Change of Threat Expectancy as Mechanism of Exposure-Based Psychotherapy for Anxiety Disorders: Evidence From 8,484 Exposure Exercises of 605 Patients

Andre Pittig[1,2], **Ingmar Heinig[3]**, **Stephan Goerigk[3,4]**,
Jan Richter[5,6], **Maike Hollandt[5]**, **Ulrike Lueken[7,8]**, **Paul Pauli[1]**,
Jürgen Deckert[7], **Tilo Kircher[9]**, **Benjamin Straube[9]**, **Peter Neudeck[10]**,
Katja Koelkebeck[11], **Udo Dannlowski[11]**, **Volker Arolt[11]**,
Thomas Fydrich[8], **Lydia Fehm[8]**, **Andreas Ströhle[12]**, **Christina Totzeck[13]**,
Jürgen Margraf[13], **Silvia Schneider[13]**, **Jürgen Hoyer[3]**, **Winfried Rief[14]**,
Michelle G. Craske[15], **Alfons O. Hamm[5]**, and **Hans-Ulrich Wittchen[3,4]**

[1]Department of Psychology (Biological Psychology, Clinical Psychology, and Psychotherapy), University of Würzburg; [2]Translational Psychotherapy, Department of Psychology, Friedrich–Alexander University Erlangen–Nürnberg; [3]Institute of Clinical Psychology and Psychotherapy, Technische Universität Dresden; [4]Department of Psychiatry and Psychotherapy, University Hospital, Ludwig Maximilians University; [5]Department of Biological and Clinical Psychology/Psychotherapy, University of Greifswald; [6]Department of Experimental Psychopathology, University of Hildesheim; [7]Department of Psychiatry, Psychosomatics, and Psychotherapy, Center of Mental Health, University Hospital of Würzburg; [8]Department of Psychology, Humboldt-Universität zu Berlin; [9]Department of Psychiatry and Psychotherapy & Center for Mind, Brain and Behavior, Philipps-University Marburg; [10]Protect-AD Study Site Cologne, Cologne, Germany; [11]Institute for Translational Psychiatry, University of Muenster; [12]Department of Psychiatry and Psychotherapy, Campus Charité Mitte, Charité – Universitätsmedizin Berlin; [13]Mental Health Research and Treatment Center, Ruhr-Universität Bochum; [14]Department of Clinical Psychology and Psychotherapy, Faculty of Psychology & Center for Mind, Brain and Behavior, Philipps-University of Marburg; and [15]Department of Psychology and Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles

Abstract

Individual responses to behavioral treatment of anxiety disorders vary considerably, which requires a better understanding of underlying processes. In this study, we examined the violation and change of threat beliefs during exposure. From 8,484 standardized exposure records of 605 patients with different anxiety disorders, learning indicators were derived: expectancy violation as mismatch between threat expectancy before exposure and threat occurrence, expectancy change as difference between original and adjusted expectancy after exposure, and prediction-error learning rate as extent to which expectancy violation transferred into change. Throughout sessions, high threat expectancy but low occurrence and adjusted expectancy indicated successful violation and change of threat beliefs by exposure. Expectancy violation, change, and learning rate substantially varied between patients. Not expectancy violation itself, but higher learning rate and expectancy change predicted better treatment outcome. Successful exposure thus requires expectancy violation to induce actual expectancy change, supporting learning from prediction error as transdiagnostic mechanism underlying successful exposure therapy.

Corresponding Author:
Andre Pittig, Translational Psychotherapy, Department of Psychology, Friedrich–Alexander University Erlangen–Nürnberg
Email: andre.pittig@fau.de

Anxiety disorders, including panic disorder (PD), agoraphobia (AG), social anxiety disorder (SAD), and specific phobias (SP), belong to the most frequent mental disorders (Jacobi et al., 2014; Kessler et al., 2010). Without treatment, anxiety disorders tend to persist with a waxing and waning course, resulting in increasing individual and socioeconomic burden (Beesdo et al., 2007; Craske et al., 2017; Gustavsson et al., 2011; Wittchen et al., 2011). Timely and effective treatment is therefore a pivotal goal for mental-health care. Exposure-based cognitive-behavioral therapy (CBT) consistently yields large within-groups treatment effect sizes for symptom reduction, superiority to placebo control and other active treatments and shows comparable positive outcomes in routine care (e.g., Carpenter et al., 2018; Hofmann & Smits, 2008; Hoyer et al., 2017; Marcus et al., 2014; Tolin, 2010; Watts et al., 2015). In contrast to consistently positive effects on the group level, individual responses to exposure-based CBT vary substantially. Although some patients show full remission of symptoms, others do not fully benefit (Loerinc et al., 2015). Moreover, average treatment success tends to stabilize or improve in the long run (e.g., Carpenter et al., 2018; Gloster et al., 2013; Loerinc et al., 2015), but some individuals show a return of symptoms after successful treatment (Craske & Mystkowski, 2006; Pittig et al., 2021). Thus, although exposure-based CBT is highly effective, these findings highlight a need for optimization on the individual level. This calls for a better understanding of its underlying processes and mechanisms of change.

Inhibitory-learning models have highlighted fear-extinction learning as one potential mechanism of exposure (Craske et al., 2014, 2018; Pittig et al., 2016, 2018). Exposure involves repeated confrontation with feared stimuli or contexts (e.g., objects, situations, interoceptive stimuli, or fear memories) that are associated with an anticipated threat. For example, a patient may associate public speaking with the threat of social rejection and thus experience anxiety when confronted with giving a presentation at work. Such threat associations can be a result of direct experience, instruction, or vicarious learning (Rachman, 1977). "Fear extinction" refers to the process of learning that the anticipated threat does not occur or no longer occurs (Craske et al., 2018; Hermans et al., 2006) and is thus initiated by prediction error (Rescorla & Wagner, 1972). Inhibitory-learning theory assumes that this learning process does not erase the original threat association but prompts a novel association that the anticipated threat does not occur under specific circumstances. This novel learning actively inhibits the original threat association and thereby down-regulates fear and anxiety responses (Bouton, 2002, 2004). Consequently, recent basic and clinical research has aimed to enhance inhibitory learning during exposure to optimize treatment outcome.

To this end, a variety of strategies have been proposed during the last years (Craske et al., 2014, 2018; Pittig et al., 2016, 2018). Clinically, the most frequently proposed strategy is the maximization of threat-expectancy violation (Craske et al., 2014; Hofmann, 2008; Pittig et al., 2016). In humans, threat associations entail expectancies about the occurrence of the perceived threat (e.g., "When I give a presentation, the audience will judge me as incompetent and reject me"). "Expectancy violation" (EV) refers to the mismatch between this threat expectancy and the actual occurrence. This mismatch is believed to be the clinical indicator for the occurrence of a prediction error and is thus assumed to create novel inhibitory learning (e.g., "The audience did not reject me after this presentation"). The important role of a mismatch between an expected outcome and its absence was highlighted by Rescorla and Wagner (1972), who stated that "organisms only learn when events violate their expectations." Associated learning models mathematically formalize that the rate of learning crucially depends on the magnitude of EV: the larger the violation of expectancies, the more learning should occur. As a clinical implication, it is thus assumed that the more threat expectancies are violated during exposure, the better the treatment outcome (e.g., Craske et al., 2014; Hofmann, 2008; Pittig et al., 2016).

Despite strong basic science (see Craske et al., 2014, 2018; Pittig et al., 2016, 2018), only a few clinical studies have evaluated EV and its role for treatment success. A higher proportion of anticipated threats not occurring during exposure, relative to the overall number of anticipated threats, was associated with better treatment response to large-group one-session exposure in fear of heights (Wannemueller et al., 2019). Imaginal exposure in posttraumatic stress disorder resulted in significant violation of expectancies (de Kleine et al., 2017), which was not associated with treatment outcome. Expectancies, however, partly referred to the anticipation of strong fear responses during exposure (e.g., "I fear to panic"). Fear during exposure is likely to occur and does not represent an aversive unconditioned stimulus (US) or consequence. In a conditioning model, fear during exposure is the conditioned response in anticipation of an aversive US. Strictly, testing the occurrence of fear responses during exposure does not fit with the inhibitory expectancy-violation model. Moreover, intensified interoceptive exposure focusing on maximizing EV was more effective than standard exposure in subclinical PD (Deacon et al., 2013). However, comparison between the two types of exposure was confounded by a higher dose of intensified exposure. In addition, this past research was mostly based

on small and/or subclinical samples (Deacon et al., 2013; Wannemueller et al., 2019) or accounted only for parts of the conducted exposure exercises (e.g., imaginal but not in vivo exposure; de Kleine et al., 2017). Moreover, varying indicators of EV were used, for example, the proportion of anticipated threats not occurring during exposure, the proportion of testable compared with untestable beliefs, or average mismatch between threat expectancy and threat occurrence. A more comprehensive investigation of the mismatch between threat beliefs and threat occurrence, their course across treatment, and the association of different indicators with treatment success is thus needed.

Crucially, EV can be distinguished from expectancy change (EC). In humans, the amount of learning from the same experience typically differs between individuals. For example, a lack of adjusting expectancies on the basis of experience has been associated with various psychopathologies, including anxiety disorders (e.g., Beck & Haigh, 2014; Duits et al., 2015; Korn et al., 2014; Rief et al., 2015). These findings highlight that EV does not necessarily result in EC (of equal magnitude). In learning models, this is typically addressed by individual learning rates (i.e., the amount EC depends on the overall EV weighted by a learning rate). This learning rate thus quantifies the individual extent to which EV is transferred into EC. During exposure, a patient may, for example, test the specific threat expectancy "When I give a presentation, I will make a mistake and the audience will laugh at me" by giving a presentation in front of an audience and experience no laughing (i.e., EV). However, if the threat expectancy remained unchanged for a repeated presentation, the learning rate would be 0. If the threat expectancy became half as high as before, the learning rate would be 0.5. It is thus important to evaluate the rate of learning from EV, the relationship between these different learning indicators, and their link to treatment outcome.

