Available online at www.sciencedirect.com

ScienceDirect

Behavior Therapy 52 (2021) 1325–1338

Behavior Therapy

www.elsevier.com/locate/bt

Skill Acquisition During Transdiagnostic Treatment With the Unified Protocol

Shannon Sauer-Zavala
University of Kentucky

Anthony J. Rosellini
Boston University

Kate H. Bentley
Massachusetts General Hospital/Harvard Medical School

Amantia A. Ametaj
Harvard School of Public Health

James F. Boswell
University at Albany, State University of New York

Clair Cassiello-Robbins
Duke University Medical Center

Julianne Wilner Tirpak
Todd J. Farchione
David H. Barlow
Boston University

The Unified Protocol (UP) for Transdiagnostic Treatment of Emotional Disorders is an emotion-focused, cognitive-behavioral intervention developed to address the full range of anxiety, depressive, and related disorders. The UP consists of core therapeutic skills that, though unique in focus, are each designed to promote an approach-oriented stance toward emotional experiences. The goal of the present investigation was to characterize changes in these skills for patients that received a course of treatment with the UP, as well as to examine associations between skills and symptoms changes. Patients with principal anxiety disorders, assigned to receive treatment with the UP ($N = 88$) as part of a randomized controlled trial, were included in this study. They completed validated self-report measures of UP skills (Understanding Emotions, Mindful Emotion Awareness, Cognitive Flexibility, Countering Emotional Behaviors, and Interoceptive Awareness and Tolerance), as well as clinician-rated measures of psychological symptoms. Skill measures improved significantly over the course of 12 to 16 UP treatment sessions and changes in these skills measures were associated with improvements in anxiety symptoms. Determining whether improvement on all the skills learned during a course of treatment with UP is associated with symptom remission is critical to establishing the most streamlined and efficient interventions that may ultimately be best suited to widespread dissemination.

Address correspondence, including for requests for de-identified study data and code, to Shannon Sauer-Zavala, Department of Psychology, 106A Kastle Hall, University of Kentucky, Lexington, KY 40506. e-mail: ssz@uky.edu.

Keywords: transdiagnostic; Unified Protocol; cognitive-behavioral therapy; skills

The predominant scheme for the classification of mental disorders has emphasized grouping psychopathology into thinly sliced categories based on observed differences, a method exemplified by the *Diagnostic and Statistical Manual* (*DSM*; American Psychiatric Association, 2013). Emerging research has revealed problems with this approach, including high rates of comorbidity and substantial phenotypic overlap across categories; accordingly, it has been suggested that focusing on fine-grained differences, though enhancing diagnostic reliability, may come at the expense of validity (see: Barlow et al., 2014). In other words, there is increasing evidence that some *DSM* disorders do not represent unique constructs, but are relatively trivial variations of a common underlying syndrome (Andrews, 1990, 1996; Brown & Barlow, 2009). As a result of these findings, recent research initiatives highlight the importance of identifying shared features that are implicated in the development of a range of disorders (e.g., the National Institute of Mental Health's Research Domain Criteria; Insel et al., 2010). One overarching goal of this more dimensional approach to classification is to ultimately craft more efficient transdiagnostic treatments that simultaneously address comorbid conditions by directly targeting shared maintaining factors.

The term "emotional disorders" represents a grouping of mental disorders informed by a more dimensional or functional basis for classification (Bullis et al., 2019). As the name suggests, this broad category of disorders is chiefly characterized by the propensity to experience strong emotions, accompanied by perception of one's inability to cope with such experiences; this negative emotionality has also been referred to as a neurotic temperament (Barlow et al., 2014). Functionally, emotional disorders are maintained by aversive reactions to frequently occurring negative emotions that, in turn, result in problematic attempts to avoid or suppress these experiences (Bullis et al., 2019). Avoidant coping has been associated with rebound effects whereby suppressed emotions return with greater frequency and intensity (Abramowitz et al., 2001; Gross & John, 2003), exacerbating a range of *DSM* disorder symptoms (Purdon, 1999). Although inclusion in the emotional disorders group is made on the basis of this functional model and is not tied to diagnostic status, disorders generally included within this purview are anxiety, depressive, and related disorders (i.e., internalizing disorders; Barlow, 1991). Indeed, there are extremely high rates of comorbidity among these common disorders (Allen et al., 2010; Brown et al., 2001) and a substantial literature underscores the contribution of negative emotionality in accounting for this overlap (e.g., Brown, 2007; Brown & Barlow, 2009; Griffith et al., 2010; Zinbarg et al., 2016).

Given the increasing evidence to suggest that common processes are responsible for the development and maintenance of a range of disorders, transdiagnostic interventions that directly target shared deficits have gained prominence in recent years (Sauer-Zavala, Gutner et al., 2017). For example, the Unified Protocol for the Transdiagnostic Treatment of Emotional Disorders (UP) (Barlow et al., 2018) is a cognitive-behavioral intervention that was explicitly developed to address the full range of anxiety, depressive, and related disorders. The overall aim of the UP is to extinguish distress in response to the experience of strong emotions such that patients reduce their reliance on the avoidant coping responsible for the rebound effects that maintain symptoms.

Indeed, the UP consists of core skills modules that, though distinct, are each designed to reduce aversive reactions to emotions. Following a motivational enhancement session (Module 1), the UP first focuses on improving one's understanding of the adaptive, functional nature of emotions and teaches patients to break emotional experiences down into three interacting components: thoughts, physical sensations, and behaviors, as well as antecedents (i.e., triggers) and short- and long-term consequences of responses to emotion (Module 2). Next, mindful emotion awareness (Module 3) facilitates willingness to engage with emotional experiences, rather than automatically pushing them away; specifically, patients practice three experiential exercises that encourage them to approach their emotions with a present-focused, nonjudgmental stance.

Then, patients are taught specific skills that map onto the three interacting components of an emotional experience. First, patients are encouraged to question their automatic interpretations of situations that elicit strong emotions (Module 4). Consistent with the UP's emphasis on understanding and tolerating all aspects of an emotional experience (including cognitions), the goal of this module is to foster flexibility in thinking, rather than changing maladaptive cognitions. Countering emotional behaviors (Module 5) involves identifying each patient's idiosyncratic behavioral attempts to avoid full exposure to strong emotions, followed by encouragement to engage in alternative actions that encourage contact with emotional experiences. Next, greater tolerance of physical sensations is cultivated through the use of interoceptive exercises (e.g., hyperventilation,

breathing through a thin straw) that deliberately provoke the physiological sensations associated with strong emotions that often trigger emotions even in the absence of situational cues (Module 6).

Treatment with the UP culminates with an emotion exposure module (Module 7), during which patients engage in a series of activities explicitly designed to elicit strong or uncomfortable emotions; once again, the goal of these exercises is to gradually extinguish aversive reactions to emotions by amassing new evidence that these experiences are temporary and can be tolerated. Finally, a relapse prevention session (Module 8) is conducted to discuss progress and delineate ways to practice skills in the future.

THE PRESENT STUDY

There is strong evidence to suggest that the UP is efficacious for treating heterogeneous anxiety disorders (e.g., Barlow et al., 2017). However, the extent to which skill acquisition is related to symptom reduction has yet to be evaluated. This is an important question as its answer may provide preliminary evidence that the acquisition of specific skills may serve as a mechanism of action during treatment with the UP. For example, if certain skills targeted by the UP change across treatment but are not associated with symptom improvement, this may suggest that the UP module(s) associated with that skill are not active ingredients in this treatment. Ultimately, shorter and more-streamlined interventions have a better chance of being widely disseminated.

