A SMART approach to personalized care: preliminary data on how to select and sequence skills in transdiagnostic CBT

Shannon Sauer-Zavala, Matthew W. Southward, Nicole E. Stumpp, Stephen A. Semcho, Caitlyn O. Hood, Anna Garlock & Alex Urs

Published online: 27 Apr 2022.

Submit your article to this journal ⤤

View related articles ⤤

View Crossmark data ⤤

Routledge
Taylor & Francis Group

Check for updates

A SMART approach to personalized care: preliminary data on how to select and sequence skills in transdiagnostic CBT

Shannon Sauer-Zavala ⓘ, Matthew W. Southward ⓘ, Nicole E. Stumpp ⓘ, Stephen A. Semcho ⓘ, Caitlyn O. Hood ⓘ, Anna Garlock ⓘ and Alex Urs ⓘ

Department of Psychology, University of Kentucky, Lexington, Kentucky 40506, US

ABSTRACT

Given that over 20 million adults each year do not receive care for their mental health difficulties, it is imperative to improve system-level capacity issues by increasing treatment efficiency. The present study aimed to collect feasibility/acceptability data on two strategies for increasing the efficiency of cognitive behavioral therapy: (1) personalized skill sequences and (2) personalized skill selections. Participants (N = 70) with anxiety and depressive disorders were enrolled in a pilot sequential multiple assignment randomized trial (SMART). Patients were randomly assigned to receive skill modules from the Unified Protocol in one of three sequencing conditions: standard, sequences that prioritized patients' relative strengths, and sequences that prioritized relative deficits. Participants also underwent a second-stage randomization to either receive 6 sessions or 12 sessions of treatment. Participants were generally satisfied with the treatment they received, though significant differences favored the Capitalization and Full duration conditions. There were no differences in trajectories of improvement as a function of sequencing condition. There were also no differences in end-of-study outcomes between brief personalized treatment and full standard treatment. Thus, it may be feasible to deliver CBT for personalized durations, though this may not substantially impact trajectories of change in anxiety or depressive symptoms.

Introduction

Cognitive behavioral therapy (CBT) is an efficacious treatment for anxiety and depressive disorders (Hofmann & Smits, 2008; Van Straten et al., 2010). Most CBT protocols involve at least 12–16 sessions (e.g. Craske & Barlow, 2006), in line with dose-response evaluations suggesting that approximately 13 treatment sessions are needed to observe improvement in two-thirds of patients (Garfield, 1994; Hansen et al., 2002). In contrast, patients in community practice attend less than five sessions on average (Harnett et al., 2010), suggesting that many individuals may not fully benefit from treatment. Indeed, despite the availability of efficacious intervention protocols, prevalence rates for anxiety and depressive disorders remain persistently high (Bandelow & Michaelis, 2015) and may be rising (Weinberger et al., 2017).

CONTACT Shannon Sauer-Zavala ✉ ssz@uky.edu 🖳 Department of Psychology, University of Kentucky, Lexington, Kentucky 40506, USA

Given the high rates of comorbidity among anxiety and depressive disorders (Kessler et al., 1996, 1998), transdiagnostic interventions that simultaneously address symptoms across a range of conditions may represent a more efficient treatment approach (McHugh et al., 2009). Gold-standard CBT protocols typically focus on a single diagnosis (e.g. Craske & Barlow, 2006), often requiring co-occurring conditions to be treated sequentially and thereby extending the length of care. In contrast, transdiagnostic interventions concurrently address symptoms of multiple disorders by targeting shared, underlying processes that maintain symptoms (Sauer-Zavala, Gutner, et al., 2017). Ample research suggests that aversive reactions to frequently occurring negative emotions may explain the pervasive pattern of comorbidity among anxiety and depressive disorders (e.g. Barlow et al., 2014; Bullis et al., 2019). Rather than treat heterogeneous expressions of this shared vulnerability (i.e. discrete disorder symptoms such as panic attacks, worry episodes, social withdrawal), aversive reactivity to emotions can serve as the primary target in transdiagnostic interventions (Sauer-Zavala, Southward, et al., 2022).

The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders (UP; Barlow et al., 2018) represents one of the best known treatments for emotional disorders (i.e. anxiety, depressive, and related disorders). The UP includes five discrete CBT skills designed to directly target the aversive, avoidant responses to emotional experiences that maintain symptoms across emotional disorders. There is promising empirical support for the UP across a variety of settings (e.g. inpatient, outpatient), formats (e.g. individual, group), and disorders (Cassiello-Robbins et al., 2020), with large improvements demonstrated for anxiety and depressive symptoms (Sakiris & Berle, 2019). Additionally, the UP is associated with reductions in aversive reactivity to emotions (Boswell et al., 2013; Eustis et al., 2019; Sauer-Zavala et al., 2012). However, the majority of support for the UP's efficacy comes from trials that provided 12–20 sessions (though some studies have tested versions of the UP with as few as 5 sessions [e.g. Bentley et al., 2017; Sauer-Zavala et al., 2019], underscoring the need for novel approaches to increase treatment efficiency (Southward et al., 2020).

Personalized CBT skill selection/sequencing

Personalizing the delivery of an intervention such that patients receive only the treatment components that best fit with their presentations, referred to as a modular approach (Chorpita et al., 2005), may represent a way to further increase the efficiency of transdiagnostic CBT. Modular interventions circumvent the need to work through an entire treatment protocol, given that some components may not apply to a given patient; modular approaches are associated with steeper trajectories of improvement compared to traditional manualized care (Weisz et al., 2012). The greatest gains, however, may result from the integration of a personalized, modular approach with a transdiagnostic intervention. For example, given that the UP consists of multiple skills designed to target the same core vulnerability of aversive reactivity to emotions, it is possible that some skills may be more or less robust at targeting this vulnerability for a particular patient.

Recent data indicates that each component of the UP independently engages its associated skill when presented in isolation (Sauer-Zavala, Cassiello-Robbins et al., 2017; Sauer-Zavala, Cassiello-Robbins, et al., 2020), suggesting that the UP need not be presented in its standard order to enact change—a prerequisite for modular delivery. In

light of these findings, researchers have begun to explore methods for personalizing the selection and sequencing of UP skill modules. Using data from 30 daily assessments collected prior to Fisher and Boswell (2016), Fisher et al. (2019)) conducted person-specific factor analyses to determine predominant pathological dimensions (e.g. worry, behavioral avoidance) and temporal relations among these dimensions for each patient. UP modules corresponding to each patient's presentation were then selected and sequenced to prioritize symptoms that preceded other difficulties. Although treatment personalized using these methods resulted in symptom improvements, a control condition in which the UP was presented in its standard order was not utilized, making it difficult to determine if a dynamic assessment and modeling approach to personalization is more efficient than treatment as usual.

Another approach to personalizing the sequencing of UP modules is to focus on pre-existing skill capacities rather than psychopathological deficits. In other words, evaluating each patient's relative strengths and deficits in specific skills at baseline could be used to individualize the sequence of those skills (Cheavens et al., 2012). A capitalization model suggests that skill sequences should prioritize a patient's relative strengths, whereas a compensation model prioritizes areas of greatest deficit. Preliminary data suggests that, for patients with major depressive disorder receiving CBT skills (e.g. cognitive restructuring, behavioral activation, mindfulness), capitalizing on existing strengths is associated with steeper trajectories of improvement relative to compensating for deficits (Cheavens et al., 2012). In a recent pilot evaluation of the capitalization/compensation approach to personalization with the UP, Sauer-Zavala et al. (2019) found that clinically significant intraindividual strengths and deficits at baseline could be determined using empirically validated questionnaires corresponding to each UP skill, and that patients were satisfied with personalized treatment sequences. This pattern of results indicates that sequencing UP modules according to skill level is feasible, yet studies with larger samples and a standard-order comparison group are necessary to determine whether this approach to personalization leads to more efficient improvements.