To this end, in the present study, we analyzed 8,484 standardized records of exposure exercises from 651 patients with PD, AG, SAD, and multiple SP. Patients completed multiple exposure exercises during a 14-session manualized exposure-based CBT and standardized diagnostic assessments before ($n = 651$) and after treatment ($n = 605$; Heinig et al., 2017; Pittig et al., 2021). For each exposure exercise, patients indicated the specific anticipated threat and rated their subjective probability that this threat would occur during exposure (i.e., threat expectancy). After exposure, patients rated actual threat occurrence and their new threat expectancy for repeating the same exercise (i.e., adjusted threat expectancy). Data were used to calculate three process-based learning indicators: the extent of EV and EC and each patient's exposure-related learning rate.

Our two major aims were to examine (a) the intensity, violation, and change of threat expectancies across exposure-based CBT in a large patient sample with severe anxiety disorders and (b) the association of these process variables with treatment outcome. We hypothesized that higher EV is associated with higher EC, which is associated with better treatment outcome. In addition, higher learning rate was expected to be associated with better treatment outcome.

Method

General study design

Data were taken from a multicenter controlled clinical trial for optimizing exposure therapy for adults with anxiety disorders (Heinig et al., 2017; Pittig et al., 2021). This trial was preregistered at the National Institute of Mental Health Protocol Registration System (01EE1402A) and the German Register of Clinical Studies (DRKS00008743). In the trial, 726 patients with anxiety disorders were recruited and treated at eight study sites across Germany using a standardized treatment manual. All patients provided written informed consent to study procedures approved by the local ethic committee (Ethics Committee of Technische Universität Dresden, EK 234062014). Patients were randomly assigned to two treatment groups and completed diagnostic and experimental add-on procedures before, during, and after treatment (Heinig et al., 2017).

Patients

For the main trial, inclusion criteria were a principal diagnosis of SAD, PD, AG, or multiple SP (i.e., at least two different SP diagnoses) based on the fifth edition of the *Diagnostic Statistical Manual of Mental Disorders* (*DSM-5*; American Psychiatric Association, 2013). Diagnoses were established by trained interviewers using a standardized computer-assisted version of the Composite International Diagnostic Interview (CIDI DIA-X; see Wittchen, 1994). Additional baseline severity criteria were clinician-rated anxiety symptom severity of 19 or higher on the Hamilton Anxiety Rating Scale (SIGH-A; Shear et al., 2001) and clinician-rated overall severity of 4 or higher on the Clinical Global Impression (CGI) Scale, indicating at least a moderate severity (Guy, 1976). Additional inclusion criteria were age 15 to 70, able to attend outpatient treatment sessions, and sufficient German language skills (determined during the CIDI and defined as being able to understand and answer the interview questions). Exclusion criteria were any psychotic disorder, current substance use disorder (except nicotine dependence), primary mood or bipolar disorders, acute

suicidality, parallel psychotherapy, any medical contra-indication for exposure or psychotherapy in general, and monosymptomatic specific phobia. Other comorbid disorders, such as depressive disorders or other anxiety disorders, were not excluded.[1] Patients in psychopharmacological treatment could be included when medication was stable for at least 3 months. In total, 726 patients were included in the main trial (Pittig et al., 2021).

In the present study, analysis focused on novel data from standardized records of exposure exercises as part of the treatment, which have not been reported before. No records were available from 63 patients (8.6%) because they dropped out before exposure, and 12 patients (1.7%) did not file any record during treatment. As a result, 651 patients (89.7%) were included for the analyses of exposure records. Six hundred seven patients were born in Germany (93.2%; 14.6% had at least one parent born in another country). In a second step, the data obtained from the exposure records were related to symptom reduction from before to after treatment. For 46 of the 651 patients, no postassessment data were available (either because of dropout or missing the postassessment measure; see Pittig et al., 2021). Thus, 605 patients (83.3%) were included for the analyses of outcome associations. Five hundred sixty-five patients were born in Germany (93.3%; 14.4% had at least one parent born in another country). Demographic and clinical data of the samples are shown in Table 1.

Treatment and therapists

In the main trial, patients were randomly assigned to two treatment groups. Both groups were treated with a content-identical treatment manual but differed in the temporal spacing of sessions. Treatment consisted of 14 sessions (100 min each) divided into four phases (see Heinig et al., 2017). Components of the cognitive-preparation phase (Sessions 1–4) were (a) general psychoeducation (e.g., information on anxiety, functional behavior analysis); (b) developing individual models for disorder development and maintenance; (c) identification of individual threat beliefs; (d) understanding dysfunctional effects of threat beliefs, avoidance, safety behavior, and other maladaptive anxiety-control strategies; and (e) the exposure rationale. The exposure phase (Sessions 5–10) began with two sessions of therapist-guided exposure in session (Sessions 5 and 6) and self-guided exposure between sessions (i.e., homework exposure). The next session (Session 7) included an interim evaluation to recapitulate experience from prior exposures, planning subsequent exposures, and a

homework exposure. For Sessions 8 through 10, further therapist-guided and self-guided exercises were conducted to challenge most prominent threat beliefs. The self-management phase (Sessions 11 and 12) did not include therapist-guided exposure but emphasized self-exposure between sessions (i.e., continuing exposure in daily life) and relapse prevention. Summarizing the main treatment phase (Sessions 5–12), all patients were asked to complete five therapist-guided exposures (in Sessions 5, 6, 8, 9, 10) and at least one between-session self-guided exposure after Sessions 5 to 11, respectively. Patients could, however, complete more than one self-guided exposure between sessions. Postassessment measures were completed after Session 12. The final booster phase (Sessions 13 and 14) included booster sessions 2 and 4 months after Session 12 without therapist-guided exposure but with emphasis on self-guided exposure.

Treatment groups received the same treatment content, which differed only in the temporal spacing of Sessions 5 to 10. Whereas one group received weekly sessions (nonintensified prediction-error-based exposure [PeEx-S]), the other group received three sessions per week (temporally intensified prediction-error-based exposure [PeEx-I]), resulting in a duration of 6 versus 2 weeks for the exposure phase (Heinig et al., 2017). Because there were no significant group differences for the used primary and secondary outcome measures at posttreatment (Pittig et al., 2021), we combined all patients for the present analyses. However, we included treatment group as control factor in the statistical models.

Treatment was delivered by trained and certified study therapists who were either licensed cognitive-behavioral psychotherapists or postgraduate trainees in advanced training for cognitive-behavioral psychotherapy. All therapists received a comprehensive 2-day training of the treatment manual and subsequently had to treat one case to become certified as study therapist. Certification required successful adherence and competence to deliver the treatment manual, which was evaluated by experienced CBT therapists on the basis of video recordings of five standardized treatment sequences. For more details on adherence, treatment integrity, and outcome, see Pittig et al. (2021).

Regarding exposure, therapists were trained to define individual threat beliefs for each exposure together with the patient and design exposure to violate these threat beliefs. Following the inhibitory-learning approach, therapists were specifically trained to define threat beliefs about a testable feared outcome (e.g., "I will get a heart attack and die!")[2] but not about fear responses during exposure (e.g., "I will panic!"). When a testable outcome

Table 1. Sociodemographic and Clinical Data of the Patients With Anxiety Disorders

	Exposure records (N = 651)	Exposure records and outcome data (n = 605)
Age	33.36 (11.59)	33.21 (11.56)
Female sex	359 (55.2%)	334 (55.2%)
Socioeconomic status (lower/lower-middle/ middle/upper-middle/upper class)	37/151/374/85/2	31/142/350/78/2
Primary disorder		
PD/AG[a]	393 (60.3%)	363 (60.0%)
SAD	195 (30.0%)	184 (30.4%)
SP	63 (9.7%)	58 (9.6%)
Any PD/AG	442 (67.9%)	406 (67.1%)
Any SAD	267 (41.0%)	252 (41.7%)
Any multiple SP	235 (36.1%)	218 (36.0%)
Any depressive disorder (MDD/dysthymia)	301 (46.2%)	276 (45.6%)
Years since onset of primary diagnosis	13.73 (12.19)	13.85 (12.20)
Number of diagnoses[b]	3.98 (1.91)	3.94 (1.90)
On stable medication	170 (26.1%)	157 (26.0%)

Symptom measures (unstandardized)	Pretreatment (Exposure records, N = 651)	Pretreatment (Exposure records and outcome data, n = 605)	Posttreatment (Exposure records and outcome data, n = 605)
Clinician-rated anxiety (HAMA)	24.44 (5.33)	24.43 (5.30)	12.70 (8.09)
Clinical Global Impression	5.01 (0.68)	5.02 (0.68)	3.20 (1.28)
Anxiety across diagnoses (DSM-5 Cross-D)	15.65 (7.37)	15.64 (7.38)	6.15 (5.27)
Panic and agoraphobia (PAS)	20.52 (10.85)	20.43 (10.95)	8.37 (8.34)
Social anxiety (LSAS)	47.17 (30.56)	47.42 (30.59)	30.21 (25.46)
Specific phobia (DSM-5 SP)	15.58 (10.41)	15.49 (10.47)	6.91 (6.69)
Depression (BDI)	16.47 (9.67)	16.53 (9.71)	9.00 (9.03)

Note: Means (with standard deviation) or number (with percentage) for samples included in analyses of exposure records and outcome association. PD/AG = panic disorder and/or agoraphobia; SAD = social anxiety disorder; SP = specific phobias (SP); MDD = major depressive disorder; HAM-A = Hamilton Anxiety Rating Scale; DSM-5 Cross-D = *Diagnostic and Statistical Manual of Mental Disorders* Cross-Cutting Dimensional Scale for Anxiety Disorders; PAS = Panic and Agoraphobia Scale; LSAS = Liebowitz Social Anxiety Scale; DSM-5 SP = DSM-5 Dimensional Specific Phobia Scale; BDI = Beck Depression Inventory.
[a]Because of the small number of cases diagnosed with independent agoraphobia and panic disorder (see Pittig et al., 2021), we included these cases in the PD/AG group. Results remained the same when separating cases with these primary disorders.
[b]Including primary disorder.

was unclear or ambiguous for a specific threat belief (e.g., "I will go crazy" or "I will lose control"), therapists were trained to determine concrete and testable outcomes or behavioral responses associated with these beliefs (e.g., by asking patients what would happen or how it would look if they go crazy). Moreover, therapists were trained to conduct exposure for the most prominent threat beliefs and encouraged to switch to novel stimuli or contexts and target different threat beliefs across exposure exercises. In line with this, exposure exercises could also address threat beliefs associated with different diagnoses of a patient (e.g., SAD and PD). As a result, exposure exercises were typically not repeated but considerably varied in terms of content, stimulus, and context for an individual patient.

