Behaviour Research and Therapy 137 (2021) 103797

Contents lists available at ScienceDirect

Behaviour Research and Therapy

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

Incentive-based, instructed, and social observational extinction of avoidance: Fear-opposite actions and their influence on fear extinction

Andre Pittig [a,b,*], Alex H.K. Wong [a]

[a] Department of Psychology (Biological Psychology, Clinical Psychology, and Psychotherapy), University of Würzburg, Würzburg, Germany
[b] Center of Mental Health, University of Würzburg, Würzburg, Germany

ARTICLE INFO

Keywords:
Avoidance
Safety behavior
Fear extinction
Anxiety
Fear conditioning
Fear-opposite action

ABSTRACT

Avoidance is a transdiagnostic symptom of clinical anxiety and its reduction a major focus of cognitive-behavioral treatments. This study examined the instrumental extinction of goal-directed avoidance by means of incentives, verbal instruction, and social observation and their influence on fear extinction. Participants acquired conditioned fear and instrumental avoidance responses (N = 160). In four randomized groups, the reduction of avoidance by incentives for non-avoidance, instructions to refrain from avoidance, and social observation of non-avoidance was compared to no intervention before removing the aversive outcome. Conditioned fear when avoidance became unavailable subsequently was tested. Incentives, instruction, and observation all reduced avoidance better than no intervention, however, with different degrees and influence on conditioned fear. Incentives and instructions strongly reduced avoidance despite high levels of fear (i.e., fear-opposite actions). This initiated fear extinction, thereby reducing conditioned fear when avoidance became unavailable. Social observation directly reduced conditioned fear, presumably because it conveyed additional information about the absence of the aversive outcome. However, observation only moderately reduced avoidance and resulted in higher fear when avoidance became unavailable. The effects of social observation may depend on the nuances of the demonstrator's behavior. The clear effects of incentive and instructions provide support for clinical interventions to reduce avoidance during exposure therapy and can serve as experimental models for their controlled investigation.

1. Introduction

Pathological avoidance and safety behavior is a transdiagnostic symptom of clinical anxiety and related disorders (Craske et al., 2017). Although avoidance per se is an adaptive response to threat, pathological avoidance is out of proportion to actual threat and linked to serious impairments. Importantly, avoidance is not only a consequence of fear, but also contributes to the persistence and even increase of fear (Pittig et al., 2020). Clinically, persistent avoidance has, for example, been linked to the maintenance of anxiety, increasing daily impairments, and the development of secondary psychopathology (Beesdo et al., 2007; Craske et al., 2017; Wittchen et al., 2000, 2014). Research thus aims to uncover the mechanisms of pathological avoidance and interventions to reduce it.

In laboratory experiments, avoidance is typically examined as a goal-directed response to a conditioned warning signal (CS+; Krypotos et al., 2018; Pittig et al., 2020). After learning that a formally neutral CS is repeatedly followed by an aversive unconditioned stimulus (US), participants learn to perform a designated avoidance response that prevents the aversive outcome (e.g., button press). Persistent avoidance to a CS+ that is no longer followed by an aversive outcome has been shown to maintain threat beliefs, i.e., the expectancy that an aversive event will occur when avoidance is not performed (Lovibond et al., 2009; Pittig, 2019). This effect is commonly explained by protection from extinction: The absence of an aversive event is attributed to the avoidance response, thereby prohibiting novel inhibitory learning that the CS no longer predicts an aversive outcome irrespective of avoidance. Protection from extinction may help to explain how pathological avoidance contributes to the maintenance of clinical fear and anxiety. Research on how avoidance can be reduced may thus help to better understand and inform clinical interventions targeting anxious psychopathology.

Past experimental conditioning research has greatly contributed to this understanding by examining instrumental extinction of avoidance. Dymond (2019) suggested that traditional instrumental extinction

* Corresponding author. Department of Psychology I, University of Würzburg, Marcusstrasse 9-11, 97070, Würzburg, Germany.
E-mail address: andre.pittig@uni-wuerzburg.de (A. Pittig).

https://doi.org/10.1016/j.brat.2020.103797
Received 4 September 2020; Received in revised form 15 December 2020; Accepted 18 December 2020
Available online 8 January 2021

procedures such as removing the aversive outcome, making the aversive outcome ineliminable, or making the termination of the aversive outcome response independent make avoidance unnecessary and have thus been shown to partially reduce avoidance. We have, however, recently discussed that their prerequisites may be inherently difficult for clinical translation (Pittig et al., 2020). Some of these experimental procedures require control over the actual occurrence of the aversive outcome. This control may not always be feasible or ethical as a clinical intervention. For example, social rejection independent of avoidance may be punishing for a patient with social anxiety. Heart attack or suffocation, as feared in panic disorder, prohibits any controlled occurrence. Finally, removing the aversive event seems to have limited effects on the reduction of avoidance as low-cost avoidance tends to persist after removal of the aversive outcome (Lovibond et al, 2008, 2009; Pittig, 2019; Pittig & Dehler, 2019; Solomon et al., 1953; Vervliet & Indekeu, 2015). In sum, these instrumental extinction procedures may have limited clinical value.