Present study

The purpose of the present study was to conduct a pilot sequential multiple assignment randomized trial (SMART; Collins et al., 2007) to collect (a) feasibility/acceptability data on personalized sequencing/skill selection of the UP, along preliminary data on whether (a) personalizing the sequences of UP treatment modules results in steeper trajectories of symptom improvement and (b) a personalized selection of modules leads to similar symptom reduction as a full course of care. In the first-stage randomization, patients were assigned to one of three UP sequencing orders: (1) Standard, (2) Capitalization, or (3) Compensation. Participants in the Standard condition received each UP module in the order described in the published manual, whereas the Capitalization condition prioritized modules that focused on patients' relative strengths and the Compensation condition prioritized modules that focused on patients' relative deficits. We sought to expand on previous work on personalized UP sequencing (Fisher et al., 2019; Fisher & Boswell, 2016; Sauer-Zavala et al., 2019) by comparing patients' module orders to the standard sequence. Given previous data favoring a capitalization approach (Cheavens

et al., 2012; Sauer-Zavala et al., 2019), we hypothesized that patients in the Capitalization condition would demonstrate steeper (i.e. more efficient) trajectories of symptom improvement than patients in the Compensation and Standard conditions.

In order to test whether a personalized selection of UP modules exerts more robust effects for individual patients relative to the standard delivery of all possible UP skills, participants in this study underwent a second-stage randomization following session in which they were randomly assigned to discontinue treatment immediately (i.e. Brief Condition) after their next session or after receiving the full UP treatment (i.e. all possible UP modules; Full Condition). Finally, regardless of whether this innovation (i.e. personalized skill sequencing) increased treatment efficiency, the treatment must be acceptable to patients to be widely implemented in community practice. Accordingly, we also compared patient ratings of acceptability/satisfaction among sequencing (i.e. Standard, Capitalization, Compensation) and discontinuation conditions (i.e. Brief, Full).

Methods & materials

Participants

A sample of treatment-seeking adult patients was recruited from the Clinic for Emotional Health at the University of Kentucky. Individuals were eligible for the study if they met DSM-5 (American Psychiatric Association, 2013) criteria for at least one of the following emotional disorders: panic disorder (PD), generalized anxiety disorder (GAD), social anxiety disorder (SAD), obsessive-compulsive disorder (OCD), posttraumatic stress disorder (PTSD), major depressive disorder (MDD), or persistent depressive disorder (PDD). Exclusion criteria included diagnoses or symptoms requiring clinical prioritization or hospitalization: specifically, individuals who endorsed mania within the past year (i.e. uncontrolled bipolar disorder), acute suicide risk (i.e. imminent intent), substance use disorder within the last 3 months, or psychotic features. Individuals were also excluded if they received five or more sessions of CBT within the last five years. Anyone receiving other psychotherapy focused on an emotional disorder agreed to discontinue their treatment before participating in the study. Individuals taking psychotropic medication were asked to maintain their current dosages during study participation. The study was approved by the University of Kentucky Institutional Review Board, and informed consent was obtained prior to any research activity.

A total of 70 participants consented to participate in the study (see, Figure 1 for study flow). Participants were 33.74 ($SD = 12.64$) years old on average and the majority (67.1%) of the sample identified as female ($n = 47$), white ($n = 49$; 70.0%), and heterosexual ($n = 52$; 74.3%). The most common clinically significant diagnosis (i.e. rated as most distressing/interfering) was GAD ($n = 45$; 64.3%), followed by SAD ($n = 35$; 50.0%), and MDD ($n = 35$; 50.0%). On average, participants met criteria for three concurrent diagnoses at baseline (Table 1).

Of the initial 70 participants, 26 (37.1%) were assigned to the Standard condition, 23 (32.9%) were assigned to the Capitalization condition, and 21 (30.0%) were assigned to the Compensation condition (Figure 1). Eleven participants (15.7%) did not complete study procedures. Reasons for withdrawal included being lost to contact following baseline assessment ($n = 2$; 2.9%), an inability to attend regular
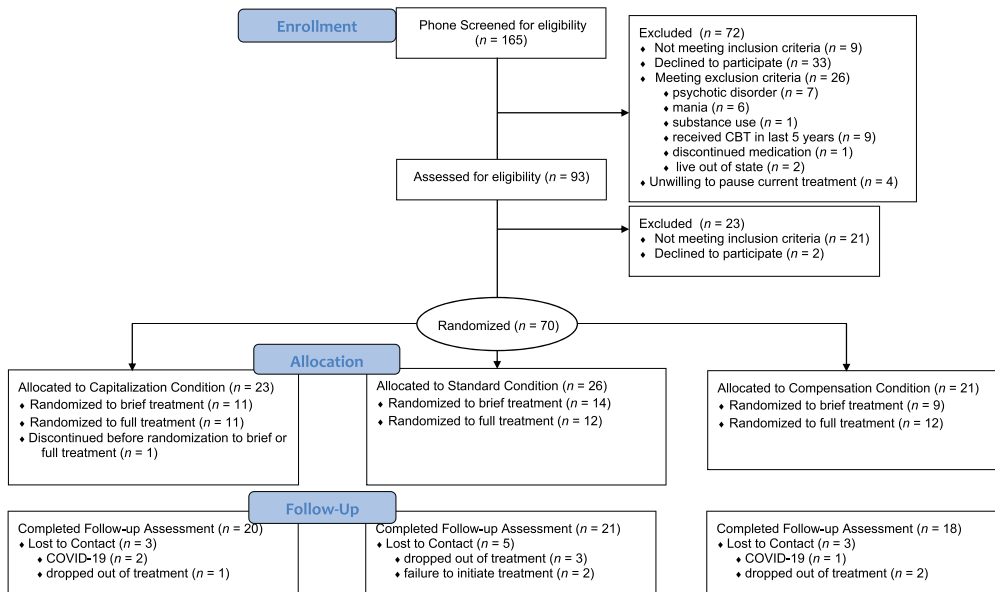
Figure 1. Recruitment flow diagram.

therapy sessions ($n = 6$; 8.6%), and discontinuation due to COVID-19 (i.e. patient moved out-of-state and could not receive telehealth sessions; $n = 3$; 4.3%). Of these withdrawals, five were assigned to the Standard condition, and three each were assigned to the Capitalization and Compensation conditions, respectively. Thus, complete posttreatment data is available for 59 participants: 21 (35.6%) assigned to the Standard condition, 20 (33.4%) assigned to the Capitalization condition, and 18 (30.5%) assigned to the Compensation condition. There were no demographic differences at baseline between study completers and those that were withdrawn or dropped out, $ps > .05$. Finally, due to the COVID-19 pandemic, participants recruited prior to 15 March 2020 ($n = 29$; 41.4%) completed at least some of their study visits in-person, whereas those enrolled after that date ($n = 41$; 58.6%) participated in all aspects of the study remotely. The only demographic difference between those who completed the study partially in-person or fully remotely was a larger proportion of participants who identified as heterosexual completed the study fully remotely, ($n = 34$; 82.9%) than partially in-person ($n = 18$; 62.1%), $\chi^2(1) = 3.87$, $p = .049$.

Study treatment

The treatment modules provided in the present study were drawn from the UP (Barlow et al., 2018). We included five core skills designed to engage the UP's putative mechanism, aversive reactivity to emotions: Understanding Emotions, Mindful Emotion Awareness, Cognitive Flexibility, Countering Emotional Behaviors,[1] and Confronting Physical Sensations. For a full description of the UP modules, see, Payne et al. (2014).

Table 1. Baseline demographic and diagnostic characteristics.

