapists to limit behaviors that might promote avoidance/distraction, such as talking about non-exposure topics (unrelated talk) or providing reassurance (accommodation). Consultants continued to use group viewing of video to provide general feedback (as in the baseline phase), but also incorporated EG feedback. Therapists completed the intervention phase once they reached *a priori* behavior benchmarks (rated by masked EPCS coders as described below) for three consecutive weeks (excluding weeks without a session, when behaviors could not occur; see “missing data”).

MEASURES

Independent Observer-Rated: Therapist Behaviors
Therapist behaviors were observed on video for each weekly treatment session throughout the study period.

Exposure Process Coding System (EPCS; Benito et al., 2012). EPCS is a microanalytic coding system measuring therapist, parent, and child behaviors, and fear levels during exposure. Variables of interest in this study were fear-increasing and fear-decreasing therapist behaviors ([Table 1](#)) and patient habituation. To measure habituation, coders judge whether each instance of observed fear decrease (on a 0–5 scale) is associated with an antecedent likely to explain the reduction (e.g., rituals, avoidance, reduction of task difficulty, unplanned change in exposure stimulus, focus on anticipation of exposure end). Decreases not better explained by such antecedents (i.e., occurring “on their own”) are more likely to signal an internal learning process and are judged as *habituation*. Given the complexity of habituation measurement considerations, we refer interested readers to [Benito et al. \(2018\)](#) for in-depth rationale and discussion. EPCS has demonstrated good inter-coder reliability and predictive validity with treatment outcome in a small sample of young children with OCD ($N = 18$; [Benito et al., 2012](#)) and in a larger sample of children and adolescents ($N = 111$; [Benito et al., 2018](#); [Benito et al., 2020](#)).

Table 1
Therapist Behaviors Included in the Exposure Guide (EG) and Exposure Process Coding System (EPCS)

Behavior	Definition
Fear Increasing^a	
Encourage Approach	Facilitates contact with the exposure, including redirection to the task (e.g., “Keep looking at the video”), discouraging avoidance/rituals (e.g., “Try to resist asking for reassurance”), requests to describe fear content (e.g., “What are the worries saying?”), discussion of the stimulus (e.g., “The trash can is out”).
Intensify	Changes the exposure task to make it more difficult (e.g., contaminant on both hands instead of one) or makes a statement likely to increase fear by providing <u>new</u> relevant information (e.g., “This carpet hasn’t been washed in forever”) or opinion (e.g., “Wow, that is a lot of vomit”).
Reduce Parent Accommodation	Discourages family accommodation during the exposure; either independent of or in response to actual occurrence of accommodation (e.g., “Your mom is going to resist answering questions”).
Fear Decreasing^a	
Accommodation	Changes the exposure task to make it less difficult (e.g., instructs decrease contact with the stimulus) or makes a statement likely to decrease fear by providing <u>new</u> relevant information (e.g., “I eat this all the time and nothing bad happened to me”) or opinion (e.g., “I think you’ll be ok”).
Unrelated Talk	Engages in talk not related to the current exposure and could serve to decrease focus on the exposure task. Content may be general (e.g., “What are you guys up to this weekend?”) or include CBT components with limited relevance for patient experience in current task (e.g., planning homework).
Encourage Relaxation	Encourages relaxation techniques, such as deep breaths, muscle relaxation, or pleasant imagery (e.g., “Remember how to control your breathing?”).
Fear Neutral^b	
Teaching	Therapist provides information about a CBT principle or symptom that relates to the current exposure
Externalizing	Therapist refers to OCD or anxiety as being separate (e.g., “You’re the boss of Mr. Worry”).
Changing Anxious Thoughts	Therapist leads the child to use a cognitive or coping strategy to manage fear during the exposure (e.g., evaluating the likelihood of a feared consequence, using coping statements).

Note. Table adapted from Benito et al. (under review).

^a Bolded items indicate inclusion in benchmarks.

^b Therapist training in current study did not focus on neutral behaviors; we have included them in this table because they appear as items on the EG.

- *Coder training and reliability.* Coders were bachelor’s-level research assistants who were unaware of therapist training phase/baseline length and clinical outcomes. Coder training included guided reading of the manual, observing trained coders, coding with supervision, and independent coding of training tapes to criterion ($K > .70$ on all variables). Trained coders attended weekly meetings to discuss EPCS implementation and prevent drift. Ten percent of tapes were double-coded for inter-coder reliability, which was excellent in the present sample ($K = .85$ to 1.00). To prevent drift, 10% of tapes were subject to ongoing review by the first author.

- *Coding procedures.* EPCS uses Noldus Observer software (Noldus Observer, Version XT 11), which links observations to video and generates a time-stamp for each variable coded. Using this time-stamp, variables of interest can be calculated for their frequency and/or duration. EPCS is applied during the exposure task itself and is not applied during other portions of an exposure ses-

sion. Session tapes were first screened to determine the presence or absence of exposures and to mark the beginning and end of each exposure. An exposure was said to have started if any of the following were observed: therapist statement that an exposure is starting, clear presentation of an exposure stimulus, and/or at least two SUDS ratings with reference to the present activity.

