Study Title:

Optimizing Extinction Using Intensified Psychological Interventions (IPI) for Adult Anxiety Disorders (AD)

- Study Protocol (Version 3.1) -

Research Network on Mental Disorders (2014)

granted by the German Federal Ministry of Education and Research

Principal Investigator:

Prof. Dr. phil. habil. Hans-Ulrich Wittchen

Institut für Klinische Psychologie und Psychotherapie
Technische Universität Dresden
Chemnitzer Str. 46
01187 Dresden
Tel.: 49 351 463 36985
E-Mail: wittchen@psychologie.tu-dresden.de

Study Centers:

1. Wittchen, H.-U: Institut für Klinische Psychologie und Psychotherapie, Technische Universität Dresden
2. Arolt, V.: Klinik und Poliklinik für Psychiatrie und Psychotherapie, Westfälische Wilhelms Universität Münster
3. Schneider, S.: Forschungs- und Behandlungszentrum für Psychische Gesundheit, Ruhr Universität Bochum
4. Hamm, A.: Physiologische und Klinische Psychologie / Psychotherapie, Ernst-Moritz-Arndt Universität Greifswald
5. Kircher, T.: Klinik und Poliklinik für Psychiatrie und Psychotherapie, Philipps-Universität Marburg
6. Deckert, J.: Klinik und Poliklinik für Psychiatrie, Psychosomatik und Psychotherapie, Julius-Maximilians-Universität Würzburg
7. Ströhle, A.: Klinik und Poliklinik für Psychiatrie und Psychotherapie, Charité-Universitätsmedizin Berlin

0. Summary

APPLICANT/ COORDINA-TING INVES-TIGATORS	Professor Dr. Hans-Ulrich Wittchen & Prof. Dr. Volker Arolt (Co-PI) Institute of Clinical Psychology and Psychotherapy & CELOS Technische Universität Dresden Chemnitzer Strasse 46, 01187 Dresden Tel.: 0351-463-36985; Fax: 0351-463-36984 wittchen@psychologie.tu-dresden.de
TITLE	**Optimizing Extinction Using Intensified Psychological Interventions (IPI) for Adult Anxiety Disorders (AD)**
CONDITION	DSM-5: Panic Dis. (PD), Generalized Anxiety Dis. (GAD), Agoraphobia (AG), Social Anxiety Disorder (SAD), Specific Phobias (SP), Separation Anxiety Dis.(SepAD), and associated comorbid conditions
OBJECTIVE	To examine the efficacy of intensified psychological interventions (IPI) aiming to augment extinction learning during exposure. In IPI we increase the number of exposure trials and decrease the spacing between sessions providing more "extinction trials" over a shorter duration. To stabilize treatment effects we provide spaced exposure training during follow-up (booster) in different contexts. We also based exposure on a positive prediction error model. We hypothesize that IPI results in stronger, faster and more pervasive effects in subjective, behavioural, physiological, neural and epigenetic indices at post and follow-up via optimizing extinction learning. To test the hypothesis, we propose a multicentre randomized clinical trial (RCT). Patients will be randomly allocated to either IPI or *standard research exposure treatment as usual* (TAU) as the control condition. Primary hypotheses: (1) IPI will be superior to TAU at post and follow-up on primary and secondary outcome measures and will recover faster. Secondary hypotheses are: (2) IPI is associated with enhanced positive prediction error to increase extinction learning during and after exposure and (3) will result in more pronounced changes in behavioural proxy measures (between session reduction of fear, anxiety, avoidance) of extinction learning. We also explore the effect of type of diagnosis and comorbidity as potential moderators of outcome and extinction. We further expect that IPI will be associated with lower direct and indirect health care costs.
INTERVENTION	Eligible patients are randomized to either IPI or TAU. Using a well established modular 12-sessions exposure-based treatment manual template [1,2], we compare two content-wise almost identical versions. However, IPI provides the manual`s behavioural exposure module (BE, sessions 6-10) in one week allowing for a higher number of exposure trials over a short time period, and provides "spaced trials" during the booster and follow-up period. In the control condition (TAU), the BE module will be provided over 4-5 weeks as is usual in research settings with one to two sessions/week and no spaced exposure trials during the follow-up phase. Follow-up period per patient: 6 month Duration of intervention per patient, excluding 6 months follow-up: 5-6 weeks (IPI), and 7-9 weeks (TAU)
INCLUSION/ EXCLUSION CRITERIA	Stratified sampling of patients in anxiety clinics. Key inclusion criteria: (1) Outpatients, (2) age: 15-70 years, (3) current primary DSM-5 anxiety

	disorder, (4) severity at baseline (HAMA>18) and CGI>3, (5) written informed consent, (6) able to attend on his/her own or accompanied by significant other, (7) language competence. Key exclusion criteria: (1) any DSM-IV/5 psychotic, primary mood disorders (bipolar I, recurrent or chronic major depression), current substance use dependence, (2) concomitant psychological/psychiatric treatment, (4) acute suicidality, (5) general medical contraindications, (5) mono-symptomatic specific phobia.
OUTCOME(S)	Primary efficacy endpoint is the clinician-rated Hamilton anxiety score (SIGH-A). Key secondary endpoint(s): CGI total score, patient rated generic (BSI) and diagnosis-specific symptom scores (i.e. FFS, PAS, MI), impairment/ disability days (CIDI Harvard Index, SDS), depression (dimensional: BDI, categorical: CIDI) quality of life (EQ-5D) and social functioning (SAS). As an objective and ecological valid measure for exposure trials we use a novel combined mobile mobility-EMS tool (see proposal ESPRIT). The range of measures is meant to demonstrate that IPI is associated with more pervasive changes in various domains. As proxy measure for extinction learning we use subjective measures of within-session and between-session exposure effects (anxiety ratings, expectancy ratings of central concerns). (Note: see further assessments in P3-P6.) Assessment of safety: suicidality assessment and adverse behavioural effects.
STUDY TYPE	Randomised controlled clinical trial with two active conditions associated with mediator and moderator analyses (see projects P3-6)
STATISTICAL ANALYSES	Efficacy: HAMA total score (standardized interview, SIGH-A) Description of primary efficacy: Superiority of IPI against TAU will be tested with t-tests for independent samples using an alpha level of 0.5. Intent-to-treat (ITT), completer analyses will be conducted and mixed models (saturated for the joint effects of treatment group and time) fit to address dropout. Safety: N/A Secondary endpoints: ITT, completer analysis and mixed models as above but with two-sided tests in case of explorative analyses. Pearson correlations between dimensional comorbidity measures and outcomes will be calculated, Associations with categorical predictors will be analyzed with linear regressions.
SAMPLE SIZE	To be assessed for eligibility n=1.400 To be allocated in a stratified way (by diagnosis) to trial (n = 720) To be analysed (n = 620)
TRIAL DURATION	First patient in: 6/1st year – last patient out (fup): 1/4th year Recruitment and enrolment: start: 6/1st year Duration of the entire trial: 36 months Recruitment period: 6/1st year – 5/3rd year
CENTERS	7: Greifswald, Berlin, Münster, Marburg, Dresden, Bochum, Würzburg

1. Evidence and study rationale

1.1 Extinction learning

A widely accepted view on the nature of extinction holds that extinction is conveyed by forming a second, inhibitory CS-US association in addition to the formerly acquired excitatory CS-US association (fear memory). Fear memories are thus not erased, but can be retrieved under certain circumstances. This phenomenon, also known as return of fear, clinically presents as a reoccurrence of anxiety symptoms or relapse. A great body of preclinical evidence has investigated the conditions under which return of fear can be observed and emphasises the key role of context variables during fear extinction [1,2]. Context information is critically linked to the extinction memory and acts as a gating mechanism as to whether the fear memory is inhibited or not. From a neurobiological viewpoint, context information during extinction is stored in the hippocampus and fear inhibition is signalled by concerted action of the ventromedial prefrontal cortex, the amygdala and the hippocampus. Potential context variables do not only encompass external or internal states, but also temporal characteristics of the treatment. One prominent example is the spontaneous recovery of fear where the time between treatment and relapse can be interpreted as a gradually changing context, making it difficult to recall the original inhibitory memory trace in this seemingly new context. *In this view, the timing and spacing of exposure sessions (inter-trial intervals (ITI)) seems to be particular relevant and thus we focus this chapter on this topic.*

Extinction learning comprises several distinct phases: 1. The individual learns that a fear cue is no longer followed by the expected aversive consequences under the current circumstances or in the current context. 2. This learning experience is consolidated, and inhibitory associations are reinforced by repetitive exposure sessions [3]. 3. In case of successful long-term extinction learning, extinction memory is recalled even if the previously feared cues occur in different contexts. Under critical circumstances recall of the extinction memory can be blocked and return of fear and even relapse can occur. Comparable learning and extinction processes are also important for contextual fear conditioning which occurs when aversive events appear unpredictable so that the context becomes the best, albeit imprecise, predictor of aversive events. Such contextual fear conditioning is an experimental model of sustained anxiety and explains avoidance and safety behavior at least in panic disorder patients [4].

We acknowledge that there are various constructs and paradigms to assess extinction-related mechanisms of change (e.g. enhanced discrimination learning, enhanced safety learning, reduced vulnerability to reinstatement, or context renewal). However, we refrain from providing here a complete critical review of evidence. Instead we present a review of evidence and highlight the following critical state of the art reviews [1,2,5,6] (see also P3-P5).