Accordingly, the goal of the present report was to explore associations between change in the UP skills and improvements in anxiety and depressive symptoms in patients with a range of anxiety disorders. Data were drawn from a recently completed, large randomized controlled trial (Barlow et al., 2017). Although the primary purpose of the parent study was to determine whether the transdiagnostic UP approach results in equivalent symptom improvement to gold-standard single-disorder interventions, measures of each UP skill were included specifically to explore the effects of this intervention on these targets. First, we hypothesized that treatment with the UP would lead to increases in measures of the skills taught during treatment (i.e., increased emotional understanding [Module 2], increased mindful emotion awareness [Module 3], increased cognitive flexibility [Module 4], and decreased behavioral avoidance [Module 5] and interoceptive sensitivity [Module 6]). Modules 1 (Motivation and Goal Setting) and 8 (Relapse Prevention) were not included in the present study because they are not considered core skills that address reactions to negative emotions. Furthermore, Module 7 (Emotion Exposures) was not assessed as a separate skill as it represents a culmination of treatment in which the practice of the previous 5 skills (Modules 2–6) is encouraged. Additionally, we hypothesized that initial change in each skill would predict overall improvements in anxiety and depressive symptoms.

Method

PARTICIPANTS

Participants in the present report were drawn from a sample (N = 223) of treatment-seeking individuals at the Center for Anxiety and Related Disorders at Boston University (CARD) who enrolled in a trial comparing two active treatment conditions and a waitlist control condition. Given our focus on understanding the role of skill improvement during the course of treatment with the UP, only individuals in the UP condition (n = 88) were included in this investigation. Given that the study's active treatment comparison condition utilized single disorder protocols (e.g., Mastery of Your Anxiety and Panic; Barlow & Craske, 2007), participants were primarily recruited on the basis of principal diagnoses that aligned with these treatments. Specifically, individuals were eligible for the trial if they were (1) assigned a principal (most interfering and severe) diagnosis of panic disorder (PD), generalized anxiety disorder (GAD), obsessive-compulsive disorder (OCD), or social anxiety disorder (SOC), as assessed using the Anxiety and Related Disorders Interview Schedule (ADIS; see description below); (2) 18 years or older; and (3) fluent in English. Following long-standing procedures in our clinical trials (see: Barlow et al., 2000), individuals taking psychotropic medications were asked to remain stable on the same dose for the duration of the study.

Exclusion criteria consisted primarily of conditions that required prioritization for immediate or simultaneous treatment that could interact with the study treatment in unknown ways: specifically, a current diagnosis of bipolar disorder, schizophrenia, schizoaffective disorder, or organic mental disorder, current high suicide risk, or, recent (within 3 months) history of substance use disorder. Individuals were also excluded if they previously received at least eight sessions of CBT within the past 5 years. For more information on study inclusion/exclusion and flow via the CONSORT diagram, see Barlow et al. (2017).

The majority of the sample was female (54.5%) with a mean age of 30.95 (*SD* = 11.54). The majority of participants identified as Caucasian (83%), whereas 9.1% identified Black or African-American, 6.8% as Asian, and 1.1% as Native Hawaiian or Pacific Islander. A small subset of the sample (3.4%) identified as Hispanic or Latino/a. Principal diagnoses included GAD (*n* = 22), PD (*n* = 25), SOC (*n* = 23), and OCD (*n* = 18). Most patients (*n* = 72) had at least one additional comorbid condition; on average, participants had 2.22 (*SD* = 1.70) comorbid diagnoses (for a breakdown, see: Jarvi Steele et al., 2018). Finally, 47 individuals were taking psychotropic medications.

MEASURES

Validated self-report and clinician-rated measures were utilized to assess variables of interest. Following a diagnostic assessment, measures of UP skills and symptom severity were administered at five time points: prior to Session 1 (pretreatment) and immediately following Sessions 4, 8, 12, 16 (posttreatment).

DIAGNOSTIC ASSESSMENT

The Anxiety and Related Disorders Interview Schedule (ADIS; Brown & Barlow, 2014; Brown et al., 1994) was used to confirm study inclusion/exclusion criteria. The ADIS is a semi-structured clinical interview that focuses on *DSM* diagnoses of anxiety, mood, somatic symptom, and substance use disorders, along with screening items for other disorders. Diagnoses are assigned a clinical severity rating (CSR) on using an 9-point Likert scale ranging from 0 (no symptoms) to 8 (extremely severe symptoms), with a rating of 4 or above (definitely disturbing/disabling) representing the clinical threshold for *DSM* diagnostic criteria.

UP Skills

Psychometric studies support the reliability and validity of the skills and symptoms scales/scoring described here (see citations below). We nevertheless evaluated internal consistency in our sample by calculating McDonald's ω[1] using the Hayes and Coutts (2020) procedure with a macro for SPSS.

The awareness subscale of the *Difficulties in Emotion Regulation Scale- Awareness Subscale* (DERS-A; Gratz & Roemer, 2004) was used to assess improvement associated with UP Module 2 (Understanding Emotions). The DERS-A consists of 6 items (e.g., "I pay attention to how I feel," "when I'm upset, I believe my feelings are valid and important," "I take time to figure out what I'm really feeling) that are rated on a 5-point Likert-scale with response choices ranging from 1 (almost never) to 5 (almost always). The items are then reverse scored and summed, such that lower DERS-A total scores indicate higher levels of emotional understanding. Validation of the full DERS suggests that this measure has strong psychometric properties (Gratz & Roemer, 2004). In this sample, ω = 0.83 for the awareness subscale of the DERS (DERS-A) at baseline. The DERS-A will be referred to as a measure of emotional understanding through the manuscript.

The Southampton Mindfulness Questionnaire (SMQ; Chadwick et al., 2008) was included to measure improvements in nonjudgmental, present-focused awareness of emotions associated with UP Module 3 (Mindful Emotion Awareness). The SMQ is a 16-item self-report measure with items beginning with "Usually, when I have distressing thoughts or images" and continuing with a mindfulness-related response, such as "I try just to experience the thoughts or images without judging them." Participants rate these items on a 7-point Likert scale ranging from 0 (*strongly disagree*) to 6 (*strongly agree*), with higher scores reflecting greater levels of mindful responding. The scale consists of a single factor structure (items are summed to form a total score) and has demonstrated good internal consistency and validity (Chadwick et al., 2008). In this sample, ω = .88 at baseline.

The Cognition Checklist - Anxiety subscale (CCL-A; Beck et al., 1987; Steer et al., 1994) is typically used to assess the frequency of anxious automatic thoughts. On a 5-point Likert scale ranging from 0 (*Never*) to 4 (*Always*), individuals identify how frequently they have 16 different thoughts (e.g., "Something awful is going to happen," "What if no one reaches me in time to help?"). The CCL-A has documented good internal consistency, and test-retest reliability (Beck et al., 1987; Steer et al., 1994; Taylor et al., 1997). In order to assess change that corresponds more closely to Module 4 (Cognitive Flexibility) of the UP, individuals were also asked to rate the degree to which they believe the thoughts associated with each item to be true (0 = *Not at all* to 4 = *Very much*), modeled after the format of the Automatic Thoughts Questionnaire (Hollon & Kendall, 1980). Lower total scores (i.e., a sum of

[1] Emerging recommendations suggest using McDonald's ω for internal consistency estimates because Cronbach's alpha (the more traditional approach) assumes that scales are a single factor (see Hayes & Coutts, 2020).

item scores) on this measure reflect greater cognitive flexibility. In this sample, ω = .95 at baseline.

The behavioral avoidance subscale of the Multidimensional Experiential Avoidance Questionnaire (MEAQ-BA; Gámez et al., 2011) measures the tendency to avoid situations that evoke discomfort and distress, corresponding to Module 5 (Countering Emotional Behaviors). Individuals rate (on a 6-point Likert Scale) the extent to which they agree or disagree (1 = *strongly disagree*, 6 = *strongly agree*) with 11 items (e.g., "I go out of my way to avoid uncomfortable situations"), with higher sum scores reflecting greater emotional avoidance. The MEAQ and its subscales have established good internal consistency, strong convergent associations, and discriminant validity with associated higher-order temperamental constructs (e.g., neuroticism; Gámez et al., 2011). In this sample, ω = .91 at baseline.

The Anxiety Sensitivity Index (ASI; Reiss et al., 1986) is a widely accepted measure of interoceptive sensitivity, which refers to the fear and distress accompanying the physical and cognitive symptoms associated with emotional responses. This measure was included to evaluate improvements in the experience of interoceptive sensitivity, mapping on to UP Module 6. In this 16-item measure, individuals rate the extent to which they agree with statements (e.g., "Unusual body sensations scare me") on a 5-point Likert-type scale (0 = *very little*, 5 = *very much*), with higher sum scores indicative of more anxiety sensitivity. The ASI has established good internal consistency and convergent validity (Peterson & Heilbronner, 1987; Vujanovic et al., 2007). In this sample, ω = .91 at baseline.