Therapist Self-Report

Therapists completed self-report measures at pre-training, postbaseline training, and post-EG training.

Therapist Beliefs about Exposure Scale (TBES; Deacon et al., 2013). The TBES assesses negative beliefs about exposure, including ethics, acceptability, and safety of implementing exposure. The TBES contains 21 items using a 5-point scale, with higher scores indicating more negative beliefs. It has excellent psychometric properties with internal consistency ranging from .90 to .96 and good test-retest reliability ($r = .89$) over a 6-month interval (Deacon et al., 2013).

Exposure Process Knowledge: Case Vignette (adapted from Deacon et al., 2013). This case vignette was used to assess therapist knowledge of exposure process. The measure provides a brief case description followed by various decision points during a hypothetical exposure; therapists rate whether they would use a variety of responses on a 0 (“very unlikely”) to 4 (“very likely”) scale. Response options were modified to include therapist behaviors on the EG. Mean endorsement of increasing, decreasing, and neutral behaviors was used in analyses.

Training Acceptability Rating Scale (TARS; Milne & Noone, 1996). The TARS was used to evaluate training acceptability. The first six items (TARS-1) assess general acceptability, perceived effectiveness, negative side effects, appropriateness, and social validity using a 6-point scale (“strongly disagree” to “strongly agree”). The second portion (TARS-2) assesses the teaching process and whether it was helpful on a 4-point scale (“not at all” to “a great deal”).

Independent Evaluator (IE) Rated: Clinical Symptoms

Youth participants completed clinical assessments with IEs at pretreatment (all measures) and post-treatment (CY-BOCS and CGI only; see below). IEs were not aware of therapist training phase/baseline length or in-session behaviors and were otherwise not involved with this study. They were psychology residents or licensed psychologists who were trained to criterion on all measures ($K > .80$).

Mini International Neuropsychiatric Interview (MINI-KID; Sheehan et al., 2010). The MINI-KID is a widely used structured diagnostic interview that assesses for DSM-IV and ICD-10 psychiatric diagnoses in children and adolescents ages 6 to 17 years and has demonstrated adequate reliability and validity. The MINI was used to establish inclusionary diagnostic criteria.

Children’s Yale-Brown Obsessive Compulsive Scale (CY-BOCS; Scahill et al., 1997). The CY-BOCS is a semistructured interview that includes a symptom checklist and 10 items that assess symptom severity. Psychometric properties of the subscales and total score are excellent.

Clinical Global Impressions (CGI) Scale: Improvement and Severity (CGI-I and CGI-S; Guy et al., 1976). This scale assesses overall clinical impressions of the youth’s current presentation based on symptoms observed and impairment reported using a 7-point scale, with higher scores indicating worse severity/improvement. Although a transdiagnostic measure, the CGI-I and CGI-S

have been used successfully in pediatric OCD (Storch et al., 2010).

Clinical Trial Benchmarks

Prior to setting benchmarks, we confirmed the relevance of each therapist behavior for youth outcomes in a prior study, which analyzed 459 EPCS-coded exposures for 111 youth who received CBT in one of three clinical trials for pediatric OCD (POTS trials; Benito et al., 2020). Twenty-six therapists delivered exposures in that sample; each was trained on the treatment manual, received weekly cross-site supervision, and was highly adherent (>80%). Results indicated that a lower proportion of fear-decreasing behaviors (unrelated talk, accommodation) and a higher proportion of fear-increasing behaviors (encourage approach) were associated with greater habituation and better clinical outcomes. Some behaviors were too infrequently observed for analysis (reducing parent accommodation, relaxation) or exhibited varied relationships by analysis level (intensifying); we included those in benchmark categories based on consensus among exposure specialists and strong theoretical rationale (i.e., relaxation in the fear decreasing category; reducing parent accommodation and intensifying in the fear increasing category). Final benchmarks were generated *a priori* using the mean percentage of exposure time during which therapists used each behavior category (increasing or decreasing). Table 2 displays each benchmark with the means for individual behaviors that comprise each category.

ANALYTIC PLAN

Therapist behaviors were observed starting at the beginning of the baseline phase (i.e., first week with an exposure). To evaluate whether the EG is associated with changes in therapist behaviors (Aim 1), we determined the number of therapists meeting both benchmarks (*above* the benchmark for fear increasing behaviors and *below* the benchmark for fear decreasing behaviors) for three consecutive weeks, *only after the introduction of the EG*. We also calculated the percentage of weeks during which both benchmarks were met in each phase (all therapists combined) and examined the significance of the difference in percentage between study phases.

