

Treatment sensitivity of implicit threat evaluation, avoidance tendency and visual working memory bias in specific phobia

Andrea Reinecke^{a,*}, Christian Soltau^b, Jürgen Hoyer^c, Eni S. Becker^d, Mike Rinck^d

^a University of Oxford, Oxford, UK

^b Universitaet Bremen, Bremen, Germany

^c Technische Universitaet Dresden, Dresden, Germany

^d Radboud University Nijmegen, Nijmegen, The Netherlands

ARTICLE INFO

Article history:

Received 20 July 2009

Received in revised form 3 December 2011

Accepted 18 December 2011

Keywords:

Attention

Bias

Cognitive-behaviour therapy

Implicit association

Phobia

Working memory

ABSTRACT

Cognitive theories of anxiety postulate that negative processing biases play a causal role in the pathogenesis of a disorder, while a normalisation of bias drives recovery. To test these assumptions it is essential to investigate whether biases seen in anxiety are treatment-sensitive, or whether they instead represent enduring vulnerability factors. Twenty-nine spider fearfuls were tested before and after brief cognitive-behaviour therapy (CBT), with half of them additionally being tested before a waiting period to control for retest effects. Using three cognitive bias tasks, we measured implicit threat evaluation (Extrinsic Affective Simon Task), avoidance tendency (Approach-Avoidance Task), and working memory for threat. CBT significantly enhanced negative implicit evaluation and avoidance. This indicates that these cognitive biases are no stable risk factors and provides further evidence for their potential key role in the development and remission of anxiety.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Cognitive theories of anxiety consider biases in automatic emotional processing as key agents in the pathogenesis and remission of anxiety disorders (e.g. Foa, Huppert, & Cahill, 2006; Williams, Watts, MacLeod, & Mathews, 1997). This assumption is supported by a large number of studies showing an association between heightened states of anxiety and bias, using paradigms that tap into different aspects of information processing: compared to non-anxious controls, anxious individuals show threat-favouring processing biases in attention (for a review, see Mathews & MacLeod, 2005) and visual working memory (VWM; Reinecke, Becker, & Rinck, 2009; Reinecke, Rinck, & Becker, 2006; Reinecke, Rinck, & Becker, 2008), more negative implicit evaluation of fear material (Huijding & de Jong, 2009; Reinecke, Becker, Hoyer, & Rinck, 2010; Teachman, Marker, & Smith-Janik, 2008), and stronger avoidance tendencies in reaction-time based approach-avoidance tasks (Reinecke, Becker, Hoyer, & Rinck, 2010).

Visual working memory bias has recently been studied using a cueing paradigm that allows the experimental manipulation of attention and working memory load (Reinecke et al., 2009; Reinecke et al., 2006). Originally developed in cognitive psychology,

this task has been shown to reliably measure VWM storage capacity and serial position effects for neutral stimuli (Wolfe, Reinecke, & Brawn, 2006). However, in the fear bias version of the task (Reinecke et al., 2006), one of the images shows a spider, while all other images are non-threatening. Participants are asked to memorise a subset of images out of a multiple-picture display, and these images are loaded into memory by rapidly being cued one after another. Results show that spider fear is associated with enhanced visual working memory for spider images compared to non-anxious controls. Interestingly, this effect not only occurred for cued spider pictures but also for uncued spiders visible within the display, suggesting that threat stimuli are automatically monitored in VWM (Reinecke et al., 2006).

Furthermore, fear bias has been reliably demonstrated using implicit evaluation or association tasks such as the Extrinsic Affective Simon Task (EAST; De Houwer, 2003). The theoretical concept underlying *implicit evaluation* is that fear is encoded in memory as a network of fear-related situations, fear triggers, and fear responses (Foa et al., 2006; Lang, Cuthbert, & Bradley, 1998). It is assumed that anxiety disorders are underpinned by particularly elaborate or strong connections within this fear network. In the EAST, participants first categorise single valence words as positive or negative, thus implicitly giving a positive versus negative meaning to two response keys. In the main experiment, the same two keys are used to categorise anxiety-relevant stimuli, for instance images of spiders, with respect to a fear-irrelevant dimension, such as the gaze direction of the animal. The implicit fear value of an image can

* Corresponding author at: University of Oxford, Department of Psychiatry, Oxford OX3 7JX, UK.

E-mail address: andrea.reinecke@psych.ox.ac.uk (A. Reinecke).

be measured by assessing the difference in reaction time with the positive versus with the negative key, as a reaction is faster if the valence of the target matches the valence of the response key (De Houwer, 2003). Recent experimental research has confirmed that compared to healthy controls, patients with anxiety disorders show more negative implicit evaluations of threat stimuli (generalised anxiety disorder: Reinecke, Becker, Hoyer, & Rinck, 2010; panic disorder: Teachman et al., 2008; spider phobia: De Jong, van den Hout, Rietbroek, & Huijding, 2003).

The association of anxiety with processing biases has also been established using an approach-avoidance-task (AAT), which specifically measures *threat avoidance tendencies*. Similarly to implicit evaluation tasks, this paradigm is based on the concept of compatibility versus incompatibility between target stimulus and response. However, this task taps into the behavioural rather than the cognitive component of anxiety (Lang, 1994): in a study addressing bias in spider phobia, participants were required to pull pictures of spiders or butterflies with a joystick whenever a landscape format image was presented, and to push images that were presented in portrait format, or vice versa. To visually support the response movement, pictures shrank in size when pushed and grew when pulled. Participants experience the pushing response as pushing the stimulus away (avoidance), while pulling is experienced as pulling the stimulus closer towards themselves (approach), making it more difficult for spider anxious participants to pull spider images towards them than to push them away (Reinecke, Becker, & Rinck, 2010; Rinck & Becker, 2007).