Exposure records to assess mechanisms of change

Before and after each exposure exercise, patients were asked to complete a structured record. Before exposure, patients indicated their specific threat belief that was targeted during exposure and provided a detailed description of the planned exercise (e.g., "Take a crowded bus during hot temperature without taking my phone with me"). They also rated their subjective

expectancy of their feared threat to occur (0%–100%) and the level of anticipatory anxiety (0 = *not at all*, 100 = *maximal anxiety*). After the exercise, patients indicated whether they completed the exercise and whether it was conducted as therapist-guided or self-guided exposure (as part of between-session homework exercises). Moreover, patients rated the extent to which their expected threat actually occurred (0%–100%) and their adjusted expectancy if they would repeat the same exercise (0%–100%). The expectancy ratings were used as measures for (a) threat expectancy, (b) threat occurrence, and (c) adjusted threat expectancy.

Overall, we collected 13,287 exposure records. For the present analyses, we focused on exercises during the main treatment phases (Sessions 5–12) because of inconsistent completion of records during the subsequent booster phase. For this main treatment phase, we included records of 8,484 exposure exercises from 651 patients, of which 605 patients completed the postassessment measures. On average, patients completed 13 exposure exercises ($M = 13.09$, $SD = 6.09$), which did not differ between treatment groups, $t(645.24) = 0.98$, $p = .329$, $d = 0.07$.

Outcome measures

Patients completed a comprehensive baseline assessment to determine inclusion/exclusion criteria, which included clinical interviews and self-report questionnaires. Further comprehensive assessments were conducted after Session 12 (postassessment measure) and 6-month follow-up (Heinig et al., 2017; Pittig et al., 2021).

Transdiagnostic primary outcomes were clinician-rated anxiety symptoms on the Hamilton Anxiety Rating Scale (HAM-A; Shear et al., 2001; Cronbach's α for the current sample at baseline = .41) and clinician-rated overall severity on the CGI (Guy, 1976; Cronbach's α = .76). For secondary outcome measures, all patients completed the same questionnaire battery assessing self-reported transdiagnostic and disorder-specific symptoms and global functioning. For the present analyses, we included the six most important symptom measures for treatment outcome. For transdiagnostic symptom measures, we included both clinician-rated primary outcomes (SIGH-A and CGI) and self-reported anxiety symptoms across disorders assessed with the DSM-5 Cross-Cutting Dimensional Scale for Anxiety Disorders (DSM-5 Cross-D) as the only transdiagnostic self-report measure included in the trial (Lebeau et al., 2012; Cronbach's α = .87). In addition, the main self-report measures for each primary diagnosis were included (as indicated in Heinig et al., 2017): (a) The Liebowitz Social Anxiety Scale (LSAS; Fresco et al.,

2001; Liebowitz, 1987; Cronbach's α = .97) served as self-reported symptom measure for SAD, (b) the Panic and Agoraphobia Scale (PAS; Bandelow, 1995; Cronbach's α = .87) served as self-reported symptom measure for PD and AG, and (c) the DSM-5 Dimensional Specific Phobia Scale (DSM-5 SP; Craske et al., 2013; Lebeau et al., 2012; Cronbach's α = .93) served as self-reported symptom measure for multiple SP. The DSM-5 SP was used as a combined measure for all specific phobias of a patient (i.e., a single symptom score for multiple phobias). Patients were provided with examples of phobias (e.g., animals, height, blood, injections, storms, bridges) and instructed to answer the items for all relevant stimuli and situations combined. At both baseline and postassessment, these six outcome measures were combined into a standardized composite score to form a more reliable index of symptom severity (Steketee & Chambless, 1992). First, each measure was z transformed using mean and standard deviation of the full sample at baseline: ($X − M$ before treatment) / SD before treatment (Pittig et al., 2021). Thus, negative values after treatment indicate symptom reduction, and positive values indicate symptom increase. Next, a composite symptom score was separately calculated at baseline and postassessment by averaging across outcome measures (i.e., SIGH-A, CGI, DSM-5 Cross-D, LSAS, PAS, DSM-5 SP).

Data reduction and statistical analyses

Main aims were to examine (a) the intensity, violation, and change of threat expectancies during exposure-based CBT and (b) the association of these process variables with treatment outcome. For the first aim, we first compared the raw expectancy ratings across sessions (i.e., threat expectancy, threat occurrence, adjusted expectancy if a patient would repeat the same exercise). To this end, we used linear mixed models, which allowed us to include all available data and are not refrained to complete cases as traditional repeated measures analyses of variance (Singmann & Kellen, 2019). Models were calculated in R (R Core Team, 2020) using the *afex* package (Singmann et al., 2020). The analyzed exposure records were collected from therapist-guided and self-guided exposure exercises from Sessions 5 to 12. We therefore included rating (threat expectancy vs. threat occurrence vs. adjusted expectancy), session (5 to 12 as continuous variable), and exercise (therapist- vs. self-guided) as fixed factors. Because the three-way interaction was not significant, $F(2, 22574.09) = 0.96$, $p = .383$, we included only two-way interactions of the fixed factors in the final model

for simplicity. Exclusion of the three-way interaction did not change results for the remaining effects. In addition, treatment group was included as fixed factor without interactions. Patients were entered as random factor with random intercept and random slopes for session. The final model did not include rating as random slope or correlation between random components because models failed to converge (Singmann & Kellen, 2019). For comparison of estimated marginal means in each session, a similar model with session as factor was conducted. For these analyses, effect sizes were calculated as Cohen's *d* using raw data (i.e., means and standard deviation).

Next, we calculated different learning indicators for EV, EC, and the patients' prediction-error learning rate (α). Because operationalization of EV has been inconsistent and there is no clear standard, we calculated different indicators and their associations to provide a better understanding of EV and EC during exposure therapy. For EV and EC, absolute and relative indicators were calculated. In line with previous research, absolute EV was calculated as the average mismatch between threat expectancy and threat occurrence (i.e., threat expectancy – threat occurrence averaged across exposure exercises of an individual patient; Craske et al., 2014; de Kleine et al., 2017). Absolute EV can range from –100 to 100; 100 indicates maximal EV in each exposure exercise (–100 indicates the hypothetical case of threat expectancy = 0 but threat occurrence = 100). Because the expected threat does typically not occur during exposure, absolute EV is strongly determined by the initial level of threat expectancy. This is problematic for translating the within-subjects prediction of the EV model to a between-subjects analyses (e.g., see Fisher et al., 2018). The within-subjects prediction can be phrased as "Higher expectancy violation during a specific exposure exercise should result in better outcome for an individual patient." For example, a patient should respond better from a specific exposure, in which no heart attack occurred (threat occurrence = 0), if threat expectancy could be raised to 100 (EV = 100) compared with 50 (EV = 50). Typically, this is translated to the between-subjects level as "A patient with higher expectancy violation should show better outcome compared with a patient with lower expectancy violation." However, many factors can contribute to differences in initial threat expectancies between patients, such as the specific threat belief, the number of repetitions of the same exposure, primary diagnoses, and many more. Most importantly, one patient compared with another patient may simply show less severe or strong threat expectancy for the same exposure: Whereas one patient may be 100% certain to get a heart attack, another patient may be only 50% certain. When

experiencing the nonoccurrence of a heart attack, direct translation of the within-subjects prediction would assume better outcome for the former patient although this patient shows more severe threat expectancy. In addition, absolute EV does not distinguish between a patient reporting threat expectancy of 100 and threat occurrence of 50 and a patient reporting threat expectancy of 50 and threat occurrence of 0 (EV would be 50 for both). Thus, absolute EV may be biased when predicting treatment outcome on a between-subjects level. In addition, accounting for the individual level of threat expectancy is important. The same applies to absolute EC, which we calculated as average difference between threat expectancy and adjusted expectancy after exposure.

Therefore, we also calculated relative indicators that account for the individual level of threat expectancy. Relative EV was calculated as the proportion of mismatch between threat expectancy and threat occurrence relative to threat expectancy: (threat expectancy – threat occurrence) / threat expectancy. Relative EV thus represents the proportion of violating a patient's overall threat expectancies across exposure sessions. Likewise, relative EC was calculated as the proportion of difference between threat expectancy and adjusted expectancy relative to threat expectancy: (threat expectancy – adjusted expectancy) / threat expectancy. Both relative indicators can range from a maximum of 1 to negative infinity: 1 indicates that 100% of the patient's overall threat expectancy was violated or changed by exposure, 0 indicates that 0% were violated or changed, and negative values indicate that threat occurrence or adjusted expectancy was higher than original threat expectancy. In sum, the indicators represent the individual proportion of violated or changed threat expectancy. Relative indicators do not distinguish between patients reporting threat expectancy of 100 and threat occurrence of 0 and patients reporting threat expectancy of 50 and threat occurrence of 0 (for both, relative EV would be 1). Again, accounting for the individual level of threat expectancy for outcome prediction is important.

In addition, we computed the exposure-related learning rate for each patient. Individual learning rates were estimated as free parameter via maximum likelihood estimation using the following formula[3]:

$$\text{Adjusted Expectancy}_i = \text{Threat Expectancy}_i$$
$$+ \alpha \times (\text{Threat Occurrence}_i - \text{Threat Expectancy}_i),$$

where α represents the patient's prediction-error learning rate across exposure exercises and (Threat Occurrence$_i$ – Threat Expectancy$_i$) represents the magnitude of EV of exercise *i* (i.e., the prediction error).