The primary study goal was to test whether therapist behaviors show initial evidence of change with EG training; in general, the purpose of testing *initial target engagement* is to probe whether an early signal is detected before devoting additional resources for continued research. Therefore, num-

Table 2
Therapist Behaviors in the Current Study vs. Clinical Trial Benchmarks: M(SD) Percent Time Observed Per Exposure

Therapist Behavior	Current Study		Clinical Trial Benchmark ^a
	Baseline Phase	Exposure Guide Phase	
Fear Increasing Behaviors^b	35.58 (18.49)	43.00 (16.16)	26.36 (16.81)
Encourage Approach	29.96 (14.89)	34.43 (14.82)	23.12 (14.89)
Intensify	5.60 (9.15)	8.45 (7.97)	3.24 (5.83)
Decrease Parent Accommodation	0 (0)	0.05 (0.19)	0.04 (0.28)
Fear Decreasing Behaviors^b	6.43 (9.80)	1.37 (3.07)	5.59 (10.08)
Accommodation	0.77 (1.46)	0.36 (0.80)	1.19 (2.69)
Unrelated Talk	5.60 (9.89)	0.99 (2.94)	4.10 (9.67)
Relaxation	0 (0)	0 (0)	0.04 (0.46)
Fear Neutral Behaviors^c	3.09 (4.43)	2.72 (2.52)	7.23 (9.44)
Teaching	1.70 (4.12)	0.93 (1.86)	2.89 (4.92)
Change Anxious Thoughts	0.46 (1.10)	0.08 (0.19)	1.38 (4.74)
Externalizing	0.82 (1.13)	1.66 (1.93)	2.96 (6.71)
Patient Habituation (sum of all instances) ^c	0.86 (0.74)	1.62 (2.03)	0.90 (1.31)

^a Benito et al. (2020).

^b Bolded items indicate inclusion in benchmarks.

^c Therapist training in the present study did not explicitly target neutral behaviors or habituation; we have included them in this table for comparison and because they appear as items on the EG.

ber of observations (i.e., weeks) was intentionally low, ending at three consecutive weeks meeting *a priori* benchmarks in the EG phase. As such, analyses of change in level/trend of therapist behaviors are likely underpowered and should be considered exploratory. Nevertheless, we explored these changes using Simulation Modeling Analysis (SMA; Borckardt & Nash, 2014). SMA uses bootstrapping methods to evaluate level and slope changes (given the degree of autocorrelation and phase length in the observed sample) and is well-suited for analyzing short autocorrelated data streams in a multiple baseline design. Given the presence of variability and linear trend in some of our baseline data, we did not use traditional visual analysis; when variability or linear trend are present, visual analysis may be inappropriate and statistical analysis such as SMA is preferred (Borckardt & Nash, 2014; Tate et al., 2016).

To evaluate change in nontargeted therapist factors (Aim 2), we used within-subjects *t*-tests to determine significance of change in therapist beliefs about exposure (TBES) and endorsement of fear-increasing and decreasing behaviors on a case vignette (separate analyses for pre- vs. post-baseline and for pre- vs. post-EG). To describe youth outcomes (Aim 3), we used within-subject *t*-tests to determine significance of change in symptom severity (CY-BOCS) and global severity (CGI-S) from pre- to posttreatment and characterized the rate of treatment response (CGI-I). To address Aim 4, we calculated descriptive statistics by study phase for training acceptability (TARS), EG completion time, and rate of consultation attendance.

MISSING DATA

Therapist behavior data were missing for some weeks (35.6%; 31 of 87 total weeks across therapists) due predominantly to a session not being scheduled for a given week (64.5%) or being cancelled by the family (22.6%). Other reasons included sessions that occurred but did not include exposure (6.5%) and sessions with exposures that were conducted out of the office (6.5%). This is consistent with rates of attendance in CMHA settings (e.g., Ingoldsby, 2010). Although missed sessions were anticipated, we opted to use data points by week (rather than session) because it is important to the internal validity of multiple baseline studies to use a real-time metric on the horizontal axis (Tate et al., 2016). This is a conservative approach given that data by session would appear nearly identical but with many fewer missing data points. With the exception of exposures conducted out of the office, all exposures that occurred were recorded and coded. Thus, nearly all targeted therapist behaviors that occurred were included in analysis.

Posttreatment assessment data were missing for three youths; two of these finished treatment with subjective improvement (per therapist report of treatment progress during consultation) but were lost to study follow-up, and one discontinued treatment (lost to both the therapist and the study team). One additional youth discontinued treatment to receive a higher level of care but completed a post-treatment assessment. Therapist six did not reach the EG phase due to two assigned

youth discontinuing treatment prior to the end of the baseline period.

Results

AIM I: INITIAL TARGET ENGAGEMENT (THERAPIST BEHAVIORS)

Primary Outcomes

Three of six therapists did not meet one or both benchmarks in the baseline phase but were able to meet both benchmarks for three consecutive weeks during the EG phase (Figure 2). Of the other

three therapists, one did not reach the EG phase (due to youth treatment discontinuation) and two had already met both benchmarks for three consecutive weeks during the baseline phase. Combining all weeks across therapists, there was a significant benchmark rate difference of 33.23% (95% CI [20.67%, 45.79%]) across study phases, indicating a higher percentage of weeks meeting benchmarks in the EG vs. baseline phases.