Characteristic	Total(N = 70)	Compensation Condition (n = 21)	Standard Condition (n = 26)	Capitalization Condition (n = 23)
Age (Mean, SD)	33.74 (12.64)	33.71 (13.96)	32.88 (12.85)	34.74 (11.59)
Gender				
Female	47 (67.1)	13 (61.9)	18 (69.2)	16 (69.6)
Male	22 (31.4)	8 (38.1)	7 (26.9)	7 (30.4)
Genderqueer/Non-binary	1 (1.4)	0 (0.0)	1 (3.8)	0 (0.0)
Racial/Ethnic Background[a]				
Caucasian	52 (74.3)	17 (81.0)	18 (69.2)	17 (73.9)
African-American	9 (12.9)	2 (9.5)	5 (19.2)	2 (8.7)
Arab/Middle-Eastern American	2 (2.9)	2 (9.5)	0 (0.0)	0 (0.0)
East Asian	3 (4.3)	0 (0.0)	1 (4.0)	2 (8.7)
Latinx	2 (2.9)	0 (0.0)	1 (4.0)	1 (4.2)
South Asian	2 (2.9)	0 (0.0)	1 (4.0)	1 (4.2)
Heterosexual/Straight	52 (74.3)	13 (61.9)	21 (80.8)	18 (78.3)
Bachelor's Degree or Higher	42 (60.0)	13 (61.9)	10 (38.5)	19 (82.6)
Married	23 (32.9)	4 (19.0)	7 (26.9)	12 (52.2)
Current Psychotropic Medication	16 (22.9)	5 (23.8)	7 (26.9)	4 (17.4)
Clinically Significant Diagnoses[b]				
Obsessive-Compulsive Disorder	5 (12.9)	4 (19.0)	4 (15.4)	1 (4.3)
Social Anxiety Disorder	35 (50.0)	11 (52.4)	16 (61.5)	8 (34.8)
Generalized Anxiety Disorder	45 (64.3)	9 (42.9)	18 (69.2)	18 (78.3)
Panic Disorder	11 (15.7)	4 (19.0)	5 (19.2)	2 (8.7)
Agoraphobia	6 (8.6)	3 (14.3)	2 (7.7)	1 (4.3)
Major Depressive Disorder	35 (50.0)	10 (47.6)	16 (61.5)	9 (39.1)
Persistent Depressive Disorder	17 (24.3)	7 (33.3)	9 (34.6)	1 (4.3)
Acute Stress Disorder	1 (1.4)	1 (4.8)	0 (0.0)	0 (0.0)
Posttraumatic Stress Disorder	8 (11.4)	2 (9.5)	4 (15.4)	2 (8.7)
Diagnoses Met (M, SD)	3.01 (1.81)	3.14 (2.29)	3.42 (1.68)	2.43 (1.31)
Clinical Severity Rating (CSR; M, SD)	4.70 (1.01)	4.48 (1.12)	4.96 (1.00)	4.61 (0.89)
OASIS (M, SD)	9.13 (3.67)	9.67 (3.53)	9.00 (4.59)	8.78 (2.63)
ODSIS (M, SD)	8.28 (5.05)	9.00 (4.96)	9.32 (5.03)	6.48 (4.88)

Data are presented as number (percentage) of patients unless otherwise indicated. [a]Values may not sum to total in each column because participants could select multiple racial/ethnic backgrounds. [b]Values may not sum to total in each column because participants could be diagnosed with multiple clinically significant diagnoses. OASIS = Overall Anxiety Severity and Impairment Scale. ODSIS = Overall Depression Severity and Impairment Scale.

Modules were delivered in weekly, individual, 45–60 minute sessions. All modules consisted of two individual sessions except Countering Emotional Behaviors, which was delivered across four sessions. Four study therapists provided the treatment: a licensed clinical psychologist, a post-doctoral fellow, and two advanced clinical psychology graduate students who were certified in the provision of the UP by one of its developers. All sessions were audio recorded, and 20% were randomly selected to be rated for competence on a 5-point scale. Average competence, which consisted of fidelity to the treatment protocol and therapeutic skill (e.g. time management, empathy) was high ($M = 4.26$, $SD = .54$). There were no differences in competence between study sessions completed in-person ($M = 4.26$, $SD = .50$) or via telehealth ($M = 4.26$, $SD = .59$), $t(44) = .05$, $p = .96$, 95% CI [–.33, .35].

Study design

We conducted a SMART with a two-stage randomization. The first-stage randomization occurred following the baseline assessment. Patients were assigned to one of three sequencing conditions: Standard, Compensation, or Capitalization. In the Standard condition, patients received the UP modules according to its published

order: Understanding Emotions, Mindful Emotion Awareness, Cognitive Flexibility, Countering Emotional Behaviors, and Confronting Physical Sensations. In the Compensation condition, patients received modules in an order that prioritized relative skill deficits, whereas participants in the Capitalization condition received modules in an order that prioritized their relative skill aptitudes. Procedures for module ordering are described in the *Module Sequencing* section.

Patients underwent a second randomization between their fifth and sixth sessions in which they were assigned to receive Brief treatment or Full treatment. Those in the Brief treatment condition discontinued care after their next session for a total of six sessions of treatment, whereas those in the Full condition completed a total of 12 sessions. During consenting, patients were informed about the timing of the second-stage randomization (i.e. they knew they would be alerted to whether they were in the Brief or Full condition following session 5), and study therapists also were masked to condition until this point. The randomization occurred between sessions 5 and 6 so that knowledge of treatment duration would not influence patient or therapist behaviors, and to allow for a final session to complete the current skill for those in the Brief condition. Patients were notified of their second-stage condition at the beginning of their sixth session. Participants in the Brief condition received either two UP modules, if one module received was Countering Emotional Behaviors, or three, if Countering Emotional Behaviors was not assigned early, and participants in the Full condition received all five modules.

Assessment

Participants completed three major assessments that included clinician-rated[2] and self-report measures. These assessments occurred at baseline, prior to the second-stage randomization (i.e. between sessions 5 and 6), and at the end of the 12-week treatment window. All participants also completed a self-report battery before each session and, for those in the Brief condition, each week between weeks 7–12. Self-report measures were completed via REDCap, a secure online survey platform.

Diagnostic interview

The Diagnostic Interview for Anxiety, Mood, and Obsessive-Compulsive and Related Neuropsychiatric Disorders (DIAMOND; Tolin et al., 2018) is a semi-structured diagnostic interview for DSM-5 disorders. Assessors assign categorical DSM-5 diagnoses and dimensional ratings of the subjective distress and/or degree of functional impairment of each diagnosis using a seven-point (1–7) clinical severity rating (CSR) scale; ratings of 3 or higher represent clinically significant distress/impairment. The DIAMOND was administered at baseline, prior to the second-stage randomization, and at the end of the 12-week treatment window. Graduate students certified in the DIAMOND administered all diagnostic assessments. Assessors demonstrated excellent reliability on categorical ratings of clinically significant diagnoses (Krippendorff's αs: .91–1.00; median = 1.00)[3] and dimensional severity ratings (CSRs) of each disorder (Krippendorff's αs: .83–1.00; median = .92).

Emotional disorder symptoms

The Overall Anxiety Severity and Impairment Scale (OASIS; Norman et al., 2006) is a 5-item self-report questionnaire designed to measure anxiety symptoms over the prior week. Total scores range from 0 to 20 with a clinical cutoff score of 8. Participants completed the OASIS at baseline, mid-treatment, post-treatment, and weekly before each session. In the current sample, OASIS items demonstrated good internal consistency at baseline (McDonald's ω = .84).

The Overall Depression Severity and Impairment Scale (ODSIS; Bentley et al., 2014) is a 5-item self-report questionnaire designed to assess depressive symptoms over the prior week. Total scores also range from 0 to 20 with a clinical cutoff score of 8. In the current sample, ODSIS items demonstrated excellent internal consistency at baseline (ω = .94).

Module sequencing

To sequence modules in the strengths and weaknesses conditions, self-report measures corresponding to each UP skill were administered as part of the baseline self-report battery. Specifically, the 12-item Beliefs about Emotions Scale (BES; Rimes & Chalder, 2010) was used to assess competence in the skills associated with the Understanding Emotions module; items are rated on a 0–6 Likert-type scale with scores ranging from 0–72 (higher scores indicate greater negative beliefs about emotions). Mindful Emotional Awareness skills were measured using the 16-item Southampton Mindfulness Questionnaire (SMQ; Chadwick et al., 2008); items are rated on a 0–6 Likert-type scale with scores ranging from 0–96 (higher scores indicate greater ability to apply a present-focused, nonjudgmental stance toward emotions). Cognitive Flexibility skills were assessed using the 7-item UP Cognitive Skills Questionnaire (UP-CSQ; Sauer-Zavala et al., 2019); items are rated on a 1–5 Likert-type scale with total scores ranging from 7–35 (higher scores indicate a greater ability to consider other perspectives in emotion-generating situations). Countering Emotional Behavior skills were assessed using the 11-item MEAQ—Behavioral Activation subscale (MEAQ-BA; Gámez et al., 2011); items are rated on a 1–6 Likert-type scale with scores ranging from 11–66 (higher scores indicate a greater tendency to avoid situations and activities that generate emotions). Finally, the Confronting Physical Sensation skills were assessed using the 16-item Anxiety Sensitivity Index (ASI; Reiss et al., 1986); items are rated on a 0–4 Likert-type scale with total scores ranging from 0 to 64 (higher scores indicate greater fear associated with emotion-related physical sensations). Items from all measures demonstrated good-to-excellent internal consistency in the present sample at baseline (ωs ranging from .85 to .91).