Preclinical evidence: Although being far from consistent, animal and human studies on the timing of extinction trials suggest that massed extinction can result in enhanced fear inhibition, albeit possibly more susceptible to return of fear in the long-term. Early *animal research* suggested that massed extinction resulted in more rapid extinction learning. However, some studies found no effects, opposite effects, or have found that massed extinction even results in greater return of fear [7]. In mice, the most effective method of enhancing extinction is found with initial massed trials followed by subsequent spaced trials. In *human subjects*, evidence on the inter-trial interval in extinction is scant, and many of the

studies provide contradictory results. One study found that massed extinction trials result in higher conditioned responding compared to spaced extinction trials but have similar rates of spontaneous recovery, and another study found no difference in extinction based on *ITI duration*. Similar to the finding reported in mice above, there is some evidence that initial massed extinction trials followed by spaced trials may provide an advantage for humans at the end of extinction, but the benefits of this "expanding" practice might be washed out at renewal tests compared to subjects extinguished with constant ITI durations [8]. There is further increasing evidence for a core role of "prediction error", implying that focussing on violating patients central concerns about aversive outcomes is an important ingredient of successful extinction learning [2].

Clinical evidence: Clinical application studies have attempted to discern the most effective length of time between exposure sessions and have varied in their results. Comparing different exposure schedules (massed one-session treatment, uniform time intervals, gradually expanding time intervals), massed treatments result in more return of fear than uniform or expanding schedules. Other studies find no differences in extinction levels for massed versus spaced exposure. Some researchers find a benefit of massed extinction for behavioral measures while others report a benefit of spaced trials, but none of these tested for long-term follow up. However, studies that include follow-up and generalization tests tend to find that while massed exposures result in immediate enhanced extinction learning, spaced exposures result in long term inoculation against return of fear on generalization tests. One possible explanation for a benefit of spaced exposures is that the spaced sessions allow for changes in context (e.g. the internal state of the participant, changes in the environment or time of day, etc.) that increase retrieval cues for the inhibitory memory and therefore make the extinction learning more robust and less context-dependent, which could produce enhanced long-term extinction learning [9]. Also, the variability in the timing of exposure (e.g. condensed as well as spaced) may uncouple extinction memories from their temporal context [2].

To conclude, condensed exposure with short ITI results in faster extinction than spaced. This might however be at the expense of return of fear and clinically a higher chance of relapse. Therefore in the initial active treatment phase more trials with shorter ITI should be most promising for a rapid and more effective extinction learning. To stabilize this, subsequent spaced trials in different contexts is preferred.

1.2 Moderators for treatment outcome and extinction learning in AD

Type of anxiety disorder: There are no systematic studies that directly compare outcomes and mediator processes across various forms of anxiety disorders [5]. With the possible exception of Generalized Anxiety Disorder (GAD) where results seem to be characterized by slightly poorer outcomes than for Panic, Agoraphobia and Specific Phobia, we are also unaware of evidence and studies that have found substantial differences between AD when examining exposure-based treatments. Nevertheless the content of manuals and interventions tested are quite variable, making comparisons in terms of outcome problematic. Focussing on extinction learning as a common shared core mechanism of action for exposure therapy, we expect that our treatment manual will result in similar outcomes and patterns of mediations irrespective of diagnosis. We might speculate that GAD is less influenced by extinction based mechanisms, however evidence is lacking.

Comorbidity: Clinically, the failure to use exposure therapy for AD is justified with the assumption that these procedures do not work in comorbid diagnostic patterns [5]. Specific research in this area is scarce [10], but there is no evidence from RCT on exposure-based treatments that depression and other comorbid constellations are associated with poorer outcomes. There seems to be also no evidence that depression moderates extinction-based mechanisms. Yet, it is fair to state that this hypothesis has so far only reviewed and examined for Panic Disorder and Agoraphobia [11] and that there is a strong need to specifically test this question in appropriate large scale studies across the full spectrum of AD [12].

1.3 Study rationale

There is considerable evidence that exposure-based treatments for AD reveal higher effect-sizes and greater persistence of improvement than cognitively focussed treatments without explicit exposure. However, relapse by return of fear is frequently observed, and the mechanisms of action underlying exposure remain debated and are yet to be studied in clinical samples. Novel preclinical research evidence suggests extinction learning as the core mechanism of action and provides according strategies to improve the effectiveness of treatment by optimized extinction. However, the preclinical evidence has not yet been systematically translated into clinical application and has never been rigorously tested. A translational research agenda is suggested to examine whether enhanced extinction learning components derived from preclinical research, applied within an "intensified" exposure-based treatment, improves outcomes. In a multicentre randomised clinical trial linked to mechanistic subprojects, we test in n=720 patients with primary AD allowing for comorbidity whether intensified psychological interventions based on augmented extinction learning (IPI) result in faster, stronger and more persistent outcomes on subjective, clinical, behavioral, physiological and neural indices as compared to an, otherwise identical, standard research treatment without explicit enhanced extinction (TAU). We hypothesize that enhanced extinction elements (IPI) will result in (a) higher effect sizes, faster recovery, and (b) more pronounced changes in an array of systems, including elements of extinction learning and in objective behavioral measures assessed in intersession exposure trials. We also examine moderators of outcome (i.e. type of diagnosis, comorbidity) and explore whether IPI is associated with lower health care costs.

As recently reviewed by us [5], clinical research and numerous meta-analyses of cognitive-behavioral therapy (CBT) trials for the various forms of AD have shown - with little diagnostic variations - impressive effect sizes and persistence of improvement beyond the end of the therapy in 6 and 12 month follow-ups. Considerable agreement also exists that exposure-based treatments typically reveal higher effect-sizes and greater persistence of improvement than cognitive treatments without explicit exposure. However, there are significant gaps on various levels: (1) Despite agreement that exposure, a procedure derived from extinction theory [6], is an essential core component, the debate related to the mechanism of action as well as the form, duration and density of effective exposure in treatment persists; clinical trial evidence systematically assessing these issues is lacking [1,5] with a few notable exceptions [13]. (2) The considerable body of basic and preclinical advances regarding extinction learning [14,15] has not yet been sufficiently translated into clinical interventions. It remains unclear to what degree supplemental elements of enhanced extinction learning also translate into improved outcomes in clinical anxiety, and a stepwise translational research agenda on fear extinction and extinction learning has been proposed [1]. (3) Most clinical research in this field has been conducted either with a single measure of conditioned fear and at best in selected mono-

symptomatic clinical cases. It remains unclear whether enhanced extinction is beneficial in all AD as well as comorbid patterns. (5) Patients with AD - at least in Germany - rarely receive appropriate anxiety treatment. In fact, studies have shown that exposure treatments are rarely applied at all, and if they are applied they seem to be provided without alignment with research standards or theory. This suggests indirectly that they might fail to stimulate appropriately presumed core mechanisms of action. Routine treatments in Germany are typically much longer (>50 weeks vs. RCT: 4-16), contain little to no exposure, are not manualized, and have considerably lower effect sizes [5]. This might be at least partially due to the historically established practice, reinforced by insurance regulations favoring psychotherapy sessions of 50 min. once a week irrespective of method and diagnosis applied. This practice is unlikely to promote effectively essential mechanisms of fear inhibition. Clinical research has so far had little impact to change this situation.

The proposed trial will address these critical issues by conducting a multicenter randomised clinical trial (RCT) in typical patients of anxiety clinics to test whether efficacy adjustments in exposure-based treatment packages, targeting elements of enhanced extinction learning, lead to superior results. At the same time the RCT will provide the data to address - within projects P3, P4 and P5 - the key question of extinction as a core mediating process as well as questions regarding moderators (i.e. diagnostic pattern, comorbidity). A resolution of these questions is expected to have significant effects on patients with AD and the way exposure-based treatments are used in the psychotherapy sector.

2. Multicenter randomised controlled clinical trial (P1)

2.1 Intervention scheme/trial flow

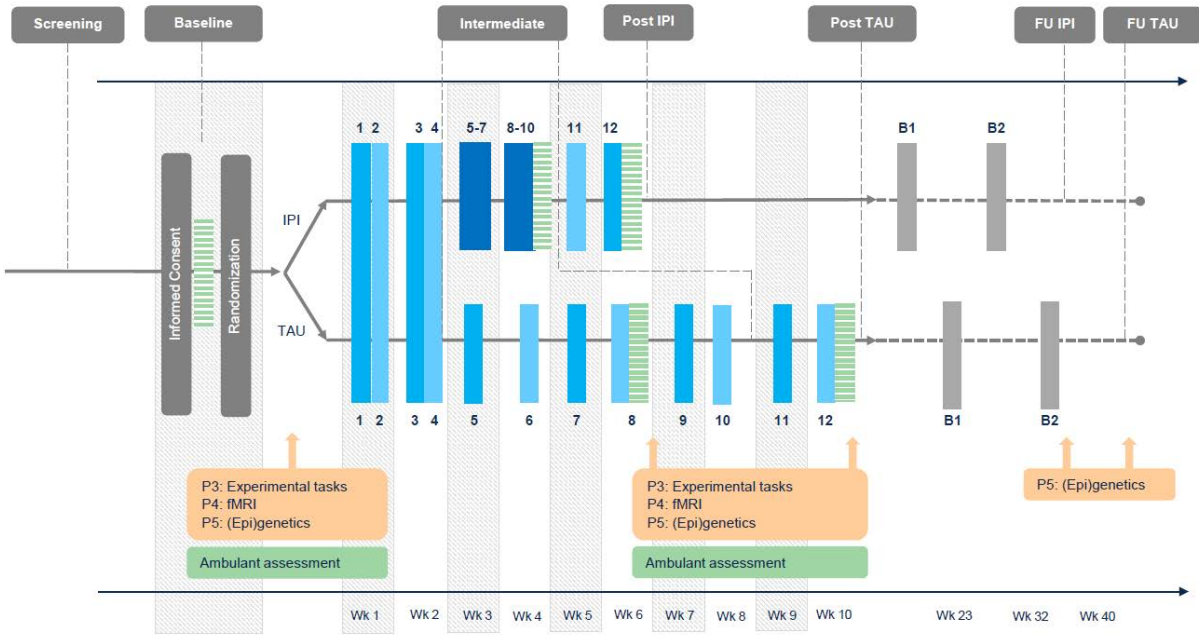

Figure 1. Intervention scheme/trial flow

2.1.1 Frequency and scope of study visits

Baseline diagnostic (2-3h) assessment visits will be scheduled on two days. In both groups 14 therapeutic sessions according to an established modular manual, of same length and content is provided. In IPI, however, the exposure module (sessions 6-10) is provided intensified over

a shorter duration of two weeks, ending 4 weeks earlier than TAU as the comparator. In TAU, the exposure module is provided as usual over a 6 weeks period. Post assessment will occur at approx. week 10 for both groups. The two booster sessions occur in the follow-up period. Patients are instructed to engage in spaced exposure trials across different contexts at post and during the two booster sessions. If eligible, additional visits are scheduled for the mechanistic projects (P3-5; see Figure 1 for details).