Exploratory Outcomes

Therapist 4 had a significant change in level of fear-increasing ($R = .90, p = .02$) and decreasing

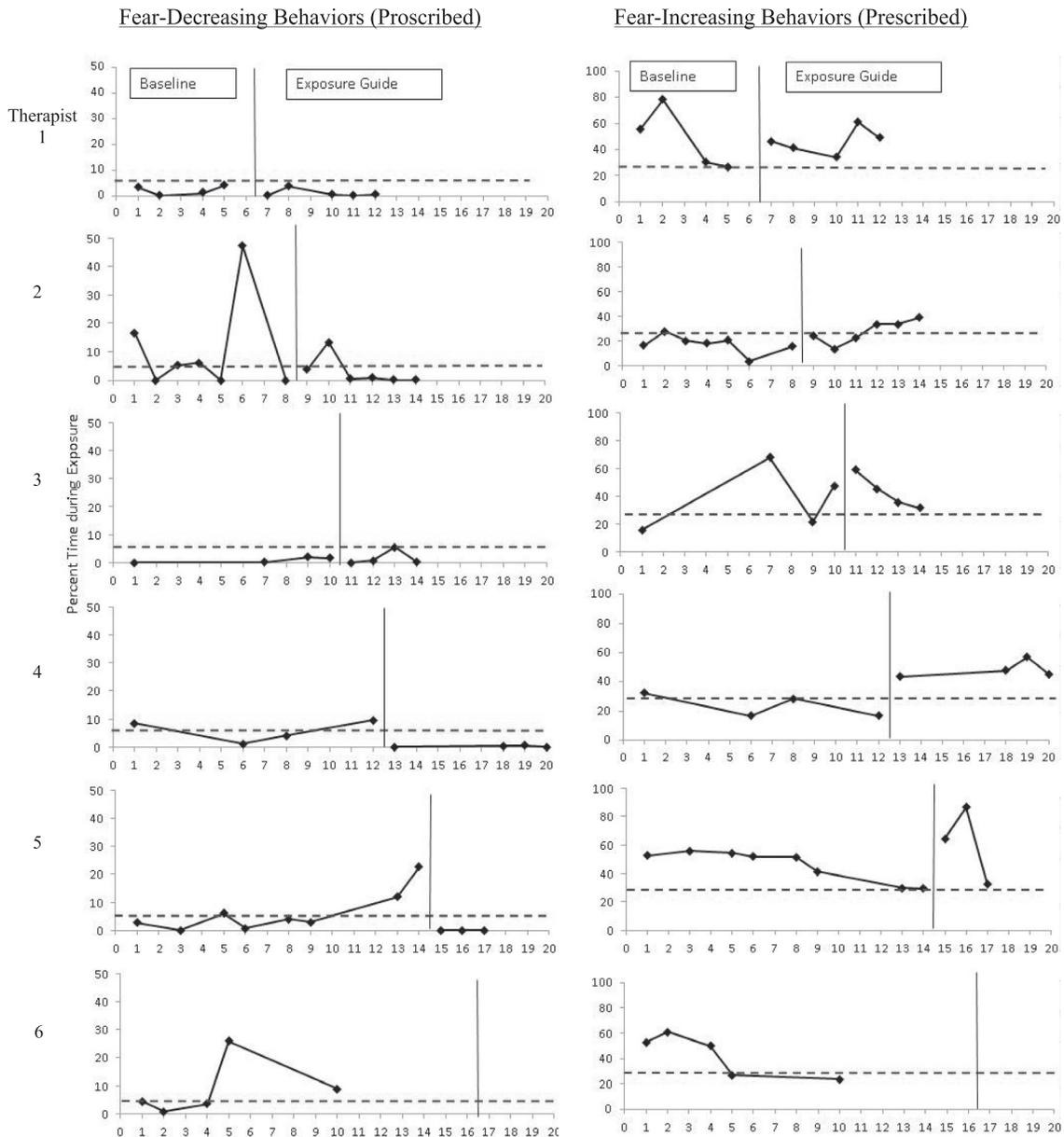


FIGURE 2. Therapist behaviors: Multiple baseline data in relation to clinical trial benchmarks. Note. Dashed lines represent the clinical trial mean; therapists met benchmarks when at or below the mean for decreasing behavior and at or above the mean for increasing behavior, for three consecutive weeks.

behaviors ($R = -.76$, $p = .03$). Therapist 5 met benchmarks early in the baseline phase but then had a decline in fear-increasing behaviors and an uptick in fear-decreasing behaviors immediately prior to EG introduction (Figure 2). Holding these baseline trends constant, this therapist had a significant in level of increasing ($R = .71$, $p = .03$) and decreasing behaviors ($R = -.78$, $p = .02$). There were no other significant changes in level or slope of behaviors. Full SMA results are available in Supplemental Figure 3.

AIM 2: TARGET SPECIFICITY (THERAPIST ATTITUDES)

Therapist negative beliefs about exposure on the TBES significantly improved during the baseline phase but did not change further during the EG phase (Table 3). Similarly, therapist endorsement of fear-decreasing behavior on a case vignette declined from initial training to the end of the baseline phase but did not change further during the EG phase (Table 3). There were no significant changes in case vignette endorsement of fear-increasing behavior during either study phase.