Symptom Outcomes

The Hamilton Anxiety Rating Scale (HAM-A; Hamilton, 1960) is a commonly used measure of symptoms of anxiety with demonstrated good levels of interrater and test-retest reliability, in addition to convergent validity with similar clinician-rated and self-report measures (Shear et al., 2001). Independent evaluators administered the measure accordance with the Structured Interview Guide for the Hamilton Anxiety (SIGH-A; Shear et al., 2001). Common symptoms of anxiety disorders (e.g., anxious mood) are rated on a 5-point Likert scale (0 = *no anxious mood*, 4 = *nearly constant anxiety*); responses for the 16 items are summed to create a total score. In this sample, ω = .84 at baseline.

Similarly, the Hamilton Depression Rating Scale (HAM-D; Hamilton, 1960) is an accepted measure of depressive symptoms administered by independent evaluators in accordance with the Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH-D; Williams, 1988). Common depressive symptoms (e.g., guilt) are rated on a 5-point Likert scale (0 = *no guilt*, 4 = *hears accusatory or denunciatory voices*); responses for the 17 items are summed to create a total score. The HAM-D has demonstrated good levels of inter-rater and test-retest reliability (Williams, 1988), as well as concurrent validity with similar clinician rated and self-report measures of depression symptoms (Bech et al., 1992). In this sample, ω = .81 at baseline.

PROCEDURE

All study procedures were approved by a university institutional review board (IRB) and written informed consent was obtained prior to any research activity. Data were derived from a randomized controlled trial comparing the UP to single-disorder protocols for the four principal anxiety conditions included in this study (e.g., *Mastery of Your Anxiety and Panic* – 4th Edition [MAP-IV]; Barlow & Craske, 2007), and a waitlist control (see Barlow et al., 2017 for details).

The present report focuses on participants who were randomized to the UP condition. Treatment with this intervention lasted 16 sessions (to be completed within a 21-week window) for individuals with a primary diagnosis of GAD, SOC, or OCD, and 12 sessions (to be completed within a 16-week window) for those with PD in order to be consistent with the duration of the single-disorder protocol associated with this diagnosis (i.e., MAP-IV). Regardless of treatment length, patients in the UP condition received all eight modules. Study therapists were doctoral students with 2 to 4 years of experience, postdoctoral fellows with 5 to 6 years of experience, and licensed doctoral-level psychologists with 10 or more years of experience, all of whom underwent training and certification in the UP before treating study patients. Treatment was monitored in weekly supervision meetings and 20% of session tapes were rated for adherence using a measure developed for this study by a protocol developer (TJF). Adherence ratings (M = 95.78, SD = 8.94, out of 100) indicated high levels of protocol fidelity.

Following the diagnostic assessment administered prior to randomization, participants completed the study assessments, including both self-report and clinician-rated measures, at pretreatment, following Sessions 4, 8, 12, and posttreatment. Although individuals who dropped out of UP treatment (n = 9) were contacted to complete the posttreatment assessment, none provided data

at this time point. Independent evaluators (IEs), who were blind to the patient's study condition, completed the Hamilton interviews. IEs were doctoral students with 2 to 4 years of experience and postdoctoral fellows with 5 to 6 years of experience. To test rater agreement, 20% of assessment tapes were re-rated and interrater reliability (i.e., Kappa) for the Hamilton measures was .92 (HAM-D) and .85 (HAM-A) in this sample.

DATA ANALYSIS

Latent growth curve models (LGM) were estimated using Mplus 7.2 (Muthén & Muthén, 1998). All LGMs were estimated in the intent-to-treat sample including 88 patients randomized to the UP. Missing data were accommodated using robust maximum likelihood estimation under a missing at random assumption (see Supplemental Tables 1–4 for tests of differences between individuals with complete and noncomplete outcome data). Model fit was evaluated using the Hu and Bentler (1999) guidelines. Unconditional LGMs were first estimated for each of the five UP skills under study, anxiety symptoms, and depression symptoms to determine whether there were individual differences in trajectories of change during UP treatment. In these single-process LGMs, the latent intercept was centered on the baseline assessment, the pretreatment slope loading was fixed at 0, the Session 4, 8, and 12 slope loadings were freely estimated (to fit nonlinear change), and the posttreatment slope loading was fixed at 1.0. Using this specification, the latent slope represents average change (slope mean) and individual differences in change (slope variance) in the UP skills/symptoms measure between baseline and end of treatment.

Next, we used parallel-process LGMs to explore the associations between initial changes in skills and overall changes in symptoms. Given our interest evaluating how initial changes in skills predict symptom improvement, latent slopes for the skill measures in the parallel-process LGMs were specified such that the Session 12 assessment was the final slope loadings (fixed at 1.0). This decision was made because (a) new treatment strategies were not introduced after Session 12, (b) most of the changes in skills occurred between baseline and Session 12 (see Table 3 slope factor loadings).

Results

OBSERVED CHANGE IN SKILLS AND SYMPTOMS

Means and standard deviations for each study variable at each assessment time point can be seen in Table 1. Additionally, change scores and effect sizes for each variable at the major assessment points (i.e., pretreatment, posttreatment) can be seen in Table 2; the standardized mean gain (ES_{sg}) was chosen as the effect size estimate as it includes a correction for repeated measures assessments. Consistent with our hypotheses, the skills associated with each UP module improved from pre- to posttreatment and demonstrated effects that were moderate to large in magnitude. The one exception was change in the DERS-A, reflecting emotional understanding (the skill targeted in Module 2), which worsened slightly from pretreatment to Session 4, then improved steadily thereafter, evidencing a small effect by posttreatment. Improvements in anxiety and depressive symptoms were also large.

SINGLE-PROCESS LGMS

A single-process unconditional LGM was estimated for each of the five skills, anxiety symptoms, and depression symptoms. In these models, the latent intercept was centered on the baseline assessment, the mid-treatment slope loadings were freely estimated (to fit nonlinear change, e.g., in emotional understanding), and the posttreatment slope loading was fixed at 1.0. Using this specification, the latent slope represents average change (slope mean) and individual differences in change (slope variance) in the skill/symptom measure between baseline and the end of treatment. The fit of the unconditional LGM for HAM-D was questionable (e.g., CFI = .91 but TLI = .87 and RMSEA = .13). The other six unconditional single-process LGMs provided acceptable model fit (e.g., nonsignificant chi-square and RMSEA CFit, CFI range = .97 to 1.0, TLI range = .95 to 1.0).

As shown in Table 3, there were significant improvements in all five skills (SMQ is scored opposite the other measures, hence the positive slope mean, SMQ slope mean = 19.35, $p < .001$; CCL-A = −5.60, $p < .001$; MEAQ-BA = −9.61, $p < .001$; ASI = −12.09, $p < .001$; DERS = A = −1.81, $p < .001$), anxiety symptoms (HAM-A slope mean = −7.45, $p < .001$), and depression symptoms over the course of UP treatment (HAM-D slope mean = −4.75, $p < .001$)[2].

[2] We also estimated conditional single-process LGMs to compare the eects of UP to waitlist-control (n = 44) on improvements in skills and symptoms (total n = 132). In each LGM, the latent slope was regressed onto a dummy variable representing treatment condition. The slope dierence scores indicated that patients randomized to receive UP had significantly greater improvements in each skill, anxiety symptoms, and depression symptoms than patients randomized to waitlist control (see Supplemental Table 5).