Learning rate was restricted to values from 0 to 1. We used a Gaussian likelihood distribution with the predicted adjusted expectancy as mean and $SD = 30$. In sum, we estimated a single learning-rate value per patient, which describes the best fitting value for the extent to which EV was transferred into EC; $\alpha = 1$ indicates that the full extent was transferred into EC, and $\alpha = 0$ indicates that EV did not result in any EC.

For the second aim, we used these learning indicators to predict treatment outcome. For treatment outcome, we used residual-gain scores (RGSs). RGSs better control for initial differences in symptom severity and for repeated-measurement error (Steketee & Chambless, 1992). RGSs were calculated by multiplying composite symptom scores at baseline with the correlation between pretreatment and posttreatment scores and subtracting composite symptom scores at posttreatment (Steketee & Chambless, 1992). A higher RGS thus indicates higher symptom reduction. We used robust linear regression to predict symptom reduction (RGSs) while accounting for variables of potential influence using the *robustbase* package (Maechler et al., 2020). As a first step, we therefore defined a baseline model including predictors of interest. The baseline model included the number of comorbid disorders, the individual number of exposure exercises, intake of psychotropic medication (0 = *no*, 1 = *yes*), and treatment group and, importantly, accounted for the individual level of threat expectancy. In the next step, we separately added each learning indicator and compared the resulting models to the baseline model. We repeated this approach in each primary disorder category (PD/AG, SAD, multiple SP) using the RGSs calculated from the standardized disorder-specific symptom measure (PAS, LSAS, DSM-5 SP).

Results

Intensity, violation, and change of threat expectancy during exposure-based CBT

Threat expectancy, threat occurrence, and adjusted expectancy. Average self-reported threat expectancy, threat occurrence, and adjusted expectancy are shown in Figure 1 (and see Table S1 in the Supplemental Material available online). On average, there was a strong mismatch between high threat expectancy and low threat occurrence for each therapist- and self-guided exercise, indicating successful EV by exposure (Tukey-adjusted comparison of estimated marginal means: zs > 30.06, ps < .001, ds > 1.19). Moreover, average adjusted expectancies were consistently lower than original threat expectancies in each therapist- and self-guided exercise, indicating successful EC (zs > 19.30, ps < .001, ds > 0.66). Average

adjusted expectancies were higher than threat-occurrence ratings in each therapist- and self-guided exercise, indicating that EC did not occur at the same magnitude as EV (zs > 10.77, ps < .001, ds > 0.46).

Moreover, the three expectancy ratings distinctly differed between therapist-guided and self-guided exposure; Rating × Exercise interaction: $F(2, 21471.85) = 148.23$, $p < .001$. Threat expectancies were higher for therapist-guided compared with self-guided exercises (z ratio = 17.84, $p < .001$, $d = 0.33$), whereas threat occurrence was slightly lower for therapist-guided compared with self-guided exercises (z ratio = 5.11, $p < .001$, $d = 0.07$). Adjusted expectancies did not differ between therapist-guided and self-guided exposure (z ratio = 0.28, $p = .999$, $d = 0.03$; see Fig. S1 in the Supplemental Material).

In addition, change across sessions differed between threat expectancy, threat occurrence, and adjusted expectancy; Rating × Session interaction: $F(2, 21482.65) = 20.75$, $p < .001$. Threat expectancies slightly decreased across sessions (slope estimate = −1.17, 95% confidence interval [CI] = [−1.45, −0.90]), which was mainly linked to higher threat expectancies at the beginning (see Session 5 in Fig. 1). This significant but small decrease in threat expectancy was larger than the small decrease in threat occurrence (slope estimate = −0.32, 95% CI = [−0.59, −0.04]) and no change in adjusted expectancies (slope estimate = −0.26, 95% CI = [−0.53, 0.01]; threat expectancy vs. threat occurrence: z ratio = 5.39, $p < .001$; threat expectancy vs. adjusted expectancies: z ratio = 5.74, $p < .001$). Change across sessions did not differ between threat occurrence and adjusted expectancy (z ratio = 0.35, $p = .934$).

Summarized, averaged across patients, there was a consistent mismatch between high threat expectancy and low threat occurrence and low adjusted expectancy for each therapist- and self-guided exercise. These results indicate that exposure successfully induced threat EV and EC. However, EC did not occur at the same magnitude as EV. In addition, threat expectancy before exposure was higher during therapist-guided compared with self-guided exercises and for the initial exercises in Session 5.

Individual learning indicators during exposure-based CBT. The distribution of absolute and relative EV, EC, and the patients' exposure-related learning rate (α) are displayed in Figure 2. Average EV was 32.14 ($SD = 17.54$), and average EC was 21.11 ($SD = 12.25$). Both indicators substantially varied between patients. Regarding relative EV, 98.5% of the patients showed a violation of positive magnitude (i.e., relative EV > 0). The average proportion of EV was $M = 0.71$, indicating that exposure, on average, violated 71% of a patient's overall threat
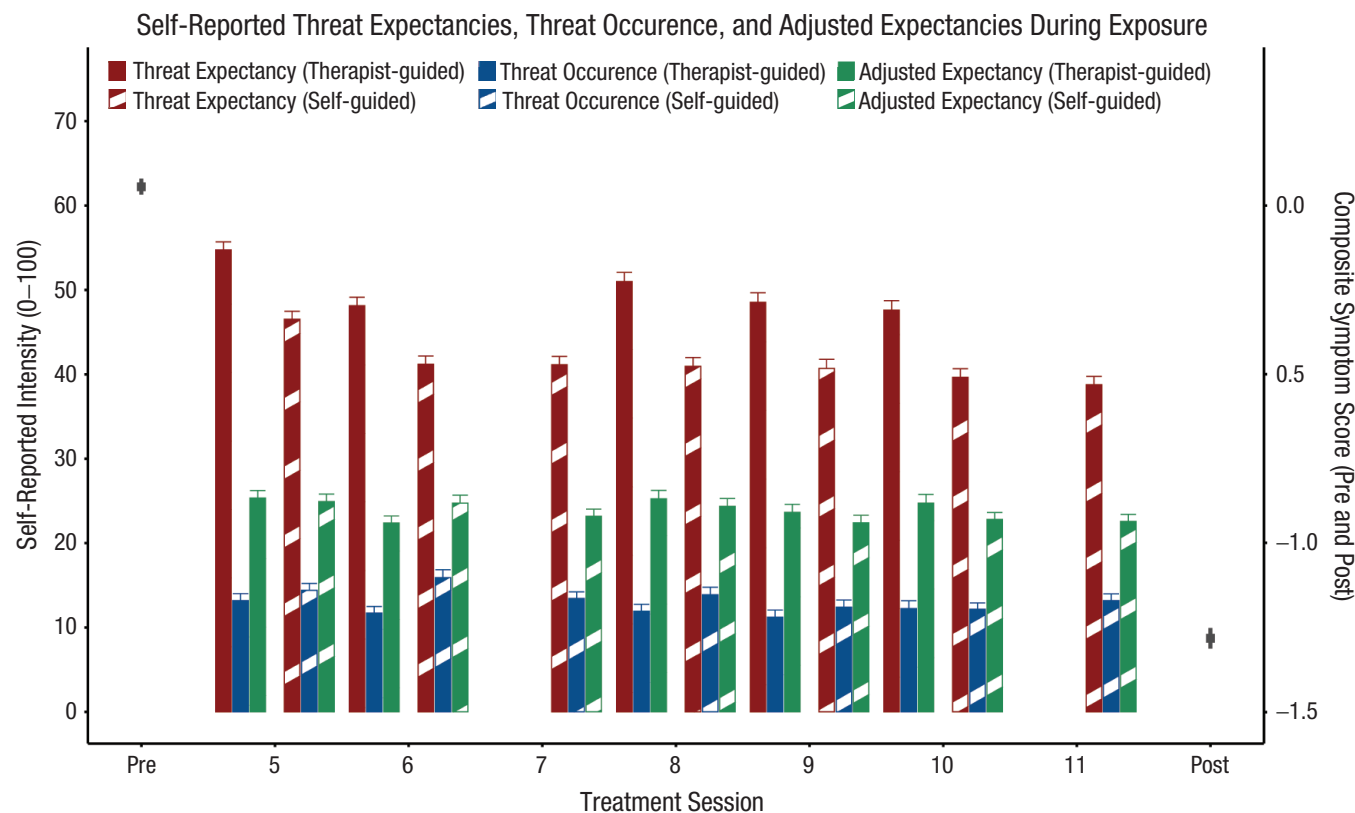
Self-Reported Threat Expectancies, Threat Occurence, and Adjusted Expectancies During Exposure

Fig. 1. Average ratings for threat expectancy before exposure and threat occurrence and adjusted expectancy after exposure for therapist-guided and self-guided exposure (i.e., as homework between sessions). In the treatment manual, no therapist-guided exposure was scheduled in Sessions 7 and 11, but patients completed self-guided exposure after both sessions. Gray squares show pretreatment and posttreatment composite symptom scores (see outcome measures) to indicate average symptom reduction.

expectancies. EV substantially differed between patients (see Fig. 2). Regarding relative EC, 97.4% of the patients showed a reduction of threat expectancy of positive magnitude (i.e., relative EC > 0). The average proportion of EC was $M = 0.48$, indicating that a patient's overall threat expectancies, on average, changed by 48%. EC also substantially differed between patients (see Fig. 2). Combined, these findings highlight that exposure, on average, induced successful EV and EC, which substantially differed between patients.

The average learning rate was $M = 0.65$. Thus, approximately two thirds of EV were, on average, transferred into EC. Again, learning rates substantially varied between patients ($SD = 0.23$, minimum = 0, maximum = 1; see Fig. 2, right), indicating that EC following EV was higher for some compared with other patients. For example, 25% of patients showed a learning rate below 0.5, whereas the highest 25% of patients showed a learning rate above 0.83.