Little research exists on alternative procedures for the extinction of avoidance. A series of studies recently investigated incentive-based reduction of avoidance. Corresponding paradigms link non-avoidance responses to a positive outcome, thereby establishing a conflict between avoidance and non-avoidance. In healthy individuals, avoidance responses are effectively reduced by a variety of incentives, for example, hypothetical and real monetary reward, positive social stimuli, or arbitrary points (Aupperle et al., 2011; Kirlic et al., 2017; Pittig, 2019; Pittig et al., 2018; Rattel et al., 2017; Sierra-Mercado et al., 2015; Talmi et al., 2009). More specifically, incentives reduce avoidance response but do not directly reduce conditioned fear (Pittig, 2019; Pittig & Dehler, 2019). In this regard, incentives for non-avoidance trigger fear-opposite actions, i.e., approach of the CS + despite high levels of fear. Fear-opposite actions in turn enable fear extinction learning as the non-occurrence of the aversive outcome is not associated to avoidance. This stepwise process that initiates fear extinction by reducing avoidance in the first place mimics the process of exposure therapy: Patients have to refrain from avoidance and confront their feared situation to enable fear extinction learning. Oftentimes, therapists will emphasize on the advantages of reducing avoidance for the patient's daily life. Therefore, incentive-based reduction of avoidance is a promising laboratory model to examine the mechanisms of pathological avoidance and its reduction under controlled conditions.

Besides learning by direct experience, past research also provided evidence for other learning pathways. For example, verbal instruction or social observation are indirect pathways for the learning of conditioned fear (see Rachman, 1977). Verbal instruction typically refers to instructing individuals about the CS-US contingency (e.g., "the CS will be followed vs. no longer be followed by a US"). For observational learning, individuals typically observe a confederate (i.e., the demonstrator) experiencing the CS-US contingencies. Past research showed that conditioned fear can be acquired through both verbal instruction and social observation, for example, indicated by an increase in US expectancy and skin conductance responses (SCRs) (Bublatzky et al., 2014; Javanbakht et al., 2017; Olsson & Phelps, 2004, 2007). Likewise, fear extinction can be acquired or facilitated by instructing or observing the CS-NoUS contingency (Dymond et al., 2012; Golkar et al., 2013; Javanbakht et al., 2017). Importantly, a recent study also provided evidence that avoidance can be acquired by verbal instruction and social observation (Cameron et al., 2016). After all participants learned that a CS+ was followed by an aversive US, participants either had to learn the correct avoidance response by trial and error (direct experience), were instructed about the correct response, or observed a demonstrator performing the correct response. All three groups of participants successfully acquired avoidance (Cameron et al., 2016). Thus, avoidance can be learned by means of direct experience, verbal instruction, and social observation. However, to the best of our knowledge, no study examined whether instructed and observational learning are also effective pathways for the extinction of avoidance. The current study therefore sought to examine whether these higher-order pathways could effectively reduce avoidance, and thereby reduce protection from extinction.

Importantly, previous research on instructed and observational learning of avoidance not only provided information about the behavioral response, but also about the corresponding CS-US contingency. For example, participants were instructed to perform a specific response *to cancel the upcoming US* or to observe the demonstrator performing the response *cancelling the upcoming US*. Thus, two types of information were delivered: i) behavioral information, i.e., which behavior to perform, and ii) contingency information, i.e. information about the occurrence of the US. In this regard, fear and avoidance processes are simultaneously targeted. Clinically, the distinct effect on instruction and social observation on the extinction of avoidance are, however, highly relevant. Inhibitory learning approaches highlight the mismatch between threat expectancy and actual outcomes as a mechanism underlying exposure therapy (Craske et al., 2014; Pittig et al., 2016). The rationale for this kind of exposure can be simplified as "stop avoiding to test whether your threat expectancy matches the actual outcome", i.e., patients are asked to refrain from avoidance and safety behavior to directly test whether their feared outcome actually occurs during exposure (instead of instructing them that their outcome will not occur). Testing a laboratory proxy that only provides behavioral information may help to better understand the impact of verbal instruction and social observation on extinction of avoidance and their clinical implications for developing an exposure rationale during treatment.

To this end, the present study examined incentive-based, instructed, and social observational extinction of acquired avoidance without explicitly mentioning the consequences of non-avoidance responses. In a single cue conditioning paradigm, participants learned that a formerly neutral CS was repeatedly followed by an aversive US (fear acquisition training). Next, participants learned to prevent the aversive US by performing a goal-directed avoidance response to the CS (US-avoidance acquisition training). Avoidance remained available in the subsequent phase, but no more USs occurred regardless of avoidance (US-avoidance extinction phase). This phase differed between four randomized groups. The *Incentive* group received a small monetary incentive for non-avoidance responses and participants had to learn the avoidance-reward contingencies by trial and error. The *Instruction* group was instructed to perform non-avoidance responses to test whether their US expectancy matches the actual outcome. This served as a laboratory proxy for the exposure rationale described above. The *Observation* group watched a demonstrator repeatedly performing non-avoidance responses. Importantly, all instructions and the video provided information on (non-) avoidance responses, but no explicit information about US occurrence was given. This served as a laboratory proxy for specific exposure exercises in which the therapist models the designated behavior (e.g., interoceptive exposures) or experience that other individuals show non-avoidance responses in daily life (e.g., taking the bus). Finally, the *No-Intervention* group served as a control group, in which participants did not receive specific intervention before extinction and were merely instructed to take a short break and keep paying attention. We expected a higher reduction of avoidance in the *Incentive*, *Instruction*, and *Observation* compared to the *No-Intervention* group. In addition, we expected that reduced avoidance would enable fear extinction in the former three groups as indicated by a decrease in US expectancy ratings and SCRs. In the final phase, the CS was presented alone and with avoidance being unavailable. Successful extinction learning was expected to result in a lower return of conditioned fear in this phase (i.e., less protection from extinction). As the present study is the first to examine the three interventions using the same paradigm, we did not have a priori hypothesis on potential differences between the intervention groups.

2. Methods

Preregistration and original data are available at https://osf.io/7en2h/ .