Table 1
Means and Standard Deviations for UP Skills and Outcomes

	Pretreatment	Session 4	Session 8	Session 12	Posttreatment
(Lack of) Emotional Understanding [DERS-A]	13.68 (4.30) $n = 73$	14.46 (4.39) $n = 68$	13.38 (4.58) $n = 66$	12.45 (4.47) $n = 68$	11.60 (4.44) $n = 60$
Mindful Emotion Awareness [SMQ]	36.45 (14.90) $n = 87$	37.23 (14.60) $n = 80$	46.58 (14.62) $n = 76$	52.60 (15.56) $n = 75$	55.92 (12.78) $n = 63$
(Lack of) Cognitive Flexibility [CCL-A]	13.56 (9.56) $n = 88$	11.20 (7.32) $n = 80$	9.80 (5.85) $n = 76$	8.67 (6.21) $n = 75$	8.31 (6.72) $n = 62$
Behavioral Avoidance [MEAQ-BA]	43.18 (9.96) $n = 73$	41.20 (9.62) $n = 68$	37.77 (9.80) $n = 66$	36.04 (10.67) $n = 68$	32.87 (9.45) $n = 60$
Anxiety Sensitivity [ASI]	30.81 (13.89) $n = 88$	29.24 (11.82) $n = 80$	24.08 (10.75) $n = 76$	22.40 (11.15) $n = 75$	19.63 (9.81) $n = 63$
Anxiety Symptoms [HAM-A]	17.08 (8.54) $n = 88$	15.31 (7.25) $n = 81$	14.20 (7.55) $n = 78$	12.48 (7.71) $n = 75$	9.34 (6.46) $n = 63$
Depressive Symptoms [HAM-D]	11.55 (7.06) $n = 88$	10.30 (5.60) $n = 81$	9.64 (5.95) $n = 78$	8.83 (6.36) $n = 75$	6.69 (4.98) $n = 63$

Note: DERS-A: Difficulties in Emotion Regulation Scale - Awareness Subscale, SMQ: Southampton Mindfulness Questionnaire, CCL-A: Cognition Checklist - Anxiety subscale, MEAQ-BA: Multidimensional Experiential Avoidance Questionnaire - Behavioral Avoidance subscale, ASI: Anxiety Sensitivity Index, HAM-A: Hamilton Anxiety Rating Sale, HAM-D: Hamilton Depression Rating Scale. For DERS-A, CCL-A, MEAQ-BA, and ASI, lower scores reflect greater skill use, whereas for SMQ, high scores reflect greater skill use. Higher scores reflect greater symptoms for HAM-A and HAM-D. DERS-A and MEAQ-BA were added to the questionnaire battery shortly after the start of data collection, underscoring lower *n*s for these variables.

Table 2
Mean Change Scores and Effect Sizes for UP Skills and Outcomes From Baseline to Posttreatment

	Change Score (Standard Deviation)	Standardized Mean Gain Effect size	Confidence Interval
Emotional Understanding [DERS-A]	−1.69 (3.99) $n = 54$	$ES_{sg} = -0.37$	(−0.61, −0.13)
Mindful Emotion Awareness [SMQ]	19.52 (19.61) $n = 63$	$ES_{sg} = 1.40$	(1.83, 0.98)
Cognitive Flexibility [CCL-A]	−5.64 (8.78) $n = 62$	$ES_{sg} = -0.75$	(−1.08, −0.43)
Behavioral Avoidance [MEAQ-BA]	−10.16 (11.34) $n = 54$	$ES_{sg} = -1.01$	(−1.37, −0.66)
Anxiety Sensitivity [ASI]	−12.76 (12.41) $n = 63$	$ES_{sg} = -1.09$	(−1.41, −0.77)
Anxiety Symptoms [HAM-A]	−7.32 (8.99) $n = 63$	$ES_{sg} = -0.96$	(−1.30, −0.62)
Depressive Symptoms [HAM-D]	−4.01 (6.08) $n = 63$	$ES_{sg} = -0.77$	(−1.05, −0.48)

Note: DERS-A: Difficulties in Emotion Regulation Scale - Awareness Subscale, SMQ: Southampton Mindfulness Questionnaire, CCL-A: Cognition Checklist - Anxiety subscale, MEAQ-BA: Multidimensional Experiential Avoidance Questionnaire - Behavioral Avoidance subscale, ASI: Anxiety Sensitivity Index, HAM-A: Hamilton Anxiety Rating Sale, HAM-D: Hamilton Depression Rating Scale. For DERS-A, CCL-A, MEAQ-BA, and ASI, lower scores reflect greater skill use (i.e., negative effect sizes are desirable), whereas for SMQ, high scores reflect greater skill use (i.e., positive effect sizes are desirable). Higher scores reflect greater symptoms for HAM-A and HAM-D (i.e., negative effect sizes are desirable). These analyses (i.e., calculation of change scores, ES_{sg}) apply only to posttreatment assessment completers.

Significant individual differences in trajectories of change were evident for all of the skill measures (SMQ slope variance = 284.11, $p < .001$; CCL-A = 43.30, $p = .006$; MEAQ-BA = 72.74, $p < .001$; ASI = 88.52, $p = .002$) except DERS-A (slope variance = 1.00, $p = .498$). There were also significant individual differences in trajectories of anxiety symptoms (HAM-A slope variance = 47.25, $p = .002$), but not depression symptoms (HAM-D slope variance = 13.19, $p = .478$). The unstandardized slope factor loadings suggest that 78–86% of total change in SMQ, CCL-A, MEAQ-BA, and ASI, and 56% of total change in DERS-A, occurred between baseline and Session 12 (cf. 52–61% of change in depression/anxiety symptoms). The significant negative intercept-slope correlations estimated for all measures except the DERS-A suggests that patients with the lower skills/greater symptoms at baseline (e.g., lower scores on SMQ; higher scores on CCL-A, MEAQ-BA, HAM-A, HAM-D) evidenced the greatest improvements in skills/symptoms over the course of UP.

PARALLEL-PROCESS LGMS

Given our interest in exploring whether change on UP skills predicts symptom change (i.e., initial changes in skills predicting subsequent and overall

Table 3
Estimates of Temporal Variation in Skills and Symptoms Over UP Treatment From Single-Process Univariate Latent Growth Models

Parameter estimate	DERS-A		SMQ		CCL-A		MEAQ-BA		ASI		HAM-A		HAM-D	
Intercept														
Mean (SE)	13.69***	(.53)	36.32***	(1.53)	13.65***	(1.07)	42.82***	(1.11)	30.98***	(1.45)	16.99***	(.93)	11.47***	(.91)
Variance (SE)	12.01***	(2.28)	163.39***	(30.47)	67.06***	(15.53)	77.77***	(13.97)	127***	(20.32)	49.55***	(9.65)	27.33***	(7.79)
Slope														
Mean (SE)	−1.81***	(.51)	19.35***	(2.23)	−5.60***	(1.01)	−9.61***	(1.35)	−12.09***	(1.43)	−7.45***	(1.09)	−4.75***	(.83)
Variance (SE)	1	(1.48)	284.11***	(69.99)	43.30**	(15.67)	72.74***	(19.84)	88.52**	(29.21)	47.25**	(15.21)	13.19	(18.57)
Factor loadings														
Baseline	0		0		0		0		0		0		0	
Session 4	− .47	(.39)	.06	(.05)	.33***	(.09)	.18*	(.08)	.20*	(.08)	.20**	(.08)	.20	(.26)
Session 8	.14	(.26)	.56***	(.06)	.66***	(.09)	.49***	(.10)	.57***	(.08)	.37***	(.1)	.37**	(.12)
Session 12	.56**	(.20)	.86***	(.06)	.81***	(.07)	.78***	(.10)	.78***	(.07)	.61***	(.11)	.52**	(.17)
Post-tx	1		1		1		1		1		1		1	
Intercept-Slope														
Covariance	.92	(1.17)	−128.61***	(40.5)	−40.26**	(14.84)	−30.91*	(14.73)	−63.40**	(22.06)	–32.36**	(11.52)	−12.97	(8.09)
Correlation	.27	(.43)	− .60***	(.1)	−.75***	(.1)	−.41**	(.14)	−.60***	(.11)	−.67***	(.11)	−.68*	(.32)