Because threat expectancy and occurrence differed between therapist-guided and self-guided exposure, we compared EV and EC between therapist-guided and self-guided exposure (see Fig. 3). Absolute and relative EV

were higher for therapist-guided compared with self-guided exposure; absolute: Wilcoxon's $W = 164,124.50$, $p < .001$, $r = .69$; relative: $W = 105,890.00$, $p < .001$, $r = .45$. EC was also higher for therapist-guided compared with self-guided exposure; absolute: $W = 154,820.00$, $p < .001$, $r = .61$; relative: $W = 124,811.00$, $p < .001$, $r = .34$.

Associations between expectancy ratings, learning indicators, and symptom severity

Descriptive data of the average expectancy ratings and bivariate correlations between expectancy ratings, learning indicators, and symptoms are provided in the Supplemental Material (see Figs. S2 and S3 in the Supplemental Material). As expected, there were various significant correlations between the different indicators. Absolute and relative indicators were correlated only to a moderate degree (rs = .49 and .55). There were only few significant associations with the overall number of exposure exercises. Absolute EV ($r = .67$) and absolute EC ($r = .51$) were most strongly correlated with threat expectancy. These findings support the importance to

Threat Expectancy Violation (EV), Threat Expectancy Change (EC), and Learning Rate During Exposure Therapy
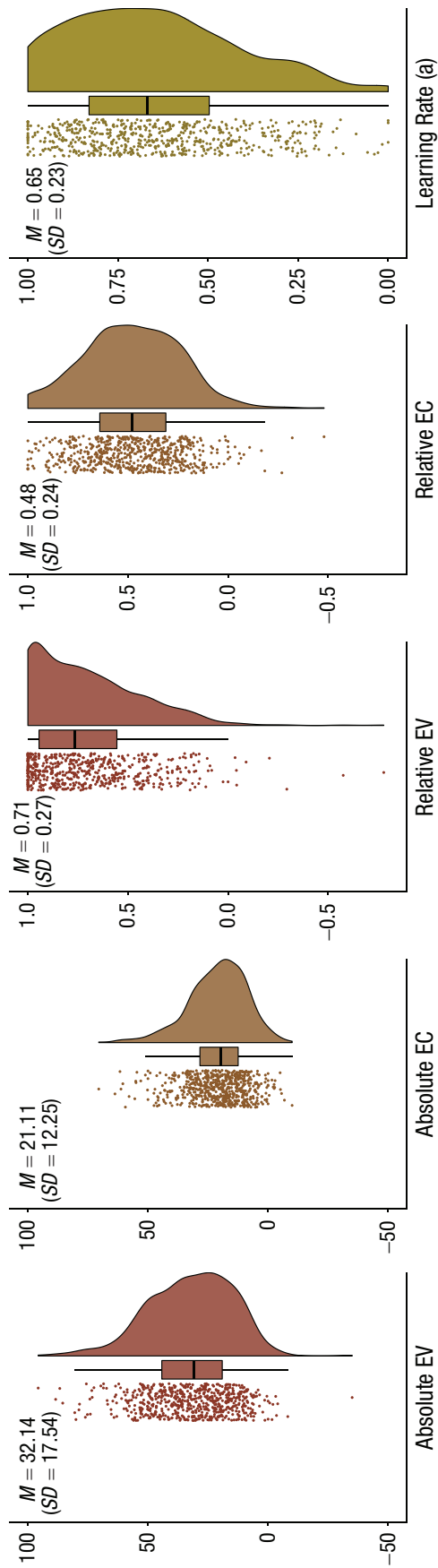

Fig. 2. Distribution of the learning indicators during exposure-based cognitive-behavioral therapy of 651 patients with anxiety disorders. For each patient, absolute and relative expectancy indicators were averaged across exposure exercises. Learning rates were estimated on the basis of all available exposure exercises of a patient. Dots represent individual patients' means. Their distribution is illustrated by density plots. Box plots illustrate median and interquartile range.

Expectancy Violation (EV) and Expectancy Change (EC) in Therapist-Guided (Blue) and Self-Guided Exposure (Red)

Absolute Expectancy Violation and Change

Relative Expectancy Violation and Change
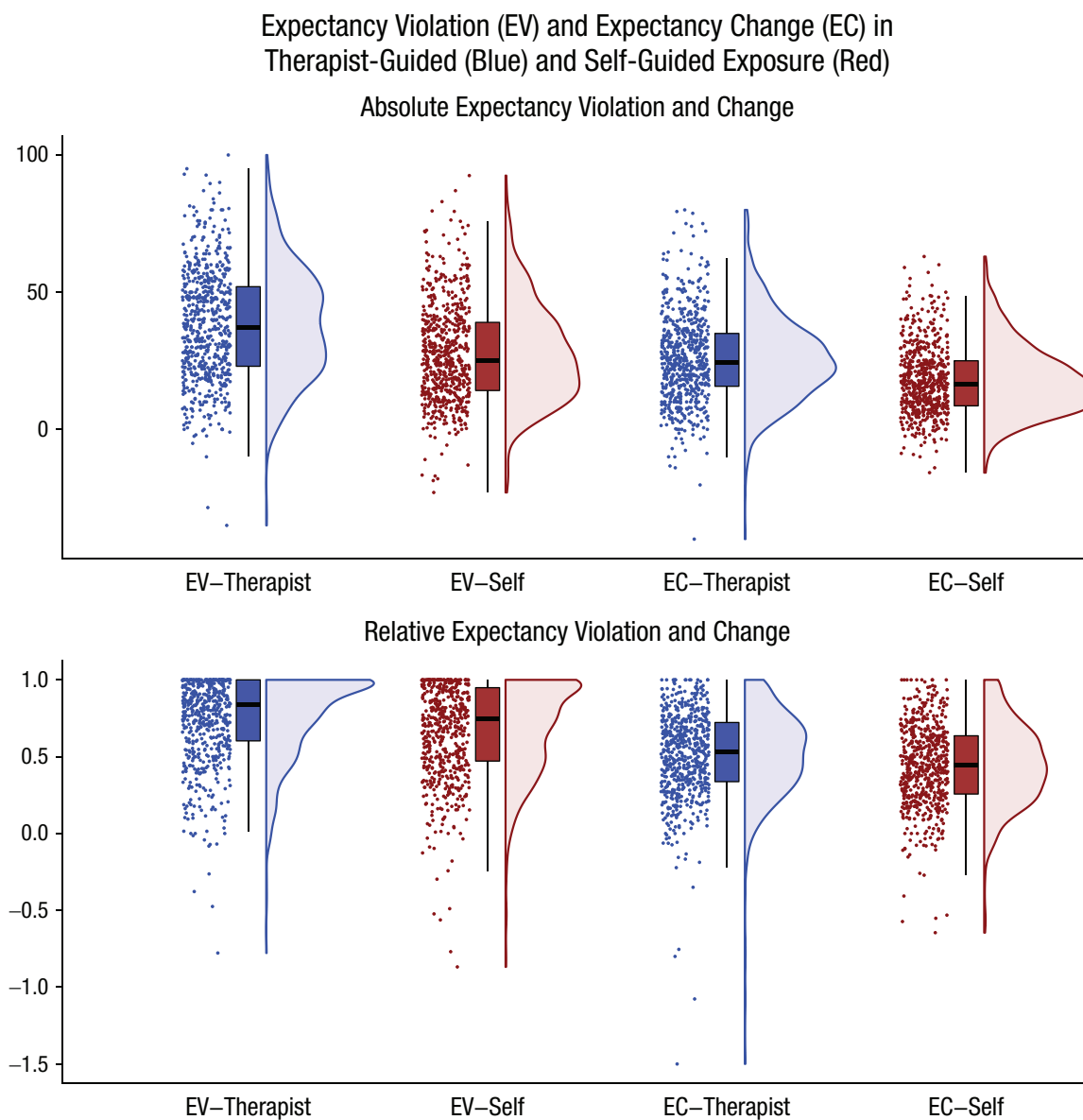
Fig. 3. Comparison of absolute (upper) and relative (lower) expectancy violation and change during therapist-guided (blue) and self-guided exposure (red). Dots represent individual patients' mean expectancy violation/change averaged across exposure exercises. Their distribution is illustrated by density plots. Box plots illustrate median and interquartile range.

control for the level of threat expectancy in prediction analyses. EV and change were positively correlated (*r*s = .62 and .45), indicating that higher EV was linked to higher EC. There were almost no significant associations with composite symptom severity at pretreatment.

Associations with treatment outcome

Composite symptom scores significantly improved from pretreatment to posttreatment with a very large effect size, $t(604) = 42.79$, $p < .001$, Cohen's $d = 1.75$ (see gray squares in Fig. 1). For the baseline model of the robust

linear regression, a higher number of comorbid disorders before treatment, stable intake of medication, and higher average threat expectancy were associated with less symptom reduction (see Table 2). Coefficients for the process-based learning indicators (see Table 2 and Fig. 4, top) showed no significant association for EV (βs = 0.05 and 0.04), but higher EC (βs = 0.24 and 0.19) and especially higher learning rate (βs = 0.36) predicted higher symptom reduction. Similar results were found in each primary disorder subgroup (see Fig. 4, bottom; for full statistics, see Table S2–S4 in the Supplemental Material): There were no significant associations between EV
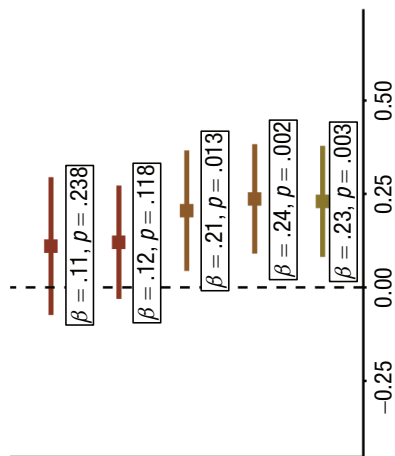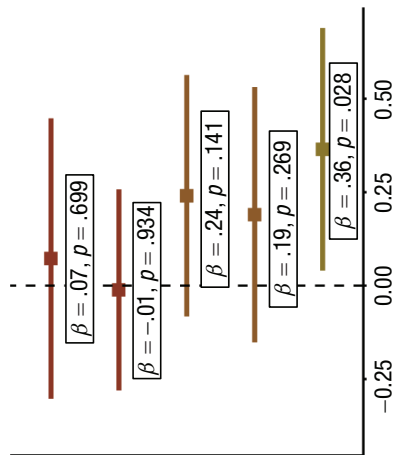
Fig. 4. Robust linear regression coefficients (and 95% confidence intervals) for the association between learning indicators (learning rate, expectancy change, expectancy violation) and symptom reduction (residual gain scores) for the whole sample (top) and subgroups with different anxiety disorders according to the principal diagnosis (bottom). Coefficients were calculated by adding the corresponding predictor to a predefined baseline model with number of diagnoses, number of exposure exercises, intake of psychotropic medication, and treatment group as predictors. PAS = Panic and Agoraphobia Scale; LSAS = Liebowitz Social Anxiety Scale; DSM-5 SP = DSM-5 Dimensional Specific Phobia Scale.

