Available online at www.sciencedirect.com

ScienceDirect

Behavior Therapy 55 (2024) 93–105

Behavior Therapy

www.elsevier.com/locate/bt

Assessing Theories of State and Trait Change in Neuroticism and Symptom Improvement in the Unified Protocol

Nicole E. Stumpp
Matthew W. Southward
Shannon Sauer-Zavala
University of Kentucky

Researchers have shown neuroticism decreases with treatment (Roberts et al., 2017), although it is unclear if this reflects fleeting state-level changes (state-artifact position) or trait-level change (cause-correction hypothesis). These theories further propose that changes in neuroticism predict symptom change (cause-correction hypothesis) or are predicted by symptom change (state-artifact position). We compared these theories in a clinical trial of the Unified Protocol (UP). Participants (N = 38; M_{age} = 34.55, 71.1% female, 78.9% Caucasian) meeting DSM-5 criteria for a primary emotional disorder completed up to 12 weekly sessions of the UP. Neuroticism exhibited state-level changes by Session 6 but trait-level changes by Session 12. Within-person reductions in neuroticism exhibited bidirectional relations with anxiety symptom change but predicted unidirectional session-to-session reductions in depression. These findings provide relatively more nuanced support for the cause-correction hypothesis that the UP leads to trait changes in neuroticism that tend to precede symptom change.

Keywords: neuroticism; Unified Protocol; change mechanisms; anxiety; depression

M.W.S.'s efforts on this project were partially supported by the National Institute of Mental Health under award number K23MH126211. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute of Health.
S.S.Z. receives royalties from Oxford University Press in her role as an author of the Unified Protocol. All other authors declare no conflicts of interest.
Address correspondence to Nicole Stumpp, University of Kentucky, 343 Waller Ave., Suite 205, Lexington, KY, 40504. e-mail: Kentuckynicole.stumpp@uky.edu.

NEUROTICISM IS DEFINED as the tendency to experience frequent and intense negative emotions, along with the belief that one lacks adequate emotional coping resources (Barlow, Ellard, et al., 2014). A neurotic temperament is theorized to develop from the transaction between (a) heritable genetic contributions (i.e., general biological vulnerability) and (b) early life experiences that promote a heightened sense of unpredictability and uncontrollability (i.e., general psychological vulnerability; Barlow, Ellard, et al., 2014).

Neuroticism has consistently emerged as a transdiagnostic risk factor for various forms of psychopathology (Andrews, 1996; Barlow, Sauer-Zavala, et al., 2014; Clark et al., 1994; Khan et al., 2005; Krueger & Markon, 2006; Sher & Trull, 1994; Weinstock & Whisman, 2006). A meta-analysis of 33 population-based samples found large associations between this trait and anxiety, mood, somatoform, schizophrenia, and eating disorders (Malouff et al., 2005). In the context of prospective studies, there is evidence that neuroticism predicts the onset of major depressive episodes (Fanous et al., 2007; Kendler et al., 1993), as well as generalized anxiety disorder, social phobia, and specific phobia (Goldstein et al., 2018). Moreover, neuroticism as an underlying predisposition may explain patterns of comorbidity among these disorders, accounting for 20–40% of the covariance among internalizing disorders (Brown, 2007; Khan et al., 2005) and demonstrating a robust association (r = .88) with general psychopathology (Southward et al., 2022).

THEORIES OF NEUROTICISM CHANGE

Although personality traits have long been considered stable and inflexible (American Psychiatric Association [APA], 2013), there is increasing

evidence that neuroticism changes over time and is responsive to treatment. For example, age-related decreases in neuroticism have been observed in the general population (Eaton et al., 2011; Roberts & Mroczek, 2008), although the degree of change varies by person (Helson et al., 2002; Mroczek & Spiro, 2003; Small et al., 2003). Additionally, the results of a recent meta-analysis suggest that personality traits change over the course of relatively brief treatment, with nearly all reductions in neuroticism occurring within eight sessions of therapy, regardless of the specific treatment (Roberts et al., 2017). Given that neuroticism has historically been considered relatively stable, especially compared to symptoms of psychopathology, these findings have prompted questions about the stability of changes in neuroticism and the relation between changes in neuroticism and changes in symptoms.

State-Artifact Position

Item-content overlap. Some researchers have argued that reductions in neuroticism during treatment reflect relatively fleeting state-level changes rather than more enduring trait-level changes. In this theory, known as the state-artifact position,[1] short-term state-level changes in neuroticism (i.e., over hours, days, or weeks) are thought to be primarily accounted for by fluctuations in symptoms (Roberts et al., 2017). According to the state-artifact position, this is in part because measures of neuroticism (putatively assessing a trait) often include items that overlap, content-wise, with more state-level measures of psychopathology (e.g., "I rarely feel fearful or anxious" from the NEO-Neuroticism scale; "In the past week, how often have you felt anxious?" from the Overall Anxiety Impairment and Severity Scale). Thus, when a treatment produces symptom improvement, reductions on measures of neuroticism may be driven by shared item content and not because people have actually experienced long-lasting change to their personality (i.e., trait change; see Magidson et al., 2014, for a review of what constitutes trait change). Measuring changes in neuroticism without accounting for this overlapping content may thus obscure the degree to which changes in treatment reflect state or trait change.

To begin to address this concern, Sauer-Zavala et al. (2021) examined changes in neuroticism in a clinical trial of the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders (UP; Barlow et al., 2018) while covarying measures of anxiety and depression. The UP led to significant reductions in neuroticism even when controlling for anxiety and depression, suggesting that the variance in neuroticism not accounted for by symptom measures may reflect stable trait change.

Change in symptoms predicting change in neuroticism. To the extent that changes in neuroticism are considered artifacts resulting from item content overlap with symptom measures, neuroticism may appear to change more slowly than symptoms because measures of neuroticism are often written from a trait perspective (i.e., assessing how people behave generally), whereas measures of symptoms are often written from a state perspective (i.e., assessing symptoms over the past week). Another perspective on the state-artifact position is that changes in neuroticism *result* from changes in symptoms. In other words, if changes in neuroticism are not simply due to item-content overlap, reductions in symptom severity during treatment may allow people to gradually return to their trait levels of neuroticism established before experiencing an episode of psychopathology. Reducing how frequently and intensely people experience negative emotions in their daily life (i.e., symptom improvement) is thought to lead to subsequent reductions in how frequently and intensely people perceive themselves to experience negative emotions in general. Others have proposed that changes in symptoms occur concurrently with observed changes in neuroticism and that these observed changes in neuroticism are a by-product of symptom improvement.

