Behaviour Research and Therapy 96 (2017) 47–56

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

From avoidance to approach: The influence of threat-of-shock on reward-based decision making

Florian Bublatzky [a, *, 1], Georg W. Alpers [a], Andre Pittig [a, b, **, 1]

[a] Clinical Psychology and Biological Psychology and Psychotherapy, Department of Psychology, School of Social Sciences, University of Mannheim, Germany
[b] Institute of Clinical Psychology and Psychotherapy, Technische Universität Dresden, Dresden, Germany

A R T I C L E I N F O

Article history:
Received 29 September 2016
Received in revised form
2 January 2017
Accepted 6 January 2017
Available online 10 January 2017

Keywords:
Decision making
Threat-of-shock
Behavioral avoidance
Skin conductance responses
Heart rate responses
Fear learning

A B S T R A C T

Potential threat can prime defensive responding and avoidance behavior, which may result in the loss of rewards. When aversive consequences do not occur, avoidance should, thus, be quickly overcome in healthy individuals. This study examined the impact of threat anticipation on reward-based decisions. Sixty-five participants completed a decision-making task in which they had to choose between high- and low-reward options. To model an approach-avoidance conflict, the high-reward option was contingent with a threat-of-shock cue; the low-reward option was contingent with a safety cue. In control trials, decisions were made without threat/safety instructions. Overall, behavioral data documented a typical preference for the profitable option. Importantly, under threat-of-shock, participants initially avoided the profitable option (i.e., safe, but less profitable choices). However, when they experienced that shocks did actually not occur, participants overcame initial avoidance in favor of larger gains. Furthermore, autonomic arousal (skin conductance and heart rate responses) was elevated during threat cues compared to safety and non-threatening control cues. Taken together, threat-of-shock was associated with behavioral consequences: initially, participants avoided threat-related options but made more profitable decisions as they experienced no aversive consequences. Although socially acquired threat contingencies are typically stable, incentives for approach can help to overcome threat-related avoidance.

1. Introduction

Individuals often need to choose between behavioral options which are linked to either positive or negative outcomes. If one's choice, however, can result in rewards and aversive events at the same time, an approach-avoidance conflict emerges (Cacioppo & Berntson, 1994; Corr, 2013; Miller, 1959). Adaptive action selection then requires balanced decisions between approaching rewards and avoiding harm (Lejuez et al., 2002). In this approach-avoidance framework, the anticipation of consequences is important to organize goal-directed behavior and a priori information about potential threat versus safety is crucial to decide which behavior is most functional. Decisions may therefore be guided by emotional stimuli that convey information about potential threat. Fundamental motivational neural circuits are assumed to organize this influence of emotional information on approach and avoidance behavior (Lang & Bradley, 2010). This model received much support from studies measuring physiological response parameters (e.g., reflex-based motor and autonomic nervous system activity; Bradley, Codispoti, Cuthbert, & Lang, 2001); as well as neuroimaging studies (e.g., Lang & Davis, 2006). However, surprisingly little is known about potential avoidance biases on more complex behavioral decision-making.

In behavioral decision-making tasks there is often either one positive or one negative outcome. In reward-based decisions, for example, individuals typically show increased selections of profitable options, which are associated with higher or more frequent rewards (e.g., Balleine & Dickinson, 1998; Richards, Plate, & Ernst, 2013). In contrast, individuals will consistently avoid options associated with a single aversive outcome (e.g., an aversive electrical stimulation; Dymond, Schlund, Roche, De Houwer, & Freegard, 2012; Glotzbach, Ewald, Andreatta, Pauli, & Mühlberger, 2012; Lovibond, Mitchell, Minard, Brady, & Menzies, 2009; Ly & Roelofs, 2009).

---

* Corresponding author. Clinical Psychology and Biological Psychology and Psychotherapy, Department of Psychology, School of Social Sciences, University of Mannheim, Germany.

** Corresponding author. Institute of Clinical Psychology and Psychotherapy, Technische Universität Dresden, Dresden, Germany.

E-mail addresses: f.bublatzky@uni-mannheim.de (F. Bublatzky), andre.pittig@tu-dresden.de (A. Pittig).

[1] Contributed equally.

http://dx.doi.org/10.1016/j.brat.2017.01.003

In approach-avoidance conflicts, however, rewards and aversive consequences directly compete. There is growing interest in how competing reward- and threat-related consequences are integrate to guide behavioral decision making (Aupperle, Melrose, Francisco, Paulus, & Stein, 2015; Botvinick & Braver, 2015; Hayes, Duncan, Xu, & Northoff, 2014; Pittig, Brand, Pawlikowski, & Alpers, 2014; Pittig, Schulz, Craske, & Alpers, 2014; Schlund et al., 2016; Sierra-Mercado et al., 2015; Talmi & Pine, 2012). For example, whereas healthy individuals will avoid aversive stimuli when competing rewards are absent, too small, or uncertain, they may tolerate the same aversive stimuli and switch towards approach behavior when sufficiently rewarded (Aupperle, Sullivan, Melrose, Paulus, & Stein, 2011; Sierra-Mercado et al., 2015; Talmi, Dayan, Kiebel, Frith, & Dolan, 2009). A reversed switch or “tipping point” towards consistent avoidance has been found for decisions associated with stable rewards, but increasing threat. Healthy individuals switch from approaching the reward to threat avoidance when the increasing threat exceeds the reward value (e.g., Schlund et al., 2016).

Similar decision making has been observed in anxious individuals when more profitable options were linked to individually fear-relevant stimuli. Spider fearful individuals, for example, initially avoided options associated with the presentation of spider pictures, but tolerated such confrontations when gaining higher rewards with these choices (similar with socially anxious individuals in response to angry facial expressions; see Pittig, Alpers, Niles, & Craske, 2015; Pittig, Brand, et al., 2014; Pittig, Pawlikowski, Craske, & Alpers, 2014). Thus, considering both competing rewards and aversive outcomes is crucial for adaptive goal-directed behavior and imbalances may be associated with psychopathology.

Decision making crucially depends on the anticipation of consequences, which in turn requires that individuals effectively learn about environmental contingencies. This is particularly true for the learning and anticipation of aversive events, as these may harm the organism's physical integrity (Johnson, Blumstein, Fowler, & Haselton, 2013). As a model of such learning processes, much research has employed Pavlovian fear conditioning paradigms, in which formerly neutral stimuli acquire emotional properties through pairing with aversive events such as electric stimulations, heat pain, or monetary loss (Craske, Hermans, & Vansteenwegen, 2006; Duits et al., 2015). Importantly, a recent study provided first evidence that fear conditioning experiences may guide subsequent decisions and result in the development of pathological avoidant decisions. Specifically, a former neutral stimulus was paired with an aversive outcome during fear conditioning. In a subsequent decision task, participants avoided options that were linked to this fear conditioned stimulus, even if these decisions resulted in monetary costs and were not anymore linked to the aversive consequences (Pittig, Schulz, et al., 2014; Experiment 1). In addition, this costly avoidance was elevated in individuals with high trait anxiety (Pittig, Schulz, et al., 2014; Experiment 2). These findings demonstrate how behavioral decisions are biased towards costly avoidance by direct fear learning experience.

However, human fear learning may also occur without direct experience of an aversive event. The mere verbal instruction about potential aversive outcomes has been shown to establish a fear-relevant association that reliably provokes defensive responding (i.e., elevated skin conductance responses, heart rate deceleration, and potentiated startle reflexes; e.g., Bradley, Moulder, & Lang, 2005; Bublatzky, Guerra, Pastor, Schupp, & Vila, 2013; Grillon, Ameli, Woods, Merikangas, & Davis, 1991) and facilitates sensory processing of environmental information (e.g., Baas, Milstein, Donlevy, & Grillon, 2006; Bublatzky & Schupp, 2012; Bublatzky, Flaisch, Stockburger, Schmälzle, & Schupp, 2010; Cornwell et al., 2007). Importantly, the acquisition of human avoidance behavior may be similarly triggered by stimuli that acquired threat associations either by direct experience or mere verbal instructions (Cameron, Roche, Schlund, & Dymond, 2016; Dymond et al., 2012). Despite increasing evidence showing the relevance of instructed threat learning for anxiety and stress-related disorders (e.g., Muris & Field, 2010; Robinson, Vytal, Cornwell, & Grillon, 2013), surprisingly little is known about its impact on the individuals' behavior and decisions in approach-avoidance conflicts.