Table 2. Robust Linear Regression Results for Symptom Reduction

Predictor	Predictor statistics				Model statistics	
	β[a]	SE	t	p	R^2	χ^2
Baseline model						
Number of diagnoses	**−0.09**	0.04	2.08	.037		
Number of exposure exercises	0.01	0.04	0.10	.918		
Medication (no/yes)	**−0.10**	0.04	2.18	.030		
Treatment group	−0.04	0.04	0.99	.321		
Threat expectancy	**−0.22**	0.04	5.23	< .001	.07	
Baseline model + single predictor						
Expectancy violation (absolute)	0.05	0.04	0.81	.416	.07	0.66[b]
Expectancy violation (relative)	0.04	0.04	0.87	.387	.07	0.75[b]
Expectancy change (absolute)	**0.24**	0.05	5.00	< .001	.11	24.95[b]
Expectancy change (relative)	**0.19**	0.04	4.39	< .001	.10	19.27[b]
Learning rate (α)	**0.36**	0.04	8.27	< .001	.18	68.44[b]

Note: Dependent variable = residual gain score. Significant coefficients are highlighted in bold. For the "Baseline Model," the predictors were added to one multiple regression model. "Baseline Model + Single Predictor" refers to three separate models in which the corresponding predictor was added to the baseline model. Threat expectancy represents the average level of threat expectancy before exposure.
[a]Estimated from robust regression.
[b]Wald test for comparison to baseline model.

and disorder-specific symptom reduction. Higher learning rate was associated with larger disorder-specific symptom reduction in each subgroup (βs = 0.23–0.36). Higher EC was significantly associated with larger disorder-specific symptom reduction in PD/AG (βs = 0.16 and 0.15) and SAD (βs = 0.21 and 0.24) but missed significance in multiple SP (βs = 0.24 and 0.19), presumably because of the lower sample size.

In sum, higher EV was not associated with treatment outcome. Instead, a higher learning rate and higher EC were associated with better treatment outcome. This pattern of results was similar in each disorder-specific subgroup.

Discussion

In the present study, we examined the intensity, violation, and change of threat expectancies across exposure-based treatment for anxiety disorders and the association of these process variables with treatment outcome. Main findings showed that (a) high threat expectancies were, on average, successfully violated and changed throughout exposure exercises; (b) EV was not fully transferred into EC and the exposure-related learning rate substantially differed between patients; and (c) not EV itself but the exposure-related learning rate and EC were associated with better treatment outcome. The latter associations were comparable for all anxiety-disorder subgroups (i.e., PD/AG, SAD,

and multiple SP) and different indicators of EV and EC. Thus, successful exposure not only requires a mismatch between threat expectancy and actual occurrence, but most importantly, this mismatch needs to trigger an actual change of threat expectancies. These findings provide evidence from a large clinical sample that changing irrational threat expectancies seems to represent a transdiagnostic mechanism underlying successful exposure (Craske et al., 2014; Pittig et al., 2016).

Exposure exercises were successful to activate individual threat beliefs throughout treatment. This was indicated by high threat expectancies before exposure. On average, threat occurrence and adjusted expectancy after exposure were consistently lower compared with original threat expectancy before exposure. The mismatch between high threat expectancy and low threat occurrence highlights that exposure, on average, induced EV in therapist- and self-guided exercises. Likewise, the difference between high threat expectancy and lower adjusted threat expectancy highlights successful change of threat expectancy. These findings expand recent demonstrations of EV (e.g., Deacon et al., 2013; de Kleine et al., 2017) by also revealing EC in a large sample that included a wider spectrum of anxiety disorders and comorbid conditions undergoing comprehensive exposure-based CBT. Combined, these findings support theoretical assumptions that exposure induces violation and change of threat expectancies (Craske et al., 2014; Pittig et al., 2016). Moreover,

average EV and EC were higher for therapist-guided compared with self-guided exposure. These differences resulted from higher threat expectancy and lower threat occurrence for therapist-guided exposure. If EV and EC are mechanisms underlying successful exposure, this finding favors therapist-guided over self-guided exposure and does not support giving exposure as homework alone (see also Gloster et al., 2011; Hipol & Deacon, 2013; Pittig et al., 2019; Waller & Turner, 2016).

As expected, exposure was highly effective in reducing clinical anxiety symptoms. As previously reported, patients in the present trial showed substantial improvement in anxiety symptoms with very large effect sizes (Cohen's ds = 1.50–2.34; Pittig et al., 2021). For the present study, composite symptom score strongly decreased from pretreatment to posttreatment (d = 1.75). Most importantly, individual learning indicators were significantly associated with treatment outcome. Whereas EV itself was not associated with symptom reduction, a higher exposure-related learning rate (α) was consistently linked to better treatment outcome with moderate to strong effect sizes (rs = .23–.36). In the whole sample, learning rate accounted for approximately 15% of the variance in symptom reduction. The learning rate positively correlated with EC, which was also positively associated with better treatment outcome (rs = .15–.24). These associations were evident in each primary diagnosis subgroup (i.e., PD/AG, SAD, multiple SP). It is important to highlight that these findings do not negate the role of EV in exposure-based CBT as derived from the inhibitory-learning model. EV is likely the first important step to change threat beliefs. In support, higher EV was associated with higher EC. However, successful EV does not seem to suffice, because it does not automatically transfer into EC. Our findings thus suggest that not EV per se, but the rate of learning resulting in actual change of threat expectancy is a transdiagnostic mechanism of successful exposure for anxiety disorders. Clinical research should thus not only target interventions to maximize EV but also more strongly focus on interventions to optimize EC.

In this regard, the exposure-related learning rate may be a valuable indicator for quantifying learning during comprehensive exposure therapy. The average learning rate of 0.65, for example, indicated that approximately two thirds of EV resulted in EC. Given that most patients suffered from their anxiety disorder for several years (on average, 13–14 years) and thereby most likely held their individual threat beliefs for several years, exposure-based CBT can be seen as highly effective in changing individual threat expectancies in session. Importantly, individual learning rates substantially differed such that some patients showed strong learning, but others showed limited learning. Because the individual learning rate

was most strongly associated with symptom reduction, these findings can be seen as support for recent assumptions that the therapeutic effect of exposure seems to depend on the relationship between EV and factors influencing whether this experience actually changes threat expectancies (Kube et al., 2020; Rief et al., 2015).

In this regard, a variety of factors have been discussed. For example, successful change of threat expectancies may depend on low cognitive immunization (Kube et al., 2020; Rief et al., 2015). Cognitive immunization, such as "This was the exception of the rule" or "I was only successful because of the therapist's support," may prohibit change of threat expectancies despite their violation. Other factors involve the content and quality of exposure, individual difference factors, or external factors (Hermans et al., 2006; Lonsdorf & Merz, 2017; e.g., Pittig et al., 2018, 2020; Rief et al., 2015). For example, the relationship between learning from EV and treatment outcome may be moderated by the specific content of the patient's threat beliefs (e.g., "I will die" vs. "I will do something embarrassing") or the amount of safety behavior used during exposure (Pittig et al., 2020; Telch & Lancaster, 2012). Most importantly, an accurate definition of individual threat beliefs is assumed to be crucial for successful exposure (Craske et al., 2014). In this regard, therapists were trained to define testable beliefs about observable feared outcomes (e.g., "When I stutter during the presentation, people will notice and make depreciating comments") but not about emotional responses ("I will be very anxious") or untestable outcomes (e.g., "Something bad will happen"). Although therapist training specifically focused on accurate definitions of threat beliefs, future research may examine how adherence to this training may affect the association between EC and treatment success. In addition, individual-difference factors, such as concomitant medication or comorbid disorder, and external factors, such as external stressors, context changes, or distraction, may also influence learning during exposure (Hermans et al., 2006; Lonsdorf & Merz, 2017). In these examples, the exposure-related learning rate is conceptualized as a flexible parameter that can be modulated by the specific context or intervention. However, it could also be seen as a more stable, trait-like individual-difference factor itself, describing the individual capacity of learning from exposure, which, for example, could help to predict treatment outcome. In this regard, the association of the exposure-related learning rate with neurobiological learning circuits involved in the regulation of anxiety may be important. In sum, future research should examine the characteristics and moderators of the exposure-related learning rate, which may help to better optimize a crucial mechanism underlying successful exposure.

In the present study, we also found consistently high threat expectancies before exposure throughout treatment. Specifically, self-reported threat expectancy slightly decreased after the first session but remained stable for the remaining sessions. These findings may be seen as somewhat surprising given that a reduction of threat expectancy is typically reported in experimental and clinical-analogue studies (Deacon et al., 2013; Hermans et al., 2006). However, in experimental and clinical-analogue studies, exposure to the same stimulus or context is typically repeated multiple times. A change of context or stimulus typically results in little reduction or a return of threat expectancy (Bouton, 2002; Maren et al., 2013; Vervliet et al., 2005, 2006). In the present trial, therapists were trained to conduct exposure for the most prominent threat belief and encouraged to switch to those stimuli or contexts for which high threat expectancies were (still) present. High threat expectancies throughout treatment were thus in line with the individual focus of the treatment and the inhibitory-learning-model approach to exposure that aims to repeatedly increase threat expectancy (Craske et al., 2014). Stable threat expectancies thus do not indicate failure to reduce threat expectancy for a specific context or stimulus. They may better be interpreted as a lack of generalizing the experience of non-occurring threat from one to another exposure exercise. These clinical findings relate to experimental research on deficits in generalization of fear extinction: Learning that a feared stimulus is no longer followed by an aversive outcome (i.e., fear extinction) does not easily generalize to stimuli or contexts resembling the original fear stimulus or context (Vervliet et al., 2005, 2006; Wong & Lovibond, 2020). Although anxiety symptoms markedly decreased as high threat expectancies were repeatedly changed, higher average threat expectancy across exposure exercises was associated with lower symptom reduction. These findings may tentatively support the notion that a failure to generalize the experience of EV across exposures is associated with worse treatment outcome. Future research may thus examine the association between symptom reduction and individual generalization gradients in more detail.