2.1. Participants

Overall, 160 participants were recruited from the University of Würzburg and the general community.[1] Before the experiment began, participants provided informed consent. All procedures were approved by the university's ethics committee. Exclusion criteria were current or history of psychosis, bipolar disorder, traumatic brain injury, intellectual disability, substance dependence, current use of psychotropic medication, any serious medical conditions, and pregnancy. Participants were randomized to four equally sized groups (Incentive vs. Instruction vs. Observation vs. No-Intervention). Groups did not differ in age, sex, trait anxiety, symptoms of depression, intolerance of uncertainty, or general risk-taking (see Table 1). Groups also did not differ in self-reported consumption of caffeine, nicotine, and alcohol, or amount of physical activity per week, $Fs < 1.51$, $ps > .15$, $BF_{01} > 3.79$.

2.2. Materials and procedures

After providing informed consent, electrodes for skin conductance assessment were attached and participants completed questionnaires on various individual differences that might influence task performance. Questionnaires assessed trait anxiety (anxiety facet of the NEO-PI-R; Costa & McCrae, 1992), symptoms of depression (PHQ-9; Kroenke et al., 2001), general risk taking (short-scale risk-taking-1; Beierlein et al., 2014), intolerance of uncertainty (Dugas et al., 2004; Gerlach et al., 2008), and basic sociodemographic data. Next, the US electrode was attached and the intensity of the aversive US was individually calibrated. The US was an electrical stimulation to the non-dominant forearm consisting of 125 consecutive 5-ms stimulations delivered through a bar-electrode. US intensity was increased in a stepwise manner until participant's rate the intensity as being "unpleasant and causing discomfort, but not painful". Groups did not differ in perceived unpleasantness of the last US delivered in the paradigm (see Table 1). Finally, participants completed the single cue fear and US-avoidance conditioning paradigm.

2.3. Single-cue fear and US-avoidance paradigm

The paradigm was based on a previous study (Pittig, 2019). It consisted of 44 trials subdivided into five consecutive phases (see Table 2): i) CS habituation, ii) fear acquisition training, iii) US-avoidance acquisition, iv) US-avoidance extinction, v) test. Only the specific manipulations before the US-avoidance extinction phase and trials during this phase differed between groups. In each trial, the same geometrical shape was presented for 8s as the CS. Inter-trial-intervals (ITIs) varied from 19 to 25s. We used a single-cue conditioning paradigm, because we recently showed an increase in low-cost avoidance to a conditioned safety stimulus (CS-) despite low level of conditioned fear (Wong & Pittig, 2020). In this study, CS + avoidance was closely linked to conditioned fear, which was not true for CS- avoidance. This finding suggests that other processes besides conditioned fear are involved in CS- avoidance, which would have biased group comparison in the present study.

2.3.1. Fear and US-avoidance acquisition

At the beginning, participants were instructed that geometrical shapes and aversive USs will be presented and that they should keep paying attention (no contingency instructions, see Mertens et al., 2020). During *CS habituation* (4 trials), the CS was presented without any outcome. *Fear acquisition training* (12 trials) consisted of three blocks in which three out of four CSs per block were followed by the US (i.e., 75% reinforcement). During *US-avoidance acquisition training* (8 trials), participants were instructed that they can prevent all outcomes of the upcoming CS by pressing an avoidance button (e.g., blue button) or not prevent any outcome by pressing a non-avoidance button (e.g., white button) on a dual-button box (counterbalanced). Participants had to decide which button to press at the beginning of each trial. The CS was presented regardless of the participant's response and participants were prompted to indicate their US expectancy after the non-/avoidance response. The US was omitted or delivered in line with the response to avoid versus not to avoid. US-avoidance was operationalized as pressing the avoidance button.

2.3.2. US-avoidance extinction

Before US-avoidance extinction, all groups received additional instructions, which occurred only once before the phase and differed between groups. The *No-Intervention* group was instructed that they may take a short break, continue when they are ready, and should keep paying attention. The *Incentive* group was instructed that they can win a small amount of money during the subsequent trials, three trials will randomly be selected, and the amount of rewards gained in these trials will be paid (see Pittig, 2019). These instructions aimed to provide an incentive for non-avoidance responses, however, participants had to learn the avoidance-reward contingencies by trial and error via feedback whether the reward was received or not at the end of each trial. The *Instruction* group was encouraged to perform non-avoidance response via the instruction "Please press the non-avoidance button to test your expectancy whether an electrical stimulus will occur". This instruction served as a laboratory proxy for the exposure rationale in exposure therapy. The *Observation* group was instructed that a short video will follow showing another participant who already completed the same procedures as they completed until now and that the part shown in the video will follow directly after the video. Next, a video was presented showing a confederate demonstrating non-avoidance responses (8 trials). In the video, participants saw the demonstrator from a left-diagonally behind perspective so that the demonstrators left side of the face, the button box in front of the demonstrator, and the computer screen were visible. The video showed the CS presentations on the computer screen and the demonstrator's non-avoidance responses in the same trial sequence as for the participant. The video was recorded with the same setup as used by the participant. Importantly, all instructions targeted avoidance, but no explicit information about US occurrence was given. Likewise, the video did not provide information whether a US will be delivered in the US-avoidance extinction phase (i.e., there was no visual feedback of US presentation).

During the subsequent *US-avoidance extinction phase* (12 trials), all participants still had to press the avoidance or non-avoidance button. However, no USs were presented irrespective of the response. In the *Incentive* group, a feedback of winning a fixed small reward was presented for 2s when participants pressed the non-avoidance button ("Gained reward: 0.10€" displayed as green text). When participants pressed the avoidance button, a feedback of missing the reward was presented ("Missed reward: 0.10€" in red text). In all other groups, the CS was presented without any outcomes.

2.3.3. Test

In the final *Test* (8 trials), the CS was presented in the absence of any outcomes and avoidance responses being unavailable in all groups. Thus, the final test phase tested for protection from extinction. At the end, participants provided subjective ratings of aversiveness of the last US and their motivation to avoid the US when imagining to continue the paradigm (0–100). Participants in the Incentive group also rated their motivation to approach the rewards (0–100). Participants of the Observation group were asked which button the demonstrator pressed and how likely a US was administered after the button press.