Note. DERS-A: Difficulties in Emotion Regulation Scale- Awareness subscale, SMQ: Southampton Mindfulness Questionnaire. CCL-A: Cognition Checklist - Anxiety subscale; MEAQ-BA: Multidimensional Experiential Avoidance Questionnaire- Behavioral Avoidance subscale, ASI: Anxiety Sensitivity Index, HAM-A: Hamilton Anxiety Rating Scale, HAM-D: Hamilton Depression Rating Scale. Overall fit of unconditional single-process growth models: DERS-A, $\chi 2$ (7) = 4.69, p = .698, root-mean-square error of approximation (RMSEA) = 0 (90% confidence interval [CI] = 0.00 - 0.11, p = .796), Tucker–Lewis index (TLI) = 1.00, comparative fit index (CFI) = 1.00; SMQ, χ 2 (7) = 11.11, p = 0.134, RMSEA = 0.08 (90% CI = 0.00 - 0.17, p = 0.245), TLI = 0.95, CFI = 0.97; CCL-A, χ 2 (7) = 10.06, p = 0.185, RMSEA = 0.07 (90% CI = 0.00 - 0.16, p = 0.312), TLI = 0.98, CFI = 0.99; MEAQ-BA, χ 2 (7) = 11.98, p = 0.101, RMSEA = 0.09 (90% CI = 0.00 - 0.18, p = 0.189), TLI = 0.96, CFI = 0.97; ASI, $\chi 2$ (7) = 8.97, p = 0.255, RMSEA = 0.06 (90% CI = 0.00 - 0.15, p = 0.395), TLI = 0.98, CFI = 0.99; HAM-A, χ 2 (7) = 3.91, p = 0.79, RMSEA = 0 (90% CI = 0.00 - 0.09, p = 0. 869), TLI = 1, CFI = 1; HAM-D, χ 2 (7) = 16.77, p = 0.019, RMSEA = 0.13 (90% CI = 0.05 - 0.21, p = 0. 053), TLI = .87, CFI = .91. S4 = assessment following Session 4; S8 = assessment following Session 8; S12 = assessment following Session 8; Post = assessment following the end of treatment.
* p < .05. ** p < .01. *** p < .001.

change in anxiety symptoms), the parallel-process latent slopes for the skills measure were specified such that the Session 12 assessment was the final slope loadings (fixed at 1.0). This decision was made because (a) new skills were not introduced after Session 12 (i.e., all skills were "taught" by Session 12), (b) most of the change in skills occurred between baseline and Session 12 (see Table 2 slope factor loadings). In comparison, change in anxiety symptoms was (still) specified to reflect change between baseline and posttreatment (i.e., posttreatment assessment treated as the final slope loading). Residual covariances between the skill measure and HAM-A and HAM-D at each assessment point were specified to capture time-specific covariations between the measures (e.g., CCL-A scores at Session 4 with anxiety symptoms at Session 4).

As there was not significant individual variation in trajectories of change for the DERS-A or HAM-D, and questionable fit for the HAM-D LGM, we do not report detailed results of parallel-process LGMs examining the associations between (a) change in emotional understanding and change in symptoms, or (b) change in any of the skills

Table 4
Completely Standardized Latent Correlations From Parallel-Processes Unconditional Latent Growth Models of Unified Protocol Skills and Anxiety Symptoms

Skill construct	HAM-A_{INT}	*p*	HAM-A_{SLP}	*p*
SMQ_{INT}	−.35**	.003	.23	.147
CCL-A_{INT}	.57***	<.001	−.51**	.009
MEAQ-BA_{INT}	.44***	<.001	−.24	.110
ASI_{INT}	.34**	.009	−.28	.199
SMQ_{SLP}	.16	.164	−.31*	.023
CCL-A_{SLP}	−.69**	<.001	.88***	<.001
MEAQ-BA_{SLP}	−.20	.278	.47*	.042
ASI_{SLP}	−.21	.371	.64*	.020

Note. HAM-A: Hamilton Anxiety Rating Scale, SMQ: Southampton Mindfulness Questionnaire, CCL-A: CognitionChecklist - Anxiety subscale, MEAQ-BA: Multidimensional Experiential Avoidance Questionnaire- BehavioralAvoidance subscale; ASI, Anxiety Sensitivity Index; INT, intercept; SLP, slope.
* p < .05. ** p < .01. *** p < .001.

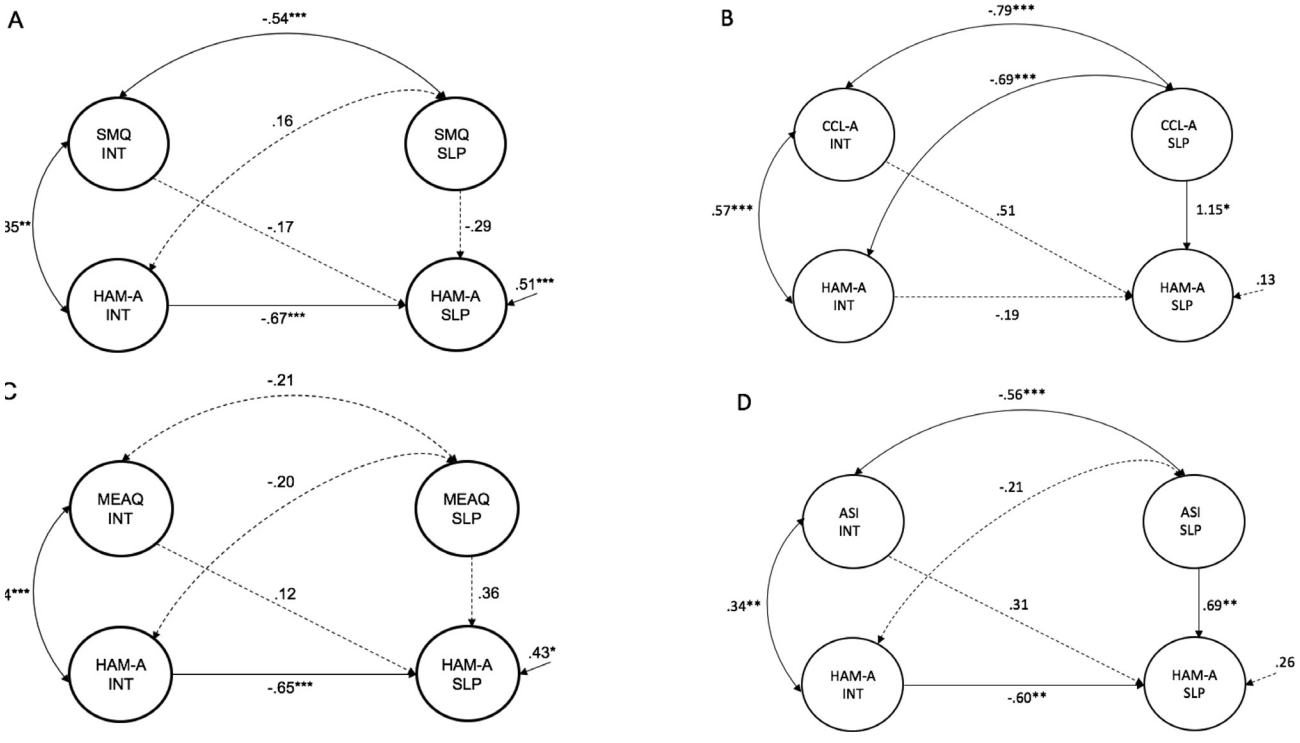

FIGURE 1 Parallel-process latent growth models of dimensions of mindful emotion awareness, cognitive flexibility, behavioral avoidance, anxiety sensitivity with anxiety symptoms. *Note*. Completely standardized estimates are presented. Nonsignificant paths and residual correlations are displayed with a dashed line. (A) Southampton Mindfulness Questionnaire (SMQ). (B) Cognition Checklist -Anxiety subscale (CCL-A). (C) Multidimensional Experiential Avoidance Questionnaire- Behavioral Avoidance subscale (MEAQ - BA). (D) Anxiety Sensitivity Index (ASI). Hamilton Anxiety Rating Scale (HAM-A). Overall fit of models: Model A, ^{2} (22) = 15.2, p = .856, root-mean-square error of approximation (RMSEA) = 0.00 (90% confidence interval [CI] = 0.00–0.05, p = .951), Tucker–Lewis Index (TLI) = 1.00, comparative fit index (CFI) = 1.00; Model B, ^{2} (22) = 20.1, p = .579, RMSEA = 0.00 (90% CI = 0.00–0.08, p = .798), TLI = 1.00, CFI = 1.00; Model C, ^{2} (22) = 9.81, p = . 988, RMSEA = 0.00 (90% CI = 0.00–0.00, p = .997), TLI = 1.00, CFI = 1.00. Model D, ^{2} (22) = 31.9, p = .08, RMSEA = 0.00 (90% CI = 0.00–0.12, p = .245), TLI = 0.95, CFI = 0.97. *p < .05, **p < .01, ***p < .001.

and change in depression symptoms. Nevertheless, parallel-process LGMs were estimated to confirm nonsignificant associations between (a) DERS-A slope and HAM-A slope, (b) DERS-A slope and HAM-D slope, and (c) slopes of the other four skills measures and HAM-D slope (see Supplemental Tables 6 and 7). The four parallel-process unconditional LGMs examining the associations between changes in the other four skills and changes in anxiety symptoms are reported below.