Taken together, such results to some degree do support cognitive models of anxiety which assume that changes in cognitive bias drive the development of and recovery from a disorder (e.g. Williams et al., 1997). However, while these comparative, cross-sectional studies allow conclusions as to whether anxiety is correlated with a certain bias, they leave open whether fluctuations in bias are in fact associated with fluctuations in anxiety states, or whether the bias instead represents an enduring vulnerability factor. To further test current theoretical assumptions regarding the functional relationship between cognitive bias and pathological states, it is essential to investigate whether the bias is susceptible to treatment. To date, the few studies that have investigated the impact of CBT on emotional information processing provide mixed evidence. While some studies do report treatment sensitivity of attention bias (generalised anxiety disorder: Mogg, Bradley, Millar, & White, 1995; spider phobia: Van den Hout, Tenney, Huygens, & de Jong, 1997) and implicit evaluation bias (spider phobia: Teachman & Woody, 2003; panic disorder: Teachman et al., 2008), no patient waiting groups had been included in these study designs, making it difficult to disentangle real treatment effects from mere test–retest effects. Another study including such a patient waiting group failed to find an effect of a single-block session of CBT on implicit bias over and above retest effects (Huijding & de Jong, 2009). Experimental post-assessment, however, took place immediately after a single session of treatment, although bias change might require a more thorough consolidation of treatment effects. Furthermore, research has not addressed yet whether visual working memory biases and automatic avoidance tendencies that have recently been associated with anxiety (Reinecke et al., 2006; Rinck & Becker, 2007) are sensitive to treatment, leaving it open what role these biases play in the development of a disorder.

The present study investigated the impact of brief CBT (Öst, 1996) on a range of emotional processing parameters in a sample of patients with spider anxiety, while taking the methodological limitations described above into account. Cognitive bias was measured using the EAST and AAT (Reinecke, Becker, & Rinck, 2010) previously shown to reliably assess implicit threat associations and automatic threat avoidance tendencies in spider anxiety, and the previously developed visual working memory task (VWMT;

Reinecke et al., 2006) to assess monitoring of threat in working memory. All participants were tested three times, with two weeks between assessments. However, they were randomly allocated to a treatment group versus a waiting group: the waiting group received CBT after the second assessment to control for test–retest effects, while the treatment group received treatment between the first two assessments. The following hypotheses, derived from the assumptions made by cognitive models of anxiety (Foa et al., 2006; Williams et al., 1997), were tested: (i) CBT affects self-report, behavioural and cognitive bias measures of anxiety over and above mere practise effects, resulting in stronger reduction on these measures in the treatment group than the waiting group from the first to the second test, (ii) compared to the treatment group, bias reduction will be stronger in waiting group patients, who by this point have received CBT as well, between the second to the third test.

2. Materials and methods

2.1. Participants

Twenty-nine spider-fearful participants were recruited into the study through advertisements in Dresden University of Technology lectures and local newspapers. Exclusion criteria were panic disorder, depression, psychosis, and alcohol or drug abuse. Participants were screened for mental disorders, using the International Diagnosis Checklist for DSM-IV (ICDL; Hiller, Zaudig, & Mombour, 1997). Screenings were applied by trained interviewers, and diagnostic decisions were supervised (ESB) based on written records. To be eligible for the study, participants had to at least fulfil the DSM-IV criteria A to D for specific phobia. As it is fairly easy to avoid spiders in Northern Europe, criterion E, which requires significant impairment in everyday life, was not required to be fulfilled. A specific phobia was diagnosed for 15 of the participants, who were equally represented in the two groups (TG: 6/14, WG: 9/15, $\chi^2(1) = 1.29$, $p = .256$).

2.2. Materials and apparatus

2.2.1. Self-report questionnaires

Subjective spider anxiety was measured using the Fear of Spiders Questionnaire (FSQ; Szymanski & O'Donohue, 1995, German version: Rinck et al., 2002). In addition, the Body Sensations Questionnaire (BSQ; Chambless, Caputo, Bright, & Gallagher, 1984; German version: Ehlers, Margraf, & Chambless, 1993) was used to assess fear of physical symptoms commonly associated with anxiety.

2.2.2. Behavioural challenge tests (BT)

Passive BT. Heartbeat per minute was measured with an ear clip device during four experimental phases: (1) during a baseline phase while completing the Beck Depression Inventory (BDI; Beck & Steer, 1987), (2) during an anticipatory phase while looking at the picture of a tarantula, after having been instructed about the imminent confrontation, (3) during an exposure phase in which the experimenter quickly approached the participant with the carapace of a tarantula, and (4) during a rebound phase while completing the BSQ. Outcome measures were pulse scores during anticipation, exposure and rebound, minus baseline pulse.

Active BT. Patients were instructed to enter a room where a cage with a living tarantula was kept on the windowsill, and they were asked to approach the cage as quickly and as closely as they felt comfortable with. Outcome parameters were speed in approaching the tarantula and final distance to the cage in cm. Recent research suggests high 1-week test–retest reliability of this test ($r = .84$, $p < .001$; Reinecke, Becker, & Rinck, 2010).

2.2.3. Experimental tasks

All stimuli were presented on a 17" screen with a resolution of 1024 × 768 pixels. Throughout all tasks, participants were instructed to react as quickly and accurately as possible. Breaks were possible whenever requested. Different stimuli were used for each task, and all task stimuli had been evaluated in previous studies (EAST/AAT: Reinecke, Becker, & Rinck, 2010; Rinck & Becker, 2007; VWMT: Reinecke et al., 2009; Reinecke et al., 2006). The psychometric properties of the EAST and the AAT have recently been characterised (Reinecke, Becker, & Rinck, 2010), and results suggest that both instruments are sufficiently reliable (1-week test–retest reliability EAST: $r = .42$, $p < .001$; AAT: $r = .35$, $p < .01$; internal consistency: EAST: Cronbach's $\alpha = .44$, AAT: Cronbach's $\alpha = .66$) and valid (correlation with FSQ EAST: $r = -.32$, $p < .01$, AAT: $r = -.59$, $p < .001$). In addition, the VWMT effects have proven to be highly replicable (Reinecke et al., 2009; Reinecke et al., 2006) and reliable (2-weeks test–retest reliability, assessed by correlating WG effects achieved during the first and second test in the current study: cued VWM bias: $r = .74$, $p < .01$; uncued VWM bias: $r = .59$, $p < .05$).