[1] Power analysis (power = .80, α error = 0.05, f = 0.25) indicated a total of 156 participants for the most critical between-subject factor in the 2 × 4 repeated measure ANOVA (planned protection from extinction analyses).

Table 1
Demographic and questionnaire data.

	Incentive (n = 40)		*Instruction* (n = 40)		*Observation* (n = 40)		*No-Intervention* (n = 40)		*F or* χ^2	p	η^2	BF_{01}
Females (%)	32	(80.0)	29	(72.5)	26	(65.0)	24	(60.0)	4.32[b]	.229		4.42
Age	24.58	(5.56)	25.35	(7.22)	24.45	(4.68)	23.90	(6.22)	0.40[a]	.755	0.01	19.26
Trait anxiety (NEO-PI-R-N1)	16.25	(5.02)	15.10	(5.07)	14.60	(5.01)	15.20	(5.89)	0.69[a]	.557	0.01	13.56
Intolerance of Uncertainty (IU)	40.48	(11.41)	40.65	(11.98)	40.65	(12.05)	40.40	(11.00)	<0.01[a]	1.00	<0.01	30.64
Depression (PHQ9)	4.78	(3.50)	4.55	(2.65)	4.98	(3.26)	5.65	(3.98)	0.79[a]	.503	0.02	12.26
Risk taking	4.00	(1.24)	3.80	(1.14)	4.08	(1.23)	4.10	(1.43)	0.46[a]	.708	0.01	17.81
Unpleasantness of last US	60.95	(23.96)	63.85	(22.60)	59.60	(26.76)	62.88	(24.47)	0.24[a]	.886	0.005	23.13

Note. Means (and standard deviations). NEO-PI-R-N1 = anxiety subscale of NEO-PI-R (Costa & McCrae, 1992); IU = Intolerance of Uncertainty Scale (Gerlach et al., 2008); PHQ-9 = Patient Health Questionnaire (Kroenke et al., 2001); Risk taking = Short-scale risk-taking-1 (Beierlein et al., 2014).
[a] $F(3,156)$.
[b] $\chi^2(3)$.

Table 2
Experimental design.

Group	CS habi-tuation	Fear acquisition	US-avoidance acquisition	Intervention	US-avoidance extinction	Test
No-Intervention	A - (4)	A + (9)	A [+] (8)	"*Keep paying attention*"	A [-] (12)	A - (8)
Incentive		A - (3)		"*You can gain rewards*"	A [-, €] (12)	
Instruction				"*Stop avoiding to test your threat expectancy*"	A [-] (12)	
Observation				Non-avoidance video of a confederate	A [-] (12)	

Note. A = CS (geometric shape), - = no US, + = US, […] = avoidance response available, € = fixed small reward for non-avoidance, number of trials is indicated in parentheses.

2.4. Indicators of conditioned fear: US expectancy ratings and SCRs

US expectancy ratings and SCRs to the CS were measured, as they are most commonly used as cognitive and physiological indicators of conditioned fear (Lonsdorf et al., 2017). For US expectancy ratings, participants indicated their subjective likelihood of US occurrence after the CS during every CS presentation (on a visual analog scale from 0% to 100%; see Pittig & Dehler, 2019). For trials in which avoidance responses were available, US expectancy was rated after participants performed the (non-)avoidance response.

Skin conductance was recorded on the hypothenar eminence of the non-dominant hand with two reusable Ag/AgCl electrodes and a constant voltage of 0.5 V using a V-Amp system (Brain Products, Germany; sampling rate = 1000 Hz). Data monitoring, acquisition, and parametrization was conducted with BrainVision Analyzer (Brain Products, Germany). A notch filter (50 Hz) and a 1 Hz FIR lowpass filter to remove high frequency noise was applied to the raw data. Biased response intervals (e.g., coughing, excessive movement) were time marked and excluded. SCRs were obtained with semi-automatic trough-to-peak scoring by calculating the maximum increase in skin conductance during CS presentation in comparison to the corresponding trough (see Boucsein et al., 2012). The square root was taken to obtain normal distribution (Dawson et al., 2007). Two participants in each group (5%) were excluded from SCR analyses due to technical failure.

2.5. Statistical analysis

Main analyses focused on group differences in the frequency of avoidance responses during the US-avoidance extinction phase and the level of conditioned fear during US-avoidance extinction and Test. The same variables were analyzed for CS habituation, fear acquisition training, and US-avoidance acquisition training to verify comparable levels between groups (manipulation check). For all measures, responses from two consecutive trials were averaged to reduce noise (see Pittig, 2019). Analyses were conducted within each phase using repeated measure ANOVAs (Group x Trial) because different trajectories were expected per phase. Moreover, a planned 2 × 4 repeated measure ANOVAs with factors Group and Trials was conducted to test for group differences in the increase of conditioned fear for the last trials of US-avoidance extinction (Trials 35–36) to the first test trials (Trials 37–38). This analysis was conducted to test for differences in protection from extinction. For all ANOVAs, Greenhouse-Geisser correction was applied whenever necessary. Follow-up analyses for significant main or interaction effects were conducted using Bonferroni-Holm corrected t tests or non-parametric U or W tests when assumptions of normal distribution were violated. Exploratory, we compared self-reported avoidance motivation when imagining to continue the paradigm using a one-way ANOVA.

In addition to frequentist analyses, Bayes Factor (BF) analyses were conducted (see van Doorn et al., 2020; Krypotos et al., 2017). BF_{10} is reported for comparing the probability of the data coming from the H1 (e.g., mean difference between groups is not zero) compared to the H0 (e.g., mean difference between groups is zero) and BF_{01} for the reversed comparison. Bayesian analyses with default priors and frequentist analyses were conducted in JASP (Version 0.13.1; JASP Team, 2020). In case of multiple factors in Bayesian ANOVAs, BFs refer to analyses of effects (across matched models) in which models including the effect are compared to equivalent models without the effect.