Researchers have provided indirect evidence for this position (i.e., symptom improvement predicts trait improvement). Naturalistically, people experiencing a major depressive episode have reported concurrent decreases in neuroticism back to baseline levels as their depressive symptoms decrease (Barnett & Gotlib, 1988; Ormel et al., 2004), suggesting overlap between these constructs. In a treatment study of selective serotonin reuptake inhibitors (SSRIs), patients who responded to treatment reported decreases in both depression and neuroticism, whereas those who did not respond did not report decreases in either construct (Du et al., 2002). These results suggest that changes in neuroticism may depend on symptom improvement rather than the direct influence of medication on personality, but no researchers to our knowledge have demonstrated that changes in depression predict changes in neuroticism.

[1] In the treatment literature, this theory has also been referred to as the state effect hypothesis (Tang et al., 2009). In the personality and developmental literatures, researchers have discussed closely related theories such as the complication or scar hypothesis (Akiskal et al., 1983) and the spectrum or continuity hypothesis (Hirschfeld & Klerman, 1979). For a review of these theories, see Clark et al. (1994).

Cause-Correction Hypothesis

In contrast to the state-artifact position, the cause-correction hypothesis posits that changes in neuroticism lead to changes in symptoms (Soskin et al., 2012). Given that neuroticism is associated with the development of a variety of emotional disorders, this theory suggests that reductions in the general frequency with which people experience and react to negative emotions allows for changes in specific emotional disorder symptoms. For example, improvements in neuroticism mediated the effects of SSRI medication on depression (Quilty et al., 2008; Tang et al., 2009). Similarly, the effects of antidepressant medication on temperament significantly contributed to improvements in depressive symptoms across several studies (Soskin et al., 2012) but changes in depression symptoms have explained relatively little change in neuroticism (De Fruyt et al., 2006; Santor et al., 1997). Finally, in the context of personality disorders, changes in personality have been shown to precede symptom change, whereas symptom improvements do not lead to subsequent change in personality dimensions (Warner et al., 2004). However, these findings have not been replicated in the context of emotional disorders (i.e., anxiety and depressive disorders).

Emerging research suggests that behavioral interventions specifically designed to target neuroticism may be associated with more robust reductions in this trait (e.g., Armstrong & Rimes, 2016; Sauer-Zavala et al., 2021). The UP is a cognitive-behavioral intervention that was developed to engage neuroticism and has shown efficacy in reducing symptoms of anxiety and depressive disorders (Cassiello-Robbins et al., 2020; Sakiris & Berle, 2019). It has also been associated with significant improvements in neuroticism from pre- to posttreatment (Carl et al., 2014). Using multiple timepoints within treatment, researchers have shown that specific exposure-based modules of the UP led to greater reductions in neuroticism than symptom-focused cognitive behavioral therapy (CBT; Sauer-Zavala et al., 2021). However, it is important to note that the measure of neuroticism used in this study demonstrated relatively low internal consistency, and the researchers only measured constructs at five timepoints across treatment, making it difficult to determine the reliability of these results.

CURRENT STUDY

Given the ample evidence underscoring the relation between neuroticism and psychopathology, along with growing support for the notion that neuroticism can be engaged in treatment, it is important to understand whether neuroticism demonstrates state or trait change and the temporal sequence of changes between neuroticism and symptoms. Indeed, researchers have rarely measured neuroticism and symptoms frequently enough to examine temporal precedence regarding order of change, particularly in the context of emotional disorders. In a recent trial testing personalized skill sequencing with the UP (Sauer-Zavala et al., 2022b), participants completed measures of neuroticism and anxiety and depressive symptoms prior to weekly therapy sessions. In this secondary analysis, we first evaluated whether neuroticism demonstrated state-artifact or trait changes over the course of the UP. The state-artifact position would predict that any changes observed in neuroticism are accounted for by changes in symptoms, whereas the cause-correction hypothesis would predict that changes in neuroticism predict changes in symptoms and that these changes in neuroticism occur above and beyond symptom change. We hypothesized that participants would demonstrate significant reductions in neuroticism across treatment and that these reductions would remain significant when including anxiety and depressive symptoms as covariates. We then compared evidence for the state-artifact position and cause-correction hypothesis regarding the direction of change between neuroticism and symptoms. To test the state-artifact position, we explored whether reductions in anxiety and depressive symptoms predicted improvements in neuroticism. To test the cause-correction hypothesis, we explored whether reductions in neuroticism predicted improvements in anxiety and depressive symptoms. We hypothesized that changes in neuroticism would predict changes in anxiety and depressive symptoms, in line with the cause-correction hypothesis. Finally, we conducted a sensitivity analysis to explore the relations between neuroticism and symptom change across different lag lengths.

Method

PARTICIPANTS

A subsample[2] of participants ($N = 38$, $M_{age} = 34.55$, 71.1% female, 78.9% Caucasian,

[2] To reduce participant burden, half of participants were assigned to complete weekly measures of neuroticism and half were assigned to complete a measure of UP skill use (Southward & Sauer-Zavala, 2022). Seventy participants were randomized in the parent trial, of which 11 did not complete study procedures. The present subsample was drawn from the 70 randomized participants.

84.2% heterosexual, median income = $75,000–$99,999, median education level = undergraduate degree) were drawn from a sequential multiple assignment randomized trial (SMART) of the UP (Sauer-Zavala et al., 2022b) for secondary data analyses related to the present aims. Individuals were eligible for the parent trial if they met *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5;* APA, 2013) criteria for an emotional disorder (i.e., anxiety, depressive [major depression, persistent depression], obsessive-compulsive and related, or trauma and stressor-related disorder). The most common primary diagnosis was generalized anxiety disorder ($n = 15$; 39.5%), followed by social anxiety disorder ($n = 9$; 23.7%), major depressive disorder ($n = 9$; 23.7%), and persistent depressive disorder ($n = 9$; 23.7%). Individuals were excluded from participation if they endorsed diagnoses or symptoms requiring clinical prioritization or hospitalization (i.e., mania within the past year, acute suicide risk, substance use disorder not in early remission, or the lifetime presence of psychotic features [i.e., hallucinations or delusions]). Potential participants were also excluded if they had received five or more sessions of CBT within the last 5 years. Anyone receiving other psychotherapy focused on an emotional disorder agreed to discontinue while participating in the study. Individuals taking psychotropic medication ($n = 9$ in the present subsample) were asked to maintain their current dosages while participating in the study and be stable on their current medications for 1 month prior to starting the study. Participants who completed at least two weekly self-report measures of neuroticism, anxiety symptoms, and depressive symptoms within the first seven weeks of the parent trial were included in this subsample.

TREATMENT

Participants received the five core skill modules of the UP: Understanding Emotions, Mindful Emotional Awareness, Cognitive Flexibility, Countering Emotional Behaviors, and Confronting Physical Sensations (Barlow et al., 2018; Payne et al., 2014; Wilamowska et al., 2010). These modules are designed to teach skills of self-monitoring, nonjudgmental present-moment awareness, reevaluation of overly negative thoughts, approach-oriented behaviors, and tolerance of uncomfortable physical sensations, respectively. Each module consisted of two weekly, individual, 50-minute virtual telehealth sessions, except Countering Emotional Behaviors, which was delivered across four weekly 50-minute sessions, resulting in a total of 12 possible sessions. Therapists providing treatment included a licensed clinical psychologist, a postdoctoral fellow, and two advanced clinical psychology graduate students (two men, two women). All therapists were certified in the provision of the UP by one of its developers. Each session was audio-recorded, and average competence was high ($M = 4.26$ on a 5-point scale; Sauer-Zavala et al., 2022b).