AIM 3: EFFECTIVENESS (YOUTH OUTCOMES)

Including three youth participants without post-treatment data (using pretreatment values carried forward), there was a significant pre-post change in OCD symptom severity ($n = 8$; CY-BOCS; $t = 2.49$, $p = .041$, $d = 1.14$) and in global illness severity (CGI-S; $t = 2.39$, $p = .048$, $d = 1.19$). Global improvement was only assessed at posttreatment; of participants who completed posttreatment assessment ($n = 5$), 80% ($n = 4$) were considered treatment responders (CGI-I score indicating “much” or “very much” improved).

AIM 4: FEASIBILITY AND ACCEPTABILITY OF EG TRAINING

Consultation attendance was 87.5% during the baseline phase and 91.1% during the EG phase. EG completion rate was high (91.7%) and therapists reported needing less than 5 minutes to complete, on average. Therapists rated training acceptability between “moderate” and “strong” ($M = 4.86$, $SD = 0.31$) and training helpfulness near “a great deal” ($M = 2.72$, $SD = 0.15$; TARS-2). Informal interviews with therapists after study completion suggests that therapists found the EG user-friendly (“simple and easy to use”) and a useful addition to clinical practice (“concrete”; “adds precision”). Therapists also reported using the EG with nonstudy clients. Therapist suggestions for improving training with the EG were relatively minor (e.g., expanding the response scale for one item).

Discussion

This study provides initial support for the EG as an implementation strategy that can target specific therapist behaviors associated with exposure quality. Specifically, addition of the EG produced a change in therapist behaviors above and beyond gold-standard training. Three of five therapists who completed training met behavioral benchmarks only after using the EG. The other two met benchmarks during baseline; one of these moved away from benchmarks at the end of the baseline phase but met them again immediately after EG introduction. This therapist also had a level change in behaviors from baseline to the EG phase, suggesting that future tests of the EG might examine whether it improves consistency

Table 3
Change in Therapist-Report Measures by Study Phase

Measure	Change: Study Entry – End of Baseline ($n = 6$)	Change: End of Baseline– End of EG ($n = 5$)
Therapist Beliefs about Exposure	$t = 2.75^*$ $M (SD) = 11.17(9.95)$ 95% CI = [0.73, 21.61] $d = 1.13$	$t = 0.07$ $M (SD) = 0.20(6.18)$ 95% CI = [−7.47, 7.87] $d = 0.03$
Case Vignette: Fear Increasing Behaviors	$t = 2.22$ $M(SD) = 0.22(0.23)$ 95% CI = [−0.03, 0.46] $d = 0.90$	$t = -2.00$ $M (SD) = -0.55(0.62)$ 95% CI = [−1.31, 0.21] $d = 0.76$
Case Vignette: Fear Decreasing Behaviors	$t = 5.15^*$ $M(SD) = 1.04(0.50)$ 95% CI = [0.53, 1.57] $d = 2.01$	$t = 0.05$ $M (SD) = 0.19(0.97)$ 95% CI = [−1.19, 1.22] $d = 0.02$

* $p < .05$; Negative values indicate an increase over time.

of behaviors over time. Notably, all therapists met benchmarks by the end of the study. Across therapists, the EG phase had a higher percentage of weeks meeting benchmarks vs. the baseline phase. Together, findings show early evidence that therapist behaviors during exposure can be successfully targeted in a CMHA setting. However, this study was designed to detect a preliminary signal of change and results will need to be confirmed in larger studies that also examine durability of change over time.

Results also provide early evidence that EG training may have a *specific* effect on therapist behaviors. Consistent with previous research (Deacon et al., 2013), therapist beliefs about exposure improved with gold-standard training; however, no additional change was found after introduction of the EG despite changes in therapist behavior during this same period. A similar pattern was observed for endorsement of fear-decreasing behaviors on a case vignette; prior to training, therapists endorsed more of these behaviors than at the end of the baseline phase, but no additional changes were found at the end of the EG phase. Endorsement of fear-increasing behaviors did not change over time (although the direction and effect size would suggest that endorsement declined during baseline). This provides initial support for *target specificity* (i.e., that the EG only changes targeted behaviors) and increases confidence that the change in therapist behaviors can be attributed to introduction of the EG rather than to other relevant factors (i.e., improvements in knowledge or comfort with exposure over time).

Exploratory examination of effectiveness outcomes suggests that treatment response (80%) was comparable to prior trials demonstrating efficacy (68%; McGuire et al., 2015) and effectiveness (72.6%; Torp et al., 2015). Youth also experienced significant pre- to posttreatment reductions in OCD severity. These outcomes appear promising, although we emphasize that some analyses included only youth completing posttreatment assessment, and the current design does not permit conclusions about the effect of EG training on clinical outcomes. Future studies will need to examine this and confirm whether any effects occur via therapist behavior changes. Although not an explicit focus of this study, descriptive results show that youth habituation—a theorized signal of initial exposure learning—nearly doubled in the EG vs. the baseline phase. Prior findings from this study series linked habituation with clinical outcomes (Benito et al., 2018). The current finding is consistent with the idea that

improving therapist behaviors influences the rate of habituation, although future studies will be needed to test a causal relationship.