The present study therefore combined verbal threat instructions and a reward-based decision-making task to test the impact of anticipated threat on reward-directed decisions. Participants had to choose between two decks of cards, which were differently reinforced by monetary incentives and contingent with instructed threat-of-shock or safety cues. Building upon previous research, differential positive reinforcement should favor more frequent choices of the high reward options (e.g., monetary gains; Bechara, Damasio, Tranel, & Damasio, 1997; Pittig, Schulz, et al., 2014). On the other hand, instructed threat of aversive events may lead to behavioral avoidance (Dymond et al., 2012) and enhanced defense activation when confronted with a threat cue (i.e., enhanced SCR and heart rate deceleration; Bradley et al., 2005; Olsson & Phelps, 2004). Regarding the interaction of decision making and threat-of-shock, we hypothesized that choices associated with potential threat would be avoided initially (Pittig, Schulz, et al., 2014). However, behavioral avoidance should gradually diminish with increasing experience of reward contingencies and the omission of the aversive consequence. The actual absence of the instructed aversive consequences should further help to overcome behavioral avoidance (see Pittig, Brand, et al., 2014; Pittig, Schulz, et al., 2014), and may support extinction learning (see Bublatzky, Gerdes, & Alpers, 2014).

2. Methods

2.1. Participants

Sample size was based on power analyses conducted with G-Power (Faul, Erdfelder, Lang, & Buchner, 2007), which indicated that 62 participants were required to detect all relevant behavioral effects at a medium effect size (power $= 0.80$, α error $= 0.05$, medium effect sizes; assumed correlation of repeated measures in repeated measures ANOVA $= 0.40$). Because of randomized assignment to two groups, group sizes varied slightly ($n = 31$ and 34). Sixty-five healthy participants (39 females; 60.0%) were recruited from the University of Mannheim. Their age was between 18 and 41 ($M = 24.3$, $SD = 4.2$). Participants were informed about the general study procedure before providing informed consent according to University of Mannheim ethics guidelines and received course credits for participation. Participants were assigned to two groups (i.e., initial non-threat instruction vs. initial threat/safety instruction), which did not differ in age (see Cauffman et al., 2010), sex distribution, or anxiety and depression scores.[2]

2.2. Procedure

Participants completed questionnaires on anxiety and depression (State-Trait Anxiety Inventory, Spielberger, Gorsuch, Lushene, & Vagg, 1983; Anxiety Sensitivity Index, Peterson & Reiss, 1992;

---

[2] No group differences were observed for Age: $t(63) = 0.03$, $p = 0.975$; Sex: $\chi^2(1, N = 65) = 0.04$, $p = 0.839$; Symptoms of depression (Beck Depression Inventory; BDI-II): $t(63) = 0.07$, $p = 0.945$; State anxiety (State-Trait Anxiety Inventory – State version; STAI-State): $t(63) = 0.31$, $p = 0.754$; Trait anxiety (STAI–Trait): $t(63) = 0.40$, $p = 0.687$; Anxiety sensitivity (Anxiety Sensitivity Index; ASI): $t(63) = 1.30$, $p = 0.198$.

Beck Depression Inventory, Beck, Steer, Ball, & Ranieri, 1996). Sensors for physiological recordings were attached and an electric stimulation electrode was placed at the left index finger (see Riemer, Bublatzky, Trojan, & Alpers, 2015). A brief shock work-up protocol was carried out to calibrate the stimulation intensity at a level rated as "maximally unpleasant but not yet painful" (see Bublatzky et al., 2010). Participants were then instructed that the intensity of the electric stimulation during the experiment would equal the most unpleasant test stimulus.

Experimental procedures are shown in Fig. 1. Half of the participants started with the non-threat instruction followed by a threat/safety instruction phase; the other participants started with threat/safety followed by the non-threat instruction phase. Participants rated the hedonic valence and arousal of the colored squares after each phase using the Self-Assessment Manikin (SAM; Bradley & Lang, 1994); and perceived threat was scored using a visual analog scale ranging from *not at all* to *highly threatening* (1–10) after the threat/safety instruction phase. At the end of the experiment, participants indicated whether the threat/safety instructions were perceived as convincing (yes/no) and effective (yes/no).

The computer-based card game was controlled by Presentation software (Neurobehavioral Systems, Inc., Albany, CA), and presented on a 22 inch computer screen approximately 1 m in front of the participant. Electrical stimulations (max. 10 mA, 100 ms) were generated with a stand-alone electrical stimulator (Jaeger-Toennies, Germany). Physiological data were recorded with a vAmp amplifier (BrainProducts, Munich, Germany).

2.3. Decision-making task, materials, and design

For the computerized decision-making task, participants had to choose between two card decks displayed on the screen with the instructed goal of maximizing overall gain (Fig. 1; see also Pittig, Schulz, et al., 2014). Each choice was followed by a fixation cross as inter stimulus interval (5s), a colored square (blue or yellow, 7 s), another fixation cross (5s), and a subsequent visual feedback stating whether they did or did not win a fixed amount of € 0.50 (5 s). As in previous studies (Pittig, Schulz, et al., 2014), hypothetical rewards were used, i.e., participants did not receive the rewards gained in the task at the end of the experiment. A high-reward option (left or right card deck; counterbalanced across participants) was associated with 60% reward probability and a low-reward option with 40% reward probability. For each participant, the color of the square was locked to a specific deck, which was counterbalanced across participants (e.g., the high-reward deck was always followed by the blue square for one participant, but always followed by the yellow square for another participant).

Participants completed two subsequent experimental instruction phases, each consisting of 40 decision trials. During the threat/safety instruction phase, participants were instructed that they may receive an aversive electrical stimulation whenever a certain colored square is presented (e.g., blue square as instructed threat cue), but never during the presence of the other colored square (e.g., yellow square as instructed safety cue). Importantly, to maximize the approach-avoidance conflict (i.e., high-reward and threat-of-shock), selection of the high-reward deck was 100% contingent with the instructed threat color during the threat/safety instruction phase. Thus, participants could avoid the threat cue by choosing the low-reward deck, which resulted in less gain. During non-threat instruction phase, participants were instructed that no (more) electrical stimulations will occur at all. Thus, there was no decision conflict during the non-threat instruction trials. As a between subject manipulation of the conditions under which the participants initially learned about reward contingencies, approximately half of the participants ($N = 31$) started with the non-threat instruction phase followed by the threat/safety instruction trials (= initial non-threat instruction); the other group ($N = 34$) completed sessions in reversed sequence (= initial threat/safety instruction; see Fig. 1).

2.4. Data recording and reduction

As in previous studies (e.g., Pittig, Schulz, et al., 2014), the trials of the decision task were combined into blocks of ten trials each. For each block, the number of selections of the (aversive) high-reward deck was calculated as the main behavioral outcome variable. Thus, a score of five indicates that both decks were selected equally often and a higher score indicates more frequent selections of the high-reward deck.

Skin conductance responses (SCRs) were recorded with two Ag/AgCl electrodes (constant voltage of 0.5 V; 500 Hz sampling rate) placed on the hypothenar eminence of the palm of the non-dominant hand. Noise and slow frequency/level changes were removed using a 2 Hz FIR low- and a 0.05 Hz high-pass filter in