PROCEDURE

All procedures were approved by the local university Institutional Review Board, and all participants provided informed consent prior to engaging in study procedures.[3] After an initial phone screen, likely eligible participants were invited to complete a semistructured diagnostic interview assessment to confirm eligibility and a battery of self-report measures. In line with one of the primary aims of the parent study, eligible participants were randomly assigned to receive core modules of the UP in one of three sequences: (1) the standard published order (standard condition; Barlow et al., 2018), (2) an order that prioritized patients' skill strengths (strengths condition), or (3) an order that prioritized patients' skill deficits (weaknesses condition).[4] Skill strengths and deficits were determined by evidence-based questionnaires selected to measure the skills targeted by each UP module (see Sauer-Zavala et al., 2022b, for further details).

Following Session 5, and in line with the second primary aim of the parent study, participants underwent a second-stage randomization to determine if they would terminate treatment after Session 6 (brief treatment condition) or continue for the full 12 sessions (full treatment condition).[5] Participants assigned to receive modules in the standard order completed modules targeting psychoeducation and self-monitoring, mindfulness, and cognitive flexibility. Participants assigned to one of the personalized sequencing

[3] At the beginning of the study, written informed consent was obtained. When study procedures moved online due to the COVID-19 pandemic, verbal consent was obtained.

[4] A primary aim of the parent trial was to assess whether delivering UP modules in sequences prioritizing patients' strengths, deficits, or the standard published order led to greater reductions in anxiety and depression symptoms.

[5] The second primary aim of the parent study was to test the effects of early treatment withdrawal to identify which patients demonstrated maintenance of or continued improvements in symptoms and which patients required a full course of treatment to demonstrate symptom reduction.

conditions could receive any combination of 2–3 UP modules in the first 6 sessions.[6] In the current study, 36.8% of participants received Understanding Emotions, 63.2% received Mindful Emotion Awareness, 63.2% received Cognitive Flexibility, 68.4% received Countering Emotional Behaviors, and 34.2% received Confronting Physical Sensations in the first six sessions. Participants were compensated $25 for completing additional diagnostic assessments after baseline (up to $50 total).

MEASURES

Diagnostic Assessment

The Diagnostic Interview for Anxiety, Mood, and OCD and Related Neuropsychiatric Disorders (DIAMOND; Tolin et al., 2018) is a semistructured diagnostic interview that assesses *DSM-5* criteria for anxiety, depressive, bipolar, obsessive-compulsive, trauma- and stressor-related disorders, and schizophrenia spectrum and other psychotic disorders. Modules of the DIAMOND were administered at baseline to determine the presence of an anxiety, depressive, or related disorder and the absence of mania or hypomania, substance use requiring a higher level of care, and psychotic disorders. Interrater reliability among certified graduate student assessors masked to treatment condition was excellent for categorical diagnoses in the 20% of tapes randomly selected for reliability testing (Krippendorff's αs: .91–1.00; median = 1.00).

Neuroticism

The NEO Five Factor Inventory–Neuroticism subscale (NEO-FFI-N; Costa & McCrae, 1989) is a 12-item self-report subscale of the NEO-FFI designed to measure the personality dimension of neuroticism (i.e., "I often feel tense and jittery"; "I am seldom sad or depressed"). Participants rated each item on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*) before each session. NEO-FFI-N items demonstrated good internal consistency in the present sample across all twelve weeks (McDonald's ωs: .64–.87).

Symptom Severity

The Overall Anxiety Severity and Impairment Scale (OASIS; Norman et al., 2006) is a five-item measure assessing the severity of and impairment due to anxiety symptoms over the last week. Participants use a five-point Likert-type scale ranging from 0 to 4 with unique anchors for each item. Higher scores indicate greater symptom severity and functional impairment with a clinical cutoff score of 8. Participants completed the OASIS before each session. OASIS items demonstrated good internal consistency in the present sample across all 12 weeks (ωs: .76 –.84).

The Overall Depression Severity and Impairment Scale (ODSIS; Bentley et al., 2014) is a five-item measure assessing the severity of and impairment due to depressive symptoms over the last week. Participants use a five-point Likert-type scale ranging from 0 to 4 with unique anchors for each item. Higher scores indicate greater symptom severity and functional impairment with a clinical cutoff score of 8. Participants completed the ODSIS before each session. ODSIS items demonstrated excellent internal consistency in the present sample across all 12 weeks (ωs: .91–.95).

DATA ANALYTIC PLAN

We first examined whether the demographic characteristics of the subsample of patients included in the present study ($n = 38$) differed from the remaining participants from the primary study ($n = 32$). We used an independent samples *t*-test to determine if the two groups differed in age and chi-squared tests to determine if the groups differed in gender identity, marital status, and sexual orientation. Fisher's exact test was used to address the small cell sizes in chi-squared comparisons. Finally, Wilcoxon-Mann-Whitney *U* tests were used to determine if the two groups differed in education level and family income.

We then examined whether the UP led to state- or trait-level changes in neuroticism. Given the nested structure of the data (i.e., sessions within patients), we conducted intent-to-treat analyses using hierarchical linear modeling (HLM) as implemented in *proc mixed* in SAS Version 9.4. We first regressed neuroticism scores on session number, including dummy-coded variables representing ordering condition and study therapist, respectively, as covariates. We also modeled random intercepts, used a lag-1 autoregressive covariance structure for the residuals, applied the Kenward-Roger method to calculate denominator degrees of freedom, and used restricted maximum likelihood estimation to adjust parameter estimates in response to missing data. We assessed patterns of change across the first-stage random-

[6] If patients in the personalized sequencing conditions were assigned to complete the Countering Emotional Behaviors module prior to the second-stage randomization, they received a total of two UP modules in the first 6 weeks of treatment, whereas if they were not assigned to complete this module, they received three of the other 2-session modules.

ization (i.e., the first 7 sessions/weeks[7]) because all participants were engaged in treatment during this stage and given that the majority of change in neuroticism has been shown to occur by this point (Roberts et al., 2017). As a supplemental analysis, we also examined change in neuroticism across all available sessions (i.e., data from Sessions 1–12 for full treatment and Sessions/Weeks 1–7 for brief treatment) to test the durability of these effects across the full treatment window. Finally, to test if changes in neuroticism represent state-artifact effects, we added anxiety and depressive symptoms as covariates to each of these models to examine if changes in neuroticism remained significant over and above symptom change.