2.1.2 Treatment description

Treatment is based on an established modular manual which has been initially developed primarily for Panic Disorder and Agoraphobia [16]. The manual is highly structured with word-by-word instructions for therapists and has been tested in multicenter RCTs. The compliance and treatment integrity has been investigated and is high [13].

In order to extend the scope of the manual to all included anxiety disorders, a number of elements will be traced back to their generic form (e.g. the vicious circle of anxiety will not be explained solely for panic but for the somatic and cognitive features of the anxiety reaction as such). Other elements of the manual will be adapted to the respective diagnosis (e.g., exposure trials will be adapted to the specific feared stimuli of a given patient). All therapeutic elements will be taken from validated manuals on anxiety disorders and will be aligned to extinction as the suspected operating mechanism.

Treatment will comprehend five phases: (1) In sessions 1-3, therapeutic report is established and patients will receive psychoeducation on normal and pathological anxiety. A diathesis stress model will be presented as a generic approach to anxiety disorders and models of the development of fear in a given situation will be discussed. Based on the patient's "life line", individual risk factors and antecedents of AD will be added to the diathesis stress model in order to develop a comprehensive conception of the development of AD in the patient. Patients are encouraged to monitor their fear reaction over the period of several days. Selected feared situations will be analysed via behavioral analysis, which leads to a discussion of the role of avoidance and safety behavior in the maintenance of AD.

(2) In session 4, patients will be cognitively prepared for exposure. This phase includes confrontation with imagined feared situations (exposure in sensu) or with bodily sensations that resemble those of anxiety (symptom provocation). The rationale of extinction is explained in detail to foster learning during exposure. A thought experiment on extinction will be carried out which involves the recognition that fear cannot be a permanent state and may reduce anticipatory anxiety. Finally, a hierarchy of feared situations is to be established that can be used to derive exposure situations for the following phase.

(3) In sessions 5-10, patients will undergo exposure in various feared situations with varying context conditions. During exposure trials, anxiety will be monitored continuously. Each exposure trial is introduced and summarized together with the therapist (e.g. to control the use of safety behaviours). The therapist is, however, not necessarily present during the whole exposure trial. If necessary, motivational strategies are applied prior to the trial.

(4) The purpose of sessions 11-12 is to summarize learning experiences and to facilitate further generalization of the learned behavior. To that end, an individual training schedule is developed that encourages patients to continue with self-guided exposure (spaced in IPI). To prevent relapse, the phenomenon of return of fear is discussed. Patients are sensitized for situations in which return of fear is frequent and are instructed to create an emergency plan for such cases.

(5) Two booster sessions (B1, B2) are planned in which the training schedule is controlled and a further "spaced exposure" training plan in various contexts is developed.

Therapy sessions are projected to last about 90 minutes, except for exposure sessions which may last longer. Between sessions, patients will receive homework that consists of preparing material for the following session and of self-guided exposure exercises. Treatment between IPI and TAU differs only in regard to the time schedule in sessions 5-12.

2.1.2 Assessment battery

The cross-diagnosis and diagnosis-specific process and outcome measures used in the study are summarized in Table 1.

2.1.3 Strategies for data handling

Data handling: The Coordination Center for Clinical Trials (KKS) Dresden will be responsible for the data management according to the Good Clinical Practice Guidelines. This includes the setup of a study database by means of study software REDCap 3.3. Data will be examined by programmed range checks, validity checks, and consistency checks. REDCap is a secure web application, including an audit trail, for managing online surveys and databases. The data management will be supported by Dr. Dipl.-Stat. Michael Höfler and CELOS staff and resources. The responsible data officers will be jointly coordinating centrally all aspects including data handling, management and analysis in Dresden. All participating centers are intimately knowledgeable about this set-up from previous studies and have an assigned data manager, who will manage and supervise the online transmission of data and site specific quality assurance procedures. Sampling and data collection will be continuously monitored for quality by CELOS staff and the KKS Dresden. Using the KKS data management system (REDCap) established for previous clinical trials in the BMBF Psychotherapy Funding initiative the existing comprehensive internet-based data entry platform that covers all stages and components from stratified sampling to analyses will be adapted. All procedural steps and assessments will be online with automatic recall and plausibility and quality assurance measures. The assessment packages for each visit include standardized computerised diagnostic data (DIA-X CIDI). Compliance with the trial specifications, treatment integrity and compliance, and the explicit manual specifications will be monitored by random video documentation.

Table 1. Process and outcome measures

construct	Instrument	domain	diagnostics	baseline	intermediate	post	follow-up
Timeline			**week 1**	**week 1**	**week 8**	**week 10**	**week 34**
Duration[1]			**3 h**	**2 h**	**2h**	**2h**	**5h**
primary outcomes	HAM-A / SIGH-A	somatic and psychic anxiety symptoms	**X**			**X**	**X**
over-arching	DIA-X	categorical diagnosis	**X**				**X**
	CGI	symptom severity	**X**		**X**	**X**	**X**
symptom-atology	BSI	psychological symptoms		**X**	**X**	**X**	**X**
	BDI-II	depressive symptoms		**X**	**X**	**X**	**X**
	ASI	anxiety sensitivity		**X**	**X**	**X**	**X**
diagnosis specific symptom-atology	DSM-5 scales	dimensional measures		**X**	**X**	**X**	**X**
	PAS	panic and agoraphobia		**X**		**X**	**X**
	ACQ	agoraphobic cognitions		**X**		**X**	**X**
	BSQ	bodily symptoms		**X**		**X**	**X**

	MI	avoidance behavior	**X**		**X**	**X**
	PAS	panic and agoraphobia	**X**		**X**	**X**
	GAD-7	worrying	**X**		**X**	**X**
	LSAS	social anxiety	**X**		**X**	**X**
impair-ment & quality of life	WHODAS/CIDI	quality of life	**X**		**X**	**X**
	EQ5D		**X**		**X**	**X**
	WHO-5		**X**		**X**	**X**
	EMA	behavior change	**X**	**X**	**X**	
process variables	Sess. protocols	process exposure		**after each session**		
	CTS	childhood traumata	**X**			
	C-Scale	credibility of rationale			**X**	
	AAQ-II	psychological flexibility	**X**		**X**	**X**
	PFB-K	partnership quality	**X**			**X**
	INEP	therapeutic side effects			**X**	**X**
	STA-R	therapeutic alliance		**X**	**X**	

HAM-A: Hamilton Anxiety Scale [17]; SIGH-A: structured interview guide for HAM-A [18]; CGI: Clinical Global Impression Scale [19]; DIA-X: Computerized Version of the Munich Composite International Diagnostic Interview (M-CIDI) [20]; BSI: Brief Symptom Inventory [21]; BDI-II: Beck Depression Inventory-II [22]; ASI: Anxiety Sensitivity Inventory [23]; DSM-5 Scales: Dimensional Anxiety Scales from the Diagnostic and Statistical Manual of Mental Disorders [24]; PAS: Panic and Agoraphobia Scale [25]; ACQ: Agoraphobic Cognitions Questionnaire [26]; BSQ: Body Sensations Questionnaire [27]; MI: Mobility Inventory [28]; CTS: Childhood Trauma Screener [29]; C-Skale: Credibility Scale [30]; LSAS: Liebowitz Social Anxiety Scale [31]; AAQ-II: Acceptance and Action Questionnaire [32]; PFB-K: Partnerschaftsfragebogen Kurzform [33]; EQ5D: EMA: Ecological Momentary Assessment; INEP: Inventar zur Erfassung Negativer Effekte von Psychotherapie [35]; STA-R: Skala Therapeutische Allianz-Revised [36]; WHO-5: WHO Well-Being Index [44]

2.2 Justification of design aspects

2.2.1 Controls/comparator

Given the existence of state-of-the-art exposure-based manuals with established efficacy there is no need to have an untreated control group. We thus test two variants of a highly standardized treatment manual. Both conditions are considered to be state of the art. Participants in both conditions, IPI and TAU, are expected to improve in the range of previous RCTs.

Note: We use the acronym TAU (Treatment As Usual) for the control condition. This should not be misunderstood as meaning that TAU is the type of treatment typically applied in routine care by German providers. We refer, when using TAU to established manualized intervention programs for AD used in controlled clinical studies. Thus we provide a tough and rigorous test aiming to demonstrate that effects that are observed relate directly to modifications of extinction learning elements.