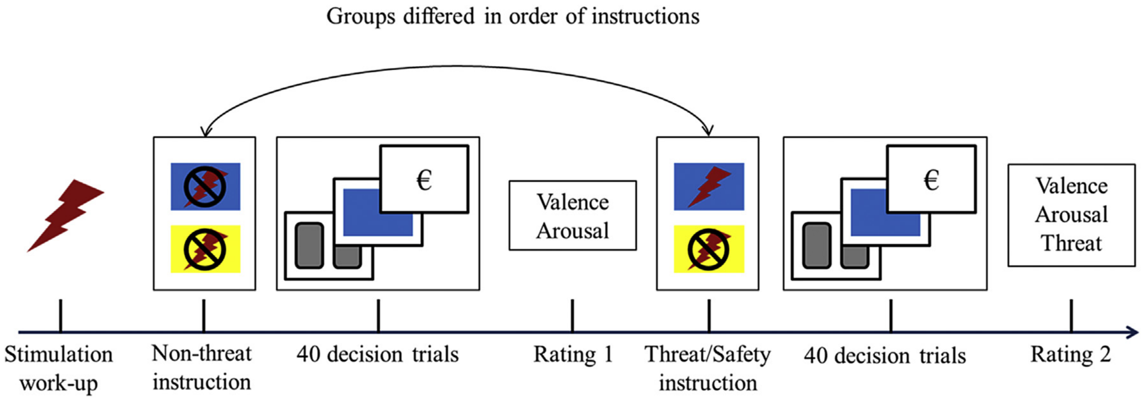

Fig. 1. Experimental procedures. After shock work-up, participants received either threat/safety or non-threat instruction (e.g., "You may receive an aversive stimulation whenever the blue square, but never when the yellow square is presented." vs. "No aversive stimuli will be delivered at all.") and completed 40 trials of the decision task. Participants had to choose between a high-reward option (e.g., left card deck, associated with 60% chance of gain) which was always followed by a particular colored square (e.g., blue); or the low-reward option (e.g., right deck, followed by a yellow square) which was associated with 40% chance of gain. Importantly, to implement an approach-avoidance conflict, the high-reward option was 100% contingent with the colored square instructed to signal threat-of-shock. In contrast, the low-reward option was associated with instructed safety square during threat/safety instruction phase. After rating the colored squares for valence and arousal (as well as threat following the threat/safety instruction phase), instructions were switched and participants completed another 40 trials. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Brain Analyzer 2.1 (BrainProducts, Munich, Germany). SCRs to the onset of the colored squares were calculated as the maximum increase in skin conductance amplitude in the interval of 1–7 s (relative to a 1 s pre-stimulus period). A threshold of 0.02 MicroSiemens (μS) was used; all SCRs below this threshold were scored as zero response and included in the analyses (i.e., SCR magnitude); range and distribution correction was applied (square root [response/maximum response]; see Pittig, Schulz, et al., 2014).

Heart rate was derived from the electrocardiogram recorded at lead II. The signal was acquired at a sampling rate of 1000 Hz and frequencies below 0.1 and above 13 Hz were filtered. R-wave detection, visual inspection of inter-beat intervals, and conversion to heart rate (beats per minute, bpm) was done with Brain Analyzer 2.1. The HR averages every second after onset of the colored squares are expressed in terms of differential scores with respect to a 1 s baseline period (see Bradley et al., 2005), and averaged across the total cue interval. Four participants were excluded from HR analyses (two from each group) due to equipment failure (n = 3) and an excessive amount of ECG artifacts (n = 1; large number of ectopic beats and movement artifacts). One participant was excluded from SCR analyses due to equipment failure.

2.5. Data analysis

Decision-making data were entered into a 2 × 4 × 2 repeated measures ANOVA with Instruction Phase (threat/safety vs. non-threat) and Block (four blocks depicting decision trials 1–10, 11–20, 21–30, 31–40) as within subject factors, and Phase Order (initial threat/safety vs. initial non-threat instruction phase) as a between subject group factor. To test the modulation of decision behavior by means of threat/safety instruction, two a-priori planned analyses were conducted. First, to investigate avoidance of instructed threat stimuli with vs. without prior experience of reward contingencies, planned comparison *t*-tests were conducted for the first block of each instruction phase between both groups and against the constant 5 (which represents equal selections from both decks). Second, the critical change from the last block of the first instruction phase (Trials 31–40) to the first block of the second instruction phase (Trials 41–50) was separately analyzed using a 2 × 2 repeated measures ANOVA with Phase Order and the two blocks as a within subject factor. This change represents the modification of decisions by switching instructions (i.e., last block of threat/safety vs. first block of non-threat instruction phase for the Initial Threat group; last block of non-threat vs. first block of threat/safety instruction for the Initial Non-Threat group). For separate *t*-tests, Cohen's *d* was calculated as effect size.

For self-reported ratings of unpleasantness and arousal, repeated measures ANOVAs were conducted with Cue (instructed threat cue vs. instructed safety cue), Instruction Phase (threat/safety vs. non-threat), and Phase Order as between-subject factor. To test the effects of threat-of-shock on skin conductance and heart rate, separate repeated measures ANOVAs were conducted including the within-subject factors Cue (blue vs. yellow square) and Instruction Phase (non-threat vs. threat/safety instruction), and the between-subject factor Phase Order (initial non-threat vs. initial threat/safety). Physiological responses were averaged across the four blocks in each instruction phase; this was done because the number of trials per factor level differed considerably as a function of individual decision behavior across experimental blocks (e.g., more approach reciprocally resulted in less avoidance trials). However, for exploration of the time course of the impact of threat instruction on SCRs, block-wise analyses were conducted for SCRs to the threat vs. safety cues during the threat/safety instruction phase.

Greenhouse-Geisser correction procedure was applied where relevant, and the partial η^2 (η^2_p) is reported as a measure of effect size. To control for type 1 error, Bonferroni correction was applied for post hoc tests.

3. Results

3.1. Self-report data

Overall, participants rated the threat/safety instruction as credible (n = 59; 90.8%) and effective (n = 52; 80%) in provoking a state of aversive apprehension. The successful verbal threat manipulation was also evident in valence and arousal ratings. Self-reported valence and arousal varied as a function of Cue by Instruction Phase, $Fs(1,62) > 8.41$, $ps < 0.006$, $\eta^2_p > 0.118$. Whereas ratings for both stimuli did not differ on both dimensions after the non-threat instruction phase, all $ts < 1.50$, all $ps > 0.145$, $d_{valence} = 0.10$, $d_{arousal} = 0.28$, the instructed threat cue was rated as more unpleasant and arousing than the safety cue after the threat/safety instruction phase, all $ts > 2.77$, all $ps < 0.008$, $d_{valence} = 0.53$, $d_{arousal} = 0.94$. In addition, the threat cue was more unpleasant and arousing after the threat/safety compared to non-threat instruction phase, all $ts > 2.49$, all $ps < 0.016$, $d_{valence} = 0.52$, $d_{arousal} = 1.56$. The instructed safety cue was more pleasant after threat/safety compared to non-threat instruction, $t(64) = 2.27$, $p = 0.027$, $d = 0.30$, but did not differ in arousal, $t(64) = 0.37$, $p = 0.715$, $d = 0.05$. Finally, the instructed threat cue compared to the safety cue was rated as more threatening following the threat/safety instruction phase, $t(64) = 5.29$, $p < 0.001$, $d = 0.88$. In sum, threat instructions specifically resulted in elevated unpleasantness, arousal, and threat ratings for the instructed threat cue.

Furthermore, self-reported arousal revealed an interaction of Instruction and Phase Order, $F(1,62) = 12.02$, $p = 0.002$, $\eta^2_p = 0.196$, which was not observed for valence ratings, $F(1, 62) = 0.06$, $p = 0.803$, $\eta^2_p < 0.01$. Across both stimuli, arousal ratings were higher after the threat/safety instruction phase compared to the non-threat instruction phase for the Initial threat group, $t(33) = 6.28$, $p < 0.001$, $d = 0.93$, but not for the Initial non-threat group, $t(30) = 0.46$, $p = 0.648$, $d = 0.08$.

3.2. Decision making behavior

As illustrated in Fig. 2, the number of high-reward choices was modulated by Instruction Phase, $F(1,63) = 6.30$, $p = 0.015$, $\eta^2_p = 0.091$, and Block, $F(3,189) = 11.32$, $p < 0.001$, $\eta^2_p = 0.152$, as well as the interaction of Instruction Phase by Block, $F(3,189) = 2.77$, $p = 0.047$, $\eta^2_p = 0.42$. Following up on this interaction, separate ANOVAs were calculated for each instruction phase (collapsed across both groups). Indicating progressively more selections of the profitable card deck, the number of high-reward choices significantly increased during the non-threat instruction phase, $F(3,192) = 3.28$, $p = 0.028$, $\eta^2_p = 0.049$. Interestingly, this behavioral pattern was even more pronounced when profits were associated with potential threat-of-shock during the threat/safety instruction phase, $F(3,192) = 12.75$, $p < 0.001$, $\eta^2_p = 0.166$.