To test the cause-correction hypothesis, we again used HLM to test if changes in neuroticism would predict session-to-session changes in anxiety and depressive symptoms across Weeks 1–7. We first disaggregated participants' neuroticism scores into between- and within-person variability in line with Wang and Maxwell's (2015) recommendations. Between-person variability in neuroticism was determined by calculating each participant's mean neuroticism score across Weeks 1–7, using these scores to calculate a grand mean of the sample, and subtracting the grand mean from each participant's mean score. Within-person variability in neuroticism was calculated by subtracting each participant's mean score across Weeks 1–7 from their raw neuroticism score each week. We also created a 1-week lagged variable for depressive and anxiety symptoms. Finally, we regressed the target symptom (e.g., anxiety) at session *t* on within- and between-person neuroticism at session *t* and the target symptom at *t*–1 to test if neuroticism predicted session-to-session changes in the target symptom. We again included session number and dummy-coded covariates representing ordering condition and study therapist, modeled random intercepts, used a lag-1 autoregressive covariance structure for the residuals, applied the Kenward-Roger method to calculate denominator degrees of freedom, and used restricted maximum likelihood estimation. To estimate R^2 effect sizes, we used the Glimmix_R2 macro (Jaeger et al., 2017) in SAS. To explore these effects at different lag lengths, we conducted exploratory sensitivity analyses by regressing anxiety or depression symptoms at session *t* on within- and between-person neuroticism and symptoms (i.e., anxiety or depression) at session *t*–1, using the same method of disaggregation and model specifications above.

To test the state-artifact position (i.e., that changes in symptoms would predict session-to-session changes in neuroticism), we regressed neuroticism scores at session *t* on within- and between-person anxiety at session *t* and neuroticism scores at session *t*–1, using the same method of disaggregation and model specifications above. We then replaced anxiety with depression symptoms in a separate model. We again used the Glimmix_R2 macro in SAS to estimate R^2 effect sizes. To explore these effects at different lag lengths, we conducted exploratory sensitivity analyses by regressing neuroticism scores at session *t* on within- and between-person anxiety and neuroticism scores at session *t*–1, using the same method of disaggregation and model specifications above. In a separate model, we replaced anxiety with depression symptoms.

We conducted sensitivity power analyses to calculate the smallest effect sizes we were powered to detect. Assuming $\alpha = .05$, power = .80, $n = 38$, with 1 predictor, we were powered to detect medium-to-large sized between-person effects ($R^2 > .18$; Faul et al., 2009). Using Lafit et al. (2021) power analysis Shiny app, we were powered to detect medium-sized within-person effects ($R^2 \geq .10$). All code is available at https://osf.io/8bhvr/.

Results

PRELIMINARY ANALYSES

The sample included in the current study did not significantly differ in age, gender identity, sexual orientation, education level, marital status, or family income from the remaining participants from the parent trial, *p*s > .05. Participants in the present sample met criteria for an average of 2.79 diagnoses and reported pretreatment scores of anxiety ($M = 9.70$, $SD = 3.35$) and depressive symptoms ($M = 8.43$, $SD = 4.79$) in the clinical range. Participants completed a total of 255 sessions that were relatively evenly distributed among the standard ($n = 62$; 24.3%), strengths ($n = 96$; 37.6%) and weakness ($n = 97$; 38.0%) ordering conditions.

STATE AND TRAIT CHANGE IN NEUROTICISM

Treatment with the UP was associated with significant decreases in neuroticism across the first 7 weeks, $B = -.39$, $SE = .15$, $p = .01$, 95% CI [−.70, −.09], $d = .23$ (Table S1). Similarly, both

[7] Given that participants in the brief treatment condition continued to fill out all measures above each week until Week 12, data from Week/Session 7 was included to model the effects of content taught in Session 6, the last session completed by all participants.

anxiety, $B = -.47$, $SE = .10$, $p < .01$, 95% CI [−.67, −.28], $d = .95$, and depression, $B = -.50$, $SE = .11$, $p < .01$, 95% CI [−.71, −.29], $d = .75$, significantly decreased across the first 7 weeks (Table 1). Significant decreases in neuroticism were also observed across all 12 study weeks, $B = -.42$, $SE = .11$, $p < .01$, 95% CI [−.64, −.20], $d = .26$, for all possible sessions (i.e., including neuroticism scores from all 12 sessions from those randomized to the full treatment condition and neuroticism scores from the first 7 weeks from those randomized to the brief treatment condition; Table S2). When adding anxiety and depressive symptoms as covariates, decreases in neuroticism across the first 7 weeks were no longer significant, $B = -.15$, $SE = .15$, $p = .30$, 95% CI [−.46, .15], $R^2 = .01$ (Table S3). However, decreases in neuroticism across all possible sessions remained significant when adding anxiety and depressive symptoms as covariates, $B = -.27$, $SE = .11$, $p = .01$, 95% CI [−.49, −.06], $R^2 = .04$ (Table S4). This pattern of results supports the notion that relatively short-term changes in neuroticism across the first seven weeks are accounted for by symptom improvement, in line with the state-artifact position, yet relatively longer-term changes in neuroticism across all 12 weeks, above and beyond symptom change, supports the cause-correction hypothesis.

TESTING THE STATE-ARTIFACT POSITION AND CAUSE-CORRECTION HYPOTHESIS

In line with the cause-correction hypothesis that neuroticism change predicts symptom improvement, within-person improvements in neuroticism (i.e., lower than one's personal average) predicted session-to-session improvements in anxiety, $B = .22$, $SE = .06$, $p < .01$, 95% CI [.11, .33], $R^2 = .08$ (Table 2a). Within-person improvements in neuroticism also significantly predicted session-to-session improvements in depression, $B = .24$, $SE = .06$, $p < .01$, 95% CI [.12, .36], $R^2 = .07$ (Table 2b). However, between-person neuroticism was not significantly related to anxiety, $B < .01$, $SE = .03$, $p > .99$, 95% CI [−.07, .07], $R^2 < .01$, or depression, $B = .21$, $SE = .11$, $p = .06$, 95% CI [−.01, .43], $R^2 = .17$. Further, in our exploratory sensitivity analyses, neither within- nor between-person anxiety or depression predicted subsequent session-to-session changes in neuroticism, *Bs*: −.03−.08, *ps* > .27.