To assess each patient's relative strengths and deficits, raw scores on the above measures were converted to z-scores and rank-ordered; normative data for z-scores was obtained using a dataset from a large reference sample (Sauer-Zavala et al., 2019). Of note, because higher scores on three measures indicate greater skill deficits (i.e. BES, MEAQ-BA, ASI) and higher scores on two measures indicate greater skill strengths (i.e. SMQ, UP-CSQ), we multiplied the z-scores from the BES, MEAQ-BA, and ASI by—1 before rank-ordering all five measures in order of greatest strength. These procedures were pilot tested by Sauer-Zavala et al. (2019) prior to their use in this study.

Treatment acceptability and satisfaction

At post-treatment (i.e. after session/week 12), participants reported their satisfaction with the treatment and how acceptable they found the treatment content and length. Satisfaction was rated on a 5-point Likert scale from 1 (*not at all satisfied*) to 5 (*extremely satisfied*) in response to the prompt, "Overall, how satisfied were you with the treatment?" Acceptability of treatment content was rated on a 5-point Likert scale from 1 (*not at all acceptable*) to 5 (*extremely acceptable*) in response to the prompt, "Overall, how acceptable was the treatment content to you? In other words, did you think that the treatment approach and activities made sense and were reasonable?" Acceptability of treatment length was rated on a 5-point Likert scale from 1 (*not long enough*) to 5 (*much too long*) in response to the prompt, "Overall, how acceptable was the treatment length to you? In other words, did you think that the length of the treatment you received was not enough, just right, or too much?"

Data analytic plan

Preliminary analyses

To determine if there was sufficient intraindividual variability to warrant personalized skill sequences, we explored the degree of variability within participants between their skill of greatest strength and greatest deficit. In line with Sauer-Zavala et al. (2019), we defined "sufficient" variability in skill strength as ≥ 1.96 points between each participants' largest and smallest standardized scores. We then examined descriptive statistics among the three module-sequencing conditions by comparing the frequency with which different modules were assigned in each condition using a series of chi-squared tests in SPSS Version 28.

Finally, we compared the frequency and severity of clinically significant diagnoses among the three module-sequencing conditions and two duration conditions using two one-way ANOVAs. Clinically significant diagnoses were defined as all inclusion diagnoses with a CSR ≥ 3 at baseline. We calculated the mean of the CSRs for each participants' clinically significant diagnoses at each timepoint to account for participants meeting criteria for different numbers of diagnoses at baseline. We refer to these means as CSR scores in all subsequent analyses. For example, if a participant was rated ≥ 3 on both PTSD and MDD at baseline, their CSR score was the average of the CSRs from these two diagnoses at each timepoint.

Acceptability and satisfaction

We conducted three two-way ANOVAs to compare end-of-treatment ratings of overall treatment satisfaction, overall acceptability, and acceptability specifically with regard to treatment length. We entered main effects of skill sequencing condition and treatment duration condition as well as the interaction between these two conditions. We used post-hoc tests to directly compare the marginal means of acceptability and satisfaction between each pair of conditions.

Diagnostic and clinical severity

Given the multi-stage randomization design used in this SMART, we applied regression splines to a hierarchical linear modeling (HLM) framework using a knot at the Assessment 2/Session 6 time point. We conducted one HLM spline model for each outcome: CSR, OASIS, and ODSIS. The fixed effects included in these models are represented by the equation below:

$$CSR_{ij} = \beta_0 + \beta_1(time_{ij}) + \beta_2(seq\ cond_i) + \beta_3(dur\ cond_i) + \beta_4(time_{ij} \times seq\ cond_i)$$
$$+ \beta_5(time_{ij} \times dur\ cond_i) + \beta_6(seq\ cond_i \times dur\ cond_i) + \beta_7(time_{ij} \times seq\ cond_i \times dur\ cond_i)$$
$$+ \beta_8(rand_{ij}) + \beta_9(rand_{ij} \times seq\ cond_i) + \beta_{10}(rand_{ij} \times dur\ cond_i) + \beta_{11}(rand_{ij} \times seq\ cond_i \times dur\ cond_i)$$
$$+ e_{ij}$$

Here, *time* indicates assessment number (in the case of CSR scores) or session/week number (in the case of OASIS and ODSIS scores); *seq cond* is a dummy-coded variable indicating whether the participant was assigned to the Capitalization, Standard, or Compensation sequencing condition; *dur cond* is a dummy-coded variable indicating whether the participant was assigned to the Brief or Full duration condition; and *rand* indicates whether the observation occurs during the first- or second-stage randomization. We included random intercepts in all models and random slopes when doing so did not lead to model nonconvergence. For each model, we chose the residual covariance structure that optimized model fit as judged by AIC, and we applied the Kenward-Roger method for calculating denominator degrees of freedom in all models. We used *proc mixed* in SAS Version 9.4 to analyze all HLM spline models.

First-Stage randomization: personalized skill sequencing. To test our hypotheses regarding the stage-one sequencing randomization, we compared the average rate of change in CSR, OASIS, and ODSIS scores among participants in the three sequencing conditions across the first six sessions. We focused on slopes across the first six sessions for several reasons. First, all participants received treatment prior to the second-stage randomization, enhancing our statistical power for effects during this study stage. Additionally, previous research comparing capitalization and compensation approaches to treatment suggests that the majority of symptom change occurs early in treatment (Cheavens et al., 2012) and, for those in the Full condition, all UP modules were provided by week 12, likely negating any effects of personalization (Sauer-Zavala et al., 2019). We first examined the Type 3 omnibus tests of the time × sequencing condition interaction term and followed these up with individual between-condition comparisons using the *estimate* command in *proc mixed* as needed.

Finally, we conducted two exploratory analyses to test alternative potential moderators. First, to test whether the effects of personalization were stronger for participants with a larger separation between their skill of greatest strength and greatest deficit, we included the degree of skill spread as a conditional main effect and moderator of all terms in the model above. We examined the Type 3 omnibus test of the time × sequencing condition × degree of skill spread interaction term and followed these up with individual between-condition comparisons using the *estimate* command in *proc mixed* as needed. Second, because similar numbers of participants engaged in in-person and telehealth treatment, we explored whether the switch to telehealth moderated changes in CSR, OASIS, and ODSIS scores. We used the *estimate* command to compare predicted

intercept values to test if there were any baseline differences in clinical outcomes between those who completed baseline procedures in-person or via telehealth. We then used the HLM spline model above, including a dummy-coded indicator variable representing whether a participant was involved in-person or via telehealth as a conditional main effect and interaction term with all other variables, to test if telehealth status moderated the slope of change of each clinical outcome.

Second-Stage randomization: personalized skill selection

To test whether a personalized selection of modules led to similar outcomes at the end of the study period as a full course of a standard selection of modules, we re-centered the intercept in each of the models above to reflect the end of study assessment for CSR models or session/week 12 for OASIS and ODSIS models. We used the *estimate* command to compare re-centered intercept values for (a) participants randomized to the Brief duration condition and either the Capitalization or Compensation sequence condition and (b) participants randomized to the Standard sequence/Full duration condition. To ensure that a lack of difference between brief personalized treatment and full standard treatment are due to the personalized selection of modules (rather than due to the fact that most patients in sample were much improved by the second stage randomization), we also compared participants randomized to the Standard sequence *and* Brief duration conditions to those that were assigned to the Standard sequence and Full duration conditions.

Results

Preliminary analyses

The majority of participants ($n = 42$; 60.0%) reported at least a 1.96-point difference between the standardized score of their greatest skill strength and greatest skill deficit. The spread of these scores ranged from .61–4.64 standard deviations ($M = 2.24$, $SD = .86$).