Importantly, decision behavior also varied as a joint function of Instruction Phase and Phase Order, $F(1,63) = 12.66$, $p = 0.001$, $\eta^2_p = 0.167$ (see Fig. 3). Following up on this of-interest interaction, the average number of high-reward choices in each phase (i.e., averaged across blocks) was compared between groups using post-hoc *t*-tests. Whereas the initial threat/safety compared to the initial non-threat group showed more frequent high-reward choices during non-threat instruction, $t(63) = 2.94$, $p = 0.005$, $d = 0.73$, there were no group differences for the threat/safety instruction phase, $t(63) = -1.35$, $p = 0.184$, $d = 0.34$. Furthermore, the initial threat/safety group showed more frequent high-reward choices during the non-threat compared to the threat/safety instruction,

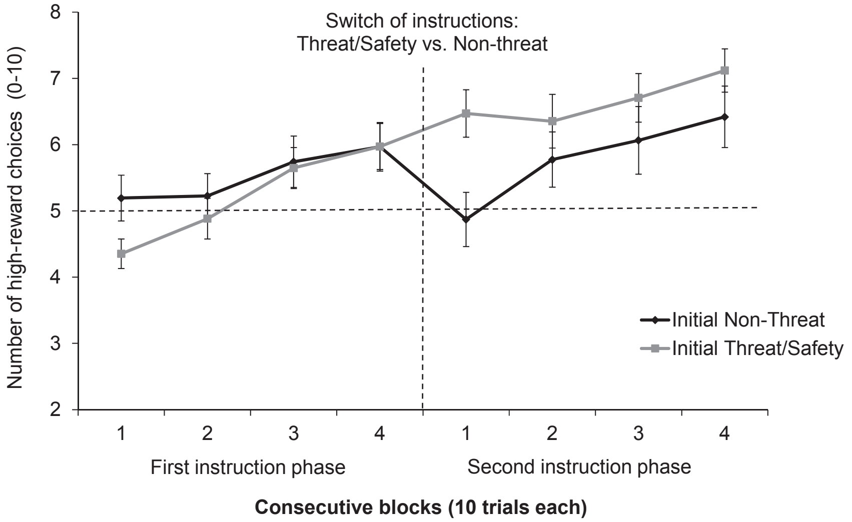

Fig. 2. Mean number of high-reward selections (and standard error) as a function of Instruction Phase, Block, and Phase Order (initial non-threat vs. initial threat/safety). Scores above 5 (dotted horizontal line) indicate more frequent high-reward than low-reward deck selections. The group that started with the non-threat instruction phase (black line) completed Trials 1–40 under non-threat instructions ("No aversive stimuli will be delivered at all.") and Trials 41–80 under threat/safety instructions (e.g., "You may receive an aversive stimulation whenever the blue square is presented, but never when the yellow square is presented."), phase order was reversed for the initial threat/safety group (grey line). The dotted vertical lines index the time when the instructions were switched.

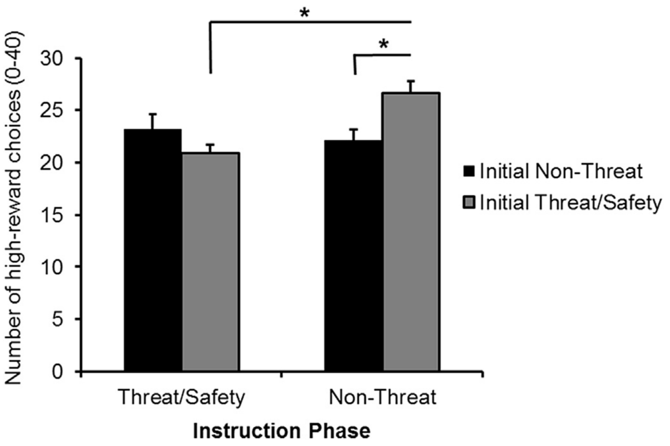

Fig. 3. Total number of selections (averaged across 4 blocks each) of the (aversive) high-reward deck (with standard errors) during threat/safety and non-threat instruction phases separated by phase order. $^{*}p < 0.05$.

$t(33) = 4.78$, $p < 0.001$, $d = 0.99$, but no differences were observed between instruction phases within the initial non-threat group, $t(30) = 0.67$, $p = 0.509$, $d = 0.14$. Thus, participants made more high-reward choices during non-threat instruction, when they already had prior experiences with reward contingencies during threat/safety instruction phase.

3.2.1. *Immediate avoidance following threat instruction*

For the first ten trials, the initial threat/safety compared to initial non-threat group showed significantly fewer high-reward choices, $t(63) = 2.08$, $p = 0.041$, $d = 0.52$, and significantly avoided the high-reward option immediately after threat-of-shock was instructed, $t(33) = -2.90$, $p = 0.007$, $d = 0.50$. The initial non-threat group, however, showed equal selections of both decks during the first ten trials, $t(30) = 0.56$, $p = 0.579$, $d = 0.10$, indicating balanced decision making in the absence of instructed threat. Regarding the change between instruction phases, planned analyses revealed a significant interaction between Phase Order and Block, $F(1,63) = 5.82$, $p = 0.019$, $\eta_p^2 = 0.085$. The follow-up *t*-test showed that the omission of threat instruction did not significantly change decision behavior in the initial threat/safety group, $t(33) = -1.28$, $p = 0.211$, $d = 0.25$. However, introducing threat-of-shock for the initial non-threat group was associated with a tendency of less frequent high-reward choices with medium effect size, which, however, missed statistical significance, $t(30) = 2.02$, $p = 0.053$, $d = 0.51$. No differences were observed between groups in the last block of the first instruction phase (Trials 31–40), $t(63) = -0.01$, $p = 0.996$, $d < 0.01$, but the initial non-threat compared to initial threat/safety group showed significantly fewer high-reward choices after instructions were switched, $t(63) = -2.95$, $p = 0.004$, $d = 0.73$. In addition, compared to an equally balanced selection of both decks (indicated by a mean number of reward choices of $M = 5$), the initial threat/safety group showed significant approach of the high-reward choices before and after instructions were switched, Trial 31–40 $t(33) = 2.80$, $p = 0.009$, $d = 0.48$, and Trials 41–50 $t(33) = 4.10$, $p < 0.001$, $d = 0.70$. For the initial non-threat group, such an approach pattern was observed before switching instructions, $t(30) = 2.64$, $p = 0.013$, $d = 0.47$, however, no approach was found after threat-of-shock was introduced, $t(30) = -0.32$, $p = 0.755$, $d = 0.06$.

To summarize behavioral data, the initial threat/safety group initially avoided the high-reward option but in the absence of aversive events steadily demonstrated more approach. At the end of the first instruction phase, they approached the aversive high-reward option with no significant differences to those participants who started with the non-threat instructions. When instructions were reversed, the initial threat/safety group continued to approach the high-reward option after omission of instructed threat, but the initial non-threat group avoided this option after the threat-of-shock instruction was introduced.

3.3. *Physiological data*

3.3.1. *Skin conductance responses (SCRs)*

Physiological data served to verify that threat instructions were linked to elevated fear responses towards the threat-cue compared to safety-cue and non-threat cues. Skin conductance responses towards each colored square averaged across both instruction phases varied as a main effect of Cue and Instruction Phase, $Fs(1,62) = 7.63$ and 17.37, $ps < 0.01$ and 0.001, $\eta_p^2 = 0.11$ and 0.22, but no interaction was observed, $F(1,62) = 0.32$, $p = 0.58$, $\eta_p^2 < 0.01$ (see Fig. 4 A). Follow-up tests for each instruction phase revealed enhanced SCRs to instructed threat compared to safety cues, $p = 0.036$. As expected, no differences between the colored cues were found in the non-threat instruction phase, $p = 0.083$. In addition, SCRs to the instructed threat and safety cue during the threat/safety instruction phase were more pronounced compared to the corresponding cues during the non-threat instruction phase, $p < 0.01$ and 0.035. The between-group manipulation Phase Order (initial non-threat vs. initial threat/safety instruction phase) did not show interactions with Instruction Phase, $F(1,62) = 0.53$, $p = 0.47$, $\eta_p^2 = 0.01$, Cue, $F(1,62) = 0.06$, $p = 0.81$, $\eta_p^2 < 0.01$, nor with Cue by Instruction Phase, $F(1,62) = 0.13$, $p = 0.72$, $\eta_p^2 < 0.01$.