In line with the state-artifact position (i.e., symptom change predicts changes in neuroticism), within-person improvements in anxiety significantly predicted session-to-session improvements in neuroticism, $B = .34$, $SE = .12$, $p < .01$, 95% CI [.11, .57], $R^2 = .06$ (Table S5a). However, in contrast to the state-artifact position, within-person improvements in depression did not significantly predict session-to-session changes in neuroticism, $B = .15$, $SE = .11$, $p = .16$, 95% CI [−.06, .36], $R^2 = .02$. Further, there were no significant between-person associations between anxiety or depression and neuroticism, *Bs*: −.17 to .35, *ps* > .19, $R^2 < .09$. In our exploratory sensitivity analyses, between-person differences in depression were significantly associated with neuroticism, $B = .18$, $SE = .09$, $p = .04$, 95% CI [.01, .35], but within-person changes in depression did not significantly predict subsequent session-to-session changes in neuroticism, $p > .29$. Neither within- nor between-person anxiety predicted subsequent session-to-session changes in neuroticism, *Bs:* −.25 to −.18, *ps* > .14.

Assessing the cause-correction hypothesis and state-artifact position using all possible sessions revealed a similar pattern of results. In line with the cause-correction hypothesis, within-person improvements in neuroticism predicted session-to-session improvements in anxiety, $B = .19$, $SE = .05$, $p < .01$, 95% CI [.10, .29], $R^2 = .07$, and depression, $B = .27$, $SE = .05$, $p < .01$, 95% CI [.16, .37], $R^2 = .08$ (Tables 3a and 3b). Between-person neuroticism was not significantly related to anxiety, $B = .03$, $SE = .03$, $p = .44$, $R^2 = .02$. However, between-person neuroticism was significantly associated with depression symptoms, $B = .24$, $SE = .10$, $p = .03$, 95% CI [.03, .45],

Table 1
Average Neuroticism, Anxiety, and Depression Scores Across First-Stage Randomization.

Construct	Session 1 *M* (*SD*)	Session 2 *M* (*SD*)	Session 3 *M* (*SD*)	Session 4 *M* (*SD*)	Session 5 *M* (*SD*)	Session 6 *M* (*SD*)	Session/Week 7 *M* (*SD*)
NEO-FFI-N	42.48 (8.31)	43.00 (6.53)	42.79 (6.59)	41.43 (5.73)	42.06 (6.85)	39.86 (6.53)	40.25 (8.16)
OASIS	9.18* (3.21)	8.36* (3.12)	8.18* (3.60)	8.11* (2.98)	7.03 (2.74)	6.50 (2.83)	6.28 (2.65)
ODSIS	7.74 (4.45)	7.34 (4.81)	7.16 (4.77)	6.76 (4.13)	6.53 (4.45)	5.14 (4.21)	4.94 (3.78)

Note. OASIS = Overall Anxiety Severity and Impairment Scale. ODSIS = Overall Depression Severity and Impairment Scale. NEO-FFI-N = NEO Five Factory Inventory – Neuroticism subscale.
* Scores above recommended clinical cutoff.

Table 2a
Neuroticism Predicting Session-to-Session Changes in Anxiety (Sessions 1–7)

Variable	*B/F*	*SE*	*df*	*p*	95% CI	R^2
Intercept	7.62	1.05	50.3	<.01	[5.52, 9.72]	
Session	−.24	.12	51	.05	[−.48, .001]	.02
Anxiety (Lag-1)	.20	.07	133	.01	[.06, .34]	.07
Ordering Condition	6.78		2, 19.6	.01		.02
Therapist	5.06		3, 19.5	.01		.02
Neuroticism (Within)	.22	.06	174	<.01	[.11, .33]	.08
Neuroticism (Between)	<.01	.03	18.8	>.99	[−.07, .07]	<.01

Note. Random Intercept $\sigma^2 = 0$, Residual $\sigma^2 = 6.53$, AR(1) $\rho = .08$, AIC = 968.7, Model $R^2 = .35$.

Table 2b
Neuroticism Predicting Session-to-Session Changes in Depression (Sessions 1–7)

Variable	*B/F*	*SE*	*df*	*p*	95% CI	R^2
Intercept	9.74	1.80	36.5	<.01	[6.10, 13.38]	
Session	−.46	.18	43.6	.01	[−.83, −.09]	.03
Depression (Lag-1)	−.18	.07	151	.01	[−.32, −.04]	.04
Ordering Condition	1.83		2, 21.5	.18		.05
Therapist	.98		3, 21.2	.42		.07
Neuroticism (Within)	.24	.06	124	<.01	[.12, .36]	.07
Neuroticism (Between)	.21	.11	22.1	.06	[−.01, .43]	.17

Note. Random Intercept $\sigma^2 = 9.37$, Residual $\sigma^2 = 12.54$, AR(1) $\rho = .55$, AIC = 1065.3, Model $R^2 = .65$.

$R^2 = .21$. In our exploratory sensitivity analyses, neither between- nor within-person changes in depression or anxiety were significantly related to subsequent session-to-session changes in neuroticism, *Bs:* −.03 to .12, *ps* > .22.

In line with the state-artifact position, within-person anxiety significantly predicted session-to-session changes in neuroticism across all sessions, $B = .25$, $SE = .10$, $p = .02$, 95% CI [.05, .46], $R^2 = .03$ (Table S6). However, within-person depression was not significantly related to neuroticism, $B = .16$, $SE = .09$, $p = .07$, $R^2 = .02$, nor were there any significant between-person associations between anxiety or depression and neuroticism, *Bs*: −.04 to .26, *ps* > .16, $R^2 < .11$. In our exploratory sensitivity analyses, neither between- nor within-person changes in depression or anxiety were significantly related to subsequent session-to-session changes in neuroticism, *Bs:* −.14 to .15, $p > .08$.

Discussion

In this study, we examined state and trait changes in neuroticism and symptoms of anxiety and depression. We found that neuroticism significantly decreased across treatment with the UP. When covarying anxiety and depressive symptoms, decreases in neuroticism were not significant across the first stage of treatment but were significant when examining a longer course of care. This pattern of results may support a model in which early state changes in neuroticism give way to later trait-level changes. The state-artifact position posits that changes in neuroticism through the

Table 3a
Neuroticism Predicting Session-to-Session Changes in Anxiety (All Available Sessions)

Variable	*B/F*	*SE*	*df*	*p*	95% CI	R^2
Intercept	7.51	.84	59.6	<.01	[5.83, 9.19]	
Session	−.17	.06	33.9	.01	[−.30, −.05]	.03
Anxiety (Lag-1)	.19	.06	186	<.01	[.07, .31]	.10
Ordering Condition	8.31		2, 19.7	<.01		.16
Therapist	6.43		3, 18.9	<.01		.15
Neuroticism (Within)	.19	.05	257	<.01	[.10, .29]	.07
Neuroticism (Between)	.03	.03	20.2	.44	[−.04, .10]	.02

Note. Random Intercept $\sigma^2 = 0.00$, Residual $\sigma^2 = 6.30$, AR(1) $\rho = .16$, AIC = 1259.2, Model $R^2 = .80$.