3. Results

US-avoidance, US expectancy, and SCRs are shown in Fig. 1.

3.1. Fear and US-avoidance acquisition

US expectancy and SCRs decreased across CS habituation trials and subsequently increased across fear acquisition training, main effect Trials: $Fs > 4.13$, $ps < .001$, $\eta s^2 > 0.006$, $BFs_{10} = 4.19$. There were no other main or interaction effects, $Fs < 1.43$, $ps > .126$, $\eta s^2 < 0.006$, $BFs_{01} = 8.71$.

During US-avoidance acquisition, all groups showed increasing US-avoidance responses to the CS+, main effect Trials: $F(3,156) = 5.44$, $p = .001$, $\eta^2 = 0.015$, $BF_{10} = 10.26$. No effects involving Group reached significance, $Fs < 0.46$, $ps > .903$, $\eta s^2 < 0.005$, $BFs_{01} > 30.81$. SCRs decreased across trials, main effect Trials: $F(3,147) = 29.20$, $p < .001$, $\eta^2 = 0.051$, $BF_{10} > 1000$. No effects involving Group reached significance, $Fs < 0.98$, $ps > .460$, $\eta s^2 < 005$, $BFs_{01} > 10.96$. No significant main or interaction effects were found for US expectancy, $Fs < 1.37$, $ps > 0.254$, $\eta s^2 < 0.003$, $BFs_{01} > 19.97$. Summarized, all groups acquired conditioned fear and US-avoidance without any group differences.

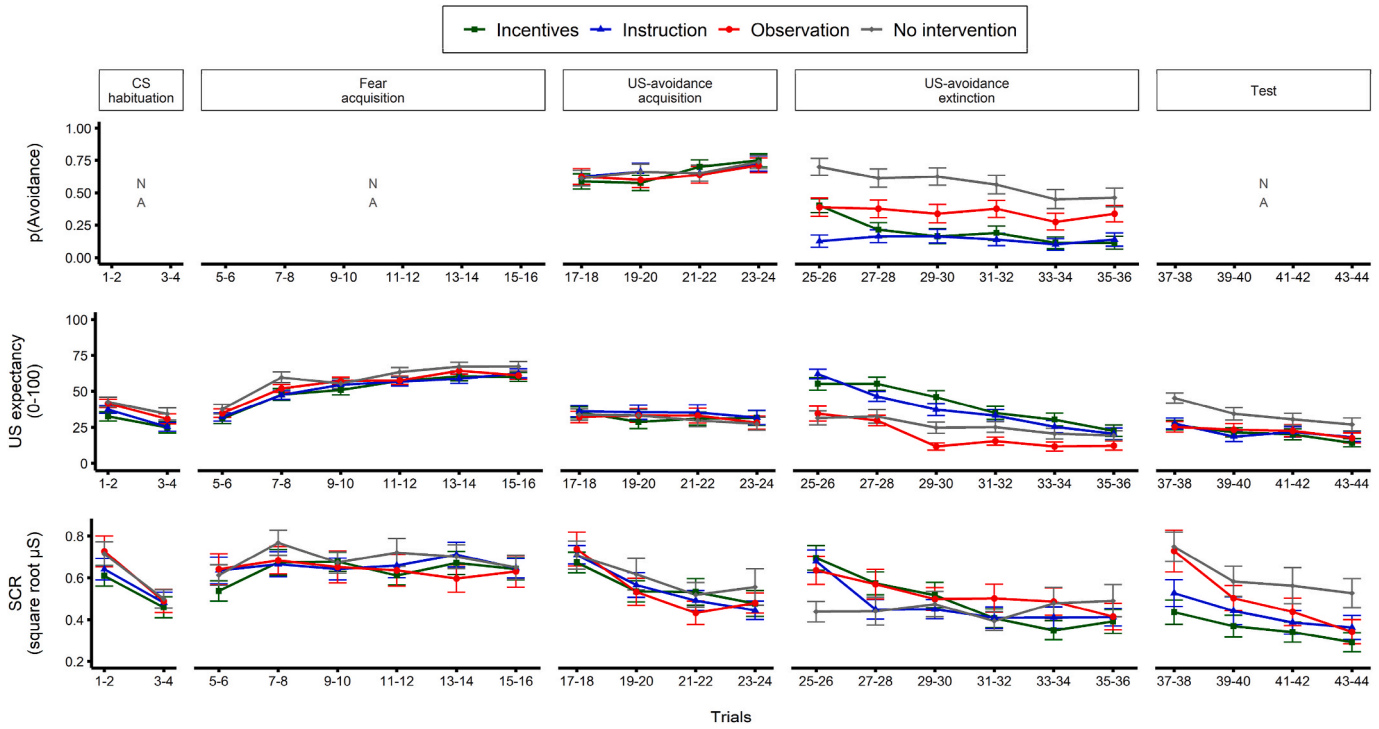

Fig. 1. Relative frequency of US-avoidance (top), US expectancy (middle), and SCRs (bottom) across phases (+/− standard error of the mean), averaged for two consecutive trials. NA = Avoidance responses not available during the phase.