IPI and TAU. Using an existing, well studied and frequently used manual template [16] modified for the purpose of this RCT, the manual has 12 therapy sessions (plus baseline, post, booster and FU) with identical therapeutic modules (i.e. rapport/diagnostic assessment, psychoeducation, cognitive- and behavioral exposure interventions), adapted to the type of AD and the purpose of the study. The experimental condition IPI differs only with regard to the temporal structure (higher number of trials in a shorter duration) in the behavioral exposure module in sessions 5-10, provided in only two weeks and (b) instructed spaced exposure trials during sessions 11-12. These modifications are associated with respective manualised instructions and assessment tools (see 3.3.5). In the control intervention (TAU)

this BE module will be provided over 6 weeks and no instructed spaced trails in different contexts are provided. One essential content modification in both conditions is to evoke a positive prediction error. Patients are asked for the aversive consequences that they expect to be associated with the respective standard exposure trial and about the probabilities that these consequences might occur prior, during and after exposure trials. The contingency expectancies are discussed with the patient including probability ratings for aversive consequences in subsequent exposure trials, in order to facilitate learning that the consequences did not occur at the rate expected or were not as aversive as expected. We provide this component in both groups, because we assume that the positive prediction error will specifically enhance extinction learning only in IPI.

2.2.2 Dose, mode and scheme of intervention

Session 1-5 of the interventions in both groups will be identical in IPI and TAU (e.g. building up therapeutic rapport, psychoeducation, monitoring, identifying antecedents, techniques to reduce anticipatory anxiety). However, the temporal structure of sessions 5-10 and 11-12 are different, while all other sessions are identical. In the follow-up phase booster sessions participants are instructed for spaced inter-trial exposure in different contexts. Due to the study-specific requirements of documentation, assessments and quality assurance (video, patient compliance ratings etc.) each session lasts about 90 min.

2.2.3 Additional treatments

Patients concomitantly treated by other psychotherapists or psychiatrists for any mental health reason are excluded. We enrol patients that prior to inclusion were on medication, but wish to stop for whatever reason after tapering out for at least 8 weeks, monitored by urine tests. In case of acute suicidality during treatment, patients will be seen by the site psychiatrists and eventually withdrawn from the trial.

2.2.4 Inclusion/exclusion criteria and recruitment strategy.

To increase relevance for routine care, we intentionally keep exclusion criteria minimal. Key inclusion criteria: (1) outpatients, (2) age: 15-70 years, (3) current primary DSM-5 anxiety disorder, (4) severity at baseline (HAM-A>18) and CGI>3, (5) written informed consent, (6) able to attend on his/her own or accompanied by significant other, (7) language competence. Key exclusion criteria: (1) any DSM-IV/5 psychotic, primary mood disorders (Bipolar I, recurrent or chronic Major Depression), current Substance Dependence, (2) concomitant psychological/psychiatric treatment, (4) acute suicidality, (5) general medical contraindications, (5) mono-symptomatic Specific Phobia as the primary diagnosis.

2.2.5 Outcome measures

- Consistent with all clinical trials and guidelines [36] the primary efficacy endpoint is the clinician-rated Hamilton Anxiety Rating Scale (interview version, SIGH-A). Although the HAM-A lacks sensitivity for reflecting changes induced by psychological treatment, we are unaware of any established alternative that would work for all AD. We currently also explore the ADIS global measure of clinical severity (0-8) as an option; given this dilemma, we opted for a broader range of “secondary” outcome measures (see below).
- Other core generic outcome measures are (ordered by relevance); the Clinical Global Impression Scale (CGI), patient rated psychological (BSI) and depressive symptoms (BDI-II), anxiety sensitivity (ASI), and impairment/disability days (CIDI Harvard Index; WHODAS 2.0), The choice of these generic measures is meant to maximize comparability with other studies.

They serve as core measures to determine clinical treatment efficacy and differences between conditions in the first place. They also will be used to determine speed of recovery, being administered every 2nd session. To test diagnosis-specific effects we use as secondary endpoint(s) the DSM-5 dimensional anxiety-scales (DSM-5) and established diagnosis specific measures (i.e. PAS, ACQ, BSQ, MI, GAD-7, LSAS).

- A more reliable, objective and ecological valid evaluation of behavioral effects in everyday life is used to assess quantity, frequency and appropriateness of prescribed and natural exposure trials within sessions and between sessions. We use an innovative e-health tool that simultaneously records geographic position, mobility (actography), together with event-related momentary assessments of mood and anxiety ratings (ecological momentary assessment; EMA). Patients will be supplied with smartphones for objective assessments in real life situations.
- Subjective assessments, directly linked to experiences during standardized exposure protocol sessions, allow for (i) the quantification of subjective levels of anxiety and expectations about central concerns and (ii) the derivation of proxy measures of within- and between-session extinction, including emotional learning as well as changes in contingency expectations. Beyond the coupled experimental paradigms (P3, P4) we relate clinical outcome data to extinction learning as the central mechanism of change. For each patient an extinction curve on multiple response output measures (i.e. cognitive, autonomic) will provide information about consolidation of extinction memory, reactivation and reinstatement of fear memory as well as recall of extinction memory for each patient. These explorative proxy measures will be related to diagnostic groups and treatment outcome variables.
- To examine the impact of the intervention on the patient life and functioning we measure the generic quality of life as well as all direct and indirect health care costs (CIDI health economic module).
- In addition the CIDI, standard sets of predictors (personality, biographical and family genetic factors; baseline CAPI M-DIAX CIDI interview) are covered.

2.2.6 Determination of primary and secondary measures

As endpoints we use the standard total scores, derived from the respective scales (i.e. HAM-A). We explore the additional use of composite scores.

2.2.7 Methods against bias

- *Randomisation:* Participants are allocated randomly to the two study conditions. No restrictions of feasibility apply; there is no reason to take any prognostic factor into account, because sites use a stratified recruitment scheme.
- *Multicenter trial*: 7 centers will equally contribute patients, all of which have similar patient populations. Centers are also all involved in behavior therapy curricula from which therapists are sampled. Thus, significant side effects are unlikely and if they occur they will be controlled for.
- *Blinding of condition* is not feasible in psychotherapy studies.
- *Assessments:* Center staff will be trained (2-days course) in the assessment procedures that are standardized, computerized and accompanied by SOPs. Initial screening is based on the M-CIDI screener, prior to the more detailed baseline psychopathological and psychological assessment with the computerized clinical diagnostic interview (M-CIDI-CL) supplemented by the assessment of various constructs (personality, anxiety sensitivity, behavioral inhibition, family genetic module) with the embedded outcome measures. The

computerization allows for immediate access to standardized data files across sites, reduces data cleaning efforts and enhances objectivity and quality.

- *Therapy and manual*: All details including the standardized exposure sessions in-vivo are specified in the treatment manual. The manual is modular, highly structured by session with verbatim instructions (see [16]), has been tested in RCTs primarily for comorbid Panic Disorder and phobic patients and is widely used. The compliance and treatment integrity has been investigated and is high [13].
- *Therapists and adherence*: Manual compliance is monitored, repeated violations lead to the exclusion of the therapist. Each center designates a minimum of 4 and up to 8 licensed therapists with at least 1 year experience. There will be a total of four 2-day training seminars, before and during the study, conducted by the central study center in Dresden and including a video adherence procedure protocol. Based on continuous adherence ratings, therapists will be licensed, non-compliant therapists will be excluded. During the trial all sessions are taped; a random 5% is checked by monitoring.
- *Monitoring*: All sites will receive regular (4) visits by study monitors to check randomly protocol adherence. Both assessment and therapy sessions will be monitored. Project violations will be documented, and lead to exclusion of patients, therapists or the center.

2.2.8 Proposed sample size and power calculations

Focusing on the HAM-A total score as the core primary outcome measure for determining the sample size, we expect - building on the meta-analytic results and pilot tests - that patients in IPI group will reduce the HAM-A-score (end of treatment) by 12 points. In the TAU group we expect a reduction by 10 points. Using a power of 80%, a test significance level of alpha= 5%, a standard deviation of 10 points and a one-sided t-test to detect differences between the IPI and TAU group, a total of 310 patients are needed per group. Based on previous such trials [13] drop-out rates during treatment including follow-up of 10%-15% are expected. Total n to be enrolled is 720 and to be analyzed 620.

2.2.9 Compliance/Rate of loss to follow-up

The rate of non-compliance and other drop-out has been determined based on a similar study that was just completed [13]. The loss of 15% to follow-up is lower than in previous studies, presumably due to the treatment density.

2.2.10 Feasibility of recruitment

Each center has to screen approx. 200 patients to enroll an equal number of approx. 110 patients over two years, to reach the goal of approx. 90 patients with post and follow-up for analysis (= 620 patients total). All centers have provided firm evidence to reach these targets, confirmed by the fact that such target numbers were actually exceeded in previous trials.

2.2.11 Stopping results

Treatment will be discontinued on request of the patient (e.g. not willing to conduct the trials, not able to attend sessions) or in case the therapist after consulting with the supervisor comes to the decision that treatment is associated with substantial risks for the patient (i.e. need for drug treatment, suicidality). These cases will be counted as drop out. Based on the results of prior such studies as well as prior research we do not see any reason for findings that make failure criteria for the trial likely. Participating centers will be closed in case of repeated failure of recruitment or continued violation of the SOP and protocol. In this case we have two backup centers (Tübingen and Chemnitz) to replace them.

2.3 Statistical analyses

Statistics for the core primary outcome measure (HAM-A): In order to demonstrate that the IPI and TAU treatments differ, we use a one-tailed t-test with the following hypotheses:
H_0: $\mu_{HAMADiff\text{-}IPI} = \mu_{HAMADiff\text{-}TAU}$; H_A: $\mu_{HAMADiff\text{-}IPI} > \mu_{HAMADiff\text{-}TAU}$.
Here $\mu_{HAMADiff\text{-}IPI}$ and $\mu_{HAMADiff\text{-}TAU}$ are in each case the differences between the HAM-A baseline and the HAM-A value after the end of therapy in the treatment groups IPI and TAU.