All four parallel-process LGMs each fit the data well (e.g., nonsignificant chi-square and RMSEA CFit, CFI range = .97 to 1.0, TLI range = .95 to 1.0, see Figure 1 note). Slope means and variances for each skill remained significant when modeled between baseline and Session 12 (i.e., the first three timepoints) and in parallel with change in HAM-A across all timepoints. Zero-order cross-sectional (intercept) and longitudinal (slope) correlations of each skill with anxiety symptoms are presented in Table 4. Change in each of the four skills was significantly correlated with change in anxiety symptoms such that improvements in each skill between baseline and Session 12 was associated with reductions in anxiety symptoms between baseline and posttreatment (SMQ and HAM-A slope $r = -.31$, $p = .023$; CCL-A and HAM-A $r = .88$, $p < .001$; MEAQ-BA and HAM-A $r = .47$, $p = .042$; ASI and HAM-A $r = .64$, $p = .020$). There were also significant intercept (baseline) correlations between each skill measure and anxiety symptoms (range of rs = −.35 to .57, ps < .01), and between CCL-A intercept and HAM-A slope ($r = -.51$, $p = .009$) and CCL-A slope and HAM-A intercept ($r = -.69$, $p < .001$).

Given the significant intercept-slope correlations within and across constructs, we wanted to evaluate the unique effects of change in skills on changes in symptoms over and above baseline levels of skills and symptoms. Accordingly, the four parallel-process LGMs were re-specified to regress the anxiety symptom slope factor onto the skills slope factor *and* skills and symptom intercept factors. The associations between the other latent factors were freely estimated. The completely standardized results of these models, which have identical fit to parallel-process LGMs with freely estimated slope factor correlations, are presented in Figure 1. Whereas significant unique associations between change in skills and change in anxiety symptoms were found for CCL-A ($\beta = 1.15$, $p = .019$) and ASI ($\beta = .69$, $p = .007$), marginally significant associations were found for SMQ ($\beta = -.29$, $p = .058$) and MEAQ-BA ($\beta = .36$, $p = .065$). Total variance (R^2) in HAM-A slope accounted for by changes in a skill (slope), and baseline levels of the skill and anxiety symptoms (intercepts), ranged from .49 (SMQ) to .87 (CCL-A).

Discussion

The present study sought to explore the extent to which skills associated with core UP modules change over the course of treatment, along with the role of skill acquisition in predicting symptom improvement. As hypothesized, we observed significant improvements on all five skills from pre- to posttreatment. Effects were large in magnitude for three of five skills, with the exception of cognitive flexibility (moderate) and emotional understanding (small). We also observed significant, large reductions in anxiety and depressive symptoms from pre- to posttreatment. Patients with greater skills deficits and more severe symptoms at baseline experienced the greatest improvements in skills and symptoms during UP treatment. Largely consistent with hypotheses, changes in four of five core UP skills (with the exception of DERS-A) were associated with changes in anxiety symptoms. Likely due to baseline floor effects and restricted variability in change over the course of UP, we could not find meaningful associations between change in skills and change in depression symptoms. Changes in cognitive flexibility and interoceptive tolerance (but not other skills, though mindful emotion awareness and countering emotional behaviors trended toward significance) were associated with changes in anxiety symptoms beyond baseline levels of skills and symptoms. Each of these findings warrants further commentary.

First, whereas earlier studies have examined changes in broader measures of emotion regulation, avoidance, etc. (e.g., Ellard, Barlow, Whitfield-Gabrieli, Gabrieli, & Deckersbach, 2017; Sauer-Zavala et al., 2012) or employed idiographic single-case experimental design methods to examine within-individual skills changes in response to selected UP modules (e.g., Bentley et al., 2017; Sauer-Zavala, Cassiello-Robbins, et al., 2017), our study is the first to demonstrate that significant changes in the therapeutic *skills specifically targeted by each core UP skills module* occur during treatment with this protocol. Moreover, changes in treatment skills were generally large in magnitude (with the exception of cognitive flexibility and emotional understanding, for which effects were moderate and small, respectively), indicating that the UP indeed engages its putative skills targets.

Next, we found that improvements in both skills and symptoms were larger in magnitude for

individuals with greater skills deficits and more severe anxiety symptoms (as measured by the clinician-rated HAM-A) at baseline. Previous research has been mixed with regard to whether pretreatment symptom severity predicts response. For example, some studies have indicated that higher baseline symptoms are not significantly associated with outcomes (e.g., Watanabe, 2010; Wolitzky-Taylor, Arch, Rosenfield, & Craske, 2012), others have demonstrated that they are predictive of *poorer* outcomes, and still others have reported steeper trajectories of improvement for individuals that start with greater severity (e.g., Brown & Barlow, 2009). Overall, the current findings suggest that at least for transdiagnostic, emotion-focused CBT, individuals with more "room for change" in both skills and anxiety symptoms before initiating treatment may experience greatest improvements in these domains, whereas a "ceiling effect" might exist for patients presenting with higher skills levels or lower anxiety symptoms.

Additionally, results from parallel-process latent growth models showed that earlier improvements in mindful emotion awareness, cognitive flexibility, behavioral avoidance, and interoceptive tolerance were significantly associated with reductions in anxiety by posttreatment. Establishing not only that the UP enacts change in its putative skills targets, but also that changes in these skills lead to subsequent changes in symptoms, is an important step in ensuring that the UP is maximally streamlined, efficacious, and thus, easily disseminated. Unfortunately, given the design of the study (i.e., the timing of measure administration) we could not definitively establish the temporal precedence of skill acquisition relative to symptoms. Additionally, given the limited variability in individual trajectories of change on symptoms of depression across treatment with the UP, we did not have the power to detect associations between skill acquisition and improvements in depression; this limitation is likely due to the fact that most participants in our trial experienced mild symptoms of depression (e.g., average pretreatment HAM-D score, 11.55, reflecting mild depression). Indeed, only 20 (of 88) individuals in the UP condition met criteria for depressive disorder (and these symptoms were relatively mild; Sauer-Zavala et al., in press). Additionally, the change score for HAM-D in the full UP sample was 4.01 points. This suggests that, for most patients who exhibited change on UP skills, improvement in depressive symptoms was limited due to restriction of range, likely accounting for the lack of relationship observed between change on UP skills and depressive symptoms in this sample.

We also observed that the magnitude of the relationship between skills changes and anxiety symptom change was the strongest for cognitive flexibility, followed by interoceptive sensitivity, behavioral avoidance, and mindful emotion awareness. Further, only changes in cognitive flexibility and anxiety sensitivity demonstrated statistically significant associations with changes in anxiety symptoms above and beyond baseline skills and symptoms. Our results also provide an initial indication that changes in cognitive flexibility and interoceptive sensitivity may be the most important UP skill targets, at least in terms of facilitating subsequent change on anxiety symptoms. Furthermore, contrary to expectations and previous research (e.g., Boswell et al., 2019), the Understanding Emotions module was not associated with changes in anxiety. This may be due to the fact that the items comprising the measure used to assess emotional understanding in the larger trial (DERS-A) (e.g., "I pay attention to how I feel") do not fully capture key tenets of Module 2, aimed to monitor the interactive aspects of emotional experience (e.g., antecedents, cognitions, physical sensations, behaviors, consequences). It is possible that a measure that gauges understanding of these nuances emphasized in Module 2 would result in significant effects (see Limitations).