Exploratory follow-up analyses focused on SCRs during threat/safety instruction phase to specifically test differential SCRs to the instructed threat vs. safety cue across the time course of the threat instruction phase (see Fig. 5). SCRs strongly decreased across blocks, $F(3,189) = 43.82$, $p < 0.001$, $\eta_p^2 = 0.41$, reflecting a decrease in SCRs to both threat and safety cues, $Fs(3,189) = 21.83$ and 17.54, $ps < 0.001$, $\eta_p^2 = 0.26$ and 0.22, which was also associated with a progressively increasing number of zero-responses. The main effect of Cue was not significant, $F(1,63) = 3.54$, $p = 0.065$ $\eta_p^2 = 0.05$, however, a significant interaction of Cue by Block was observed, $F(3,189) = 5.82$, $p = 0.001$, $\eta_p^2 = 0.09$. Elevated SCRs for the threat compared to safety cue was found in Block 1 and 3, $ps = 0.007$ and 0.01, no difference in Block 2, $p = 0.26$, and a reversed pattern showing elevated SCRs to safety cues in Block 4, $p = 0.03$. Whereas the decline in SCRs across blocks was a reliable effect, caution is needed regarding differential SCRs to threat/safety cues across blocks. Here, a varying number of trials per cue (i.e., decreasing number of safety trials) and an increasing number of zero-responses may have biased statistical comparisons.

3.3.2. Heart rate responses

Heart rate revealed a pronounced deceleration following the onset of the colored threat-, safety-, or non-threat cues (see Fig. 4 B). This deceleration varied as a function of Instruction Phase, $F(1,59) = 13.15$, $p = 0.001$, $\eta_p^2 = 0.18$, Cue, $F(1,59) = 39.92$, $p < 0.001$, $\eta_p^2 = 0.40$, and Instruction Phase by Cue, $F(1,59) = 11.39$, $p = 0.001$, $\eta_p^2 = 0.16$. Follow-up analyses for the threat/safety instruction phase revealed more pronounced deceleration for the instructed threat compared to the safety cue, $p < 0.001$. A similar, but smaller difference between both stimuli was also evident during the non-threat instruction phase (see Fig. 4 B; Non-threat (low reward) vs. Non-threat (high reward)), $p < 0.05$. In addition, HR deceleration was more pronounced for the threat cue relative to the corresponding cue during non-threat instruction phase, $p < 0.001$, but no difference was found between the safety cue and the corresponding cue during non-threat instruction, $p = 1.0$.

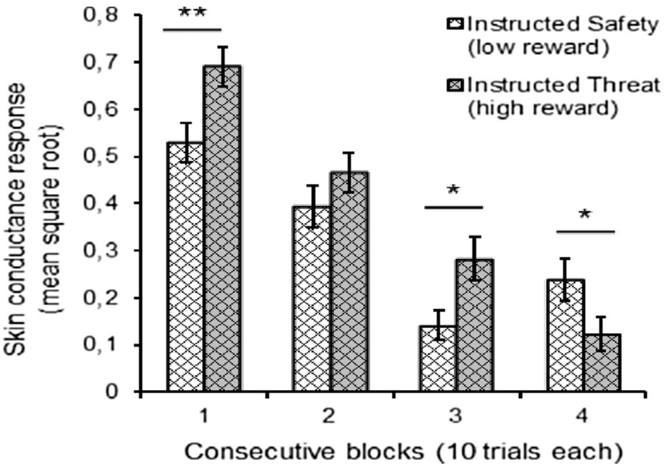

Fig. 5. Mean skin conductance responses (and SEM) to the instructed threat vs. safety cue across the time course of the threat/safety instruction phase. $^{*}p < 0.05$, $^{**}p < 0.01$.

Regarding the between-group manipulation Phase Order (initial non-threat vs. initial threat/safety instruction phase), no interactions were observed for Phase Order by Cue or Instruction Phase, $Fs(1,59) < 0.54$, $ps > 0.466$, $\eta_p^2 < 0.01$, nor for the three-way interaction Phase Order by Cue by Instruction Phase, $F(1,59) = 2.44$, $p = 0.124$, $\eta_p^2 = 0.04$. Exploratory follow-up tests separately for the initial non-threat group revealed a significant interaction of Instruction Phase by Cue, $F(1,28) = 10.67$, $p < 0.01$, $\eta_p^2 = 0.28$, indicating no difference between stimuli in the non-threat phase, $p = 0.11$, but pronounced deceleration for the threat-relative to safety-cue in the following threat/safety phase, $p < 0.001$.

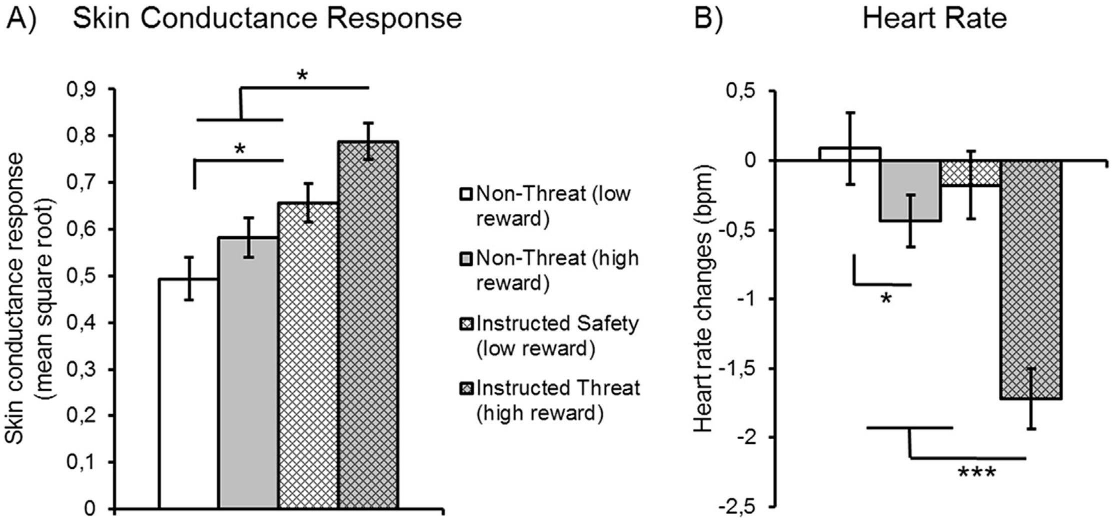

Fig. 4. Mean skin conductance responses (A) and heart rate changes (B) (and SEM) following the onset of the colored cues (within 7 s) indicating threat-of-shock, safety, or during non-threat condition (collapsed across groups and blocks within each instruction phase). $^{*}p < 0.05$, $^{***}p < 0.001$.

Interestingly, no such interaction was observed for the group starting with initial threat/safety instructions, $F(1,31) = 1.88$, $p = 0.18$, $\eta_p^2 = 0.06$. Thus, pronounced threat deceleration from the initial threat/safety instruction phase, $p < 0.01$, persisted in the following non-threat instruction phase, $p < 0.05$.

4. Discussion

The present study showed that instructed threat – although there were actually no aversive consequences– can trigger avoidance of profitable decisions. Implementing an approach-avoidance conflict, participants could choose a more profitable option which was at the same time contingent with instructed threat of aversive electrical stimulations. Another option was associated with less rewards but always followed by an instructed safety signal (indicating the absence of aversive consequences). Results showed that choices implicating potential threat were initially avoided in favor of less profitable but safe decisions. Furthermore, indicating elevated fearful arousal, threat instruction was associated with elevated skin conductance responses and heart rate deceleration, which is typically found in anticipation of aversive events (Bradley et al., 2005, 2001; Hermann, Ziegler, Birbaumer, & Flor, 2002; Pineles, Orr, & Orr, 2009). These results are in line with much previous research demonstrating the capability of verbal instructions to establish physiological and self-reported fear responses (e.g., Bublatzky et al., 2014, 2010; Olsson & Phelps, 2004; Raes, De Houwer, De Schryver, Brass, & Kalisch, 2014) and extend these findings from fear responding to behavioral avoidance. Despite the psychophysiological pattern of defense activation, avoidance behavior was transient and participants progressively overcame aversive apprehensions in favor of profitable choices. Thus, conflicting threat and reward resulted in initial avoidance of profitable decisions. However, when having the choice, the mere instruction about potential threat does not result in long-lasting behavioral costs in healthy participants.