Some of the present findings and implications require further investigation and additional limitations need to be considered. As mentioned above, we did not account for the content, quality, or individual-difference factors because the main aims focused on overall violation and change of threat beliefs and their association with treatment outcome. In addition, we used learning indicators averaged across exposure sessions, which may prohibit insights into their dynamics across sessions. The present study found a strong link between exposure, its proposed mechanism (change of threat expectancy), and symptom reduction. This link is an important criterion for a mechanism of action in psychotherapy, but additional criteria exist (see Kazdin, 2007). Whereas some criteria can be accepted as confirmed (e.g., plausibility of the mechanism; Richter et al., 2017), others require converging lines of research (e.g., replication). Most importantly, future studies can expand the present findings by adding repeated symptom measures across sessions to establish a closer timeline between EC and symptom reduction. Finally, in the present study, we measured only the change of threat expectancy directly after exposure, which represents a short-term change in threat beliefs. It may thus be interesting to expand results to the long-term course of threat expectancies after exposure. Some patients may show persistent EC, whereas others may not. Recurrence of original threat expectancies may, for example, be linked to biased processing of exposure, such as focusing on negative aspects of exposure, postevent processing, or cognitive immunization (Rachman et al., 2000; Rief et al., 2015). Long-term EC may thus be even more relevant for long-term treatment success. Finally, we did not investigate the role of ethnic or cultural background in the present study.

Conclusion

Exposure-based CBT is successful to activate, violate, and change individual threat beliefs throughout exposure sessions. Average learning from EV and EC was high but substantially differed between individual patients. Not EV itself, but a higher exposure-related learning rate and higher EC following EV was associated with better treatment outcome. Comparable results were found in all primary anxiety-disorder subgroups (PD/AG, SAD, and multiple SP). Change of threat expectancies thus seems to represent a transdiagnostic mechanism underlying successful behavioral treatment for anxiety disorders. Future research should examine which factors influence exposure-related learning rate and expand findings by establishing a closer timeline between learning and symptom reduction and examining the long-term course of expectancy after exposure.

Transparency

of the manuscript and/or added important intellectual content during the revision process. All of the authors approved the final manuscript for submission.

ORCID iDs

Andre Pittig https://orcid.org/0000-0003-3787-9576
Jan Richter https://orcid.org/0000-0002-7127-6990

Acknowledgments

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/21677026221101379

Notes

1. No bipolar disorder occurred in the present sample (assessed via the CIDI).
2. Patients were allowed to indicate two threat beliefs for each exercise, which occurred in 14.3% of all exercises. Analyses were conducted for the first threat belief. However, results were the same for analyses using the highest rated belief and using an average in case of two beliefs.
3. The formula is based on prominent learning models such as reinforcement learning or the Rescorla-Wagner model (Rescorla & Wagner, 1972; Sutton & Barto, 2018), which are often used in experimental research. In this research, participants are, for example, exposed to the same stimulus in successive trials and

expectancies probed for each trial. Because the same stimulus is encountered, EC is inferred by comparing expectancy from trial i with expectancy from trial $i + 1$. Because successive exposure exercises in naturalistic exposure treatment, however, differ considerably, it cannot be assumed that the same stimulus or context is encountered. Thus, we used adjusted expectancy rated directly after exposure as alternative indicator for EC.

References

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

Bandelow, B. (1995). Assessing the efficacy of treatments for panic disorder and agoraphobia: II. The Panic and Agoraphobia Scale. *International Clinical Psychopharmacology*, *10*(2), 73–81.

Beck, A. T., & Haigh, E. A. P. (2014). Advances in cognitive theory and therapy: The generic cognitive model. *Annual Review of Clinical Psychology*, *10*(1), 1–24. https://doi.org/10.1146/annurev-clinpsy-032813-153734

Beesdo, K., Bittner, A., Pine, D. S., Stein, M. B., Höfler, M., Lieb, R., & Wittchen, H.-U. (2007). Incidence of social anxiety disorder and the consistent risk for secondary depression in the first three decades of life. *Archives of General Psychiatry*, *64*(8), 903–912. https://doi.org/10.1001/archpsyc.64.8.903

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, *52*(10), 976–986.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, *11*(5), 485–494. https://doi.org/10.1101/lm.78804

Carpenter, J. K., Andrews, L. A., Witcraft, S. M., Powers, M. B., Smits, J. A. J., & Hofmann, S. G. (2018). Cognitive behavioral therapy for anxiety and related disorders: A meta-analysis of randomized placebo-controlled trials. *Depression and Anxiety*, 35, 502–514. https://doi.org/10.1002/da.22728

Craske, M. G., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1742), Article 20170025. https://doi.org/10.1098/rstb.2017.0025

Craske, M. G., & Mystkowski, J. L. (2006). Exposure therapy and extinction: Clinical studies. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), Fear *and learning: From basic processes to clinical implications* (pp. 217–233). American Psychological Association. https://doi.org/10.1037/11474-011

Craske, M. G., Stein, M. B., Eley, T. C., Milad, M. R., Holmes, A., Rapee, R. M., & Wittchen, H.-U. (2017). Anxiety disorders. *Nature Reviews Disease Primers*, *3*, Article 17024. https://doi.org/10.1038/nrdp.2017.24

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T. D., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Craske, M. G., Wittchen, H.-U., Bögels, S. M., Stein, M. B., Andrews, G., & LeBeau, R. T. (2013). *Severity measure*

for specific phobia—Adult version. https://www.psychia try.org/File%20Library/Psychiatrists/Practice/DSM/APA_ DSM5_Severity-Measure-For-Specific-Phobia-Adult.pdf

Deacon, B. J., Kemp, J. J., Dixon, L. J., Sy, J. T., Farrell, N. R., & Zhang, A. R. (2013). Maximizing the efficacy of inter-oceptive exposure by optimizing inhibitory learning: A randomized controlled trial. *Behaviour Research and Therapy*, *51*(9), 588–596. https://doi.org/10.1016/j.brat.2013.06.006

de Kleine, R. A., Hendriks, L., Becker, E. S., Broekman, T. G., & van Minnen, A. (2017). Harm expectancy violation during exposure therapy for posttraumatic stress disor-der. *Journal of Anxiety Disorders*, *49*, 48–52. https://doi .org/10.1016/j.janxdis.2017.03.008

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., van den Hout, M. A., & Baas, J. M. P. (2015). Updated meta-analysis of classical fear condi-tioning in the anxiety disorders. *Depression and Anxiety*, *32*(4), 239–253. https://doi.org/10.1002/da.22353

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Fresco, D. M., Coles, M. E., Heimberg, R. G., Liebowitz, M. R., Hami, S., Stein, M. B., & Goetz, D. (2001). The Liebowitz Social Anxiety Scale: A comparison of the psychometric properties of self-report and clinician-administered for-mats. *Psychological Medicine*, *31*(6), 1025–1035.

Gloster, A. T., Hauke, C., Höfler, M., Einsle, F., Fydrich, T., Hamm, A., Sthröhle, A., & Wittchen, H.-U. (2013). Long-term stability of cognitive behavioral therapy effects for panic disorder with agoraphobia: A two-year follow-up study. *Behaviour Research and Therapy*, *51*(12), 830–839. https://doi.org/10.1016/j.brat.2013.09.009

Gloster, A. T., Wittchen, H.-U., Einsle, F., Lang, T., Helbig-Lang, S., Fydrich, T., Fehm, L., Hamm, A. O., Richter, J., Alpers, G. W., Gerlach, A. L., Ströhle, A., Kircher, T., Deckert, J., Zwanzger, P., Höfler, M., & Arolt, V. (2011). Psychological treatment for panic disorder with agora-phobia: A randomized controlled trial to examine the role of therapist-guided exposure in situ in CBT. *Journal of Consulting and Clinical Psychology*, *79*(3), 406–420. https://doi.org/10.1037/a0023584

Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., Dodel, R., Ekman, M., Faravelli, C., Fratiglioni, L., Gannon, B., Jones, D. H., Jennum, P., Jordanova, A., Jönsson, L., Karampampa, K., Knapp, M., Kobelt, G., Kurth, T., . . . Olesen, J. (2011). Cost of dis-orders of the brain in Europe 2010. *European Neuro-psychopharmacology*, *21*(10), 718–779. https://doi.org/ 10.1016/j.euroneuro.2011.08.008

Guy, W. (1976). *ECDEU assessment manual for psychophar-macology*. U.S. Department of Health, Education, and Welfare, Public Health Service.