Extrinsic affective Simon task (EAST). Verbal stimuli were 10 pleasant (e.g. happiness, pleasure) and 10 unpleasant words (e.g. fear, dangerous) in font size 24. Picture stimuli were 5 spider and 5 butterfly photographs (300 × 400 pixels), each once in its original and once in mirrored form, such that the animal's head once pointed to the right and once to the left side. The experiment consisted of 2 practice blocks and 5 experimental blocks, with error feedback being provided. In the *valence practice block*, participants categorised words as either pleasant or unpleasant by pressing a left or a right key. Thereby, the two response keys received an extrinsic meaning as the positive versus the negative key. Half of the participants in each group used the left key for pleasant words and the right key for negative words, the other half responded the other way around. Each word was randomly presented 4 times, yielding a total of 80 valence practice trials. In the 20 trials of the *target practice block*, 5 photographs of neutral dragonflies were presented, each twice in the original form and twice in the mirrored form. Participants classified the animals' gaze direction as either left (left key) or right (right key). In each of each of the five *experimental blocks*, 40 valence words (each word presented twice) and 40 target trials (type of animal and gaze direction counterbalanced) were presented in pseudorandom order. Participants had to categorise words with respect to their valence and pictures with respect to gaze direction, using the same keys for both tasks. Thus, there were trials where participants had to react with the unpleasant key to spider pictures and with the pleasant key to butterfly pictures (compatible trials), but there were also trials where a spider image required a reaction with the pleasant key and a butterfly image required an "unpleasant"-reaction (incompatible trials). A negative implicit evaluation of spiders is reflected in delayed reactions in incompatible trials where the gaze direction of a spider has to be indicated with the "pleasant"-key.

Approach-avoidance task (AAT). Stimuli were 8 photographs of spiders and 8 photographs of butterflies, each once in portrait format (400 × 300 pixels) and once in landscape format (300 × 400 pixels). Control stimuli were rectangular outlines the size of the photographs. Responses were given using a standard joystick, fastened to the table. Half of the participants in each group were instructed to pull the joystick in response to landscape format pictures and to push in response to portrait pictures. The other half of participants received reverse instructions. Images gradually shrank when the joystick was pushed, and gradually grew in size when the joystick was pulled. In 18 practice trials, participants responded to rectangular outlines in landscape and portrait format. For individual calibration of joystick reactions, each of two experimental blocks also started and ended with the presentation

of 32 control trials. The remaining 160 trials per block consisted of 64 spider trials, 64 butterfly trials, and 32 control trials. Half of the trials per condition were presented in portrait versus landscape format, therefore requiring a compatible (pulling butterflies, pushing spiders) versus incompatible response (pushing butterflies, pulling spiders) in half of the trials each. The order of trials was pseudorandom.

Visual-working-memory task (VWMT). Preceded by 6 practice trials, participants performed 420 experimental trials. Stimuli were 36 colour object images (115 × 115 pixels), including 35 neutral or pleasant images and one negatively valenced spider image. Each trial started with the presentation of a fixation cross for 500 ms. Then, 16 pictures, randomly chosen from the pool of 36 images, were placed on the positions of an invisible 4 × 4 grid. After 150 ms, 5 of the 16 pictures were cued one after another for 150 ms each, by illuminating the background of the image. This was followed by one the 16 items on the screen being masked by a dark grey square for 150 ms. Immediately afterwards, the stimuli display was replaced by a grey blank screen, presented for 1000 ms, followed by a response menu including all 36 images. Participants had to report the masked image by mouse-clicking on the corresponding picture within 3 s. The masked item was one of the cued items in 85% of trials, and one of the remaining 11 uncued items in 15% of the trials. A full replication of all working memory effects explored recently (Reinecke et al., 2006) is beyond the scope of this paper, and we therefore restrict the analyses presented here to those relevant to the specific hypotheses. The most important experimental variations include (a) whether memory for a spider or non-spider was tested, and (b) whether this spider item had been among the 5 cued (*cued VWM bias*) or among the 11 uncued items (*uncued VWM bias*). For each of these combinations, participants' mean accuracy in reporting the masked item was calculated.

2.3. Treatment

Participants received brief exposure-based CBT (Öst, 1996), which is assumed to be the most effective intervention for specific phobias (e.g. Chambless & Gillis, 1993; Craske & Rowe, 1997), with clinically significant treatment gains in 71–80% of patients (Öst, 1996) and very low drop-out rates (Hellstroem & Öst, 1996). The treatment was delivered by an experienced psychologist (CS) who was continuously supervised by a licensed clinical psychologist (ESB) with particular experience in the treatment of anxiety disorders. The treatment consisted of three sessions and was delivered in groups of 4 or 5 participants. During the first session, the treatment rationale was explained. Two to three days later, patients participated in gradual exposure exercises for about 3 h. At the end of the session, participants were encouraged to plan and realise self-exposure exercises until the next session one week later. During this last session, homework was discussed and participants were given an additional opportunity to practice (1 h).

2.4. General procedure

Informed consent was taken at the beginning of the study. Each participant was invited three times for data collection, with 2 weeks between sessions. Participants were randomly assigned to one of two treatment groups in a multiple-baseline design. Participants in the *waiting group (WG)* received treatment between the second and the third test session, while participants in the *treatment group (TG)* received treatment between the first and the second test. Post-treatment assessments took place within 1–3 days after training. During each session, participants completed the three experimental tasks AAT, EAST, and VWMT after having filled out the FSQ. Afterwards, they underwent the two behavioural challenge tests. Each session lasted about 100 min.

Table 1
Questionnaire scores (means, standard deviations), active BT speed and distance, and passive BT pulse change scores (with baseline pulse having been subtracted).