Table 3b
Neuroticism Predicting Session-to-Session Changes in Depression (All Available Sessions)

Variable	*B/F*	*SE*	*df*	*p*	95% CI	R^2
Intercept	8.56	1.54	32.8	<.01	[5.42, 11.69]	
Session	−.18	.11	35.8	.11	[−.39, .04]	<.01
Depression (Lag-1)	−.16	.06	225	.01	[−.27, −.04]	.06
Ordering Condition	2.45		2, 20.8	.11		.07
Therapist	1.08		3, 19.9	.38		.07
Neuroticism (Within)	.27	.05	206	<.01	[.16,.37]	.08
Neuroticism (Between)	.24	.10	23.4	.03	[.03, .45]	.21

Note. Random Intercept $\sigma^2 = 7.45$, Residual $\sigma^2 = 11.73$, AR(1) $\rho = .52$, AIC = 1393.3, Model $R^2 = .70$.

course of treatment should largely disappear after controlling for symptom improvement (Tang et al., 2009). However, consistent with the cause-correction hypothesis, within-person reductions in neuroticism predicted session-to-session decreases in anxiety and depressive symptoms. Conversely, only within-person reductions in anxiety predicted session-to-session reductions in neuroticism. This pattern of results suggests that neuroticism exhibits a unidirectional effect on session-to-session changes in depression but a bidirectional relation with anxiety, providing somewhat stronger support for the cause-correction hypothesis than the state-artifact position.

We replicated previous findings regarding changes in neuroticism in treatments generally (Roberts et al., 2017) and the UP specifically (Carl et al., 2014; Sauer-Zavala et al., 2021) using a more reliable measure of neuroticism administered more frequently than in previous studies. We also extended previous findings that change in personality traits precede symptom improvements in the context of personality disorders (Warner et al., 2004) by assessing the temporal change among these constructs in emotional disorders. Together, these results suggest that treatment can lead to small-sized reductions in trait neuroticism over 8–16 weeks for patients with emotional disorders.

Consistent with the state-artifact position, reductions in neuroticism across the first six sessions were accounted for by changes in anxiety and depression. However, in contrast to the state-artifact position, decreases in neuroticism across the full treatment window occurred relatively independent of symptom change. These results suggest that early reductions in neuroticism may primarily reflect state-level changes in the construct that are more strongly confounded by concurrent reductions in symptoms, whereas continued treatment may lead to more trait-level reductions in neuroticism independent of symptom change. These latter reductions in neuroticism independent of symptom change suggest changes in neuroticism are unlikely to primarily be an artifact of item-content overlap. Instead, treatment may lead to reductions in initially elevated state levels of neuroticism and symptoms, as well as trait levels of neuroticism over time. Given that these novel findings provide a more nuanced elaboration of the state-artifact position, we encourage future researchers to replicate these results and apply time-varying effect models (e.g., Wright et al., 2014) to confirm how many sessions may be necessary to demonstrate more trait level changes in neuroticism independent of symptom change.

A further implication of the state-artifact hypothesis is that reductions in symptoms will predict reductions in neuroticism, whereas the cause-correction hypothesis suggests reductions in neuroticism will predict reductions in symptoms. In line with the cause-correction hypothesis, within-person reductions in neuroticism predicted session-to-session reductions in anxiety and depressive symptoms. We found less consistent support for the state-artifact position: within-person reductions in anxiety, but not depressive, symptoms predicted reductions in neuroticism. These results conceptually replicate previous findings in which reductions in neuroticism explained significant variability in symptoms (Soskin et al., 2012) and mediated reductions in symptoms (Quilty et al., 2008; Tang et al., 2009) and extend them to within-person session-to-session changes in these constructs. Together, these results suggest that treatment may act more directly on a person's sense of how frequently and intensely they react to negative emotions, which may in turn help promote reductions in the degree to which people experience these strong negative emotions (Barlow, Sauer-Zavala, et al., 2014; Bullis et al., 2019).

Of note, however, within-person reductions in anxiety, but not depression, also predicted session-to-session reductions in neuroticism. The

bidirectional relation between anxiety and neuroticism may suggest that reductions in anxiety symptoms and neuroticism exert reciprocal effects on one another. Alternatively, these results may suggest that the UP is acting more directly on functional mechanisms such as aversive reactivity, leading to reductions in both anxiety and neuroticism (Sauer-Zavala et al., 2022a, 2023) whereas other functional mechanisms (e.g., reward responsivity; Craske et al., 2019) may be more directly related to reductions in depression. Finally, it is possible that these results reflect the greater variability in anxiety symptoms compared to depression in this sample, which may have provided greater power to detect effects of anxiety symptoms compared to depression. Given that (a) researchers have found mixed results regarding differential associations between neuroticism and anxiety and depressive disorders cross-sectionally (Kotov et al., 2010) and longitudinally (Ormel et al., 2013) and (b) that our measure of neuroticism included a balanced representation of items from NEO-PI-R-N-Anxiety and -Depression subscales, we believe it is unlikely that greater overlap between anxiety items and neuroticism, compared to depression items and neuroticism, drove these effects.

Interestingly, the pattern of results found for session-to-session changes in neuroticism and symptoms did not extend to subsequent session-to-session changes. Indeed, only between-person depression predicted next-session neuroticism. These findings suggest that the effects of neuroticism on symptoms and vice versa may be occurring more rapidly throughout each week than on a week-to-week basis or that these effects are occurring relatively simultaneously and may be predicted by week-to-week changes in other constructs (e.g., aversive reactivity; Sauer-Zavala et al., 2022a, 2023). Future researchers should explore these effects at multiple lags and in the context of theorized functional mechanisms to replicate our findings and build on them to develop a more comprehensive account of the interplay between personality and symptom change in treatment.

Other limitations also warrant comment. The sample size of the current study was relatively small and composed primarily of White, heterosexual, college-educated women, limiting the generalizability of our findings to more diverse populations. Additionally, the parent trial did not include assessments of neuroticism prior to the onset of participants' anxiety and depressive episodes or follow-up assessments of these constructs after week twelve. This design would allow for the strongest conclusions regarding the timing and course of state- and trait-level changes in symptoms and neuroticism. Future researchers should consider a follow-up window to assess whether treatment gains and personality trait change are maintained for those receiving the full 12 sessions of the UP. Our measures of anxiety and depression reflect frequency of symptoms as well as distress and impairment resulting from these symptoms. Although arguably more comprehensive than symptom counts, future researchers may use measures that better distinguish these symptom dimensions to isolate their relations with neuroticism. Finally, our measure of neuroticism did not specify a timeframe for participants to consider when answering items. Future researchers should assess whether these results replicate when using a measure of neuroticism with a specific timeframe (e.g., since the previous session).