Heinig, I., Pittig, A., Richter, J., Hummel, K., Alt, I., Dickhöver, K., Gamer, J., Hollandt, M., Koelkebeck, K., Maenz, A., Tennie, S., Totzeck, C., Yang, Y., Arolt, V., Deckert, J., Domschke, K., Fydrich, T., Hamm, A. O., Hoyer, J., . . . Wittchen, H. (2017). Optimizing exposure-based CBT

for anxiety disorders via enhanced extinction: Design and methods of a multicentre randomized clinical trial. *International Journal of Methods in Psychiatric Research*, *26*(2), Article 1560. https://doi.org/10.1002/mpr.1560

Hermans, D., Craske, M. G., Mineka, S., & Lovibond, P. F. (2006). Extinction in human fear conditioning. *Biological Psychiatry*, *60*(4), 361–368. https://doi.org/10.1016/j.bio psych.2005.10.006

Hipol, L. J., & Deacon, B. J. (2013). Dissemination of evidence-based practices for anxiety disorders in Wyoming: A sur-vey of practicing psychotherapists. *Behavior Modification*, *37*(2), 170–188. https://doi.org/10.1177/0145445512458794

Hofmann, S. G. (2008). Cognitive processes during fear acquisition and extinction in animals and humans: Implications for exposure therapy of anxiety disorders. *Clinical Psychology Review*, *28*(2), 200–211. https://doi .org/10.1016/j.cpr.2007.04.009

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

Hoyer, J., Čolić, J., Pittig, A., Crawcour, S., Moeser, M., Ginzburg, D., Lin, J., Wiltink, J., Leibing, E., & Stangier, U. (2017). Manualized cognitive therapy versus cognitive-behavioral treatment-as-usual for social anxiety disorder in routine practice: A cluster-randomized controlled trial. *Behaviour Research and Therapy*, *95*, 87–98. https://doi .org/10.1016/j.brat.2017.05.012

Jacobi, F., Höfler, M., Siegert, J., Mack, S., Gerschler, A., Scholl, L., Busch, M. A., Hapke, U., Maske, U., Seiffert, I., Gaebel, W., Maier, W., Wagner, M., Zielasek, J., & Wittchen, H.-U. (2014). Twelve-month prevalence, comorbidity and cor-relates of mental disorders in Germany: The mental health module of the German Health Interview and Examination Survey for Adults (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, *23*(3), 304–319. https:// doi.org/10.1002/mpr.1439

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, *3*(1), 1–27. https://doi.org/10.1146/annurev .clinpsy.3.022806.091432

Kessler, R. C., Ruscio, A. M., Shear, K., & Wittchen, H.-U. (2010). Epidemiology of anxiety disorders. *Current Topics in Behavioral Neurosciences*, *2*, 21–35. https://doi.org/10 .1007/7854_2009_9

Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, *44*(3), 579–592. https:// doi.org/10.1017/S0033291713001074

Kube, T., Schwarting, R., Rozenkrantz, L., Glombiewski, J. A., & Rief, W. (2020). Distorted cognitive processes in major depression: A predictive processing perspective. *Biological Psychiatry*, *87*(5), 388–398.

Lebeau, R. T., Glenn, D. E., Hanover, L. N., Beesdo-Baum, K., Wittchen, H.-U., & Craske, M. G. (2012). A dimensional approach to measuring anxiety for DSM-5. *International Journal of Methods in Psychiatric Research*, *21*(4), 258–272. https://doi.org/10.1002/mpr.1369

Liebowitz, M. R. (1987). Social phobia. *Modern Trends in Pharmacopsychiatry*, *22*, 141–173.

Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., Craske, M. G., Rosen, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: Need for standardized criteria. *Clinical Psychology Review*, *42*, 72–82. https://doi.org/10.1016/j.cpr.2015.08.004

Lonsdorf, T. B., & Merz, C. J. (2017). More than just noise: Inter-individual differences in fear acquisition, extinction and return of fear in humans—Biological, experiential, temperamental factors, and methodological pitfalls. *Neuroscience and Biobehavioral Reviews*, *80*, 703–728. https://doi.org/10.1016/j.neubiorev.2017.07.007

Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M., Conceicao, E. L., & di Palma, M. A. (2020). *robustbase: Basic Robust Statistics R package* (0.93-6) [Computer software]. http://CRAN.R-project.org/package=robustbase

Marcus, D. K., O'Connell, D., Norris, A. L., & Sawaqdeh, A. (2014). Is the Dodo bird endangered in the 21st century? A meta-analysis of treatment comparison studies. *Clinical Psychology Review*, *34*(7), 519–530. https://doi.org/10.1016/j.cpr.2014.08.001

Maren, S., Phan, K. L., & Liberzon, I. (2013). The contextual brain: Implications for fear conditioning, extinction and psychopathology. *Nature Reviews Neuroscience*, *14*, 417–428. https://doi.org/10.1038/nrn3492

Pittig, A., Heinig, I., Goerigk, S., Thiel, F., Hummel, K., Scholl, L., Deckert, J., Pauli, P., Domschke, K., Lueken, U., Fydrich, T., Fehm, L., Ströhle, A., Kircher, T., Straube, B., Rief, W., Koelkebeck, K., Arolt, V., Dannlowski, U., . . . Wittchen, H.-U. (2021). Efficacy of temporally intensified exposure for anxiety disorders: A multi-center randomized clinical trial. *Depression and Anxiety*, *38*(11), 1169–1181. https://doi.org/10.1002/da.23204

Pittig, A., Kotter, R., & Hoyer, J. (2019). The struggle of behavioral therapists with exposure: Self-reported practicability, negative beliefs, and therapist distress about exposure-based interventions. *Behavior Therapy*, *50*(2), 353–366. https://doi.org/10.1016/j.beth.2018.07.003

Pittig, A., Treanor, M., LeBeau, R. T., & Craske, M. G. (2018). The role of associative fear and avoidance learning in anxiety disorders: Gaps and directions for future research. *Neuroscience & Biobehavioral Reviews*, *88*, 117–140. https://doi.org/10.1016/j.neubiorev.2018.03.015

Pittig, A., van den Berg, L., & Vervliet, B. (2016). The key role of extinction learning in anxiety disorders: Behavioral strategies to enhance exposure-based treatments. *Current Opinion in Psychiatry*, *29*(1), 39–47. https://doi.org/10.1097/YCO.0000000000000220

Pittig, A., Wong, A. H. K., Glück, V. M., & Boschet, J. M. (2020). Avoidance and its bi-directional relationship with conditioned fear: Mechanisms, moderators, and clinical implications. *Behaviour Research and Therapy*, *126*, Article 103550. https://doi.org/10.1016/j.brat.2020.103550

Rachman, S. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour Research and Therapy*, *15*(5), 375–387. https://doi.org/10.1016/0005-7967(77)90041-9

Rachman, S., Grüter-Andrew, J., & Shafran, R. (2000). Post-event processing in social anxiety. *Behaviour Research and Therapy*, *38*(6), 611–617. https://doi.org/10.1016/S0005-7967(99)00089-3

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Prokasy (Ed.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Richter, J., Pittig, A., Hollandt, M., & Lueken, U. (2017). Bridging the gaps between basic science and cognitive-behavioral treatments for anxiety disorders in routine care: Current status and future demands. *Zeitschrift Fur Psychologie / Journal of Psychology*, *225*(3), 252–267. https://doi.org/10.1027/2151-2604/a000309

Rief, W., Glombiewski, J. A., Gollwitzer, M., Schubö, A., Schwarting, R., & Thorwart, A. (2015). Expectancies as core features of mental disorders. *Current Opinion in Psychiatry*, *28*(5), 378–385. https://doi.org/10.1097/YCO.0000000000000184

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

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *afex: Analysis of Factorial Experiments* (R Package Version 0.27-2) [Computer software]. https://CRAN.R-project.org/package=afex

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. Spieler & E. Schumacher (Eds.), *New methods in cognitive psychology* (pp. 4–31). Routledge. https://doi.org/10.4324/9780429318405-2

Steketee, G., & Chambless, D. L. (1992). Methodological issues in prediction of treatment outcome. *Clinical Psychology Review*, *12*(4), 387–400. https://doi.org/10.1016/0272-7358(92)90123-P

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. The MIT Press.

Telch, M. J., & Lancaster, C. L. (2012). Is there room for safety behaviors in exposure therapy for anxiety disorders? In P. Neudeck & H.-U. Wittchen (Eds.), *Exposure therapy: Rethinking the model—Refining the method* (pp. 313–334). Springer.

Tolin, D. F. (2010). Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review*, *30*(6), 710–720. https://doi.org/10.1016/j.cpr.2010.05.003

Vervliet, B., Vansteenwegen, D., Baeyens, F., Hermans, D., & Eelen, P. (2005). Return of fear in a human differential conditioning paradigm caused by a stimulus change

after extinction. *Behaviour Research and Therapy*, *43*(3), 357–371. https://doi.org/10.1016/j.brat.2004.02.005

Vervliet, B., Vansteenwegen, D., & Eelen, P. (2006). Generalization gradients for acquisition and extinction in human contingency learning. *Experimental Psychology*, *53*(2), 132–142. https://doi.org/10.1027/1618-3169.53.2.132

Waller, G., & Turner, H. (2016). Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour Research and Therapy*, 77, 129–137. https://doi.org/10.1016/j.brat.2015.12.005

Wannemueller, A., Gruszka, P., Chwalek, S., Fröhlich, S., Mulders, M., Schaumburg, S., Schöttes, J., Wiederhold, S., & Margraf, J. (2019). Large-group one-session treatment: Feasibility in highly height fearful individuals and predictors of outcome. *Frontiers in Psychology*, *10*, Article 2411. https://doi.org/10.3389/fpsyg.2019.02411

Watts, S. E., Turnell, A., Kladnitski, N., Newby, J. M., & Andrews, G. (2015). Treatment-as-usual (TAU) is anything but usual: A meta-analysis of CBT versus TAU for anxiety and depression. *Journal of Affective Disorders*, *175*, 152–167. https://doi.org/10.1016/j.jad.2014.12.025

Wittchen, H.-U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research*, *28*(1), 57–84. https://doi.org/10.1016/0022-3956(94)90036-1

Wittchen, H.-U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., Olesen, J., Allgulander, C., Alonso, J., Faravelli, C., Fratiglioni, L., Jennum, P., Lieb, R., Maercker, A., van Os, J., Preisig, M., Salvador-Carulla, L., Simon, R., & Steinhausen, H. C. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, *21*(9), 655–679. https://doi.org/10.1016/j.euroneuro.2011.07.018

Wong, A. H. K., & Lovibond, P. F. (2020). Generalization of extinction of a generalization stimulus in fear learning. *Behaviour Research and Therapy*, *125*, Article 103535. https://doi.org/10.1016/j.brat.2019.103535