3.2. Extinction of US-avoidance

US-avoidance. The main analyses of interest regarding avoidance showed that changes in US-avoidance across trials differed between groups, interaction Group x Trial: $F(15, 156) = 2.75$, $p < .001$, $\eta^2 = 0.011$, $BF_{10} = 21.04$. At the beginning (Trials 25–26), the Instruction group showed less frequent avoidance ($M = 0.13$; $SD = 0.29$) compared to all other groups, Incentive ($M = 0.40$; $SD = 0.34$): $U = 1162.0$, $p < .001$, $r = 0.45$, $BF_{10} = 102.11$; Observation ($M = 0.39$; $SD = 0.45$): $U = 546.5$, $p = .006$, $r = 0.32$, $BF_{10} = 13.47$; No-Intervention ($M = 0.70$; $SD = 0.41$): $U = 264.0$, $p < .001$, $r = 0.67$, $BF_{10} > 1000$. Moreover, the Incentive and Observation group showed less frequent avoidance compared to the No-Intervention group, $Us > 464.0$, $ps < .001$, $rs > 0.37$, $BFs_{10} > 20.98$, but did not differ from each other, $U = 841.0$, $p = .675$, $r = 0.05$, $BF_{01} = 4.27$. Across this phase, the Incentive and No-Intervention group showed a significant reduction in US-avoidance, Incentive: $F(5,39) = 14.20$, $p < .001$, $\eta^2 = 0.267$, $BF_{10} > 1000$; No-Intervention: $F(5,39) = 7.00$, $p < .001$, $\eta^2 = 0.152$, $BF_{10} > 1000$. In contrast, the Instruction and Observation group showed no change in US-avoidance across the phase, $Fs(5,39) > 1.04$, $ps > .396$, $\eta s^2 < 0.027$, $BFs_{01} > 20.08$. At the end (Trials 35–36), due to the relatively sharp decrease in avoidance in the Incentive group, the initial group difference between Incentive and Instruction group disappeared, $U = 764.0$, $p = .582$, $r = 0.05$, $BF_{01} = 4.07$, but both groups showed less frequent avoidance compared to the No-Intervention and Observation group, $Us > 570.0$, $ps < .027$, $rs > 0.29$, $BF_{10} > 3.15$. The Observation and No-Intervention group did not differ, $U = 685.5$, $p = .468$, $r = 0.14$, $BF_{01} = 2.07$.

In sum, instructions, incentives, and observation showed a strong reduction in avoidance compared to no intervention. Instructions strongly reduced avoidance from the beginning, which remained stable across trials. For incentives, avoidance was moderately reduced at the beginning but further decreased across trials to the same level as compared to instructions. Observation resulted in a moderate reduction of avoidance which did not further decrease across trials.

US expectancy. For US expectancy, change across trials also differed between groups, interaction Group x Trial: $F(15,156) = 3.50$, $p < .001$, $\eta^2 = 0.024$, $BF_{10} > 1000$. At the beginning (Trials 25–26), the Instruction ($M = 62.00$; $SD = 21.64$) and Incentive group ($M = 55.18$; $SD = 28.17$) showed higher US expectancy compared to the Observation ($M = 34.45$; $SD = 32.33$) and No-Intervention group ($M = 31.44$; $SD = 30.67$), $Us > 1105.0$, $ps < .018$, $rs > 0.38$, $BFs_{10} > 11.82$, but did not differ from each other, $U = 696.5$, $p = .642$, $r = 0.13$, $BF_{01} = 2.27$. The No-Intervention and Observation group did not differ from each other in Trials 25–26, $U = 850.5$, $p = .627$, $r = 0.06$, $BF_{01} = 3.98$. Across the phase, all groups showed a significant reduction of US expectancy, $Fs > 3.05$, $ps < .05$, $\eta s^2 > 0.072$, $BFs_{10} > 2.12$. This reduction was larger for the Instruction and Incentive group. At the end (Trials 35–36), there were no differences between all groups, $F(3,156) = 1.53$, $p = .208$, $\eta^2 = 0.029$, $BF_{01} = 5.01$.

SCRs. For SCRs, change across trials differed between groups, interaction Group x Trial: $F(15, 146) = 3.64$, $p < .001$, $\eta^2 = 0.022$, $BF_{10} > 1000$. Similar to US expectancy, the Instruction ($M = 0.68$; $SD = 0.33$) and Incentive group ($M = 0.69$; $SD = 0.37$) did not differ from each other at the beginning (Trials 25–26), $t(74) = 0.18$, $p = .862$, $d = 0.04$, $BF_{01} = 4.13$, but both showed higher SCRs compared to the No-Intervention group ($M = 0.44$; $SD = 0.30$), $ts > 3.27$, $ps < .002$, $ds > 0.76$, $BFs_{10} > 20.17$). Although the Observation group descriptively showed higher SCRs at the beginning, $M = 0.64$; $SD = 0.41$), SCRs did not significantly differ from any other group, $ts < 0.64$, $ps > .069$, $ds < 0.15$, $BFs_{01} > 3.51$. Whereas the No-Intervention group showed no changes across trials, $F(5,37) = 1.05$, $p = 0.388$, $\eta^2 = 0.028$, $BF_{10} = 18.42$, SCRs decreased in all other groups without any group differences, $Fs > 3.25$, $ps < .008$, $\eta s^2 > 0.081$, $BFs_{10} > 3.09$.

In sum, all groups showed extinction learning to different extents as indicated by US expectancy, whereas only the three intervention groups showed extinction learning as indicated by SCRs. Specifically, the Instruction and Incentive group showed high US expectancy and SCRs at the beginning, which decreased across trials, indicating successful fear extinction. Interestingly, the Observation group showed a similar pattern for SCRs, but US expectancy was significantly lower from the beginning.

3.3. Test and protection from extinction

US expectancy. For US expectancy, the transition from US-avoidance extinction to test differed between groups, $F(3,156) = 6.41$, $p < .001$, $\eta^2 = 0.031$, $BF_{10} = 67.78$ (planned 2 × 4 ANOVA for protection from extinction). While the No-Intervention and Observation group showed an increase in US expectancy, $Ws < 77.0$, $p < .001$, $r > 0.71$, $BFs_{10} > 51.86$, there was no change in the Incentive and Instruction group, $Ws < 261.0$, $p > .174$, $r > 0.26$, $BFs_{01} > 2.77$.