Despite its limitations, in this study we replicated and extended previous findings regarding state and trait changes in both neuroticism and symptoms of anxiety and depression. By measuring self-reported neuroticism and symptoms, we were able to reduce confounds from multimethod assessments and by measuring these constructs prior to each session, increased our power to detect within-person changes and minimized confounds that could occur between longer assessment lags. We demonstrated that relatively early reductions in neuroticism may reflect more state-level variability in line with the state-artifact position whereas a full course of treatment may lead to more trait-level reductions in contrast to the state-artifact position. We also found that within-person reductions in neuroticism predicted session-to-session reductions in anxiety and depression, in line with the cause-correction hypothesis, but that anxiety and not depression exerted a bidirectional relation with neuroticism. These results provide more nuance to theories of personality and symptom change and suggest that treatments may be designed to capitalize on the relatively more unidirectional effects of personality on symptom changes.

Clinicians implementing the UP may expect to see small-sized reductions in neuroticism and large-sized reductions in anxiety and depression across the first six sessions in their patients. These early reductions in neuroticism may be more reflective of improvement in anxiety and depression than enduring personality change per se. Continued use of the UP beyond six sessions may lead to more substantial changes in trait levels of neuroticism that is not accounted for by any further reductions in anxiety and depression. Because

reductions in neuroticism predicted reductions in anxiety and depression, the UP's approach to targeting aversive reactions to frequently occurring negative emotions may represent an efficient way to address both a personality vulnerability to emotional disorders, as well as symptoms themselves. To that end, we encourage future researchers to test if the additional trait level change in neuroticism that occurs later in treatment leads to longer-lasting improvements in symptoms (compared to state-level changes observed when symptoms improve) or reduces the likelihood of relapse as neuroticism contributes to the maintenance of emotional disorders. Additionally, more research is needed to determine if the trait-level changes observed with the UP are maintained past the 12 week treatment window in this study.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.beth.2023.05.005.

References

Akiskal, H. S., Hirschfeld, R. M. A., & Yerevanian, B. I. (1983). The relationship of personality to affective disorders. *Archives of General Psychiatry, 40*, 801–810.

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596.

Andrews, G. (1996). Comorbidity and the general neurotic syndrome. *The British Journal of Psychiatry, 168*(S30), 76–84.

Armstrong, L., & Rimes, K. A. (2016). Mindfulness-based cognitive therapy for neuroticism (Stress Vulnerability): A pilot randomized study. *Behavior Therapy, 47*(3), 287–298. https://doi.org/10.1016/j.beth.2015.12.005.

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., Carl, J. R., Bullis, J. R., & Ellard, K. K. (2014). The nature, diagnosis, and treatment of neuroticism. *Clinical Psychological Science, 2*(3), 344–365.

Barlow, D. H., Sauer-Zavala, S., Farchione, T., Murray Latin, H., Ellard, K., Bullis, J., Bentley, K., Boettcher, H., & Cassiello-Robbins, C. (2018). *The unified protocol for transdiagnostic treatmetn of emotional disorders: Client workbook*. Oxford University Press.

Barnett, P. A., & Gotlib, I. H. (1988). Psychosocial functioning and depression: Distinguishing among antecedents, concomitants, and consequences. *Psychological Bulletin, 104*(1), 97–126. https://doi.org/10.1037/0033-2909.104.1.97.

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.

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.

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: Science and Practice, 26*(2), e12278. https://doi.org/10.1037/h0101755.

Carl, J. R., Gallagher, M. W., Sauer-Zavala, S. E., Bentley, K. H., & Barlow, D. H. (2014). A preliminary investigation of the effects of the unified protocol on temperament. *Comprehensive Psychiatry, 55*(6), 1426–1434. https://doi.org/10.1016/j.comppsych.2014.04.015.

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.

Clark, L. A., Watson, D., & Mineka, S. (1994). Temperament, personality, and the mood and anxiety disorders. *Journal of Abnormal Psychology, 103*(1), 103–116.

Costa, P. T., & McCrae, R. R. (1989). *NEO five-factor inventory (NEO-FFI)* (pp. 3–). Odessa, FL: Psychological Assessment ResourcesOdessa, FL: Psychological Assessment Resources.

Craske, M. G., Meuret, A. E., Ritz, T., Treanor, M., Dour, H., & Rosenfield, D. (2019). Positive affect treatment for depression and anxiety: A randomized clinical trial for a core feature of anhedonia. *Journal of Consulting and Clinical Psychology, 87*(5), 457–471. https://doi.org/10.1037/ccp0000396.

De Fruyt, F., Van Leeuwen, K., Bagby, R. M., Rolland, J. P., & Rouillon, F. (2006). Assessing and interpreting personality change and continuity in patients treated for major depression. *Psychological Assessment, 18*(1), 71–80. https://doi.org/10.1037/1040-3590.18.1.71.

Du, L., Bakish, D., Ravindran, A. V., & Hrdina, P. D. (2002). Does fluoxetine influence major depression by modifying five-factor personality traits? *Journal of Affective Disorders, 71*(1–3), 235–241.

Eaton, N. R., Krueger, R. F., & Oltmanns, T. F. (2011). Aging and the structure and long-term stability of the internalizing spectrum of personality and psychopathology. *Psychology and Aging, 26*(4), 987–993. https://doi.org/10.1037/a0024406.

Fanous, A. H., Neale, M. C., Aggen, S. H., & Kendler, K. S. (2007). A longitudinal study of personality and major depression in a population-based sample of male twins. *Psychological Medicine, 37*(8), 1163–1172. https://doi.org/10.1017/S0033291707000244.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149.

Goldstein, B. L., Kotov, R., Perlman, G., Watson, D., & Klein, D. N. (2018). Trait and facet-level predictors of first-onset depressive and anxiety disorders in a community sample of adolescent girls. *Psychological Medicine, 48*(8), 1282–1290. https://doi.org/10.1017/S0033291717002719.

Helson, R., Jones, C., & Kwan, V. S. Y. (2002). Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples. *Journal of Personality and Social Psychology*, 752–766.

Hirschfeld, R. M., & Klerman, G. L. (1979). Personality attributes and affective disorders. *The American journal of psychiatry, 136*(1), 67–70.

Jaeger, B. C., Edwards, L. J., Das, K., & Sen, P. K. (2017). An *R*2 statistic for fixed effects in the generalized linear mixed model. *Journal of Applied Statistics, 44*(6), 1086–1105. https://doi.org/10.1080/02664763.2016.1193725.

Kendler, K. S., Neale, M., Kessler, R., Heath, A. C., & Eaves, L. J. (1993). A longitudinal twin study of personality and major depression in women. *Archives of General Psychiatry, 50*, 853–862. https://doi.org/10.1001/archpsyc.1993.01820210024003.

Khan, A. A., Jacobson, K. C., Gardner, C. O., Prescott, C. A., & Kendler, K. S. (2005). Personality and comorbidity of common psychiatric disorders. *The British Journal of Psychiatry: The Journal of Mental Science, 186*, 190–196. https://doi.org/10.1192/bjp.186.3.190.

Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin, 136*(5), 768–821. https://doi.org/10.1037/a0020327.

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology, 2*(1), 111–133. https://doi.org/10.1146/annurev.clinpsy.2.022305.095213.