Results also provide preliminary support for feasibility and acceptability of the EG as evidenced by brief completion time (<5 min), high consultation attendance, and informal feedback from therapists. Our goal in creating the EG was to develop a training implementation strategy that is useful in practice. On its face, it seems uncertain that therapists would welcome specific feedback about in-session behaviors. We worked closely with CMHA partners, using stakeholder feedback through many iterations of initial development to enhance feasibility and acceptability. Several factors appear key to the value of the EG in this setting: encouraging therapist self-feedback (vs. using consultant feedback alone), including built-in support for common exposure problems on the EG, and fostering a collaborative training environment. These observations are consistent with our use of coordinated strategic action to design EG training, which emphasizes the integration of science and context to help guide decisions about “what to do when” (Chorpita & Daleiden, 2018). The EG was specifically designed to support therapist learning through practice iterations in which in-session events are linked with exposure principles (in addition to being a quality monitoring tool). Informal feedback suggests that this framework was valuable for therapists, although future studies will need to examine this further.

Although factors such as attendance were likely influenced by the context of research (e.g., compensation), most therapists reported using the EG with nonstudy youths, suggesting added value beyond the research context. While practice utility was central to the design of the EG, both phases of this study used an intensive model of therapist training that is unlikely to be feasible in most practice settings. We chose this model to create a rigorous test of whether the EG adds value to the best-known training approach. If future studies support the effectiveness of EG versus gold-standard training, subsequent work will need to confirm its utility within less burdensome models and consider implementation strategies to help sustain its use (e.g., financial incentives from payers, integrating the EG into the electronic medical record). Relatedly, we are currently testing the EG with a variety of users (e.g., agency supervisors, therapists, parents, youth) as part of a pragmatic trial evaluating its utility as a quality measure in CMHA settings. We anticipate that results will help to inform future selection of implementation strategies for sustainability.

Rates of therapist behaviors suggest very high exposure quality in this community sample, which differs from what might be expected based on the literature. Mean rates of fear-increasing behaviors exceeded those of clinical trial therapists in *both* study phases. This contrasts reports that community therapists deliver exposures in a less “intense” manner (Deacon et al., 2013). Similarly, only two sessions occurred without exposure across all therapists. This is remarkably different from the underutilization reported in practice (e.g., Becker-Haimes et al., 2017), even with intensive support (i.e., weekly specialist consultation; McLeod et al., 2017), and actually exceeds rates of exposure reported in efficacy trials (Benito et al., 2018). This high level of quality could relate to study methods, including the context of research where training was an explicit focus, use of video review, and/or voluntary enrollment of therapists who may have been highly motivated. Future studies will be essential for identifying baseline predictors of therapist behaviors so that we may understand which therapists benefit most from EG training. Finally, during the EG phase (but not during baseline) we observed fear-decreasing behaviors at a lower rate and with less variability than in the clinical trial sample. This raises the possibility that the EG could optimize and/or standardize quality in efficacy trials, although future studies would be needed to explore this use.

Importantly, this study was not designed to target therapist behaviors specific to one theory of exposure mechanism. Rather, the goal was to target behaviors that align with behavioral theory as well as clinical recommendations to promote “intense” and avoid “cautious” delivery—and that have been linked to patient outcomes. These behaviors are also relevant for habituation, a proximal outcome with high potential for utility in practice settings. When measured as fear reduction occurring “on its own,” habituation is conceptually consistent with multiple theories of mechanism. However, initial learning (i.e., within exposure) is only one component of successful long-term fear extinction and future studies should incorporate features of exposure with relevance for postexposure learning phases (i.e., consolidation and retrieval). Training that targets these might emphasize aspects of exposure task design/selection (e.g., whether to repeat exposure tasks or design tasks that target more than one type of fear) or approaches to postexposure processing (e.g., focus on self-efficacy or mismatch between expected vs. actual outcomes). Additionally, there may be other mechanisms through which therapist

behaviors have an effect (e.g., by improving motivation or parenting skills), and future studies should examine those. As these examples illustrate, there are a number of exciting areas for future work to optimize therapist exposure training. We view the present study as a *very* early step toward a larger body of work that could inform an empirically supported training “toolkit” designed to support therapists through the process of optimizing exposure delivery.

Study findings should be considered in the context of several important limitations. First, factors across multiple levels likely contribute to therapist behavior and we did not account for them all. These might include processes that naturally vary across exposure tasks (e.g., fear level, task selection, targeted fear domain), youth characteristics (e.g., age, motivation, symptom severity) or organizational factors (e.g., supports for in-office exposure). Future studies should be designed to examine the influence of such factors. Although we aimed to examine factors other than the EG that could plausibly explain therapist behavior changes with the accumulation of training over time (knowledge and beliefs about exposure), it is possible that some unmeasured variable relates to these changes. Second, missing data and high baseline variability precluded the use of traditional visual analysis. Most missing data were due to unscheduled or family-cancelled sessions (e.g., with holidays or scheduling conflicts), which is a common challenge in practice. This underscores the need for additional studies using alternative designs (e.g., randomized controlled trial). Relatedly, this study was only designed to detect a preliminary signal of change in a small sample of therapists; it will be essential to conduct a fully powered test of the EG before definitive conclusions can be drawn. Last, there are important differences in research versus practice settings—prompting researchers to recommend caution when generalizing clinical trial benchmarks to practice (McLeod et al., 2017). Future studies will need to determine whether to adjust benchmarks by setting or other relevant factors. Limitations notwithstanding, clinical trial benchmarks are useful for understanding therapist behaviors at a threshold that relates to efficacy. This study also has several strengths that include active training in the baseline phase and independent measurement of therapist behavior targets using a validated coding system. It is also unique in its application of the experimental therapeutics approach to a training intervention in a real-world setting. Explicit testing of training mechanisms could help advance the development of