Some previous studies on decision making focused on the impact of unspecific threat. For instance, verbal instructions about unspecific aversive events was linked to general decision-making biases such as elevated avoidance of risky gambling options (Clark et al., 2012) or premature and non-systematic decisions (Keinan, 1987). Moreover, previous studies also showed that behavioral avoidance may be triggered by single instructed threat cues (e.g., Dymond et al., 2012). The present study, for the first time, specifically tied instructed threat cues to more profitable behavioral choices. As expected, an increase in profitable, high-reward decisions was observed after several contingent choice-reward pairings. Importantly, when instructed threat was contingent with high-reward choices, short-term avoidance of the profitable option was observed. Such avoidance was apparent in participants who made prior experience with reward contingencies (i.e., the initial non-threat group) as well as those who did not (i.e., initial threat/safety group). Importantly, avoidance was triggered irrespective of whether threat was the sole anticipated outcome (i.e., when reward contingencies were still uncertain) or whether a decision conflict had already emerged (i.e., competing anticipation of threat and rewards). Thus, novel verbal threat information may suffice to bias decisions towards threat-driven avoidance, even when avoidance is linked to the loss of alternative rewards.

Avoidance behavior was, however, transient across time. In both groups, participants overcame initial avoidance and faced potential threat in favor of maximizing gains. This pattern of transient avoidance is in line with previous studies showing a reduction or elimination of avoidance in approach-avoidance conflicts. For example, individuals switch from avoidance of aversive or fear-relevant stimuli to approach of rewards when reward values are sufficiently high and certain (Aupperle et al., 2011, 2015; Schlund et al., 2016; Sierra-Mercado et al., 2015; Talmi et al., 2009). Likewise, increased experience with rewards may counteract avoidance of specific feared stimuli (e.g., spider pictures or angry faces in spider or social phobia; Pittig et al., 2015; Pittig, Brand, et al., 2014; Pittig, Pawlikowski, et al., 2014). Whereas decisions were linked to the actual occurrence of aversive stimuli in these studies, the present results provide first demonstration of a comparable impact of mere verbal announcement of aversive contingencies. Interestingly, having had prior experience with reward contingencies during potential threat lead to more high-reward choices in the following non-threat condition (in the initial threat/safety group). This may indicate that adequate learning about reward contingencies may have taken place during threat, but decisions were initially more guided by threat information. In the absence of threat, this prior reward learning was subsequently translated into more profitable decisions.

Switching between approach and avoidance behavior has been linked to changes of reward and threat values. Specifically, studies using the actual occurrences of aversive stimuli showed that both increasing reward and decreasing threat values may initiate a transition from avoidance to approach (see Schlund et al., 2016). As the present study used learning of stable rewards (i.e., without instructions about reward contingencies), it cannot numerically quantify the two processes. Comparing experimental groups, however, indicates that both reward and threat learning were involved in the pattern of transient avoidance. It seems likely that, during the very first decisions, the rather small difference between reward probabilities (60% vs. 40%) yielded ambiguity about which option is more profitable. Thereby, the explicitly instructed threat value was more likely to trigger avoidance (i.e., initial threat/safety group). However, this reward ambiguity was weaker or absent for participants, who could safely learn about reward contingencies before threat-of-shock was introduced, but still showed short-term avoidance of the learned profitable option (i.e., initial non-threat group). Subsequently, increasingly more frequent choices of the profitable option (despite threat) may relate to a decrease in threat value of this option. This notion is supported by exploratory analyses on the time course of physiological fear as measured by skin conductance responses. Whereas elevated SCRs were observed to the instructed threat cues (i.e., first and third threat/safety block), the magnitude of threat responses gradually decreases during subsequent decisions. Thus, introducing threat instructions may initially trigger physiological fear and avoidant decisions despite costs, but competing reward learning seems to initiate approach behavior and thereby set the stage for fear extinction.

The behavioral finding of rather short-term avoidance, and decreasing physiological fear, appears partly contradicting to previous studies that showed persistent threat-of-shock effects on the physiological expression of fear and anxiety. For instance, instructed threat effects, as indicated by subjective ratings, skin conductance level, and threat-potentiated startle responses, did not subside until repeated test days without aversive reinforcement (Bublatzky et al., 2014, 2013). In these studies, however, participants had no opportunity to actively avoid threat conditions; accordingly, diverging effects likely relate to the mitigating impact of controllability in the present design (cf. Foa, Zinbarg, & Rothbaum, 1992; Grillon, Baas, Cornwell, & Johnson, 2006; Hartley, Gorun, Reddan, Ramirez, & Phelps, 2014). Moreover, as competing rewards enhanced approach of the instructed threat cue, this self-paced approach seemed to facilitate learning about the absence of aversive consequences and fear extinction learning.

Seen from a clinical perspective, dysfunctional avoidance is a key feature of anxiety disorders (Hofmann, Alpers, & Pauli, 2009). Whereas traditional models – such as Mowrers' two-factor theory

(Mowrer, 1960) focus on fear and aversive consequences – adding the effects of competing rewards and the costs of avoidance may improve our understanding of the pathological qualities of anxiety. For instance, avoidance of a beneficial behavior, which is linked to aversive anticipations, may be particularly elevated in anxious individuals and reflect the sustained impairments of aversive apprehensions (e.g., Dymond, Schlund, Roche, & Whelan, 2014; Gillan et al., 2014; Pittig, Schulz, et al., 2014).

Furthermore, accounting for approach-avoidance conflicts is crucial for the treatment of anxiety disorders. Although exposure-based interventions are a highly effective treatment (Alpers, 2010; Hofmann & Smits, 2008; Tolin, 2010), there is ample need for optimization (e.g., Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014; Loerinc et al., 2015). Current research mainly focuses on a maximization of fear extinction during exposure (e.g., Craske et al., 2014; Pittig, van den Berg, & Vervliet, 2016). Expanding this focus, research on approach-avoidance conflicts may inform about the precursory motivational-volitional mechanisms of exposure (i.e., decision to undergo exposure exercises).

The present findings imply that incentives for approach can help to reduce threat-related avoidance and thereby facilitate exposure exercises and fear extinction. In support of this notion, better treatment outcome is reported by patients with social anxiety disorder who more often approached a fear-relevant stimulus to gain long-term rewards (Pittig et al., 2015). Providing and highlighting approach incentives may thus be an important asset to exposure therapy, especially for patients who are ambiguous about conducting exposure. For these patients, incentives may strengthen approach motivation beyond a tipping point, at which reward values exceed threat values, and thereby initiate a switch from dysfunctional avoidance to approach (see Schlund et al., 2016). Similar motivational aspects are already addressed in some therapeutic strategies such as motivational interviewing (Miller & Rollnick, 2012) or value-based exposure in Acceptance-Commitment therapy (Hayes, Strosahl, & Wilson, 2003). However, there is still little known about the involved processes, which are crucial to understand and enhance motivational outcomes. Future research is thus needed to pinpoint the underlying mechanism of approach-avoidance decisions in healthy and anxious individuals.

The present study was designed to investigate the impact of instructed threat on decision making, and some limitations and future directions need to be acknowledged. As a result of participants' free choice, the number and order of threat/safety cue presentations differed considerably and precluded a standardized trial-by-trial analysis of physiological data. Instructed threat compared to safety cues reliably enhanced overall sympathetic system activation; however, caution is needed regarding the time course of physiological threat effects. Specifically, favoring one option (e.g., profits despite threat) results in decreasing number of trials for the other option (e.g., costs on safety trials), which is likely to bias threat/safety comparisons especially when participants start to develop behavioral preferences. Here, other physiological measures of defense activation (e.g., startle response) may be better suited to test the temporal dynamics of behavioral change and its physiology to fully evaluate fear extinction and its link to behavior change (Richter et al., 2012). Moreover, increasing reward or threat probabilities (e.g., 70% vs. 30%) and/or intensities (e.g., 0.1 vs. 1€ reward, or less vs. more unpleasant shock threat) may maximize approach-avoidance conflict and potentially help identify individual differences that contribute to clinical avoidance. The present paradigm may also be useful to probe the effects of different means and mechanisms of fear learning on behavioral decision making, for example, observational learning or fear generalization (Cameron, Schlund, & Dymond, 2015; Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2015; Olsson & Phelps, 2007).