Moreover, change of US expectancy across test trials also differed between groups, interaction Group x Trial: $F(9,156) = 2.44$, $p = 0.010$, $\eta^2 = 0.007$, $BF_{10} = 1.22$. Importantly, the No-Intervention group showed higher US expectancy at the beginning compared to all other groups (Trials 37–38), $Us < 445.5$, $ps < .001$, $rs > 0.44$, $BFs_{10} > 33.11$. The remaining groups showed similar level of US expectancies, $Us < 856.5$, $ps > .590$, $rs < 0.07$, $BFs_{01} > 3.97$. All groups showed a reduction of US expectancy across test, $Fs > 3.30$, $ps < .023$, $\eta s^2 > 0.078$, $BFs_{10} > 1.51$. At the end (Trials 43–44), there were no group differences in US expectancy, $F(3,156) = 2.31$, $p = .079$, $\eta^2 = 0.042$, $BF_{01} = 2.01$.

SCRs. For SCRs, the transition from US-avoidance extinction to test differed between groups, $F(3,147) = 3.03$, $p = .031$, $\eta^2 = 0.015$, $BF_{10} = 2.46$ (planned 2 × 4 ANOVA for protection from extinction). Again, the No-Intervention and Observation group showed a significant increase across phases, $ts > 3.44$, $ps < .008$, $ds > 0.56$, $BFs_{10} > 22.18$, while there was no change in the Incentive and Instruction group, Instruction: $t(37) = 1.97$, $p = .112$, $d = 0.32$, $BF_{10} = 1.00$; Incentives: $t(37) = 1.04$, $p = .304$, $d = 0.17$, $BF_{10} = 3.42$.

Moreover, SCRs decreased across test trials, main effect Trial: $F(3,147) = 25.38$, $p < .001$, $\eta^2 = 0.041$, $BF_{10} > 1000$. There was a significant main effect of Group, $F(3,147) = 3.31$, $p = .022$, $\eta^2 = 0.045$, $BF_{10} = 2.33$, but no significant interaction, $F(9,147) = 1.55$, $p = .130$, $\eta^2 = 0.007$, $BF_{01} = 11.52$. The No-Intervention group showed higher SCRs compared to the Incentives group, $t(64) = 3.00$, $p = .019$, $d = 0.24$, $BF_{10} > 1000$. No other comparisons between groups were significant, $ts < 2.19$, $ps > .150$, $ds < 0.18$, $BFs_{10} < 2.18$. .

Summarized, the No-Intervention group showed a significant increase in conditioned fear when avoidance became unavailable, indicating protection from fear extinction. This increase was not found in the Incentives and Instruction group. Although the Observation group also showed an increase in US expectancy and SCRs, US expectancy did not differ from the Incentive and Instruction group during the first test trials.

3.4. Continued avoidance motivation

Self-reported motivation to avoid the US when imagining to continue the paradigm differed between groups (see Fig. 2), $F(3,156) = 3.17$, $p = .026$, $\eta^2 = 0.057$, $BF_{10} = 1.39$. The Incentive group reported lower avoidance motivation compared to all other groups, $Us > 591.0$, $p < .045$, $r > 0.26$, $BFs_{10} > 1.21$. The remaining groups did not differ, $Us < 870.5$, $ps > .499$, $rs < 0.09$, $BFs_{01} > 3.12$. In the Incentive group, participants reported a higher motivation to approach the rewards ($M = 79.75$, $SD = 27.54$) compared to avoid the US, $W = 597.5$, $p < .045$, $r = 0.70$, $BF_{10} = 42.23$.

After the task, all participant of the Observation group correctly indicated that the demonstrator pressed the non-avoidance button. Average rated likelihood of a US being administered after the demonstrator button press was $M = 25.7\%$ ($SD = 24.42$).

4. Discussion

This study examined incentive-based, instructed, and social observational pathways to the extinction of goal-directed avoidance and their influence on the extinction of conditioned fear. Main findings demonstrated that each of the three pathways was effective to reduce avoidance and thereby initiate fear extinction. However, the pathways reduced avoidance to different degrees. While incentives and instructions strongly reduced avoidance, social observation resulted in a moderate reduction of avoidance. As a result of the different degrees of avoidance, fear extinction gradients differed between pathways. Incentives and instructions did not directly reduce conditioned fear, but resulted in a gradual decrease of conditioned fear (i.e., fear opposite actions). In contrast, social observation directly reduced threat expectancy. When avoidance was no longer available, conditioned fear increased after social observation, but not after incentive-based and instructed extinction of avoidance. Finally, only incentives were linked to significantly lower self-reported avoidance motivation when imagining to continue the experimental paradigm. Combined, these findings highlight that incentive-based and instructional learning effectively trigger fear-opposite actions that enable extinction. Observational learning can convey additional information that more directly impact conditioned fear, but may also result in a stronger increase of fear when avoidance becomes unavailable.

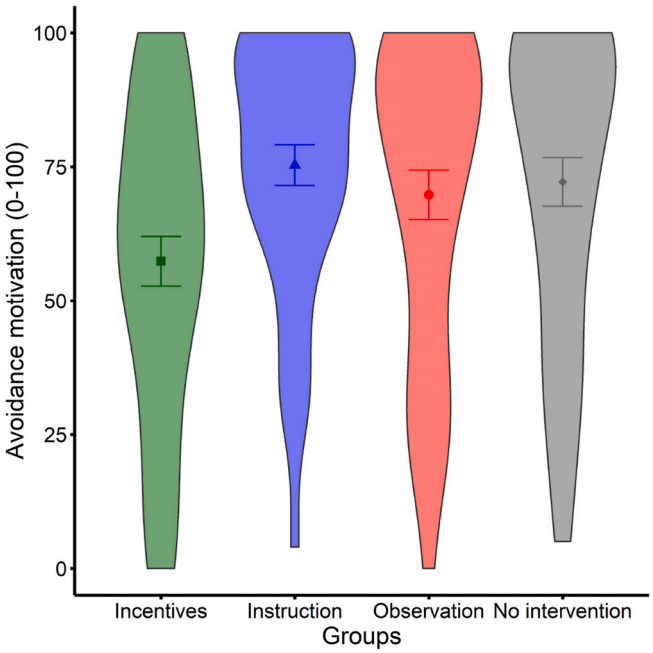

Fig. 2. Self-reported motivation to avoid the US when imagining to continue the paradigm.