new training approaches that move beyond questions of dose and format. At present, few studies have compared methods for training therapists and further research in this area is critically important.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the National Institutes of Mental Health [R21MH096828 and R33MH096828].

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.beth.2020.10.003>.

References

- Abramowitz, J. S., Deacon, B. J., & Whiteside, S. P. H. (2019). *Exposure therapy for anxiety: Principles and practice* (2nd ed.). Guilford Publications.
- Becker-Haimes, E. M., Okamura, K. H., Wolk, C. B., Rubin, R., Evans, A. C., & Beidas, R. S. (2017). Predictors of clinician use of exposure therapy in community mental health settings. *Journal of Anxiety Disorders, 49*, 88–94. <https://doi.org/10.1016/j.janxdis.2017.04.002>.
- Beidas, R. S., & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice, 17*(1), 1–30. <https://doi.org/10.1111/j.1468-2850.2009.01187.x>.
- Benito, K. G., Machan, J., Freeman, J. B., Garcia, A. M., Walther, M., Frank, H., Wellen, B., Stewart, E., Edmunds, J., Kemp, J., Sapyta, J., & Franklin, M. (2018). Measuring fear change within exposures: Functionally-defined habituation predicts outcome in three randomized controlled trials for pediatric OCD. *Journal of Consulting and Clinical Psychology, 86*(7), 615–630. <https://doi.org/10.1037/ccp0000315>.
- Benito, K. G., Machan, J., Freeman, J. B., Garcia, A. M., Walther, M., Frank, H., Wellen, B., Stewart, E., Edmunds, J., Sapyta, J., & Franklin, M. (2020). Therapist behavior during exposure tasks predicts habituation and clinical outcome in three randomized controlled trials for pediatric OCD. *Behavior Therapy, 2020.07.004*. <https://doi.org/10.1016/j.beth.2020.07.004>.
- Benito, K. G., Conelea, C., Garcia, A. M., & Freeman, J. B. (2012). CBT specific process in exposure-based treatments: Initial examination in a pediatric OCD sample. *Journal of Obsessive-Compulsive and Related Disorders, 1*(2), 77–84. <https://doi.org/10.1016/j.jocrd.2012.01.001>.
- Borckardt, J. J., & Nash, M. R. (2014). Simulation modelling analysis for small sets of single-subject data collected over time. *Neuropsychological Rehabilitation, 24*(3–4), 492–506. <https://doi.org/10.1080/09602011.2014.895390>.
- Castonguay, L. G., Boswell, J. F., Zack, S. E., Baker, S., Boutselis, M. A., Chiswick, N. R., ... Holtforth, M. G. (2010). Helpful and hindering events in psychotherapy: A practice research network study. *Psychotherapy (Chicago, Ill.), 47*(3), 327–344. <https://doi.org/10.1037/a0021164>.
- Chorpita, B. F., & Daleiden, E. (2018). Coordinated strategic action: Aspiring to wisdom in mental health service systems. *Clinical Psychology: Science and Practice, 25*(4). <https://doi.org/10.1111/cpsp.12264>.
- Chu, B. C., Crocco, S. T., Arnold, C. C., Brown, R., Southam-Gerow, M. A., & Weisz, J. R. (2015). Sustained implementation of cognitive-behavioral therapy for youth anxiety and depression: Long-term effects of structured training and consultation on therapist practice in the field. *Professional Psychology: Research and Practice, 46*(1), 70–79. <https://doi.org/10.1037/a0038000>.
- Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10–23. <https://doi.org/10.1016/j.brat.2014.04.006>.
- Deacon, B. J., Farrell, N. R., Kemp, J. J., Dixon, L. J., Sy, J. T., Zhang, A. R., & McGrath, P. B. (2013). Assessing therapist reservations about exposure therapy for anxiety disorders: The Therapist Beliefs about Exposure Scale. *Journal of Anxiety Disorders, 27*(8), 772–780. <https://doi.org/10.1016/j.janxdis.2013.04.006>.
- Gunter, R. W., & Whittal, M. L. (2010). Dissemination of cognitive-behavioral treatments for anxiety disorders: Overcoming barriers and improving patient access. *Clinical Psychology Review, 30*(2), 194–202. <https://doi.org/10.1016/j.cpr.2009.11.001>.
- Guy, W., National Institute of Mental Health (U.S.), & Early Clinical Drug Evaluation Program. (1976). ECDEU assessment manual for psychopharmacology. Rockville, MD: U. S. Dept. of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health.
- Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*(2), 147–185. <https://doi.org/10.1901/jaba.2003.36-147>.
- 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*(4), 448–466. <https://doi.org/10.1016/j.cpr.2010.02.005>.
- Higa-McMillan, C. K., Francis, S. E., Rith-Najarian, L., & Chorpita, B. F. (2016). Evidence base update: 50 years of research on treatment for child and adolescent anxiety. *Journal of Clinical Child and Adolescent Psychology, 45*(2), 91–113. <https://doi.org/10.1080/15374416.2015.1046177>.
- Ingoldsby, E. M. (2010). Review of interventions to improve family engagement and retention in parent and child mental health programs. *Journal of Child and Family Studies, 19*(5), 629–645. <https://doi.org/10.1007/s10826-009-9350-2>.
- Kazdin, A. E. (2009). Understanding how and why psychotherapy leads to change. *Psychotherapy Research, 19*, 418–428. <https://doi.org/10.1080/10503300802448899>.
- Kircanski, K., & Peris, T. S. (2015). Exposure and response prevention process predicts treatment outcome in youth with OCD. *Journal of Abnormal Child Psychology, 43*(3), 543–552.
- McGuire, J. F., Piacentini, J., Lewin, A. B., Brennan, E. A., Murphy, T. K., & Storch, E. A. (2015). A meta-analysis of cognitive behavior therapy and medication for child obsessive-compulsive disorder: Moderators of treatment