Given that data for the current study came from a large randomized controlled trial in which all individuals in the UP condition were assigned to receive the full treatment package (Barlow et al., 2017), we were unable to experimentally manipulate the number of UP modules across patients, thus preventing us from determining whether all skills modules are needed to produce significant change in skills and symptoms. This line of research is another necessary step toward ensuring that the UP is as streamlined and time-efficient as possible; for example, if large improvements in skills and symptoms occur with three rather than five core UP modules, this would allow (1) patients to engage in (potentially costly and time-intensive) treatment for shorter periods of time, (2) therapists to treat a larger number of patients in need (e.g., shorten waitlists), and (3) the UP to potentially be more easily integrated into a broader range of clinical settings (e.g., less treatment content potentially requiring less expensive and time-intensive clinician training efforts). To this end, our team has conducted a series of preliminary studies to examine the unique effects of specific UP modules in patients with emotional disorders and co-occurring conditions (e.g., nonsuicidal self-injury)

(e.g., Bentley et al., 2017; Sauer-Zavala et al., 2017). Thus far, we have generally found that briefer, more targeted courses of treatment with selected UP modules lead to positive effects; however, these studies generally use smaller-scale single-case experimental designs and must be replicated in larger, more generalizable samples.

LIMITATIONS

Several limitations of the current study must be acknowledged. As noted above, the measures used to assess treatment skills in the trial were not developed specifically with the UP in mind and thus are imprecise measures of UP skills. Future studies of treatment skills that use measures developed specifically to capture the skills taught in the UP may result in more precise and nuanced effects. Along these lines, behavioral tasks or real-time data capture (e.g., ecological momentary assessment) may provide more accurate and ecologically valid information on changes in emotion management skills. Further, although the vast majority of skills changes occurred between baseline and Session 12, and changes in skills between baseline and Session 12 at least marginally predicted pre-post change in anxiety symptoms, our sample was also underpowered and assessments too infrequent to use more sophisticated techniques (e.g., latent difference score modeling) to definitively conclude if changes in skills *precede* and predict changes in symptoms (McArdle, 2009).

Another related limitation is that because UP therapists in the trial were allowed to flexibly determine how many sessions to spend on each of the modules and skills measures were only administered approximately every 4 weeks, we were unable to determine conclusively whether the specific UP module hypothesized to enact change in a particular skill was responsible for observed changes. In other words, we could not assume that by the Session 8 assessment, all patients had received the first four UP modules, for example, and not yet begun Module 5. Other limitations of the current study include the lack of sample generalizability (e.g., participants in the parent trial were generally well educated and rates of depression were low; Barlow et al., 2017) and UP delivery by therapists (all of whom were protocol developers or trained by protocol developers), potentially limiting the generalizability of observed results.

Despite these limitations, our study also has notable strengths. We used a large, diagnostically heterogeneous sample and conducted multiple repeated assessments of both skills and outcomes. Lastly, we employed advanced statistical methods that are well-suited for modeling change over time and analyzing longitudinal designs such as treatment outcome studies in which missing data are common.

CONCLUSIONS

This study suggests that during treatment with the UP, participants showed improvements in targeted skills for adaptive emotional responding: emotional understanding, mindful emotion awareness, cognitive flexibility, countering behavioral avoidance, and interoceptive tolerance. Changes in these treatment skills are associated with reductions in anxiety symptoms, suggesting that the UP may exert its effects on anxiety through improved emotion management. Future studies aimed to delineate the degree to which (and for whom) each core UP module is needed (i.e., actively engages these skills-based targets and leads to subsequent changes in emotional disorder symptoms) are needed to maximize the potential disseminability and personalized nature of this promising transdiagnostic approach.

Conflict of Interest Statement

This study was funded by grant R01 MH090053 from the National Institute of Mental Health Awarded to Dr. Barlow. Additionally, Dr. Barlow reported receiving royalties from Oxford University Press (which includes royalties for all 5 treatment manuals included in this study), Guilford Publications Inc, Cengage Learning, and Pearson Publishing; receiving grants from the National Institute of Mental Health, the National Institute of Alcohol and Alcohol Abuse, and Colciencias (Government of Columbia Initiative for Science, Technology, and Health Innovation); serving as a consultant for and receiving honoraria from the Agency for Healthcare Research and Quality, the Foundation for Informed Medical Decision Making, the Department of Defense, the Renfrew Center, the Chinese University of Hong Kong, Universidad Católica de Santa Maria (Arequipa, Peru), New Zealand Psychological Association, Hebrew University of Jerusalem, Mayo Clinic, and various American universities. Drs. Sauer-Zavala and Farchione reported receiving royalties from Oxford University Press for one of the treatment manuals included in this study. No other disclosures were reported.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.beth.2021.03.001.

References

Abramowitz, J. S., Tolin, D. F., & Street, G. P. (2001). Paradoxical effects of thought suppression: A meta-analysis of controlled studies. *Clinical Psychology Review, 21*(5), 683–703. https://doi.org/10.1016/S0272-7358(00)00057-X.

Allen, L. B., White, K. S., Barlow, D. H., Shear, M. K., Gorman, J. M., & Woods, S. W. (2010). Cognitive-behavior therapy (CBT) for panic disorder: Relationship of anxiety and depression comorbidity with treatment outcome. *Journal of Psychopathology and Behavioral Assessment, 32*(2), 185–192. https://doi.org/10.1007/s10862-009-9151-3.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5) (5th ed.). APA.

Andrews, G. (1990). Classification of neurotic disorders. *Journal of the Royal Society of Medicine, 83*(10), 606–607. https://doi.org/10.1192/bjp.157.1.6.

Andrews, G. (1996). Comorbidity and the general neurotic syndrome. *The British Journal of Psychiatry. Supplement, 30*, 76–84. https://doi.org/10.1192/S0007125000298449.

Barlow, D. H. (1991). Disorders of emotion. *Psychological Inquiry, 2*(1), 58–71. https://doi.org/10.1207/s15327965pli0201_15.

Barlow, D. H., & Craske, M. G. (2007). *Mastery of your anxiety and panic*: Workbook. Oxford University Press.

Barlow, D. H., Farchione, T. J., Bullis, J. R., Gallagher, M. W., Murray-Latin, H., Sauer-Zavala, S., ... Cassiello-Robbins, C. (2017). The unified protocol for transdiagnostic treatment of emotional disorders compared with diagnosis-specific protocols for anxiety disorders: A randomized clinical trial. *JAMA Psychiatry, 74*(9), 875–884. https://doi.org/10.1001/jamapsychiatry.2017.2164.

Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2000). Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: A randomized controlled trial JAMA, 283(19), 2529-2536. https://doi:10.1001/jama.283.19.2529

Barlow, D. H., Sauer-Zavala, S., Carl, J. R., Bullis, J. R., & Ellard, K. K. (2014). The nature, diagnosis, and treatment of neuroticism: Back to the future. *Clinical Psychological Science, 2*(3), 344–365. https://doi.org/10.1177/2167702613505532.

Barlow, D. H., Sauer-Zavala, S., Farchione, T. J., Murray Latin, H., Ellard, K. K., Bullis, J. R., ... Cassiello-Robbins, C. (2018). *Unified protocol for the transdiagnostic treatment of emotional disorders: Patient workbook* (2nd ed.). Oxford University Press.

Bech, P., Allerup, P., Maier, W., Albus, M., Lavori, P., & Ayuso, J. L. (1992). The hamilton scales and the hopkins symptom checklist (SCL–90): A cross-national validity study in patients with panic disorders. *The British Journal of Psychiatry, 160*(2), 206–211. https://doi.org/10.1192/bjp.160.2.206.

Beck, A. T., Brown, G., Steer, R. A., Eidelson, J. I., & Riskind, J. H. (1987). Differentiating anxiety and depression: A test of the cognitive content-specificity hypothesis. *Journal of Abnormal Psychology, 96*(3), 179–183.