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 for performing power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science, 4*(1), 1–24. https://doi.org/10.1177/2515245920978738.

Magidson, J. F., Roberts, B. W., Collado-Rodriguez, A., & Lejuez, C. W. (2014). Theory-driven intervention for changing personality: Expectancy value theory, behavioral activation, and conscientiousness. *Developmental psychology, 50*(5), 1442.

Malouff, J. M., Thorsteinsson, E. B., & Schutte, N. S. (2005). The relationship between the five-factor model of personality and symptoms of clinical disorders: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment, 27*(2), 101–114. https://doi.org/10.1007/s10862-005-5384-y.

Mroczek, D. K., & Spiro, A. III, (2003). Modeling intraindividual change in personality traits: Findings from the normative aging study. *The Journals of Gerontology: Series B, 58*(3), P153–P165. https://doi.org/10.1093/geronb/58.3.P153.

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.

Ormel, J., Jeronimus, B. F., Kotov, R., Riese, H., Bos, E. H., Hankin, B., Rosmalen, J. G. M., & Oldehinkel, A. J. (2013). Neuroticism and common mental disorders: Meaning and utility of a complex relationship. *Clinical Psychology Review, 33*(5), 686–697. https://doi.org/10.1016/j.cpr.2013.04.003.

Ormel, J., Oldehinkel, A. J., & Vollebergh, W. (2004). Vulnerability before, during, and after a major depressive episode: A 3-wave population-based study. *Archives of General Psychiatry, 61*(10), 990–996. https://doi.org/10.1001/archpsyc.61.10.990.

Payne, L. A., Ellard, K. K., Farchione, T. J., Fairholme, C. P., & Barlow, D. H. (2014). Emotional disorders: A unified transdiagnostic protocol. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders: A step-by-step treatment manual* (5th ed.. New York, NY: Guilford Press.

Quilty, L. C., Meusel, L. A., & Bagby, R. M. (2008). Neuroticism as a mediator of treatment response to SSRIs in major depressive disorder. *Journal of Affective Disorders, 111*(1), 67–73. https://doi.org/10.1016/j.jad.2008.02.006.

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*. https://doi.org/10.1037/bul0000088. Advanced online publication.

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*(1), 31–35. https://doi.org/10.1111/j.1467-8721.2008.00543.x.

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.

Santor, D. A., Bagby, R. M., & Joffe, R. T. (1997). Evaluating stability and change in personality and depression. *Journal of Personality and Social Psychology, 73*(6), 1354–1362. https://doi.org/10.1037//0022-3514.73.6.1354.

Sauer-Zavala, S., Fournier, J. C., Jarvi Steele, S., Woods, B. K., Wang, M., Farchione, T. J., & Barlow, D. H. (2021). Does the unified protocol really change neuroticism? Results from a randomized trial. *Psychological Medicine, 51*(14), 2378–2387. https://doi.org/10.1017/S0033291720000975.

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

Sauer-Zavala, S., Southward, M. W., Stumpp, N. E., Semcho, S. A., Hood, C. O., Garlock, A., & Urs, A. (2022b). A SMART approach to personalized care: Preliminary data on how to select and sequence skills in transdiagnostic CBT. *Cognitive Behaviour Therapy*. https://doi.org/10.1080/16506073.2022.2053571. Advance online publication.

Sauer-Zavala, S., Southward, M. W., Terrill, D. R., Semcho, S. A., & Stumpp, N. E. (2023). Mechanism engagement as a potential evidence-based approach to personalized treatment termination. *Psychotherapy Research*. https://doi.org/10.1080/10503307.2023.2168574. Advance online publication.

Sher, K. J., & Trull, T. J. (1994). Personality and disinhibitory psychopathology: Alcoholism and antisocial personality disorder. *Journal of Abnormal Psychology, 103*(1), 92–102. https://doi.org/10.1037//0021-843x.103.1.92.

Small, B. J., Hertzog, C., Hultsch, D. F., & Dixon, R. A. (2003). Stability and change in adult personality over 6 years: Findings from the victoria longitudinal study. *The Journals of Gerontology: Series B, 58*(3), P166–P176. https://doi.org/10.1093/geronb/58.3.P166.

Soskin, D. P., Carl, J. R., Alpert, J., & Fava, M. (2012). Antidepressant effects on emotional temperament: Toward a biobehavioral research paradigm for major depressive disorder. *CNS Neuroscience & Therapeutics, 18*(6), 441–451.

Southward, M. W., Cheavens, J. S., & Coccaro, E. F. (2022). Defining the *p*-factor: An empirical test of five leading

theories. *Psychological Medicine*. https://doi.org/10.1017/S0033291722001635. Advance online publication.

Southward, M. W., & Sauer-Zavala, S. (2022). Dimensions of skill use in the unified protocol: Exploring unique effects on anxiety and depression. *Journal of Consulting and Clinical Psychology, 90*(3), 246–257. https://doi.org/10.1037/ccp0000701.

Tang, T. Z., DeRubeis, R. J., Hollon, S. D., Amsterdam, J., Shelton, R., & Schalet, B. (2009). A placebo-controlled test of the effects of paroxetine and cognitive therapy on personality risk factors in depression. *Archives of General Psychiatry, 66*(12), 1322–1330. https://doi.org/10.1001/archgenpsychiatry.2009.166.

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.

Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*(1), 63–83. https://doi.org/10.1037/met0000030.

Warner, M. B., Morey, L. C., Finch, J. F., Gunderson, J. G., Skodol, A. E., Sanislow, C. A., Shea, M. T., McGlashan, T. M., & Grilo, C. M. (2004). The longitudinal relationship of personality traits and disorders. *Journal of Abnormal Psychology., 2004*(113), 217–227.

Weinstock, L. M., & Whisman, M. A. (2006). Neuroticism as a common feature of the depressive and anxiety disorders: A test of the revised integrative hierarchical model in a national sample. *Journal of Abnormal Psychology, 115*(1), 68–74. https://doi.org/10.1037/0021-843X.115.1.68.

Wilamowska, Z. A., Thompson-Hollands, J., Fairholme, C. P., Ellard, K. K., Farchione, T. J., & Barlow, D. H. (2010). Conceptual background, development, and preliminary data from the unified protocol for transdiagnostic treatment of emotional disorders. *Depression and Anxiety, 27*(10), 882–890. https://doi.org/10.1002/da.20735.

Wright, A. G., Hallquist, M. N., Swartz, H. A., Frank, E., & Cyranowski, J. M. (2014). Treating co-occurring depression and anxiety: Modeling the dynamics of psychopathology and psychotherapy using the time-varying effect model. *Journal of Consulting and Clinical Psychology, 82*(5), 839–853. https://doi.org/10.1037/a0034430.

Received: September 19, 2022
Revised: May 12, 2023
Accepted: May 12, 2023
Available online: 24 May 2023