	Treatment group			Waiting group		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
FSQ	65.7 (22.6)	24.6 (16.5)	20.7 (19.0)	55.3 (17.0)	50.7 (17.4)	21.1 (11.4)
BSQ	2.3 (0.8)	1.2 (0.3)	1.1 (0.1)	1.9 (0.7)	1.5 (0.6)	1.2 (0.5)
Passive BT						
Anticipation	10.5 (13.0)	4.8 (4.5)	4.8 (9.3)	2.5 (7.3)	8.4 (7.5)	3.3 (6.2)
Confrontation	11.4 (14.8)	6.3 (4.7)	6.3 (5.6)	3.4 (8.4)	7.0 (11.6)	6.2 (5.1)
Rebound	8.6 (13.5)	4.4 (5.2)	6.6 (9.1)	1.9 (6.5)	5.6 (6.9)	5.6 (5.3)
Active BT						
Speed (m/s)	0.29 (0.19)	0.56 (0.27)	0.69 (0.28)	0.33 (0.18)	0.43 (0.25)	0.61 (0.22)
Distance (cm)	33.4 (52.0)	6.1 (6.8)	6.6 (7.4)	24.3 (16.5)	12.8 (13.1)	6.3 (7.4)

Note: The treatment group received CBT between tests 1 and 2, while the waiting group received CBT between tests 2 and 3.

3. Results

To establish the impact of CBT on self-report questionnaire scores, BT and experimental task scores, 2 (group) \times 3 (test time) mixed ANOVAs were run, separately for each parameter. To avoid inflation of Type I error rate, significant interactions were broken down using planned repeated contrasts, comparing outcome scores at test times 1 versus 2, and outcome scores at test times 2 versus 3, across TG and WG participants (Field, 2009). As the study was based on precise a-priori one-sided hypotheses, all these analyses were run one-tailed. In addition, to ensure that groups were not different from each other on any of the parameters at baseline already, independent-samples *t*-tests for baseline scores were run. All effect sizes are reported as Pearson's *r*.

3.1. Sociodemographic characteristics

The two groups were matched with regard to age (TG: $M = 29.1$, $SD = 11.6$; WG: $M = 33.5$, $SD = 13.1$; $t(27) = 0.94$, $p = .355$), gender (TG: 13 ♀, 1 ♂; WG: 13 ♀, 2 ♂; $\chi^2(1) = 0.30$, $p = .584$) and educational level ($\chi^2(4) = 2.26$, $p = .688$).

3.2. Self-report measures

At baseline, there were no group differences in *spider fear* (FSQ) or *fear of physical sensations* during confrontation (BSQ),¹ both $t(27) < 1.41$ both $p > .170$. Mixed ANOVAs (Table 1) revealed significant interaction effects for both questionnaires, FSQ: $F(2,54) = 28.13$, $p < .001$, $r = .59$, BSQ: $F(2,54) = 4.53$, $p < .01$, $r = .28$. The follow-up contrasts comparing TG and WG scores across the test times 1 and 2 revealed a significant interaction for both measures, both $F(1,27) > 6.57$, both $p < .01$, both $r > .44$. This suggests that while spider fear and physical sensations did not change in WG between the first two assessments, they were significantly decreased in TG after having received treatment. Comparing the groups across test times 2 and 3 shows a significant interaction for the FSQ, $F(1,27) = 38.07$, $p < .001$, $r = .76$, and a non-significant trend-interaction for the BSQ, $F(1,27) = 2.29$, $p = .072$, $r = .28$. This indicates that during test times 2 and 3 WG showed stronger improvement in spider fear than TG, after having received treatment now as well.

3.3. Behavioural challenge tests (BT)

3.3.1. Passive BT

At baseline, there were no group differences in pulse scores for the *exposure* and *rebound* phases, both $t(27) < 1.45$, both $p > .158$,

both $r < .27$, but TG showed higher pulse scores during the *anticipation* phase than WG, $t(27) = 2.06$, $p = .049$, $r = .37$. However, a group difference in this direction does not compromise the results presented here. While mixed group \times time ANOVAs (Table 1) only showed non-significant trend interactions for exposure pulse scores, $F(2,54) = 2.36$, $p = .053$, $r = .20$, and rebound pulse scores, $F(2,54) = 2.23$, $p = .059$, $r = .20$, we found a significant interaction effect for anticipation phase pulse scores, $F(2,54) = 4.50$, $p < .01$, $r = .28$. Breaking this interaction down using repeated contrasts revealed that while anticipation-phase pulse scores did not change in the WG between tests 1 and 2, they significantly decreased in TG, $F(1,27) = 11.16$, $p < .01$, $r = .54$. The contrast comparing changes between tests 2 and 3 across both groups only indicated a non-significant trend interaction, $F(1,27) = 2.82$, $p = .053$, $r = .31$, with mean scores suggesting that WG patients experienced a statistically non-significant improvement in anticipation pulse score after CBT as well.

3.3.2. Active BT

At baseline, the groups were not different from each other regarding their *speed* in approaching the spider cage and the *distance* they kept, both $t(27) < 0.70$, both $p > .489$, both $r < .13$. Mixed group \times time ANOVAs (Table 1) revealed no interaction for distance, $F(2,54) = 0.83$, $p = .441$, $r = .12$, but a significant interaction for speed, $F(2,54) = 3.40$, $p < .05$, $r = .24$. Follow-up repeated contrasts indicate that the increase in speed was significantly stronger in TG than WG between tests 1 and 2, $F(1,27) = 5.31$, $p < .05$, $r = .41$. However, for the change in speed between tests 2 and 3, no such difference was found, $F(1,27) = 0.17$, $p = .342$, $r = .08$.

3.4. Experimental tasks

3.4.1. Extrinsic affective Simon task (EAST)

Due to technical problems, complete EAST data were available for only 13 TG and 13 WG participants. Analyses reported here refer to these participants. Error rates were uniformly low and did not differ between experimental conditions (averaging 3%). Only correct trials were included in the data analysis. As the butterfly trials were only included to present positive pictures in addition to the negative spider pictures, their analysis is not reported here.² For each participant and each test time, median RT to spiders with the *unpleasant key* (compatible reaction) versus with the *pleasant key* (incompatible response) were calculated. EAST effects were computed by subtracting RT with the pleasant key from RT with the unpleasant key (Table 2 and Fig. 1). A negative score indicates a negative evaluation; a positive score suggests a positive evaluation.

Implicit spider evaluations were not different between the two groups at baseline, $t(24) = 0.15$, $p = .881$, $r = .03$. The

¹ BDI results are not reported here, as this measure only served as a filler task during the Pulse BT. However, the groups showed no difference in baseline BDI scores, TG: $M = 5.7$, $SD = 5.5$, WG: $M = 5.3$, $SD = 3.0$, $t(27) = 0.28$, $p = .785$.