Given that a plurality of participants ($n = 26$; 37.1%) were randomized to the Standard condition, there were significant differences between sequencing conditions in the frequency with which participants received modules across the first six sessions, $\chi^2(2)$s > 9.00, ps ≤ .01 (Table 2). Understanding Emotions, Mindful Emotion Awareness, and Cognitive Flexibility were delivered roughly twice as frequently in the Standard condition as in the Capitalization and Compensation conditions. Countering Emotional Behaviors and Confronting Physical Sensations were not delivered in the first six sessions of the Standard condition. Although the frequency of Countering Emotional Behaviors was similar between Compensation and Capitalization conditions, Confronting Physical Sensations was delivered twice as frequently in the Capitalization condition as in the Compensation condition.

There were no significant differences in age, gender, racial/ethnic background, income, sexual orientation, or education level at baseline among participants assigned to each sequencing condition, ps > .15. Thus, we did not include any demographic covariates in our further analyses. At baseline, there were no significant differences on

Table 2. Frequency of modules across first-stage randomization by sequencing condition.

	Module-Sequencing Condition				
	Compensation	Standard	Capitalization		
Module	n	n	n	$\chi^2(2)$	p
Understanding Emotions	11	23	9	13.51	< .01
Mindful Emotion Awareness	11	21	9	9.19	.01
Cognitive Flexibility	9	21	8	11.98	< .01
Countering Emotional Behaviors	12	0	13	22.98	< .01
Confronting Physical Sensations	6	0	14	22.16	< .01

the average CSRs of participants' clinically significant diagnoses, among module-sequencing conditions, $F(2, 67) = 1.12$, $p = .33$, $\eta^2 = .03$, 95% CI [.00, .13], or duration conditions, $F(1, 67) = 0.55$, $p = .46$, $\eta^2 = .01$, 95% CI [.00, .10].

Patient perceptions of personalized sequencing and early termination

Participants reported being moderately to very satisfied with the treatment on average (Table 3). Although satisfaction did not significantly differ by sequencing condition, $F(2, 50) = 0.30$, $p = .74$, partial $\eta^2 = .01$, participants in the Full duration ($n = 29$; 49.2%) condition reported greater satisfaction than those in the Brief duration condition ($n = 30$; 50.8%), $F(1, 50) = 8.85$, $p < .01$, partial $\eta^2 = .15$. However, this was qualified by an interaction between sequencing and duration conditions, $F(2, 50) = 3.22$, $p = .048$, partial $\eta^2 = .11$. There was a larger difference in satisfaction ratings between participants in the Brief and Full duration conditions who received modules prioritizing their deficits than between Brief and Full conditions among participants who received modules that prioritized their strengths or in the standard order.

Participants reported the treatment was very to extremely acceptable on average (Table 4). However, those in the Standard and Capitalization conditions found the treatment more acceptable than those in the Compensation condition, $F(2, 50) = 5.74$, $p < 0.01$, partial $\eta^2 = .19$. Similarly, participants in the Full duration condition found the treatment more acceptable than those in the Brief condition, $F(1, 50) = 13.53$, $p < 0.01$, partial $\eta^2 = .21$. These effects were qualified by a between-condition interaction, $F(2, 50) = 6.73$, $p < .01$, partial $\eta^2 = .21$: there was a larger difference in acceptability between participants in the Brief and Full duration conditions who received modules prioritizing their deficits than between Brief and Full conditions among participants who received modules that prioritized their strengths or in the standard order.

Table 3. Satisfaction with treatment by module-sequencing and treatment duration conditions.

	Module-Sequencing Conditions			Duration Condition
	Compensation	Standard	Capitalization	Marginal Means
Duration Condition	M (SE)	M (SE)	M (SE)	M (SE)
Brief	2.89 (.29)	3.75 (.31)	3.70 (.28)	3.44$_a$ (.17)
Full	4.44 (.29)	4.00 (.28)	4.00 (.28)	4.15$_b$ (.16)
Sequencing Condition Marginal Means	3.67$_a$ (.21)	3.88$_a$ (.21)	3.85$_a$ (.20)	3.78 (.12)

Values with different subscripts differ significantly between conditions, $ps < .05$.

Table 4. Acceptability of treatment by module-sequencing and treatment duration conditions.

	Module-Sequencing Conditions			Duration Condition
	Compensation	Standard	Capitalization	Marginal Means
Duration Condition	M (SE)	M (SE)	M (SE)	M (SE)
Brief	3.22 (.22)	4.13 (.23)	4.70 (.21)	4.02$_a$ (.13)
Full	4.67 (.22)	4.70 (.21)	4.60 (.21)	4.66$_b$ (.12)
Sequencing Condition Marginal Means	3.94$_a$ (.15)	4.41$_b$ (.15)	4.65$_b$ (.15)	4.34 (.09)

Values with different subscripts differ significantly between conditions, $ps < .05$.

Table 5. Acceptability of treatment length by module-sequencing and treatment duration conditions.

	Module-Sequencing Conditions			Duration Condition
	Compensation	Standard	Capitalization	Marginal Means
Duration Condition	M (SE)	M (SE)	M (SE)	M (SE)
Brief	1.67 (.32)	2.13 (.33)	1.90 (.30)	1.90$_a$ (.18)
Full	2.33 (.32)	2.60 (.30)	2.60 (.30)	2.51$_b$ (.18)
Sequencing Condition Marginal Means	2.00$_a$ (.22)	2.36$_a$ (.22)	2.25$_a$ (.21)	2.20 (.13)

Values with different subscripts differ significantly between conditions, $ps < .05$.

Finally, participants reported the treatment length was slightly too short to just right on average (Table 5). Although these ratings did not differ by sequencing condition, $F(2, 50) = 0.69$, $p = .50$, partial $\eta^2 = .03$, participants in the Full duration condition found the treatment length significantly more acceptable than those in the Brief condition, $F(1, 50) = 5.88$, $p = .02$, partial $\eta^2 = .11$. There was no interaction between conditions, $F(2, 50) = 0.08$, $p = .93$, partial $\eta^2 < .01$.

First-Stage randomization: the effect of personalized skills sequencing on clinical severity, anxiety, and depression

Participants demonstrated significant decreases across the first six sessions in CSR, $B = —1.93$, $SE = .18$, $p < .01$, 95% CI [–2.29,—1.57]; and ODSIS scores, $B = —.20$, $SE = .07$, $p < .01$, 95% CI [–.34,—.06]; and marginally significant decreases in OASIS scores, $B = —.13$, $SE = .06$, $p = .052$, 95% CI [–.25, .001]. However, there were no significant differences among conditions in the slopes of CSR, $F(2, 113) = 0.16$, $p = .85$, OASIS, $F(2, 484) = 1.11$, $p = .33$, or ODSIS scores, $F(2, 481) = 0.15$, $p = .87$ (Figure 2). There were nearly identical rates of responders (i.e. CSRs < 3) in the Capitalization ($n = 13$; 54.2%), Standard ($n = 12$; 48.0%), and Compensation ($n = 12$; 57.1%) conditions.

Additionally, we explored whether the degree of skill spread and switch from in-person to telehealth treatment influenced our results. At baseline, there were no significant differences in CSR, $F(1, 125) = 0.16$, $p = .73$, OASIS, $F(1, 152) = 0.15$, $p = .70$, or ODSIS scores, $F(1, 101) = 0.42$, $p = .52$, by degree of skill spread. Similarly, degree of skill spread did not moderate reductions in CSR, $F(1, 110) = 0.65$, $p = .42$, OASIS, $F(1, 506) = 0.06$, $p = .81$, or ODSIS scores, $F(1, 488) = 0.10$, $p = .75$, across the first six sessions. Finally, degree of skill spread also did not moderate differences between sequencing conditions in reductions in CSR, $F(2, 110) = 1.02$, $p = .36$, OASIS, $F(2, 508) = 0.41$, $p = .66$, or ODSIS scores, $F(2, 490) = 0.64$, $p = .53$.
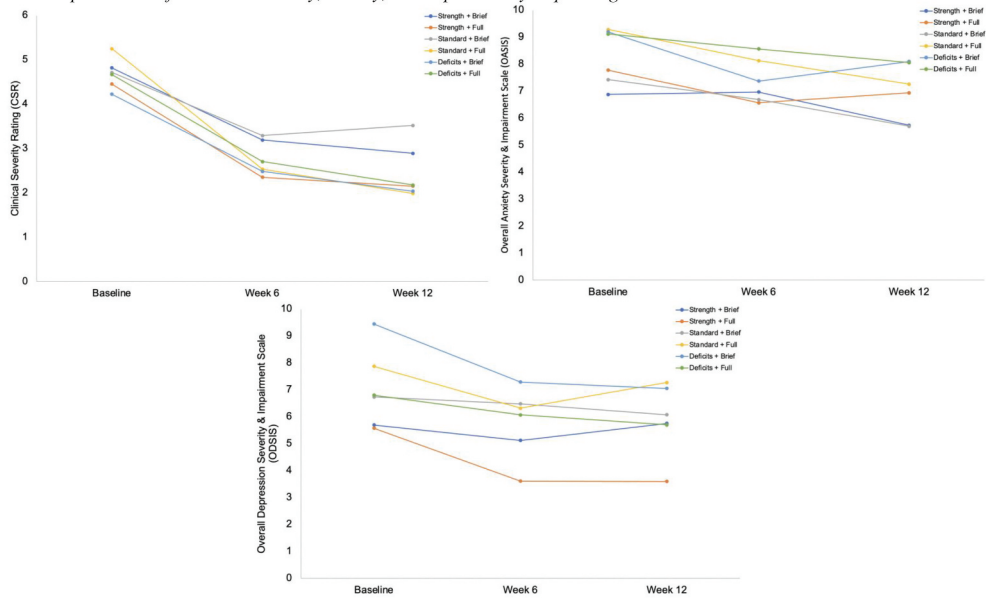
Model-Implied Values for Clinical Severity, Anxiety, and Depression by Sequencing and Duration Condition