Trial data will be analyzed both on an intent-to-treat (ITT) and per-protocol basis. To test whether groups have different outcomes for the major end points at (i) post and (ii) follow-up t-tests for independent samples will be used one-sided with an alpha level of 0.5. This will be also done for secondary and explorative analyses (in that case with two-sided tests). To quantify effect sizes for differences between groups at post-treatment and at follow up Cohen´s d (with the pooled standard deviation at baseline in the denominator) will be used. To assess temporal changes in symptoms we will compare scores at initial assessment with those at t1 through t5. Individual outcome as a function of pretreatment characteristics will be modeled with multiple linear regression and associated analyses of variance.
Data will be analyzed with STATA and, if necessary, statistical weights will be used to adjust for different sampling probabilities in different study centers. LOCF-analyses (intent-to-treat) will be carried out as well as completer analyses and mixed models (saturated for the combined effects of time and group on outcome) to assess whether dropout changes the conclusions. Moderator analyses (predictors of treatment effects within IPI and TAU patients) will address type, respectively group of disorders as well comorbidity using dimensional and categorical measures. Pearson correlations between such comorbidity measures and outcome will be calculated. Associations with categorical predictors will be analyzed with linear regressions using dummy variables and quantified with mean differences in treatment outcome (and their 95% confidence intervals).

2.4 Ethical considerations

The study will be conducted in full accordance with the Declaration of Helsinki, the German data protection act, and the GCP-Guideline and is sensitive to ethical considerations [5]. The study protocol, amendments to the protocol (if applicable), therapy contents, patient recruitment procedures, information and informed consent form will be presented to the TUD ethics committee for approval. Subsequently, the study protocol will be approved by the respective site's ethics committees.

2.5 Quality assurance and safety

2.5.1 Quality assurance and monitoring

All participating centers are highly experienced with multicenter quality assurance and management procedures specified in SOPs and used in this RCT. All staff will be trained by CELOS and KKS Dresden and will go through three 2 days training sessions for (i) assessment, (ii) the treatment manual and (iii) general procedures of data entry, quality assurance and data management using web-based system MACRO. Subsequently there will be a 2 week site implementation phase, in which the coordinating center staff will visit and monitor the site personnel (month 4). Prior to the start there will be a third 2-days training to assess reliability

for assessment instruments and final adjustment in procedures (linkages to associated projects). All sites are requested to monitor all assessment and treatment sessions by tapes. Tapes will be randomly (5%) checked by the coordinating center to detect protocol violations as soon as possible. 1-day booster training sessions for assessment and the manual will be conducted every 3 months.

2.5.2 Safety

To ensure safety and minimize risk only licensed psychotherapists will conduct the treatments and interviews. All serious adverse events will be immediately reported to the principal investigator who will discuss with its members consequences for the trial. This procedure is analogous to that of pharmacological trials according to GCP.

3. Add-on project Psychophysiology (P3): Behavioral and psychophysiological markers of extinction learning and outcome

3.1 Background and aims

The most effective treatment of anxiety disorders share exposure techniques as a core ingredient. Exposure treatment needs to be tailored in a way to promote extinction learning and prevent return of fear. The current research proposal aims to investigate the process of extinction learning across a large group of anxiety disorder (AD) patients prior to and after exposure therapy. We want to assess extinction learning at all levels of the emotional network, including verbal report (US-expectancy ratings), physiological arousal (sympathetic nervous system activity), and amygdala dependent modulation of defensive reflexes (fear potentiated startle). We expect that extinction learning is impaired in AD patients and will improve after exposure therapy. We hypothesize that extinction learning will improve to a greater degree after IPI than TAU. We expect that the response output measures are related to neural network activation (see P4). We expect extinction learning parameters being related to moderators of clinical outcome.

Furthermore, clinical anxiety is characterized by an incapacity (or reduced ability) to recognize both danger and safety signals and consequently to overestimate danger [37]. The impaired recognition of danger signals prompts a continuous search for safety and a sustained feeling of apprehension. This symptomatology is well modelled by paradigms like contextual fear conditioning. In a context-conditioning paradigm the subjects cannot predict the aversive event (US) and consequently show a sustained fear response. Sustained fear is a long-lasting state of apprehension induced by the inability to identify the source of threat [38-39]. As this technology is available at 4 sites (Dresden, Würzburg, Bochum, Marburg), we will here add a virtual reality (VR) paradigm to the assessments within P3 as an elegant and innovative tool to realize contextual fear conditioning in an ecological fashion in humans in analogy to animal models. By means of a VR context conditioning paradigm we are able to investigate the mechanisms underlying the acquisition and extinction of such sustained fear. Notably, we assess extinction learning on the cognitive (ratings), behavioral (startle responses and avoidance/approach) and physiological (skin conductance) level of responses. We expect extinction learning deficits in anxiety patients which will be improved after treatment.

Project P3 is essential for the identification of extinction learning being a mediator for outcome. Beyond it provides a range of experimental data regarding the mechanism of sustained fear and extinction in its role in AD. By applying two experimental paradigms covering sustained fear (VR) and extinction to all patients giving informed consent,

moderators of clinical outcome (e.g., type of AD, comorbidity, severity etc.) can be related to the mediating process. Moreover, by investigating the neural networks (P4) and (epi)genetic modulation (P5) of extinction learning we might provide evidence for central mediator and moderator mechanisms.

3.2 In- and exclusion criteria, recruitment strategy

Patients. Eligible patients from the RCT (P1) will be included in P3. Due to the startle procedure patients with severe hearing problems will be excluded from P3.
Controls. In conjunction with P4 we will recruit control subjects from the general population in the local area of the respective trial site. Controls will be matched for gender, age (+/- 5 years), education, handedness, and smoking. Inclusion criteria: 1) aged 15-70 years, 2) informed consent. Exclusion criteria: any lifetime psychiatric disorder according to DSM-5 criteria (excluding nicotine dependence); severe hearing problems; general medical contraindications (e.g. neurological or cardiovascular diseases).

Due to the shared experimental paradigms, patients and controls will participate in two experiments (the VR experiment and a fear conditioning procedure) in the psychophysiological lab on day 1, while extinction learning and reinstatement will be tested in the fMRI scanner on day 2. Those patients not eligible for fMRI will complete the extinction paradigm in the psychophysiological lab on day 2 (see Figure 2 for details).

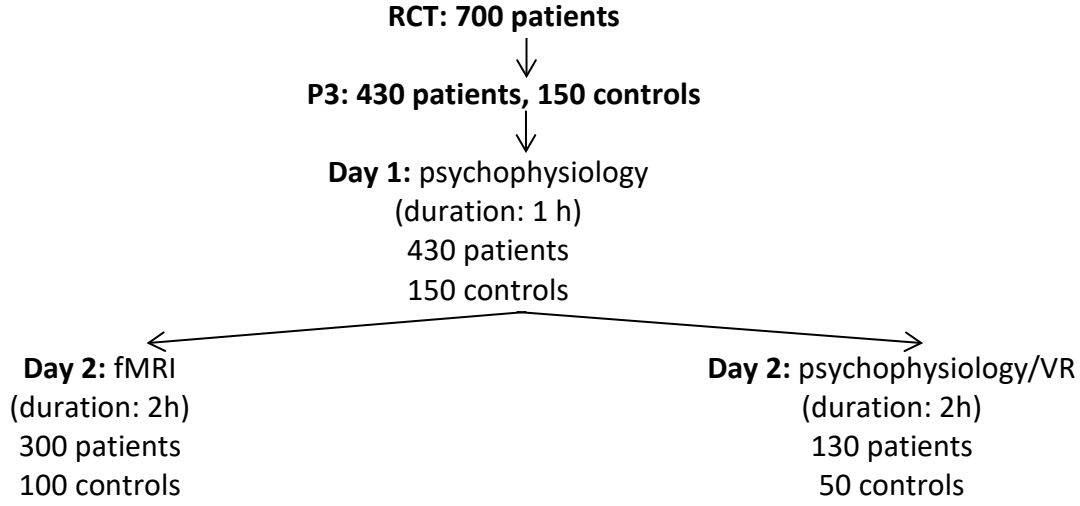

Figure 2. Patient flow for experimental add-on study

3.3 Experimental procedure

Extinction learning comprises different phases. First, the individual learns that a feared cue (or context) is no longer associated with the expected aversive outcome. With repeated extinction trials, extinction memory is consolidated and can be recalled even under stressful circumstances or in different contexts. Here we want to investigate fear conditioning and extinction learning in the laboratory. It is aimed to relate these multiple psychophysiological measures of extinction learning to neural network activation during extinction examined in P4, and its (epi)genetic modulation (see P5). We will use a delayed extinction paradigm, meaning that during day one subjects will learn that one cue (CS+ neutral face 1) is associated

with the aversive US (pulsed tactile electrostimulation on the forearm; 500ms in individually adjusted intensity reflecting "uncomfortable, but not painful" experiences) while the other cue (CS- neutral face 2) is never paired with the US. To ensure successful learning by all subjects (extinction learning can only be investigated in a meaningful way when the fear responses are reliably acquired) we will explicitly instruct all subjects that one of the two stimuli is followed by the US. To ensure successful conditioning subjective and physiological responses during acquisition were compared to responses during a short pre-conditioning phase (presenting CS+ and CS- alone for three times each) conducted immediately prior to conditioning. Duration of the task on day 1 is about 10 minutes (total duration of the assessment including electrode preparation, adjustment of electrostimulation, and expectancy ratings: 1h).