² As expected, the groups did not differ for these pictures at any of the test times.

Table 2
EAST effects, AAT effects, and visual working memory (VWM) bias for cued and uncued spiders (means, standard deviations) in the treatment versus waiting group.

	Treatment group			Waiting group		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
EAST	–45 (26)	–8 (14)	–3 (18)	–47 (38)	–43 (48)	–14 (30)
AAT	–68 (105)	41 (74)	–1 (60)	–76 (99)	–71 (116)	–8 (50)
Cued VWM bias	.46 (.12)	.40 (.13)	.34 (.11)	.48 (.19)	.47 (.15)	.41 (.15)
Uncued VWM bias	.56 (.30)	.38 (.27)	.53 (.16)	.49 (.24)	.51 (.35)	.40 (.32)

Note: The Treatment Group received CBT between Test 1 and Test 2, while the Waiting Group received CBT between Tests 2 and 3.

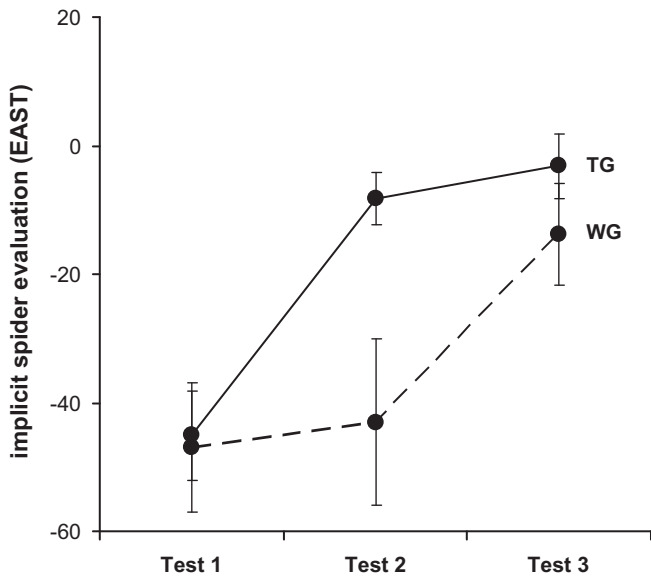


Fig. 1. Mean EAST effects for spiders, calculated as the RT in responding to spiders with the unpleasant key minus the RT in responding to spiders with the pleasant key. Error bars show the standard error of the mean.

mixed group \times time ANOVA showed a significant interaction, $F(2,48) = 2.68, p < .05, r = .23$. The follow-up contrast comparing TG and WG scores across the test times 1 and 2 revealed a significant interaction, $F(1,24) = 5.08, p < .05, r = .42$, suggesting that while implicit threat evaluation was not altered in WG, it had become significantly less negative in TG after having received CBT. Comparing the groups across test times 2 and 3 also shows a significant interaction, $F(1,24) = 3.23, p < .05, r = .34$, indicating that while during test 2 and 3 implicit evaluation scores had remained stable in TG, they had become significantly less negative in WG after having now received CBT as well.

3.4.2. Approach-avoidance task (AAT)

Due to technical problems, complete AAT data only existed for 11 TG and 14 WG participants. Analyses reported here are based on these samples. The uniformly low error rates did not differ between experimental conditions (averaging 1%). These incorrect trials were excluded from data analysis. As the positive butterfly pictures were of no relevance to the hypotheses, these analyses are not reported here.³ For each participant and each test time, threat avoidance scores were calculated by subtracting the median RT for pulling spider pictures from the median RT for pushing spider pictures, before subtracting AAT effects for control trials (pushing minus pulling) from these values to correct for general individual joystick response tendencies (Table 2 and Fig. 2). A negative AAT score indicates an avoidance reaction to spiders, while a positive score indicates an approach reaction.

Avoidance scores were not different between groups at baseline, $t(23) = 0.20, p = .844, r = .04$. We found a significant group \times time mixed ANOVA (Table 2 and Fig. 2) interaction, $F(2,46) = 3.09, p < .05, r = .25$, with follow-up contrasts suggesting that while TG showed a significant stronger reduction in threat avoidance between tests 1 and 2, $F(1,23) = 3.14, p < .05, r = .35$, WG showed a more significant reduction than TG between tests 2 and 3, $F(1,23) = 6.96, p < .01, r = .47$.

3.4.3. Visual-working-memory task (VWMT)

Due to technical problems, complete data were only available for 13 TG and 13 WG participants. Analyses reported here refer to these participants. Cued VWM bias was calculated by subtracting mean memory accuracy for cued non-spider items in spider-free displays from mean accuracy for cued spider items, averaging over the position of the tested items within the cue string. The uncued VWM bias was calculated by subtracting mean memory accuracy for uncued images in non-spider displays from mean memory accuracy for uncued spider images (Table 2 and Fig. 3).

Cued and uncued VWM bias scores at baseline were not different in the two groups, both $t(24) < 0.71$, both $p > .485$, both $r < .14$. The group \times time mixed ANOVA for cued VWM bias revealed no significant interaction and therefore no CBT effect, $F(2,48) = 0.35, p = .292, r = .11$. There was also only a non-significant trend interaction for the uncued VWM bias, $F(2,48) = 1.99, p = .07, r = .20$. Exploratory follow-up contrasts suggest that while TG participants show a particularly strong reduction of uncued VWM bias between tests 1 and 2, after having received CBT, $F(1,24) = 3.04, p < .05, r = .34$, WG

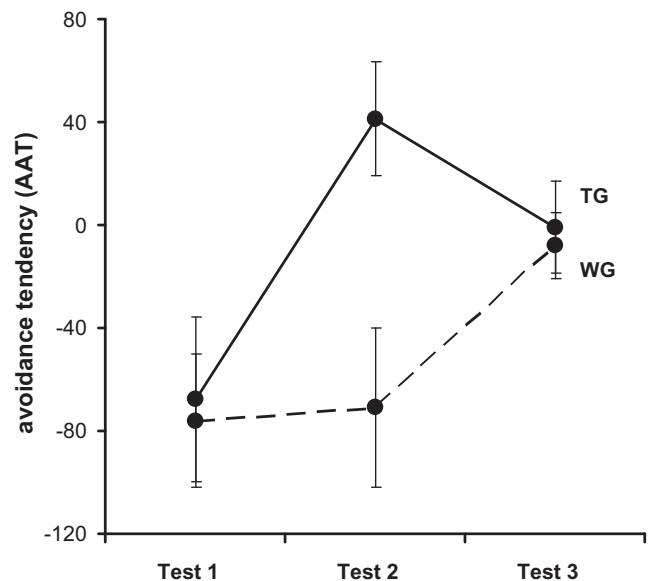


Fig. 2. Mean AAT effects for spiders, calculated as the RT for pushing spider pictures minus the RT for pulling spiders, corrected for the RT to control trials. Error bars show the standard error of the mean.