- efficacy, response, and remission. *Depression and Anxiety*, 32(8), 580–593. <https://doi.org/10.1002/da.22389>.
- McLeod, B. D., Southam-Gerow, M. A., Jensen-Doss, A., Hogue, A., Kendall, P. C., & Weisz, J. R. (2017). Benchmarking treatment adherence and therapist competence in individual cognitive-behavioral treatment for youth anxiety disorders. *Journal of Clinical Child & Adolescent Psychology*, 48(sup1), S234–S246. <https://doi.org/10.1080/15374416.2017.1381914>.
- McLeod, B. D., Southam-Gerow, M. A., Tully, C. B., Rodriguez, A., & Smith, 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, 14–32.
- Milne, D., & Noone, S. (1996). *Teaching and training for non-teachers*. Leicester: The British Psychological Society.
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., ... Aarons, G. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health*, 38(2), 65–76. <https://doi.org/10.1007/s10488-010-0319-7>.
- Scahill, L., Riddle, M. A., McSwiggin-Hardin, M., Ort, S. I., King, R. A., Goodman, W. K., ... Leckman, J. F. (1997). Children's Yale-Brown Obsessive Compulsive Scale: Reliability and validity. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(6), 844–852. <https://doi.org/10.1097/00004583-199706000-00023>.
- Sheehan, D. V., Sheehan, K. H., Shytle, R. D., Janavs, J., Bannon, Y., Rogers, J. E., ... Wilkinson, B. (2010). Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). *The Journal of Clinical Psychiatry*, 71(3), 313–326. <https://doi.org/10.4088/JCP.09m05305whi>.
- Sholomskas, D. E., Syracuse-Siewert, G., Rounsaville, B. J., Ball, S. A., Nuro, K. F., & Carroll, K. M. (2005). We don't train in vain: A dissemination trial of three strategies of training clinicians in cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology*, 73(1), 106–115.
- Smith, M. M., McLeod, B. D., Southam-Gerow, M. A., Jensen-Doss, A., Kendall, P. C., & Weisz, J. R. (2017). Does the delivery of CBT for youth anxiety differ across research and practice settings? *Behavior Therapy*, 48(4), 501–516. <https://doi.org/10.1016/j.beth.2016.07.004>.
- Storch, E. A., Lewin, A., Nadai, D. e., & Murphy, T. (2010). Defining treatment response and remission in obsessive-compulsive disorder: A signal detection analysis of the Children's Yale-Brown Obsessive Compulsive Scale. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(7), 708–717.
- Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., ... Wilson, B. (2016). The Single-Case Reporting guideline In Behavioral interventions (SCRIBE) 2016 Statement. *Aphasiology*, 30(7), 862–876. <https://doi.org/10.1080/02687038.2016.1178022>.
- Torp, N. C., Dahl, K., Skarphedinnson, G., Thomsen, P. H., Valderhaug, R., Weidle, B., ... Ivarsson, T. (2015). Effectiveness of cognitive behavior treatment for pediatric obsessive-compulsive disorder: Acute outcomes from the Nordic Long-term OCD Treatment Study (NordLOTS). *Behaviour Research and Therapy*, 64, 15–23. <https://doi.org/10.1016/j.brat.2014.11.005>.
- 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>.
- Whiteside, S. P. H., Deacon, B. J., Benito, K., & Stewart, E. (2016). Factors associated with practitioners' use of exposure therapy for childhood anxiety disorders. *Journal of Anxiety Disorders*, 40, 29–36. <https://doi.org/10.1016/j.janxdis.2016.04.001>.

RECEIVED: June 11, 2020

ACCEPTED: October 14, 2020

AVAILABLE ONLINE: 24 OCTOBER 2020