Figure 2. Model-Implied values for clinical severity, anxiety, and depression by sequencing and duration condition.

Compared to those treated in person, participants treated via telehealth reported no significant differences in their CSR, $F(1, 146) = 0.01$, $p = .92$, OASIS, $F(1, 284) = 1.61$, $p = .21$, or ODSIS scores $F(1, 263) = 0.32$, $p = .57$, at baseline. Although there was a trend, $F(1, 432) = 3.54$, $p = .06$, for participants who completed the study via telehealth to report steeper declines in depression, $B = —.60$, $SE = .22$, $p < .01$, 95% CI [$-1.03,—.17$], than those who completed the study in person, $B = —.16$, $SE = .08$, $p = .046$, 95% CI [$-.32, —.003$], there were no significant differences in the slope of change in CSR, $F(1, 126) = 0.08$, $p = .78$, or OASIS scores, $F(1, 458) = 0.84$, $p = .36$, between these groups.

Second-Stage randomization: personalized skill selection

There were no significant differences in age, gender, racial/ethnic background, income, sexual orientation, or education level at baseline among participants assigned to each treatment duration (Brief or Full) condition, ps > .15. Thus, we did not include any demographic covariates in our further analyses. There was also a similar proportion of participants in each skill sequence condition in the Brief ($n_{Standard} = 14$; $n_{Capitalization} = 11$; $n_{Compensation} = 9$) and Full ($n_{Standard} = 11$; $n_{Capitalization} = 12$; $n_{Compensation} = 12$) duration conditions, $\chi^2(2) = .82$, $p = .66$. Similarly, there was not a significant difference between treatment duration conditions on baseline CSR, OASIS, or ODSIS scores, Bs < $|.90|$, ps > .25, or CSR, OASIS, or ODSIS scores at the second stage randomization, Bs < $|.95|$, ps > .15.

We compared the final CSR, OASIS, and ODSIS scores between participants randomized to receive Brief personalized treatment and participants randomized to receive Full Standard treatment. We found no significant differences between these groups on

Table 6. Model-Implied end-of-study clinical severity ratings by randomization condition.

	Module-Sequencing Conditions			Duration Condition
	Compensation	Standard	Capitalization	Marginal Means
Duration Condition	M (SE)	M (SE)	M (SE)	M (SE)
Brief	2.37 (.72)	3.50 (.67)	3.19 (.69)	3.02 (.34)
Full	2.37 (.72)	1.45 (.72)	1.88 (.69)	1.90 (.35)
Sequencing Condition Marginal Means	2.37 (.47)	2.47 (.46)	2.54 (.46)	2.46 (.16)

Values with different subscripts differ significantly between conditions, $ps < .05$.

Table 7. Model-Implied end-of-study anxiety ratings by randomization condition.

	Module-Sequencing Conditions			Duration Condition
	Compensation	Standard	Capitalization	Marginal Means
Duration Condition	M (SE)	M (SE)	M (SE)	M (SE)
Brief	6.37 (1.53)	7.97 (1.39)	6.77 (1.50)	7.04 (.75)
Full	5.19 (1.46)	7.52 (1.41)	5.06 (1.44)	5.92 (.73)
Sequencing Condition Marginal Means	5.78 (.99)	7.74 (.94)	5.91 (.98)	6.48 (.39)

Values with different subscripts differ significantly between conditions, $ps < .05$.

Table 8. Model-Implied end-of-study depression ratings by randomization condition.

	Module-Sequencing Conditions			Duration Condition
	Compensation	Standard	Capitalization	Marginal Means
Duration Condition	M (SE)	M (SE)	M (SE)	M (SE)
Brief	5.43 (1.88)	6.52 (1.69)	4.84 (1.82)	5.60 (.93)
Full	5.63 (1.77)	5.06 (1.72)	1.94 (1.75)	4.21 (.91)
Sequencing Condition Marginal Means	5.53 (1.22)	5.79 (1.15)	3.39 (1.21)	4.90 (.53)

Values with different subscripts differ significantly between conditions, $ps < .05$.

CSR, OASIS, or ODSIS scores, $Bs < |1.35|$, $ps > .15$. We then compared the final CSR, OASIS, and ODSIS scores between participants randomized to receive Brief Standard treatment and those randomized to receive Full Standard treatment. Although those in the Full Standard condition demonstrated CSR ratings that were marginally lower than those in the Brief Standard condition, $B = —2.05$, $SE = 1.04$, $p = .051$, 95% CI [–4.12, .009], end-of-study anxiety and depression scores were relatively similar, $Bs < |1.50|$, $ps > .50$ (Tables 6–8).

Discussion

The goal of the present study was to evaluate the feasibility, acceptability, and preliminary utility of personalization approaches aimed at improving the efficiency of a modular, transdiagnostic treatment for emotional disorders. The first approach (i.e. first-stage randomization) was to personalize the sequencing of UP modules based on pre-existing skill capacities to either capitalize on patients' relative strengths or compensate for relative deficits. We expanded on prior work (Fisher et al., 2019; Fisher & Boswell, 2016; Sauer-Zavala et al., 2019) by including a control condition in which participants received UP modules in the standard order prescribed by the published manual (Barlow et al., 2018). The second approach (i.e. second-stage randomization) was to explore

whether a personalized selection of modules (i.e. 2–3 modules corresponding to patients' areas of greatest skill strength or deficit), was associated similar symptom reduction to a full course of standard care (i.e, all 5 modules).

It was important to first characterize the personalized module orders and determine whether personalized ordering is feasible to execute and satisfying and acceptable to patients. In the first-stage randomization, participants in the personalized conditions received Countering Emotional Behaviors and Confronting Physical Sensation modules more frequently and Understanding Emotions, Mindful Emotion Awareness, and Cognitive Flexibility modules less frequently than those in the Standard condition. Although expected, this result suggests the differential delivery of modules may be a source of variability in interpreting the effects of module sequencing. Consistent with Sauer-Zavala et al. (2019), our data indicate that the majority of patients possessed distinguishable strengths and deficits at baseline, suggesting that sequencing modules according to skill level is a feasible approach to treatment personalization. Although patients across all sequencing conditions reported being similarly satisfied with treatment, those in the Capitalization condition rated the treatment as significantly more acceptable than those in the Compensation condition and similarly as acceptable as those in the Standard condition. In terms of treatment duration, those in the Full duration condition reported greater satisfaction and acceptability than those in the Brief duration condition. It is also important to note that ratings in all conditions were in the moderately to extremely acceptable range, with treatment length being rated as slightly too short to just right. These results provide initial evidence that capitalizing on patient strengths early in treatment and providing the full treatment package may help optimize patient experiences with treatment *and* that it is acceptable to patients to modify these features.