³ The groups did not differ for these pictures at any of the test times.

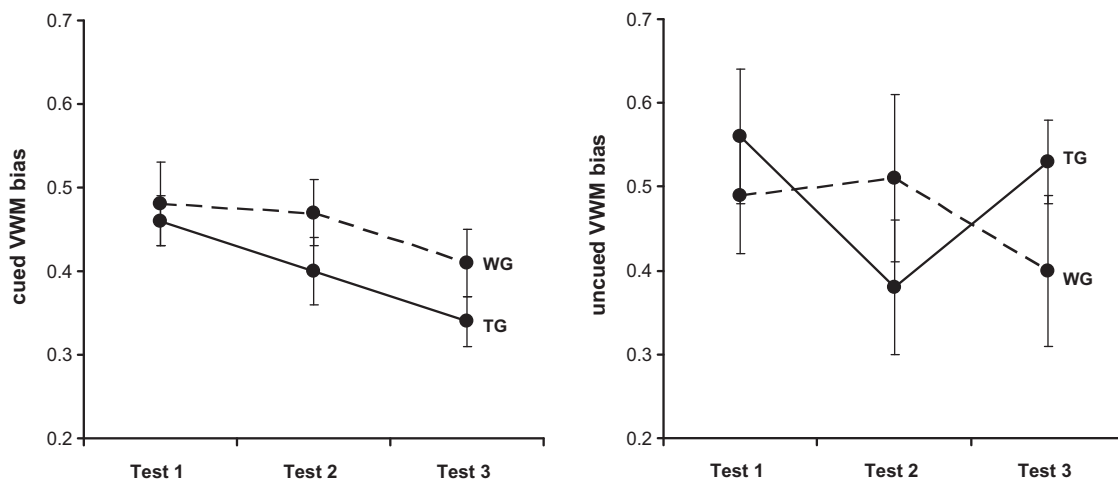


Fig. 3. Cued VWM bias, calculated as the mean accuracy in memorizing cued spiders minus the mean accuracy in memorizing cued non-spiders (left). Uncued VWM bias, calculated as the mean accuracy in memorizing uncued spiders minus the mean accuracy in memorizing uncued non-spiders (right). Error bars show the standard error of the mean.

show this treatment effect as expected between tests 2 and 3, $F(1,24) = 3.03$, $p < .05$, $r = .33$.

3.4.4. Correlation analyses

Prior to correlation analyses, residual gain scores were computed for the self-report measures (FSQ, BSQ), the BT parameters (speed, distance), and for the bias parameters (EAST, AAT, cued VWM bias, uncued VWM bias). These scores reflect the deviation of observed post-treatment scores from regression-based predicted outcome scores and have been suggested to be a more reliable index of change than simple pre-post difference scores (Senn, 1997). For TG patients, analyses included baseline and test 2 scores, for WG participants, analyses included scores at baseline and test 3. Pearson's correlation analyses were then run for the calculated residual gain scores to assess (i) the validity of the experimental tasks in reflecting therapeutic change, exploring whether improvement in cognitive bias parameters was associated with improvement in self-reported symptom severity (FSQ, BSQ) and improvement in approach behaviour (speed, distance), and (ii) whether there were any associations between changes on the different bias measures (Table 3). Exploring whether changes on different parameters of cognitive bias intercorrelated, only the cued and uncued VWM bias changes were associated, $r = .66$, $p < .001$, all other $r < .18$, all other $p > .409$. Regarding correlations of changes in cognitive bias parameters with self-reported symptoms, only the AAT was significantly correlated with the FSQ, suggesting that reductions in RT-based threat avoidance are associated with a reduction in spider fear,

Table 3
Intercorrelations of treatment-induced changes (residual gain scores) in cognitive bias, self-reported symptom severity and behavioural challenge parameters.

	EAST	AAT	Cued VWM	Uncued VWM
Bias parameters				
EAST	1	–	–	–
AAT	.18	1	–	–
Cued VWM	.01	.01	1	–
Uncued VWM	.01	.16	.66***	1
Self-report measures				
FSQ	–.01	–.44*	.08	–.14
BSQ	–.12	–.24	–.16	.14
BT parameters				
Speed	.18	.17	.23	.35
Distance	–.10	.07	–.12	.11

* .05 level of significance.

*** .001 level of significance.

$r = -.44$, $p < .05$. None of the cognitive bias measures was associated with BT parameters, all $r < .35$, all $p > .102$.

4. Discussion

Our results confirm that brief CBT for spider phobia not only reduces self-reported spider fear (Öst, 1996), but that it also affects behavioural and cognitive-emotional parameters of anxiety. With medium to large effect sizes, CBT leads to significant reductions in negative implicit threat evaluation (EAST) and threat avoidance (AAT), over and above mere test–retest effects. While changes were particularly strong in treated patients compared to waiting group patients between the first two assessments, group effects were reversed between assessments 2 and 3, which is consistent with the respective application times of treatment in the two groups. Furthermore, this study provides some limited evidence for treatment sensitivity of visual working memory biases (VWMT): while we established no effect of treatment on VWM for cued threat items, there was some indication of VWM for uncued spiders changing as a consequence of CBT in each group, reflected in significant follow-up contrast analyses. However, the omnibus test preceding these comparisons across test times and groups had only reached non-significant trend level, therefore limiting the interpretation of these effects. Nevertheless, an exploratory comparison of the observed post-treatment bias parameters with our earlier work using identical or highly similar task designs suggests a reduction of bias to a healthy level in all three tasks (EAST: -11 versus -1 in Reinecke, Becker, Hoyer, & Rinck, 2010; Reinecke, Becker, & Rinck, 2010; AAT: 17 versus $15-51$ in Rinck & Becker, 2007; uncued VWM bias: $.39$ versus $.38$ in Reinecke et al., 2006).