Bentley, K. H., Nock, M. K., Sauer-Zavala, S., Gorman, B. S., & Barlow, D. H. (2017). A functional analysis of two transdiagnostic, emotion-focused interventions on nonsuicidal self-injury. *Journal of Consulting and Clinical Psychology, 85*(6), 632–646. https://doi.org/10.1037/ccp0000205.

Boswell, J. F., Anderson, L. M., Oswald, J. M., Reilly, E. E., Gorrell, S., & Anderson, D. A. (2019). A preliminary naturalistic clinical case series study of the feasibility and impact of interoceptive exposure for eating disorders. *Behaviour Research and Therapy, 117*, 54–64.

Brown, T. A. (2007). Temporal course and structural relationships among dimensions of temperament and DSM-IV anxiety and mood disorder constructs. *Journal of Abnormal Psychology, 116*(2), 313–328. https://doi.org/10.1037/0021-843X.116.2.313.

Brown, T. A., & Barlow, D. H. (2009). A proposal for a dimensional classification system based on the shared features of the DSM-IV anxiety and mood disorders: Implications for assessment and treatment. *Psychological Assessment, 21*(3), 256–271. https://doi.org/10.1037/a0016608.

Brown, T. A., & Barlow, D. H. (2014). Anxiety and related disorders interview schedule for DSM-5 (ADIS-5L).: Lifetime version. *Client Interview Schedule*. Oxford University Press.

Brown, T. A., Barlow, D. H., & DiNardo, P. A. (1994). *Anxiety disorders interview schedule for DSM - IV (ADIS - IV): Client interview schedule*. Oxford University Press.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology, 110*(4), 585–599. https://doi.org/10.1037/0021-843X.110.4.585.

Bullis, J. R., Boettcher, H. T., Sauer-Zavala, S., & Barlow, D. H. (in press). What is an emotional disorder. *Clinical Psychology: Science & Practice*. https://doi.org/10.1111/cpsp.12278

Chadwick, P., Hember, M., Symes, J., Peters, E., Kuipers, E., & Dagnan, D. (2008). Responding mindfully to unpleasant thoughts and images: Reliability and validity of the Southampton mindfulness questionnaire (SMQ). *British Journal of Clinical Psychology, 47*(4), 451–455. https://doi.org/10.1348/014466508X314891.

Ellard, K. K., Barlow, D. H., Whitfield-Gabrieli, S., Gabrieli, J. D., & Deckersbach, T. (2017). Neural correlates of emotion acceptance vs worry or suppression in generalized anxiety disorder. *Social Cognitive and Affective Neuroscience, 12*(6), 1009–1021.

Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., & Watson, D. (2011). Development of a measure of experiential avoidance: The multidimensional experiential avoidance questionnaire. *Psychological Assessment, 23*(3), 692–713. https://doi.org/10.1037/a0023242.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41–54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94.

Griffith, J. W., Zinbarg, R. E., Craske, M. G., Mineka, S., Rose, R. D., Waters, A. M., & Sutton, J. M. (2010). Neuroticism as a common dimension in the internalizing disorders. *Psychological Medicine, 40*(7), 1125–1136. https://doi.org/10.1017/S0033291709991449.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348–362. https://doi.org/10.1037/0022-3514.85.2.348.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23*(1), 56–61.

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But... *Communication Methods and Measures, 14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629.

Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research, 4*(4), 383–395. https://doi.org/10.1007/BF01178214.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research domain criteria

(RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry, 167*(7), 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379.

Jarvi Steele, S. J., Farchione, T. J., Cassiello-Robbins, C., Ametaj, A. A., Sbi, S., Sauer-Zavala, S., & Barlow, D. H. (2018). Efficacy of the Unified Protocol for transdiagnostic treatment of comorbid psychopathology accompanying emotional disorders compared to treatments targeting single disorders. *Journal of Psychiatric Research, 104*, 211–216. https://doi.org/10.1016/j.jpsychires.2018.08.005.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605.

Muthén, L. K., & Muthén, B. O. (1998). Mplus 8.0 [Computer software]. Author

Peterson, R. A., & Heilbronner, R. L. (1987). The anxiety sensitivity index: Construct validity and factor analytic structure. *Journal of Anxiety Disorders, 1*(2), 112–121. https://doi.org/10.1016/0887-6185(87)90002-8.

Purdon, C. (1999). Thought suppression and psychopathology. *Behaviour Research and Therapy, 37*(11), 1029–1054. https://doi.org/10.1016/S0005-7967(98)00200-9.

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy, 24*(1), 1–8. https://doi.org/10.1016/0005-7967(86)90143-9.

Sauer-Zavala, S., Boswell, J. F., Gallagher, M. W., Bentley, K. H., Ametaj, A., & Barlow, D. H. (2012). The role of negative affectivity and negative reactivity to emotions in predicting outcomes in the unified protocol for the transdiagnostic treatment of emotional disorders. *Behaviour Research and Therapy, 50*(9), 551–557. https://doi.org/10.1016/j.brat.2012.05.005.

Sauer-Zavala, S., Cassiello-Robbins, C., Conklin, L. R., Bullis, J. R., Thompson-Hollands, J., & Kennedy, K. A. (2017). Isolating the unique effects of the Unified Protocol treatment modules using single case experimental design. *Behavior Modification, 41*(2), 286–307. https://doi.org/10.1177/0145445516673827.

Sauer-Zavala, S., Fournier, J. C., Steele, S. J., Woods, B. K., Wang, M., Farchione, T. J., and Barlow, D. H. (in press). Does the unified protocol really change neuroticism? Results from a randomized trial. *Psychological Medicine*, 1–10.

Sauer-Zavala, S., Gutner, C. A., Farchione, T. J., Boettcher, H. T., Bullis, J. R., & Barlow, D. H. (2017). Current definitions of "transdiagnostic" in treatment development: A search for consensus. *Behavior Therapy, 48*(1), 128–138. https://doi.org/10.1016/j.beth.2016.09.004.

Shear, M. K., Vander Bilt, J., Rucci, P., Endicott, J., Lydiard, B., Otto, M. W., ... Frank, D. M. (2001). Reliability and validity of a structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-A). *Depression and Anxiety, 13*(4), 166–178. https://doi.org/10.1002/da.1033.

Steer, R. A., Beck, A. T., Clark, D. A., & Beck, J. S. (1994). Psychometric properties of the cognition checklist with psychiatric outpatients and university students. *Psychological Assessment, 6*(1), 67–70. https://doi.org/10.1037/1040-3590.6.1.67.

Taylor, S., Woody, S., Koch, W. J., McLean, P., Paterson, R. J., & Anderson, K. W. (1997). Cognitive restructuring in the treatment of social phobia: Efficacy and mode of action. *Behavior Modification, 21*(4), 487–511. https://doi.org/10.1177/01454455970214006.

Vujanovic, A. A., Arrindell, W. A., Bernstein, A., Norton, P. J., & Zvolensky, M. J. (2007). Sixteen-item anxiety sensitivity index: Confirmatory factor analytic evidence, internal consistency, and construct validity in a young adult sample from the Netherlands. *Assessment, 14*(2), 129–143. https://doi.org/10.1177/1073191106295053.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *The Journal of Machine Learning Research, 11*, 3571–3594.

Wolitzky-Taylor, K., Arch, J., Rosenfield, D., & Craske, M. (2012). Moderators and non-specific predictors of treatment outcome for anxiety disorders: A comparison of cognitive behavioral therapy to acceptance and commitment therapy. *Journal of Consulting and Clinical Psychology, 80*(5), 786–799.

Williams, J. B. (1988). A structured interview guide for the Hamilton depression rating scale. *Archives of General Psychiatry, 45*(8), 742–747. https://doi.org/10.1001/archpsyc.1988.01800320058007.

Zinbarg, R. E., Mineka, S., Boboba, L., Craske, M. G., Vrshek-Schallhorn, S., Griffith, J. W., ... Anand, D. (2016). Testing a hierarchical model of neuroticism and its cognitive facets: Latent structure and prospective prediction of first onsets of anxiety and unipolar mood disorders during 3 years in late adolescence. *Clinical Psychological Science, 4*(5), 805–824. https://doi.org/10.1177/2167702615618162.

Received: October 7, 2020
Accepted: March 2, 2021
Available online: 11 March 2021