Taken together, the present study combined two well-established paradigms: verbal threat learning and a reward-based decision making task. Focusing on the impact of aversive anticipation on reward-directed decisions, an approach-avoidance conflict was established. Having the choice between high- or low-reward options – that were contingent with instructed threat or safety – participants initially preferred safe but non-profitable decisions. However, as instructed threat was not substantiated by actual aversive consequences, avoidance was transient and decision behavior changed in favor of profitable but potentially threatening decisions. Thus, conflicting threat and reward-based learning revealed initial avoidance of profitable decisions. However, when they have a choice, the mere instructed threat of aversive events does not result in long-lasting behavioral costs in healthy participants.

Acknowledgements

We are grateful to M. Borger, L. Laubisch, and E. Smolenskaya for their assistance in data collection. This work was supported, in part, by the "Struktur-und Innovationsfonds (SI-BW)" of the state of Baden-Wuerttemberg, Germany, and by the German Research Foundation (BU 3255/1-1) to FB.

References

Alpers, G. W. (2010). Avoiding treatment failures in specific phobias. In M. W. Otto, & S. G. Hofmann (Eds.), *Avoiding treatment failures in the anxiety disorders* (pp. 209–227). New York, NY: Springer.

Aupperle, R. L., Melrose, A. J., Francisco, A., Paulus, M. P., & Stein, M. B. (2015). Neural substrates of approach-avoidance conflict decision-making. *Human Brain Mapping, 36*, 449–462. http://dx.doi.org/10.1002/hbm.22639.

Aupperle, R. L., Sullivan, S., Melrose, A. J., Paulus, M. P., & Stein, M. B. (2011). A reverse translational approach to quantify approach-avoidance conflict in humans. *Behavioural Brain Research, 225*, 455–463. http://dx.doi.org/10.1016/j.bbr.2011.08.003.

Baas, J. M. P., Milstein, J., Donlevy, M., & Grillon, C. (2006). Brainstem correlates of defensive states in humans. *Biological Psychiatry, 59*, 588–593. http://dx.doi.org/10.1016/j.biopsych.2005.09.009.

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology, 37*, 407–419.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*, 1293–1295.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck depression inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment, 67*, 588–597.

Botvinick, M., & Braver, T. (2015). Motivation and cognitive Control: From behavior to neural mechanism. *Annual Review of Psychology, 66*, 83–113. http://dx.doi.org/10.1146/annurev-psych-010814-015044.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*, 276–298. http://dx.doi.org/10.1037//1528-3542.1.3.276.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*, 49–59. http://dx.doi.org/10.1016/0005-7916(94)90063-9.

Bradley, M. M., Moulder, B., & Lang, P. J. (2005). When good things go bad: The reflex physiology of defense. *Psychological Science, 16*, 468–473. http://dx.doi.org/10.1111/j.0956-7976.2005.01558.x.

Bublatzky, F., Flaisch, T., Stockburger, J., Schmälzle, R., & Schupp, H. T. (2010). The interaction of anticipatory anxiety and emotional picture processing: An event-related brain potential study. *Psychophysiology, 47*, 687–696. http://dx.doi.org/10.1111/j.1469-8986.2010.00966.x.

Bublatzky, F., Gerdes, A. B. M., & Alpers, G. W. (2014). The persistence of socially instructed threat: Two threat-of-shock studies. *Psychophysiology, 51*, 1005–1014. http://dx.doi.org/10.1111/psyp.12251.

Bublatzky, F., Guerra, P. M., Pastor, M. C., Schupp, H. T., & Vila, J. (2013). Additive effects of threat-of-shock and picture valence on startle reflex modulation. *PLoS One, 8*, e54003. http://dx.doi.org/10.1371/journal.pone.0054003.

Bublatzky, F., & Schupp, H. T. (2012). Pictures cueing threat: Brain dynamics in viewing explicitly instructed danger cues. *Social Cognitive and Affective Neuroscience, 7*, 611–622. http://dx.doi.org/10.1093/scan/nsr032.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115*, 401–423. http://dx.doi.org/10.1037/0033-2909.115.3.401.

Cameron, G., Roche, B., Schlund, M. W., & Dymond, S. (2016). Learned, instructed and observed pathways to fear and avoidance. *Journal of Behavior Therapy and Experimental Psychiatry, 50*, 106–112. http://dx.doi.org/10.1016/j.jbtep.2015.06.003.

Cameron, G., Schlund, M. W., & Dymond, S. (2015). Generalization of socially transmitted and instructed avoidance. *Frontiers in Behavioral Neuroscience, 9*, 1–15. http://dx.doi.org/10.3389/fnbeh.2015.00159.

Cauffman, E., Shulman, E. P., Steinberg, L., Claus, E., Banich, M. T., Graham, S., et al. (2010). Age differences in affective decision making as indexed by performance on the Iowa Gambling Task. *Developmental Psychology, 46*, 193–207. http://dx.doi.org/10.1037/a0016128.

Clark, L., Li, R., Wright, C. M., Rome, F., Fairchild, G., Dunn, B. D., et al. (2012). Risk-avoidant decision making increased by threat of electric shock. *Psychophysiology, 49*, 1436–1443. http://dx.doi.org/10.1111/j.1469-8986.2012.01454.x.

Cornwell, B. R., Baas, J. M. P., Johnson, L., Holroyd, T., Carver, F. W., Lissek, S., et al. (2007). Neural responses to auditory stimulus deviance under threat of electric shock revealed by spatially-filtered magnetoencephalography. *NeuroImage, 37*, 282–289. http://dx.doi.org/10.1016/j.neuroimage.2007.04.055.

Corr, P. J. (2013). Approach and avoidance behaviour: Multiple systems and their interactions. *Emotion Review, 5*, 285–290. http://dx.doi.org/10.1177/1754073913477507.

Craske, M. G., Hermans, D., & Vansteenwegen, D. (2006). *Fear and Learning: From basic processes to clinical implications*. Washington, DC: American Psychological Association.

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

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., … Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety, 32*, 239–253. http://dx.doi.org/10.1002/da.22353.

Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., & Hermans, D. (2015). Fear generalization in humans: Systematic review and implications for anxiety disorder research. *Behavior Therapy, 46*, 561–582. http://dx.doi.org/10.1016/j.beth.2014.10.001.

Dymond, S., Schlund, M. W., Roche, B., De Houwer, J., & Freegard, G. P. (2012). Safe from harm: Learned, instructed, and symbolic generalization pathways of human threat-avoidance. *PLoS One, 7*. http://dx.doi.org/10.1371/journal.pone.0047539.

Dymond, S., Schlund, M. W., Roche, B., & Whelan, R. (2014). The spread of fear: Symbolic generalization mediates graded threat-avoidance in specific phobia. *The Quarterly Journal of Experimental Psychology, 67*, 247–259. http://dx.doi.org/10.1080/17470218.2013.800124.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Foa, E. B., Zinbarg, R. E., & Rothbaum, B. O. (1992). Uncontrollability and unpredictability in post-traumatic stress disorder: An animal model. *Psychological Bulletin, 112*, 218–238.

Gillan, C. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Voon, V., Apergis-Schoute, A. M., … Robbins, T. W. (2014). Enhanced avoidance habits in obsessive-compulsive disorder. *Biological Psychiatry, 75*, 631–638. http://dx.doi.org/10.1016/j.biopsych.2013.02.002.

Glotzbach, E., Ewald, H., Andreatta, M., Pauli, P., & Mühlberger, A. (2012). Contextual fear conditioning predicts subsequent avoidance behaviour in a virtual reality environment. *Cognition & Emotion, 26*, 1256–1272. http://dx.doi.org/10.1080/02699931.2012.656581.

Grillon, C., Ameli, R., Woods, S. W., Merikangas, K., & Davis, M. (1991). Fear-potentiated startle in humans: Effects of anticipatory anxiety on the acoustic blink reflex. *Psychophysiology, 28*, 588–595.

Grillon, C., Baas, J. M. P., Cornwell, B., & Johnson, L. (2006). Context conditioning and behavioral avoidance in a virtual reality environment: Effect of predictability. *Biological Psychiatry, 60*, 752–759. http://dx.doi.org/10.1016/j.biopsych.2006.03.072.