Regarding behavioural correlates of spider anxiety, pulse increases during and after exposure with a tarantula carapace were not affected by CBT (with the group \times time ANOVAs minimally falling short of .05 level of significance). However, we found some evidence that treatment significantly affected pulse increase during anticipation of the imminent confrontation. While the contrast comparing pulse score changes between tests 2 and 3 across groups marginally missed statistical significance, the contrast analysis comparing group effects between tests 1 and 2 was significant, suggesting that during this period, treated patients showed a stronger pulse score decrease than waiting list participants. Furthermore, while the results gave no indication for a CBT effect on distance kept when approaching a spider, speed seemed to increase as a consequence of CBT. At least this was established for changes between

the first two tests, showing a stronger increase in speed in treated patients versus waiting list controls. The exploratory comparison of mean scores at tests 2 and 3 across the groups suggests a similar CBT effect, but this did not reach statistical significance.

Investigating intercorrelations between self-report, behavioural and cognitive-emotional parameters of treatment gain, only the AAT was significantly correlated with spider fear, suggesting that a reduction in subjective anxiety is associated with a reduction in threat avoidance tendencies.

This research provides strong evidence that biased implicit threat evaluations (EAST) and avoidance tendencies (AAT) normalise even during brief treatment, while the impact of CBT on visual working memory (VWMT) remains less clear. One reason for this discrepancy might be a slight under-powering of the study to find treatment effects in the VWMT task. However, an alternative explanation might lie in differences regarding to what degree threat stimuli need to be attended to for optimal task performance. While the VWMT instructions explicitly encourage paying attention to task stimuli to yield high accuracy performance, the EAST and AAT measure threat effects at a more incidental level, as participants react to stimuli based on threat-irrelevant dimensions, such as the format of the image. It is possible that tasks incorporating the threat value of a stimulus as a task-irrelevant rather than a task-relevant dimension are more powerful in capturing the impact of CBT. This would be in line with neuroimaging research showing that while both low and high anxious individuals respond to task-relevant threat stimuli with increased activity in the amygdala, which has been shown to play a key role in the acquisition and extinction of fear (Davis, 2002; LeDoux, 2000), only anxious participants also show amygdala hypersensitivity for task-irrelevant threat (Bishop, Duncan, & Lawrence, 2004). This explanation would also be consistent with previous studies reporting CBT induced bias changes, having used implicit evaluation tasks (Teachman et al., 2008; Teachman & Woody, 2003) or emotional stroop tasks, where participants name the colour of threat-related words (Mogg et al., 1995; Van den Hout et al., 1997).

In summary, our results showing reduced implicit threat evaluation and avoidance tendency following CBT are in line with cognitive models of anxiety which assume that biased information processing plays a key role in the aetiology and maintenance of a disorder, and that a normalisation of bias is an essential mechanism of treatment action (e.g. Williams et al., 1997). However, while our results indicate that these bias parameters do not represent enduring vulnerability markers for the development of a disorder, they do not allow conclusions regarding their causal role in anxiety reduction. A methodological caveat of such pre-post treatment studies is that by the end of treatment both processing bias and symptoms are resolved, making it impossible to disentangle these two effects and to draw conclusions regarding their causal relationship. As such, these results might indicate that a change in bias causes an improvement in anxiety, which would confirm the idea that information processing changes are relevant mechanisms of action in treatment. However, alternatively, the results might be interpreted in the reverse direction, in a way that cognitive bias is a side-product of altered anxiety states, disappearing once anxiety has been reduced. A causal relationship can be tested directly using cognitive bias modification (CBM) procedures, where the impact of experimentally induced bias reduction on anxiety symptoms is assessed. Recent CBM studies have successfully demonstrated that modifying biased attention (Amir, Beard, Burns, & Bomyea, 2009; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002), interpretation (Hayes, Hirsch, Krebs, & Mathews, 2010; Mathews & Mackintosh, 2000), or implicit fear evaluations (Clerkin & Teachman, 2010) towards a more positive direction results in reduced symptom severity in a range of anxiety disorders, supporting the idea that negative emotional processing plays

a causal in the development of anxiety and recovery. However, it remains open whether similar effects can be found in specific phobia, using the tasks presented here.

While the study presented here incorporates several significant strengths, including (i) the inclusion of a patient waiting group to evaluate CBT effects against pure retest effects, (ii) the use of a range of different measures of cognitive bias, and (iii) the measurement of symptom change not only in terms of self-reported anxiety but also through behavioural challenge tests, a few methodological limitations require highlighting as well. Most importantly, the very small sample sizes might have compromised the interpretation of the null-results seen for the VWM bias task and the correlations between different measures of change. Related to this problem of small sample sizes, it remains open to what degree pre-treatment bias, bias improvement during treatment, or bias relapse during follow-up have the potential to predict future symptom improvement and stability. Identifying such cognitive markers would be vital for predictive stratification decisions, such as whether a patient is likely to benefit from a certain type of treatment, or whether booster sessions might be required to ensure a stability of treatment effects. Moreover, the design of the pulse BT used here leaves room for improvement, in a way that the completion of a depression scale during baseline measurement might have lead to an unnecessary overestimation of baseline pulse scores. However, as all computed pulse outcome measures have relied on the same baseline value, this issue should not affect the direction of results presented here, but might want to be taken into account in future studies. In addition, it has not been recorded to what degree patients had exercised in self-management during the week between the exposure and follow-up treatment sessions, although this information might have provided valuable insights regarding the relationship between practise intensity and bias change.

Acknowledgement

Preparation of this paper was supported by a grant from the German Research Foundation (DFG) to Eni Becker and Mike Rinck.