Extinction starts at day 2 with a reactivation of the fear memory (the CS+ followed by the US once). Then CS+ and CS- will be presented 20 times each without any US. Directly after extinction learning, fear will be reinstated by presenting the US alone (3X). Recall of extinction memory will then be tested by 10-CS-trial each. In order to maximize synergies between P3 and P4 and to minimize the load for included patients, the second part of the delayed extinction paradigm will be conducted within the MRI scanner as part of project in eligible subjects (Table 2).

The VR context conditioning paradigm (delivered only at day 1) consists of six phases (exploration, two conditioning phases, a behavioral test, and two extinction phases) separated by verbal assessments. During the exploration phase, subjects explore each virtual office for 2 min by means of a joystick. During the two conditioning phases, subjects hear the desperate female scream (US) in one room (CTX+), but never in the other room (CTX-). Importantly, the US is presented unpredictably during the visit of the CTX+. A trial starts in a corridor (the inter-trial interval, ITI) and then subjects are passively guided into one virtual office. Each context (CTX+ and CTX-) is entered three times during the conditioning and extinction phases. For the behavioral test, participants are in the middle of the corridor and have to actively enter one virtual office. Subjects can make the choice by using a joystick. The extinction phases are similar the conditioning phases, except that no US is delivered. Startle-eliciting stimuli are delivered in each room as well as in the corridor. Duration of the task is approx. 1 h.

Table 2. Description of the delayed fear extinction and the VR context conditioning paradigms (shared paradigm P3 & P4)

	Day 1 (lab only)	Day 2 (fMRI; if not eligible: lab only)	
	Instructed fear conditioning	Fear extinction learning	Reinstatement test
VR paradigm:			
Conditioning phase:			
Frequency CTX+	12		
Frequency CTX-	12		
Frequency US	12		
Extinction phase:			
Frequency CTX+	20		
Frequency CTX-	20		
Conditioning & extinction paradigm:			
Frequency CS+	13	20	10

Frequency CS-	13	20	10
Frequency US	6	1	3 (CS/US)
Dependent variables:			
FPS recording	X	- (lab: X)	- (lab: X)
SC recording	X	X	X
HR recording	X	- (lab: X)	- (lab: X)
Expectancy ratings	X	X	X
Questionnaire on VR experiences	X		

FPS: fear potentiated startle; SC: skin conductance; HR: heart rate; Expectancy ratings: expectancy of US occurrence (0-100%): X: assessment; -: no assessment (technical restrictions in the MRI environment)

3.4 Assessment tools

Fear conditioning and extinction paradigm. After attaching the sensors for physiological data collection (EMG, EDA, ECG) subjects are told that two pictures were presented repeatedly. Subjects are informed that no electric stimulation will be conducted during the first phase (pre-conditioning). After this phase the experimenter attaches the electrode for electric stimulation to the participants left forearm. The intensity of the electric stimulation is individually adjusted within five warned presentations to a level that is experienced as "uncomfortable, but not painful". Then, subjects were informed that during the second phase (conditioning) one of the two shown pictures (CS+, known allocation) is be followed by an electric stimulation during most presentations (electric stimulation during 6 of ten trials). During extinction and reinstatement at day 2 subjects are told that electric stimulation might be possible again. However, electric stimulation only occurs during the fear memory reactivation (one electric stimulation during initial CS+ presentation prior to extinction) and prior the reinstatement test (three electric stimulation alone during inter-trial-interval).
Two pictures depicting male faces with neutral expressions serve as CS+ and CS-, respectively. Picture duration is about 6 sec. Electric stimulation (500 msec pulsed stimulations) starts 5500 msec after picture onset, if conducted.
Recordings of electromyographic (EMG) activity over the left orbicularis oculi muscle serve as to measure the eyeblink component of the startle response. A digitized 50-msec burst of white noise (105 dBA, rise/fall time <1 msec) is amplified by a recording mixer and is presented binaurally through headphones to serve as startle-eliciting stimulus. During the majority of trials CS presentations are followed by an acoustic startle probe either 4.5 or 5.0 s after picture onset (pre-conditioning: 2 of 3; conditioning: 8 of 10; extinction: 16 of 20; reinstatement: 8 of 10). Moreover, startle probes are also presented during the ITI that varied between 6 and 10 sec (pre-conditioning: N=3; conditioning: N=8; extinction: N=16; reinstatement: N=8). Additionally, skin conductance (recorded from the hypothenar eminence of the palmar surface of the subject's non-dominant hand) and the ECG (Einthoven-II-lead) are measured continuously. After each phase (pre-conditioning, conditioning, reactivation, first half of extinction, second half of extinction and reinstatement) subjects are rating shock expectancies, intensities of distress/unpleasantness and arousal during both pictures, and unpleasantness of perceived shocks.
VR paradigm. The VR environment is created with the Source Engine from the Valve Corporation (Bellevue, USA), which is also used for the game Half-Life 2. The VR environments consists of two offices separated by a corridor. All spaces have a gray floor. The offices have

the same square footage, but differ in the arrangement of the furniture. The corridor is empty and presents the doors to the offices. The VR environment is presented using Oculus Rift Development Kit 2 (960 x 1080 pixel resolution; Oculus VR, Inc., Menlo Park, California). The simulation is controlled by the software CyberSession (VTplus GmbH, Würzburg). In case subjects experience motion sickness within VR, recording will be interrupted and the experimenter will offer a glass of water to the subject. The experimenter will require the subject to remain in the laboratory under supervision until the motion sickness dissolves. In the VR study, the sensors for the physiological recording (EMG, EDA; ECG) are first attached and then the participants are positioned into the VR by wearing the Oculus glasses (Oculus Rift Development Kit 2, Inc., Menlo Park, California). Subjects are informed that no scream or startle-eliciting stimulus is delivered during the exploration and that they can actively explore the virtual rooms. Afterwards, subjects are informed that during the conditioning phases the scream can be presented and, if they follow carefully the experiment, they can notice the relationship among the stimuli. No further information are delivered before the test phases.
The recording of the EMG and the electrodermal activities are the same as described above. Seven startle-eliciting probes are presented before the conditioning every 7-15 s. During each phase, 18 startle probes are presented (CTX+ = 6, CTX- = 6, ITI = 6). Startle probes are separated by at least 10 s, and 7 s after room entry as well as before room exit no startle probe is delivered.

3.5 Statistical analyses

To test the effects of conditioning, extinction and reinstatement on startle blink magnitudes, skin conductance response, heart rate response, and expectancy ratings analyses of variance with repeated measurements are conducted with within-subjects factors of Phase (pre-conditioning vs. conditioning vs. extinction vs. reinstatement) and Cue (CS+ vs. CS-). To test the effect of interindividual differences between patients (e.g. primary diagnosis, severity, comorbidity, chronicity, depression) on test performance (e.g. intensity and speed of conditioning, extinction, and reinstatement) mixed models of variance or correlation analyses are conducted.
The statistical analysis for the contextual fear is similar to the above mentioned analyses. Namely, five separated analyses of variance with repeated measurements are calculated for each dependent variable (startle blink magnitude, skin conductance, hear rate response, anxiety and expectancy ratings). In parallel, the within-subjects factor phase (Conditioning 1, Conditioning 2, Extinction 1, Extinction 2) and context (CTX+, CTX-, ITI) are considered. In the same manner, the clinical diagnoses or interindividual differences are considered.

4. Add-on project Neuroimaging (P4): Neural response and fear circuitry related to extinction learning and outcome

4.1 Background and aims

Eligible RCT patients with AD from all seven P1 centers will undergo fear extinction in the MRI scanner before and after exposure-based therapy using functional Magnetic Resonance Imaging (fMRI). Using a shared paradigm with P3 we will examine the neuronal correlates of fear extinction and reinstatement. In addition, a paradigm on emotion processing; structural

scans and a resting state examination will be conducted. Brain regions of interest will be amygdala, (para-) hippocampal and anterior cingulate cortex. We hypothesize that (1) impaired extinction learning and exaggerated emotion processing in AD as compared to healthy controls relies on sustained amygdala and reduced anterior cingulate cortex activation, while enhanced reinstatement is related to (para-) hippocampal function; (2) augmented extinction learning in IPI is associated with stronger reduction in amygdala activation and enhanced ACC activation as compared to TAU, providing indirect evidence for neural mediating processes.
All eligible RCT 1 (P1) patients treated either with IPI or TAU (expected n = 300) and 100 healthy controls will be investigated before and after treatment. Maximizing synergies between P3 and P4, we will use an identical fear conditioning and extinction task. While subjects will undergo fear conditioning in P3 on day 1, extinction and a reinstatement test will be assessed in a 3T MRI scanner on day 2 (including autonomic markers of conditioning and expectancy ratings; see Table 2), thus allowing for consolidation of fear memories. Amygdala reactivity will be tested by an emotional face-matching paradigm. T1w anatomical scans will be assessed for normalization and explorative morphometric analysis. Resting state activation will be assessed as a measure of non-task related functional connectivity. Established MRI quality procedures from the "Panic-Net" (phantoms, reliability testing) will be applied.
Project P4 is conducted in all seven centers within a strictly enforced multicenter protocol. P4 is the essential cornerstone of the consortium program, because it aims at providing evidence for extinction learning being indeed a core mediating process which is reflected in pervasive changes in the neural fear circuitry. Using the RCT patients (P1), and being nested with the experimental behavioral and psychophysiological paradigms used in P3, we are to our knowledge the largest such fMRI study. This also allows exploring the potential moderating role of type of AD, comorbidity and prior psychopharmacology for extinction learning. P4 will be able to relate fMRI to psychophysiology assessed using the same methodology as in P3. P4 will be closely linked to P5 and P6 to study the impact of epigenetic variation and pharmacological treatment. Joint analysis of our data about in vivo brain processes (P4) together with psychophysiological data (P3), (epi)genetic data (P5), and clinical data (P1) joins neurobiology and observable behavior.