Hartley, C. A., Gorun, A., Reddan, M. C., Ramirez, F., & Phelps, E. A. (2014). Stressor controllability modulates fear extinction in humans. *Neurobiology of Learning and Memory, 113*, 149–156. Retrieved from liz.phelps@nyu.edu.

Hayes, D. J., Duncan, N. W., Xu, J., & Northoff, G. (2014). A comparison of neural responses to appetitive and aversive stimuli in humans and other mammals. *Neuroscience and Biobehavioral Reviews, 45*, 350–368. http://dx.doi.org/10.1016/j.neubiorev.2014.06.018.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2003). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York: Guilford Press.

Hermann, C., Ziegler, S., Birbaumer, N., & Flor, H. (2002). Psychophysiological and subjective indicators of aversive pavlovian conditioning in generalized social phobia. *Biological Psychiatry, 52*, 328–337.

Hofmann, S. G., Alpers, G. W., & Pauli, P. (2009). Phenomenology of panic and phobic disorders. In M. M. Antony, & M. B. Stein (Eds.), *Oxford handbook of anxiety and related disorders* (pp. 34–46). New York, NY: Oxford University Press.

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

Johnson, D. D. P., Blumstein, D. T., Fowler, J. H., & Haselton, M. G. (2013). The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends in Ecology and Evolution, 28*, 474–481. http://dx.doi.org/10.1016/j.tree.2013.05.014.

Keinan, G. (1987). Scanning of alternatives under controllable and uncontrollable threats. *Journal of Personality and Social Psychology, 52*, 639–644. http://dx.doi.org/10.1037/0022-3514.52.3.639.

Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology, 84*, 437–450. http://dx.doi.org/10.1016/j.biopsycho.2009.10.007.

Lang, P. J., & Davis, M. (2006). Emotion, motivation, and the brain: Reflex foundations in animal and human research. *Progress in Brain Research, 156*, 3–29. http://dx.doi.org/10.1016/S0079-6123(06)56001-7.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., … Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The balloon analogue risk task (BART). *Journal of Experimental Psychology: Applied, 8*, 75–84. http://dx.doi.org/10.1037//1076-898X.8.2.75.

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

Lovibond, P. F., Mitchell, C. J., Minard, E., Brady, A., & Menzies, R. G. (2009). Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behaviour Research and Therapy, 47*, 716–720. http://dx.doi.org/10.1016/j.brat.2009.04.013.

Ly, V., & Roelofs, K. (2009). Social anxiety and cognitive expectancy of aversive outcome in avoidance conditioning. *Behaviour Research and Therapy, 47*, 840–847. http://dx.doi.org/10.1016/j.brat.2009.06.015.

Miller, N. E. (1959). Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation and social learning. In S. Koch (Ed.), *Psychology: A study of a science, study 1* (pp. 196–292). New York: McGraw-Hill.

Miller, W. R., & Rollnick, S. (2012). *Motivational interviewing: Helping people change*. Guilford Press.

Mowrer, O. H. (1960). *Learning theory and behavior*. New York, NY: Wiley.

Muris, P., & Field, A. P. (2010). The role of verbal threat information in the development of childhood fear. *"Beware the Jabberwock!" Clinical Child and Family Psychology Review*. http://dx.doi.org/10.1007/s10567-010-0064-1.

Olsson, A., & Phelps, E. A. (2004). Learned fear of "unseen" faces after Pavlovian, observational, and instructed fear. *Psychological Science, 15*, 822–828. http://dx.doi.org/10.1111/j.0956-7976.2004.00762.x.

Olsson, A., & Phelps, E. A. (2007). Social learning of fear. *Nature Neuroscience, 10*, 1095–1102. http://dx.doi.org/10.1038/nn1968.

Peterson, R. A., & Reiss, S. (1992). *Anxiety sensitivity index manual* (2nd ed.). Worthington, OH: International Diagnostic Systems.

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology, 46*, 984–995. http://dx.doi.org/10.1111/j.1469-8986.2009.00852.x.

Pittig, A., Alpers, G. W., Niles, A. N., & Craske, M. G. (2015). Avoidant decision-making in social anxiety disorder: A laboratory task linked to in vivo anxiety and treatment outcome. *Behaviour Research and Therapy, 73*, 96–103. http://dx.doi.org/10.1016/j.brat.2015.08.003.

Pittig, A., Brand, M., Pawlikowski, M., & Alpers, G. W. (2014). The cost of fear: Avoidant decision making in a spider gambling task. *Journal of Anxiety Disorders, 28*, 326–334. http://dx.doi.org/10.1016/j.janxdis.2014.03.001.

Pittig, A., Pawlikowski, M., Craske, M. G., & Alpers, G. W. (2014). Avoidant decision making in social anxiety: The interaction of angry faces and emotional responses. *Frontiers in Psychology, 5*, 1050. http://dx.doi.org/10.3389/fpsyg.2014.01050.

Pittig, A., Schulz, A. R., Craske, M. G., & Alpers, G. W. (2014). Acquisition of behavioral avoidance: Task-irrelevant conditioned stimuli trigger costly decisions. *Journal of Abnormal Psychology, 123*, 314–329. http://dx.doi.org/10.1037/a0036136.

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

Raes, A. K., De Houwer, J., De Schryver, M., Brass, M., & Kalisch, R. (2014). Do CS-US pairings actually Matter? A within-subject comparison of instructed fear conditioning with and without actual CS-US pairings. (J. C. Perales, Ed.). *PLoS One, 9*, e84888. http://dx.doi.org/10.1371/journal.pone.0084888.

Richards, J. M., Plate, R. C., & Ernst, M. (2013). A systematic review of fMRI reward paradigms used in studies of adolescents vs. adults: The impact of task design and implications for understanding neurodevelopment. *Neuroscience & Biobehavioral Reviews, 37*, 976–991. http://dx.doi.org/10.1016/j.neubiorev.2013.03.004.

Richter, J., Hamm, A. O., Pan-Farr, C. A., Gerlach, A. L., Gloster, A. T., Wittchen, H. U., … Arolt, V. (2012). Dynamics of defensive reactivity in patients with panic disorder and agoraphobia: Implications for the etiology of panic disorder. *Biological Psychiatry, 72*, 512–520. http://dx.doi.org/10.1016/j.biopsych.2012.03.035.

Riemer, M., Bublatzky, F., Trojan, J., & Alpers, G. W. (2015). Defensive activation during the rubber hand illusion: Ownership versus proprioceptive drift. *Biological Psychology, 109*, 86–92. http://dx.doi.org/10.1016/j.biopsycho.2015.04.011.

Robinson, O. J., Vytal, K., Cornwell, B. R., & Grillon, C. (2013). The impact of anxiety upon cognition: Perspectives from human threat of shock studies. *Frontiers in Human Neuroscience, 7*, 1–21. http://dx.doi.org/10.3389/fnhum.2013.00203.

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M., et al. (2016). The tipping Point: Value differences and parallel dorsal-ventral frontal circuits gating human approach-avoidance behavior. *NeuroImage, 136*, 94–105. http://dx.doi.org/10.1016/j.neuroimage.2016.04.070.

Sierra-Mercado, D., Deckersbach, T., Arulpragasam, A. R., Chou, T., Rodman, A. M., Duffy, A., … Dougherty, D. D. (2015). Decision making in avoidance–reward conflict: A paradigm for non-human primates and humans. *Brain Structure and Function, 220*, 2509–2517. http://dx.doi.org/10.1007/s00429-014-0796-7.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., & Vagg, P. R. (1983). *Manual for the state-trait anxiety inventory (STAI)*. Palo Alto, CA: Consulting Psychologist Press.

Talmi, D., Dayan, P., Kiebel, S. J., Frith, C. D., & Dolan, R. J. (2009). How humans integrate the prospects of pain and reward during choice. *The Journal of Neuroscience, 29*, 14617–14626. http://dx.doi.org/10.1523/JNEUROSCI.2026-09.2009.

Talmi, D., & Pine, A. (2012). How costs influence decision values for mixed outcomes. *Frontiers in Neuroscience, 6*, 146. http://dx.doi.org/10.3389/fnins.2012.00146.

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