References

- Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology, 118*, 28–33.
- Beck, A. T., & Steer, R. A. (1987). *Manual for the revised beck Depression Inventory*. San Antonio, TX: The Psychological Corporation.
- Bishop, S., Duncan, J., & Lawrence, A. D. (2004). State anxiety modulation of the amygdala response to unattended threat-related stimuli. *Journal of Neuroscience, 24*, 10364–10368.
- 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*, 1090–1097.
- Chambless, D. L., & Gillis, M. M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology, 61*, 248–260.
- Clerkin, E. M., & Teachman, B. A. (2010). Training implicit social anxiety associations: an experimental intervention. *Journal of Anxiety Disorders, 24*, 300–308.
- Craske, M., & Rowe, M. (1997). A comparison of cognitive and behavioral treatment of phobias. In: G. Davey (Ed.), *Phobias: description, treatment, and theory*. Wiley series on chemical psychology. New York: Wiley.
- Davis, M. (2002). Role of NMDA receptors and MAP kinase in the amygdala in extinction of fear: clinical implications for exposure therapy. *European Journal of Neuroscience, 16*, 395–398.
- De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology, 50*, 77–85.
- De Jong, P. J., van den Hout, M. A., Rietbroek, H., & Huijding, J. (2003). Dissociations between implicit and explicit attitudes toward phobic stimuli. *Cognition and Emotion, 17*, 521–545.
- Ehlers, A., Margraf, J., & Chambless, D. L. (1993). *Fragebogen zu koerperbezogenen Aengsten, Kognitionen und Vermeidung (AKV)*. Weinheim: Beltz.
- Field, A. (2009). *Discovering statistics using SPSS*. London: Sage.
- Foa, E. B., Huppert, J. D., & Cahill, S. P. (2006). Emotional processing theory: an update. In: B. O. Rothbaum (Ed.), *Pathological anxiety: emotional processing in etiology and treatment* (pp. 3–24). New York: Guilford Press.

- Hayes, S., Hirsch, C. R., Krebs, G., & Mathews, A. (2010). The effects of modifying interpretation bias on worry in generalized anxiety disorder. *Behaviour Research and Therapy*, 48, 171–178.
- Hellstroem, K., & Öst, L.-G. (1996). Prediction of outcome in the treatment of specific phobia. A cross validation study. *Behavior Research and Therapy*, 34, 403–411.
- Hiller, W., Zaudig, M., & Mombour, W. (1997). *IDCL – Internationale Diagnosen Checklisten für DSM-IV und ICD-10 (Manual, 31 Checklisten nach DSM-IV und Screening-Blatt)*. Hogrefe: Göttingen.
- Huijding, J., & de Jong, P. (2009). Implicit and explicit attitudes toward spiders: sensitivity to treatment and predictive value for generalization of treatment effects. *Cognitive Therapy and Research*, 33, 211–220.
- Lang, P. J. (1994). The varieties of emotional experience: a meditation on James–Lange–theory. *Psychological Review*, 101, 211–221.
- Lang, P. J., Cuthbert, B. N., & Bradley, M. M. (1998). Measuring emotion in therapy: imagery, activation, and feeling. *Behavior Therapy*, 29, 655–674.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, 23, 155–184.
- MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology*, 111, 107–123.
- Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology*, 109, 602–615.
- Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology*, 1, 167–195.
- Mogg, K., Bradley, B. P., Millar, N., & White, J. (1995). A follow-up study of cognitive bias in generalized anxiety disorder. *Behavior Research and Therapy*, 33, 927–935.
- Öst, L.-G. (1996). One-session group treatment of spider phobia. *Behaviour Research and Therapy*, 34, 707–715.
- Reinecke, A., Becker, E. S., Hoyer, J., & Rinck, M. (2010). Generalized implicit fear associations in generalized anxiety disorder. *Depression and Anxiety*, 27, 252–259.
- Reinecke, A., Becker, E. S., & Rinck, M. (2009). Selective visual working memory in fear of spiders: the role of automaticity and material-specificity. *Journal of Anxiety Disorders*, 23, 1053–1063.
- Reinecke, A., Becker, E. S., & Rinck, M. (2010). Test–retest reliability and validity of three indirect tests assessing implicit threat associations and behavioural response tendencies. *Journal of Psychology*, 218, 4–11.
- Reinecke, A., Rinck, M., & Becker, E. S. (2006). Spiders crawl easily through the bottleneck: visual working memory for negative stimuli. *Emotion*, 6, 438–449.
- Reinecke, A., Rinck, M., & Becker, E. S. (2008). How preferential is the preferential encoding of threatening stimuli? Working memory biases in specific anxiety and the attentional blink. *Journal of Anxiety Disorders*, 22, 655–670.
- Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Behaviour Research and Experimental Psychiatry*, 38, 105–120.
- Rinck, M., Bunschuh, S., Engler, S., Müller, A., Wissmann, J., Ellwart, T., et al. (2002). Reliability and validity of German versions of three instruments measuring fear of spiders. *Diagnostica*, 48, 141–149.
- Senn, S. (1997). *Statistical issues in drug development*. Chichester: Wiley.
- Szymanski, J., & O'Donohue, W. (1995). Fear of spiders questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 26, 31–34.
- Teachman, B. A., Marker, C. D., & Smith-Janik, S. B. (2008). Automatic associations and panic disorder: trajectories of change over the course of treatment. *Journal of Consulting and Clinical Psychology*, 76, 988–1002.
- Teachman, B. A., & Woody, S. R. (2003). Automatic Processing in spider phobia: implicit fear associations over the course of treatment. *Journal of Abnormal Psychology*, 112, 100–109.
- Van den Hout, M., Tenney, N., Huygens, K., & de Jong, P. (1997). Preconscious processing bias in specific phobia. *Behavior Research and Therapy*, 35, 29–34.
- Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1997). *Cognitive psychology and emotional disorders*. Chichester: John Wiley.
- Wolfe, J. M., Reinecke, A., & Brawn, P. (2006). Why don't we see changes? The role of attentional bottlenecks and limited visual memory. *Visual Cognition (Special Issue: Visual Search and Attention)*, 14, 1–32.