4.2 In- and exclusion criteria, recruitment strategy

Patients. Eligible patients from the RCT (P1) who have undergone P3 will be included in P4. Based on previous experience we estimate that approximately 40% (n = 300) patients from P1 will be included for P4. Additional MRI-related exclusion criteria apply: (1) ferromagnetic objects in the body that cannot be removed, (2) tattoos or permanent makeup in the face and neck area, (3) pregnancy, (4) self-report of lifetime history of neurological disorders, (5) general medical conditions that preclude from MRI assessment.
Controls. In conjunction with P4 we will recruit control subjects (n = 100) from the general population in the local area of the respective trial site. Controls will be matched for gender, age (+/- 5 years), education, handedness, and smoking. Inclusion criteria: 1) aged 15-70 years, 2) informed consent. Exclusion criteria: any lifetime psychiatric disorder according to DSM-5 criteria (excluding nicotine dependence). Additional MRI-related exclusion criteria: (1) ferromagnetic objects in the body that cannot be removed, (2) tattoos or permanent makeup in the face and neck area, (3) pregnancy, (4) self-report of lifetime history of neurological disorders, (5) general medical conditions that preclude from MRI assessment.

4.3 Experimental paradigms and (f)MRI assessments

Extinction learning. In conjunction with P3 we will conduct the second part delayed extinction learning paradigm on day 2 in the MRI scanner, while subjects underwent the fear acquisition phase on day 1 in P3 only (see 3.2, 3.3 and Table 3 for details; task duration: 20min). Skin conductance as an autonomic marker of extinction learning will be assessed during scanning.
Emotional reactivity task. In order to assess neural reactivity in emotion processing areas (e.g. amygdala) we will conduct an emotional face matching task ("Hariri-Task) [40]. Subjects are presented three faces with two of them showing the same emotional expression. Subjects are instructed to match the two identical facial expressions. Control comparisons include the matching of identical geometrical forms (task duration: 6min).
Resting state activity. In order to identify task-unrelated neural networks associated with anxiety disorders [41] we will conduct a resting state assessment (duration: 10min).
Structural scan. For the analysis of morphometric differences, as well as for normalization and segmentation procedures (related to the fMRI task data) a high-resolution T1w MPRAGE sequence will be conducted (duration: 8min).

4.4 Assessment tools

Table 3. Assessment tools in P4.

Domain	Test	Duration (min)
Neuropsychological screening		
- Handedness	EHI	2
- Working memory	Digit span	5
- Executive functions	Trail-Making-Test	5
Questionnaires		
- MRI anxiety	Questionnaire on MRI experience[42]	5
- Contingency assessment	Interview	5

EHI: Edinburg Handedness Inventory [43]

Total duration (f)MRI assessments: approx. 60min
Total duration P4 (including neuropsychological screening (Table 3), preparation, electrode attachments etc.): approx. 2h

4.5 (f)MRI data acquisition and analysis pathway

Imaging experiments will be performed in 3-Tesla scanners at all sites. Each session will consist of a standard anatomical protocol, i.e. a sagittally acquired Magnetization Prepared Rapid Gradient Echo Imaging (MPRAGE) sequence and a series of whole-brain echo-planar imaging (EPI) scans acquired to measure blood oxygen-level dependent (BOLD) functional activity. Measures of quality control in this multicenter study have been previously established and will be applied (including reliability and phantom assessments at each site). Structural and functional image processing will be conducted with SPM8 (Welcome Trust Centre for Neuroimaging, UCL, London, UK). Functional images will be realigned and unwarped to correct for movement artefacts. Structural images are coregistered to the functional scans and all volumes are normalized to the MNI (Montreal Neurological Institute, Quebec, Canada) reference brain. Functional images are subsampled and smoothed with a Gaussian kernel (iterative smoothing will be applied for multicenter data). Individual images will be carefully

checked for excessive movement artefacts. First-level statistical analysis will be carried out for all subjects applying the general linear model (GLM). The expected blood oxygen level-dependent (BOLD) signal change will be modelled by a canonical hemodynamic response function for the regressors of interest; the six movement parameters of the rigid body transformation applied to the realignment procedure will be further introduced as covariates into the model. On the second level, random effects group analyses will be performed by entering contrast images into flexible factorial analyses as implemented in SPM8, in which subjects are treated as random variables. Significance thresholds will be established using cluster threshold approaches (Monte Carlo simulation). In addition to an exploratory whole brain analysis, an a priori region of interest analysis will be conducted on target regions of the fear network. Estimated beta values will be used for correlational analyses with performance measures and clinical symptoms.

5. Add-on project (epi)genetics (P5): (Epi)genetic effects related to extinction learning and outcome

5.1 Background and aims

Anxiety disorders and components of fear conditioning in particular are genetically determined. Several risk genes of anxiety and particularly extinction have been identified, with some of them also driving response to cognitive-behavioral therapy. First pilot studies imply epigenetic mechanisms such as DNA methylation in the pathogenesis of anxiety.
Here, for the first time the role of DNA methylation in the pathogenesis, as predictors of therapy response and as potential correlates of extinction elements in psychological interventions of anxiety disorders will be investigated accompanying the RCT (P1). The identification of (epi)genetic markers - intertwined with psychophysiological and neural network markers (P3/4) - in the etiology, course and comorbidity of anxiety disorders may aid in developing resilience-increasing preventive measures in high-risk groups. Additionally, the definition of epigenetic signatures as a core mechanism of action of fear extinction in exposure-based interventions and thereby an objective biomarker of treatment outcome is hoped to contribute to the development of a more targeted, personalized treatment of anxiety disorders based on epigenetic information.

5.2 In- and exclusion criteria, recruitment strategy

Patients. Eligible patients from the RCT (P1) will be included in P5. Additional genetic-related inclusion criteria: (1) Caucasian background. Additional (epi)genetic-related exclusion criteria: (1) impaired blood coagulation, (2) lifetime diagnosis of severe somatic disorder (such as cancer), (3) illegal drug use (including cannabis), (4) excessive alcohol and nicotine use (exceeding 14 glasses of beer/week or 20 cigarettes/day). The following characteristics do not constitute exclusion criteria, but will be documented: (1) body mass index, (2) smoking status, (3) hormonal contraception, (4) current somatic medication.
Controls. Not applicable.

5.3 Procedure and assessments

As displayed in Figure 1, blood samples will be collected at three time-points: after study inclusion (baseline), at post-assessment (after session 12) and after 6-month follow-up. Blood

samples include 10ml EDTA blood for each assessment and will be carried out by trained medical personnel supervised by a physician. Blood samples will be stored at the individual centres at -20°C and sent to the Laboratory of Functional Genomics at the Department of Psychiatry, Psychosomatics and Psychotherapy, University Hospital Würzburg, on dry ice once a month, where DNA will be extracted and stored at -80°C.
Samples will be analyzed for DNA variation/methylation patterns of candidate genes of anxiety and/or modulators of emotional-associative learning (e.g., COMT, MAO-A, 5-HTT, BDNF, CNR1, NPSR1). In case new candidate genes will be discovered, these will additionally be included in the analysis. Furthermore, a genome-wide DNA analysis and a genome-wide DNA methylation analysis will be conducted. It has to be mentioned that genome-wide DNA analyses ("fingerprint") have certain implications regarding data protection, i.e. there is a principle risk of re-identification. We will address this issue explicitly in the consenting procedure. In contrast, genome-wide DNA methylation analyses do not allow for re-identifiaction since these signatures are temporally variable due to constant environmental influences on methylation patterns.

5.4 Statistical analyses

(Epi)genetic information from candidate genes / methylome-wide data will be tested in relation to clinical, psychophysiological and neural data using t-tests, (M)ANOVA for repeated testing (including pre-post effects) or regression analyses. Correction for multiple comparisons will be applied, resulting in a corrected statistical threshold of $p < 0.05$.