Based on previous data favoring a capitalization approach (Cheavens et al., 2012; Sauer-Zavala et al., 2019), we hypothesized that patients in the Capitalization condition would demonstrate steeper trajectories of symptom improvement than those assigned to Compensation or Standard conditions. However, regardless of sequencing condition, patients exhibited similarly significant decreases in the severity of their clinically significant diagnoses, as well as in self-reported anxiety and depressive symptoms. Moreover, there were nearly identical rates of treatment responders across the sequencing conditions at the second assessment.

Our sample size, though adequate to achieve our aims related to examining the feasibility and acceptability of our personalization approaches, may have precluded our ability to detect differences between our sequencing conditions; indeed, this study was under-powered to reliably detect medium-to-large sized effects.[4] Given that personalizing module sequences according to pre-treatment skill capacities is feasible, replicating this study in a larger sample is warranted and would allow for a more thorough comparison of compensation and capitalization approaches relative to standard treatment delivery. Additionally, a larger sample would allow us to investigate more reliably for whom personalized module sequencing elicits the greatest symptom improvements. Other types of treatment personalization approaches, like shared decision-making between patient and provider, may be especially important when there is not clear

evidence to support a specific module order based on skill level. More research is needed to determine empirical thresholds for skill levels to inform when it would be clinically useful to sequence modules according to strengths and deficits.

Of course, based on these findings, it is possible that personalized skill sequencing does not increase treatment efficiency (i.e. produce steeper slopes on measures of anxiety and depression) in the UP relative to the standard sequence. The skills included in the UP are all drawn from a cognitive-behavioral tradition and are each purported to engage the mechanism of aversive reactions to emotions. Differences in trajectories of improvement as a function of capitalization and/or compensation may be more pronounced if the treatment strategies under study were more distinct (e.g. interpersonal therapy vs. cognitive therapy).

In the second-stage randomization, we explored whether participants randomized to receive a brief duration of modules delivered in a personalized order (i.e. personalized selection) reported different levels of clinical severity at the end of the study window than those randomized to receive the full duration of modules delivered in the standard order. We found no differences between these groups in terms of clinical severity, anxiety, or depression. In contrast, patients who discontinued after the first three UP modules presented in the standard order (i.e. after 6 session) reported (marginally significantly) higher clinical severity at the third assessment relative to those that received the full dose (i.e. 12 sessions) of the standard sequence. These results suggest that a personalized selection of UP modules can lead to lasting symptom change comparable to the full treatment presented in its standard order. These results add to the literature on the benefits of personalizing treatment (Cheavens et al., 2012; Fisher et al., 2019; Sauer-Zavala et al., 2019) and provide the first direct evidence that personalized treatment may be more efficient than a standard treatment package.

One notable, albeit unanticipated, clinical implication that can be drawn from the present study relates to the utility of telehealth platforms for transdiagnostic treatment delivery. Due to COVID-19 regulations, many study participants completed all or part of their therapy sessions through an online telehealth platform. Changes in our primary outcomes of interest were not affected by the switch in treatment modality, which suggests that the UP is amenable to being administered remotely. Telehealth delivery may circumvent many of the logistical barriers that limit patients' access to psychological intervention. Our study adds to the burgeoning literature base suggesting that the UP can be effectively administered through telehealth platforms (Cassiello-Robbins et al., 2020; Kennedy et al., 2021; Tulbure et al., 2018).

Findings from our study should be considered in the context of the following limitations, in addition to those listed above. First, we collected patient ratings of satisfaction and acceptability at the end of the treatment window (week 12) rather than following each module, which precludes our ability to explore patient responses to each individual module. To complement patient reports, future researchers may also assess clinicians' perceptions of personalized module sequencing and early termination. We did not assess symptoms beyond the 12-week treatment window. Future researchers should consider a longer follow-up period to evaluate whether treatment gains are sustained or deteriorate over time. Lastly, our sample was predominantly white and college-educated, limiting our ability to generalize findings to more socioeconomically, racially, and ethnically diverse patients.

In sum, participants were generally satisfied with personalized treatment sequences and early treatment discontinuation. There were no differences in trajectories of improvement as a function of module sequencing, which may suggest that personalized sequencing of the UP does not produce more efficient improvements than standard delivery though power concerns preclude definitive conclusions. However, our second-stage randomization allowed us to determine a personalized selection of modules delivered across six sessions was associated with comparable outcomes to a full course of a treatment (i.e. 12 sessions). Shorter courses of care may increase treatment efficiency, which has the potential to reduce patient costs and increase the mental health service system's capacity.

Notes

1. The Countering Emotional Behaviors module in this study consists of two modules from the standard UP: Countering Emotional Behaviors and Emotion Exposures. We linked these two modules because both address aversive reactions to emotions by engaging in behaviors explicitly designed to approach emotional experiences.
2. Participants completed the clinician-rated components of the major assessments in-person at our treatment center prior to COVID-19 modifications, and via telehealth after the implementation of COVID-19 study modifications.
3. Krippendorff's αs ≥ .80 indicate reliable variables; αs between .67 and .80 indicate tentative reliability (Krippendorff, 2004).
4. Post-hoc power analyses using the *PowerAnalysisIL* package (Lafit et al., 2021) in R (Version 3.6.1; R Core Team, 2019) suggest that a sample size of 50 per sequencing condition would be needed to detect a medium-to-large effect.

Disclosure statement

The first author (SSZ) receives royalties from Oxford University Press from sales of the Unified Protocol, the treatment under study in the present manuscript.

Funding

ORCID

Shannon Sauer-Zavala 🆔 https://orcid.org/0000-0002-7322-983X
Matthew W. Southward 🆔 https://orcid.org/0000-0002-5888-2769
Nicole E. Stumpp 🆔 https://orcid.org/0000-0001-9871-6469
Stephen A. Semcho 🆔 https://orcid.org/0000-0001-6831-6763
Caitlyn O. Hood 🆔 https://orcid.org/0000-0002-4147-5366
Anna Garlock 🆔 https://orcid.org/0000-0002-9115-4907
Alex Urs 🆔 https://orcid.org/0000-0001-8491-7278

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Bandelow, B., & Michaelis, S. (2015). Epidemiology of anxiety disorders in the twenty-first century. *Dialogues in Clinical Neuroscience*, *17*(3), 327–335. https://doi.org/10.31887/DCNS.2015.17.3/bbandelow

Barlow, D. H., Ellard, K. K., Sauer-Zavala, S., Bullis, J. R., & Carl, J. R. (2014). The origins of neuroticism. *Perspectives on Psychological Science*, *9*(5), 481–496. https://doi.org/10.1177/1745691614544528

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

Bentley, K. H., Gallagher, M. W., Carl, J. R., & Barlow, D. H. (2014). Development and validation of the overall depression severity and impairment scale. *Psychological Assessment*, *26*(3), 815–830. https://doi.org/10.1037/a0036216

Bentley, K. H., Sauer-Zavala, S., Cassiello-Robbins, C. F., Conklin, L. R., Vento, S., & Homer, D. (2017). Treating suicidal thoughts and behaviors within an emotional disorders framework: Acceptability and feasibility of the unified protocol in an inpatient setting. *Behavior Modification*, *41*(4), 529–557. https://doi.org/10.1177/0145445516689661

Boswell, J. F., Farchione, T. J., Sauer-Zavala, S., Murray, H. W., Fortune, M. R., & Barlow, D. H. (2013). Anxiety sensitivity and interoceptive exposure: A transdiagnostic construct and change strategy. *Behavior Therapy*, *44*(3), 417–431. https://doi.org/10.1016/j.beth.2013.03.006

Bullis, J. R., Boettcher, H., Sauer-Zavala, S., Farchione, T. J., & Barlow, D. H. (2019). What is an emotional disorder? A transdiagnostic mechanistic definition with implications for assessment, treatment, and prevention. *Clinical Psychology (New York, N.Y.)*, *26*(2), E12278–N/a. https://doi.org/10.1111/cpsp.12288

Cassiello-Robbins, C., Southward, M. W., Tirpak, J. W., & Sauer-Zavala, S. (2020). A systematic review of unified protocol applications with adult populations: Facilitating widespread dissemination via adaptability. *Clinical Psychology Review*, *78*, 101852. https://doi.org/10.1016/j.cpr.2020.101852

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). *The British Journal of Clinical Psychology*, *47*(Pt 4), 451–455. https://doi.org/10.1348/014466508X314891

Cheavens, J. S., Strunk, D. R., Lazarus, S. A., & Goldstein, L. A. (2012). The compensation and capitalization models: A test of two approaches to individualizing the treatment of depression. *Behaviour Research and Therapy*, *50*(11), 699–706. https://doi.org/10.1016/j.brat.2012.08.002

Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research*, *7*(1), 5–20. https://doi.org/10.1007/s11020-005-1962-6

Collins, L. M., Murphy, S. A., & Strecher, V. (2007). The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): New methods for more potent eHealth interventions. *American Journal of Preventive Medicine*, *32*(5), S112–S118. https://doi.org/10.1016/j.amepre.2007.01.022

Craske, M. G., & Barlow, D. H. (2006). *Mastery of your anxiety and worry: Workbook*. Oxford University Press.