6. Trial time flow and milestones

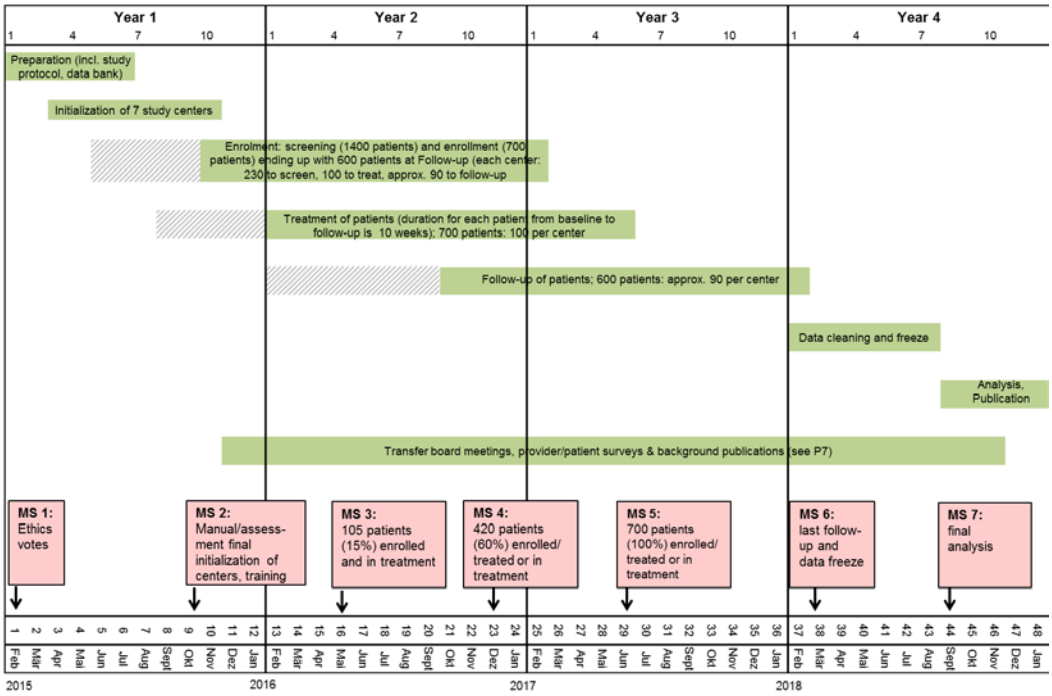

Figure 3. Project schedule and associated milestones

7. References

1. Vervliet, B., Craske, M. G. & Hermans, D. (2013). Fear extinction and relapse: State of the art. *Journal of Experimental Psychopathology*, *3*, 322-245.
2. Craske ,M., Liao, B., Brown, L. & Vervliet, B. (2012). Role of inhibition in exposure therapy. *Journal of Experimental Psychopathology*, *3*, 322-245.
3. Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, *52*, 976-986.
4. Davis, M., Walker, D. L., Miles, L. & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacoly, 35*, 105-135.
5. Neudeck, P. & Wittchen, H.-U. (2012). *Exposure therapy. Rethinking the model – refining the method* (Eds.). New York: Springer.
6. Boschen, M. J., Neumann, D. L. & Waters, A. M. (2009) Relapse of successfully treated anxiety and fear: Theoretical issues and recommendations for clinical practice. *Australian & New Zeeland Journal of Psychiatry*, *43*, 89-100.
7. Moody, E. W., Sunsay, C. & Bouton, M. E. (2006). Priming and trial spacing in extinction: Effects on extinction performance, spontaneous recovery, and reinstatement in appetitive conditioning. *Quarterly Journal of Experimental Psychology*, 59, 809-829.
8. Orinstein, A. J., Urcelay, G. P. & Miller, R. R. (2010). Expanding the intertrial interval during extinction: Response cessation and recovery. *Behavior Therapy, 41*, 14-29.
9. Rowe, M. K. & Craske, M. G. (1998). Effects of an expanding-spaced vs massed exposure schedule on fear reduction and return of fear. *Behavior Research and Therapy, 36*, 701-717.
10. Craske, M. G., Waters, A. M., Nazarian, M., Mineka, S., Zinbarg, R. E., Griffith, J. et al. (2009). Does neuroticism in adolescents moderate contextual and explicit threat cue modulation of the startle reflex? *Biological Psychiatry, 65,* 220-226.
11. Emmrich, A., Beesdo-Baum, K., Gloster, A. T., Knappe, S., Hofler, M., Arolt, V., Deckert, J. et al. (2012). Depression does not affect the treatment outcome of CBT for panic and agoraphobia: Results from a multicenter randomized trial. *Psychotherapy and Psychosomatics, 81*, 161-172.
12. Craske, M. G., Rauch, S. L., Ursano, R., Prenoveau, J., Pine, D. S. & Zinbarg, R. E. (2009). What is an anxiety disorder? *Depression and Anxiety, 26*, 1066-1085.
13. Gloster, A. T., Wittchen, H.- U., Einsle, F., Lang, T., Helbig-Lang, S., Fydrich, T. et al. (2011). Psychological treatment for panic disorder with agoraphobia: A randomized controlled trial to examine the role of therapist-guided exposure in situ in CBT. *Journal of Consulting and Clinical Psychology, 79*, 406-420.
14. Bouton, M. E., Westbrook, R. F., Corcoran, K. A. & Maren, S. (2006). Contextual and temporal modulation of extinction: Behavioral and biological mechanisms. *Biological Psychiatry,* 60, 352-360.
15. Milad, M. R. & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: Ten years of progress. *Annual Review of Psychology, 63,* 129-151.
16. Lang, T., Helbig-Lang, S., Westphal, D., Gloster, A. T. & Wittchen, H.-U. (2011). *Expositionsbasierte Therapie der Panikstörung mit Agoraphobie*. Göttingen: Hogrefe.
17. Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32*(1), 50-55.

18. Shear, M. K., van der Bilt, J., Rucci, P., Endicott, J., Lydiard, B., Otto, M. W. et al. (2001). Reliability and validity of a structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-A). *Depression and Anxiety, 13*(4), 166-178.
19. Guy, W. (1976). *ECDEU Assessment Manual for Psychopharmacology*. U.S. Department of Health, Education, and Welfare. Washington: DC.
20. Wittchen, H.-U. & Pfister, H. (Eds.) (1997). *DIA-X-Interviews: Manual für Screening-Verfahren und Interview; Interviewheft Längsschnittuntersuchung (DIA-X-Lifetime); Ergänzungsheft (DIA-X-Lifetime); Interviewheft Querschnittuntersuchung (DIA-X-Monate); Ergänzungsheft (DIA-X-Monate); PC-Programm zur Durchführung des Interviews (Längs- und Querschnittuntersuchung); Auswertungsprogramm.* Swets and Zeitlinger, Frankfurt.
21. Derogatis, L. R., & Spencer, P. M. (1993). *Brief Symptom Inventory: BSI*. Upper Saddle River, NJ: Pearson.
22. Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
23. Reiss, S., Peterson, R. A., Gursky, D. M. & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency, and the prediction of fearfulness. *Behavior Research and Therapy, 2*, 1–8.
24. American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Arlington, VA, American Psychiatric Association.
25. Bandelow, B. (1999). *Panic and Agoraphobia Scale (PAS)*. Hogrefe & Huber Publishers.
26. Chambless, D. L., Caputo, G. C., Bright, P., & Gallagher, R. (1984). Assessment of fear of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire. *Journal of Consulting and Clinical Psychology, 52*(6), 1090.
27. Chambless, D. L., Caputo, G. C., Bright, P., & Gallagher, R. (1984). Assessment of fear of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire. *Journal of Consulting and Clinical Psychology, 52*(6), 1090.
28. Chambless, D. L., Caputo, G. C., Jasin, S. E., Gracely. E. J., & Williams, C. (1985). The mobility inventory for agoraphobia. *Behavior Research and Therapy, 23*, 305-310.
29. Grabe, H. J., Schulz, A., Schmidt, C. O., Appel, K., Driessen, M., Wingenfeld, K., Barnow, S., Spitzer, C., John, U., Berger, K., Wersching, H., & Freyberger, H. J. (2012). Ein Screeninginstrument für Missbrauch und Vernachlässigung in der Kindheit: der Childhood Trauma Screener (CTS). *Psychiatrische Praxis*, *39*, 109-115.
30. Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry*, *3*(4), 257-260.
31. Liebowitz, M. R. (1987). Social phobia. *Modern Problems in Pharmacopsychiatry, 22*, 14-173.
32. Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., Waltz, T., & Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*(4), 676-688.
33. Kliem, S., Job, A. K., Kröger, C., Bodenmann, G., Stöbel-Richter, Y., Hahlweg, K., & Brähler, E. (2012). Entwicklung und Normierung einer Kurzform des Partnerschaftsfragebogens (PFB-K) an einer repräsentativen deutschen Stichprobe. *Zeitschrift für Klinische Psychologie und Psychotherapie, 41*(2), 81-89.
34. Rabin, R. & de Charro, F. (2001). EQ5D: A measure of health status from the EuroQol Group. *Annals of Medicine, 33,* 337-343.

35. Ladwig, I., Rief, W. & Nestoriuc, Y. (in press). Hat Psychotherapie auch Nebenwirkungen? – Entwicklung des Inventars zur Erfassung Negativer Effekte von Psychotherapie (INEP). *Verhaltenstherapie*.
36. Brockmann, J., Kirsch, H., Hatcher, R., Andreas, S., Benz, S. & Sammet, I. (2011). Dimensionen der therapeutischen Beziehung aus Patienten-Perspektive - Entwicklung der "Skala Therapeutische Allianz-Revised STA-R". *Psychotherapie, Psychosomatik, Medizinische Psychologie, 61*, 208-215.
37. Baldwin, D. S., Rafiq, R. & Talat, B. (2013). Overcoming barriers in the pharmacological treatment of anxiety disorders. *Neuropsychiatry, 3,* 33-44.
38. Lohr, J.M., Olatunji, B.O. & Sawchuck, C.N. (2007). A functional analysis of danger and safety signals in anxiety disorders. *Clinical Psychology Reviews, 27*, 114-126.
39. Davis, M., Walker, D.L., Miles, L. & Grillon, C. (2010). Phasic vs. sustained fear in rats and humans: Role if the extended amygdala in fear vs. anxiety. *Neuropsychopharmacology, 35,* 105-135.
40. Hariri, A. R., Mattay, V. S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D. et al. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science, 297*, 400-403.
41. Sylvester, C. M., Corbetta, M., Raichle, M. E., Rodebaugh, T. L., Schlaggar, B. L., Sheline, Y. I. et al. (2012). Functional network dysfunction in anxiety and anxiety disorders. *Trends in Neurosciences, 35*, 527-535.
42. Lueken, U., Muehlhan, M., Wittchen, H.-U., Kellermann, T., Reinhardt, I., Konrad, C. et al. (2011). (Don't) panic in the scanner! How panic patients with agoraphobia experience a functional magnetic resonance imaging session. *European Neuropsychopharmacoly, 21*, 516-525.
43. Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*, 97-113.
44. World Health Organization: Regional Office for Europe (1998). *Well-Being measures in primary health care: The DepCare Project.* Consensus meeting, Stockholm.