Eustis, E. H., Ernst, S., Sutton, K., & Battle, C. (2019). Innovations in the treatment of perinatal expression: The role of yoga and physical activity interventions during pregnancy and postpartum. *Current Psychiatry Reports*, *21*(12), 133. https://doi.org/10.1007/s11920-019-1121-1

Fisher, A. J., Bosley, H. G., Fernandez, K. C., Reeves, J. W., Soyster, P. D., Diamond, A. E., & Barkin, J. (2019). Open trial of a personalized modular treatment for mood and anxiety. *Behaviour Research and Therapy*, *116*, 69–79. https://doi.org/10.1016/j.brat.2019.01.010

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment*, *23*(4), 496–506. https://doi.org/10.1177/1073191116638735

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

Garfield, S. L. (1994). Psychotherapy: Then and now. *Clinical Psychology & Psychotherapy*, *1*(2), 63–68. https://doi.org/10.1002/cpp.5640010202

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology*, *9*(3), 329–343. https://doi.org/10.1093/clipsy.9.3.329

Harnett, P., O'Donovan, A., & Lambert, M. J. (2010). The dose response relationship in psychotherapy: Implications for social policy. *Clinical Psychologist*, *14*(2), 39–44. https://doi.org/10.1080/13284207.2010.500309

Hofmann, S. G., & Smits, J. A. J. (2008). Cognitive-behavioral therapy for adult anxiety disorders: A meta-analysis of randomized placebo-controlled trials. *The Journal of Clinical Psychiatry*, *69*(4), 621–632. https://doi.org/10.4088/JCP.v69n0415

Kennedy, S. M., Lanier, H., Salloum, A., Ehrenreich-May, J., & Storch, E. A. (2021). Development and implementation of a transdiagnostic, stepped-care approach to treating emotional disorders in children via telehealth. *Cognitive and Behavioral Practice*, *28*(3), 350–363. https://doi.org/10.1016/j.cbpra.2020.06.001

Kessler, R. C., Nelson, C. B., McGonagle, K. A., Liu, J., Swartz, M., & Blazer, D. G. (1996). Comorbidity of DSM–III–R major depressive disorder in the general population: Results from the US National Comorbidity Survey. *British Journal of Psychiatry*, *168*(S30), 17–30. https://doi.org/10.1192/S0007125000298371

Kessler, R. C., Stein, M. B., & Berglund, P. (1998). Social phobia subtypes in the National Comorbidity Survey. *American Journal of Psychiatry*, *155*(5), 613–619. https://doi.org/10.1176/ajp.155.5.613

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, *30*(3), 411–433. https://doi.org/10.1111/j.1468-2958.2004.tb00738.x

Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly Shiny app and tutorial to perform power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–24. Advance online publication. https://doi.org/10.1177/2515245920978738

McHugh, R. K., Murray, H. W., & Barlow, D. H. (2009). Balancing fidelity and adaptation in the dissemination of empirically-supported treatments: The promise of transdiagnostic interventions. *Behaviour Research and Therapy*, *47*(11), 946–953. https://doi.org/10.1016/j.brat.2009.07.005

Norman, S. B., Cissell, S. H., Means-Christensen, A. J., & Stein, M. B. (2006). Development and validation of an overall anxiety severity and impairment scale (OASIS). *Depression and Anxiety*, *23*(4), 245–249. https://doi.org/10.1002/da.20182

Payne, L., Ellard, K. K., Farchione, T. J., Fairholme, C., & Barlow, D. H. (2014). Emotional disorders: A unified, transdiagnostic approach. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders* (pp. 237–274). Guilford Press.

R Core Team. (2019). *R: A language and environment for statistical computing.* (Version 3.6.1). [Computer software]. https://www.r-project.org/

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

Rimes, K. A., & Chalder, T. (2010). The beliefs about emotions scale: Validity, reliability and sensitivity to change. *Journal of Psychosomatic Research*, *68*(3), 285–292. https://doi.org/10.1016/j.jpsychores.2009.09.014

Sakiris, N., & Berle, D. (2019). A systematic review and meta-analysis of the unified protocol as a transdiagnostic emotion regulation based intervention. *Clinical Psychology Review*, *72*, 101751. https://doi.org/10.1016/j.cpr.2019.101751

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., Ametaj, A. A., Wilner, J. G., & Pagan, D. (2019). Transdiagnostic treatment personalization: The feasibility of ordering unified protocol modules according to patient strengths and weaknesses. *Behavior Modification*, *43*(4), 518–543. https://doi.org/10.1177/0145445518774914

Sauer-Zavala, S., Cassiello-Robbins, C., Conklin, L., Bullis, J. R., Thompson-Hollands, J., & Kennedy, K. (2017). Isolating the unique effects of the unified protocol treatment modules using single-case experimental design. *Behavior Modification*, *40*(2), 286–307. https://doi.org/10.1177/0145445516673827

Sauer-Zavala, S., Cassiello-Robbins, C., Woods, B. K., Wilner Tirpak, J. W., Curerri, A., & Rassaby, M. (2020). Countering emotional behaviors in the treatment of borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment*, *11*(5), 328–338. https://doi.org10.1037/per0000379

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

Sauer-Zavala, S., Southward, M. W., & Semcho, S. A. (2022). Integrating and differentiating personality and psychopathology in cognitive behavioral therapy. *Journal of Personality*, 90 (1), 89–102. Advance online publication. https://doi.org/10.1111/jopy.12602 .

Southward, M. W., Cassiello-Robbins, C., Zelkowitz, R. L., & Rosenthal, M. Z. (2020). Navigating the new landscape of value-based care: An example of increasing access, improving quality, and reducing costs using the unified protocol. *The Behavior Therapist*, *43*(4), 134–137. https://doi.org/10.31234/osf.io/rvbzn

Tolin, D. F., Gilliam, C., Wootton, B. M., Bowe, W., Bragdon, L. B., Davis, E., Hannan, S. E., Steinman, S. A., Worden, B., & Hallion, L. S. (2018). Psychometric properties of a structured diagnostic interview for DSM-5 anxiety, mood, and obsessive-compulsive and related disorders. *Assessment*, *25*(1), 3–13. https://doi.org/10.1177/1073191116638410

Tulbure, B. T., Rusu, A., Sava, F. A., Sălăgean, N., & Farchione, T. J. (2018). A web-based transdiagnostic intervention for affective and mood disorders: Randomized controlled trial. *JMIR Mental Health*, *5*(2), E36. https://doi.org/10.2196/mental.8901

Van Straten, A., Geraedts, A., Verdonck-de Leeuw, I., Andersson, G., & Cuijpers, P. (2010). Psychological treatment of depressive symptoms in patients with medical disorders: A meta-analysis. *Journal of Psychosomatic Research*, *69*(1), 23–32. https://doi.org/10.1016/j.jpsychores.2010.01.019

Weinberger, A. H., Gbedemah, M., Martinez, A. M., Nash, D., Galea, S., & Goodwin, R. D. (2017). Trends in depression prevalence in the USA from 2005 to 2015: Widening disparities in vulnerable groups. *Psychological Medicine*, *48*(8), 1308–1315. https://doi.org/10.1017/S0033291717002781

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. J., Miranda, J., Bearman, S. K., Daleidan, E. L., Ugueto, A. M., Ho, A., Martin, J., Gray, J., Alleyne, A., Langer, D. A., Southam-Gerow, M. A., & Gibbons, R. D., & the Research Network on Youth Mental Health. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archives of General Psychiatry*, *69*(3), 274–282. https://doi.org/10.1001/archgenpsychiatry.2011.147