Incentives, instruction, and social observation reduced avoidance more strongly compared to no intervention. The mere absence of an aversive outcome resulted in little reduction of avoidance. These findings are in line with recent studies showing that acquired avoidance tends to persist in the absence of an aversive outcome, which results in protection from extinction (Lovibond et al., 2009; Pittig, 2019; Pittig & Dehler, 2019; Vervliet & Indekeu, 2015). Although incentives, instruction, and social observation all resulted in lower avoidance, their degree and time course of avoidance reduction differed. Instructions immediately reduced avoidance, which remained stable across trials. Like the acquisition of avoidance (see Cameron et al., 2016), these findings demonstrate the effectiveness of verbal instruction for the extinction of goal-directed avoidance. Incentives initially reduced avoidance to a moderate degree, most likely because reward contingencies were still unclear and had to be learned by trial and error. After learning that non-avoidance was linked to rewards, incentives also decreased avoidance to the same degree as instructions. These findings replicate our previous findings on incentive-based extinction of avoidance (Pittig, 2019; Pittig et al., 2018). It seems likely that instructing participants about reward contingencies from the beginning would have resulted in an immediate elimination of avoidance. Thus, laboratory avoidance is best reduced by unambiguous interventions.

Incentives and instructions did not directly change fear. At the beginning of avoidance extinction, levels of conditioned fear were as high as at the end of fear acquisition. This suggests that incentives for non-avoidance and the instruction to stop avoiding did not convey information about the absence of the aversive outcome. Subsequently, threat expectancy and SCRs gradually decreased when individuals experienced that no more aversive outcome occurred. This reduction of avoidance cannot be explained by a prior reduction of conditioned fear in the US-avoidance acquisition phase. Instead, non-avoidance was performed *despite* high levels of conditioned fear. In other words, incentives and instructions triggered fear-opposite actions. These actions were necessary to enable fear extinction learning. These findings replicate our recent study showing the same stepwise process of incentives reducing avoidance and subsequently initiating fear extinction (Pittig, 2019; Pittig & Dehler, 2019). The present study expands this stepwise process to verbal instructions. Both interventions are thus effective to evoke non-avoidance despite high levels of fear.

Importantly, the clear reduction of avoidance by incentives and instructions prevented an increase of conditioned fear when avoidance responses became unavailable (i.e., during test). Without intervention, persistent high levels of avoidance after the removal of the aversive outcome were associated with a strong return of conditioned fear. These findings are consistent with past studies showing that persistent avoidance prohibits fear extinction (see Lovibond et al., 2009; Pittig, 2019). No return of conditioned fear was found following incentive-based and instructional extinction of avoidance. Interestingly, incentive-based extinction, but not instruction, was associated with a lower self-reported avoidance motivation when imagining to continue the paradigm. This finding may hint at an additional benefit of incentives to reduce the return of avoidance later on. Future research should test the effects of the different pathways on the return of avoidance. For instance, instructions may be susceptible for a return of avoidance responses considering their high level of avoidance motivation.

For social observation of non-avoidance behavior, the degree and time course of avoidance reduction and fear extinction differed. First, the frequency of avoidance was still moderate and higher compared to incentives and instruction. Second, threat expectancy, but not SCRs, was significantly lower at the beginning of US-avoidance extinction. This immediate lower threat expectancy could be due to two different pathways. First, the immediate lower threat expectancy may result from the moderate level of avoidance at the beginning of US-avoidance extinction. Second, although the video of the demonstrator did not provide explicit information about the absence of the aversive outcome, the demonstrator's responses may have conveyed such information, e.g., by not acting anxious. In support, participants in the Observation group rated a low likelihood of US administration after the demonstrator pressed the non-avoidance button (M = 25.7%). Such socially transmitted information has been shown to decrease conditioned fear (Golkar et al., 2013) and may thus have caused immediately lower threat expectancy. Despite this information, participants still exhibited a moderate level of avoidance. One potential explanation is that avoidance was a low-cost response in the current study, i.e., performing avoidance was not linked to cost or efforts (see Krypotos et al., 2018; Pittig et al., 2020). Participants may still engaged with a moderate level of low-cost avoidance despite low levels of US expectancy (e.g., a "why not" strategy; Wong & Pittig, 2020). In sum, the low level of conditioned fear across the US-avoidance extinction phase could be partially due to observational extinction learning and partially due to the moderate level of avoidance. Although avoidance in the Instruction group was also low-cost, instructions resulted in a strong decrease of avoidance. Explicit instructions to 'test one's threat expectancy' may thus have encouraged active disengagement from avoidance irrespective of its cost. As previous research highlighted that social observational learning depends on various characteristics of the demonstrator (e.g., skills, social group; Golkar et al., 2013; Navarrete et al., 2009; Olsson et al., 2005; Selbing & Olsson, 2017), future research may examine how these characteristics shape the extinction of avoidance.

Interestingly, there was an increase in both threat expectancy and SCRs in the Observation group when avoidance became unavailable. This increase of fear is not in line with previous research on socially transmitted fear extinction. For example, Golkar et al. (2013) showed that the return of differential SCRs was lower after observing a calmly acting demonstrator. This pattern could be due to the moderate level of avoidance in the Observation group, which may have limited extinction learning. As noted above, it also seems likely that participants inferred some information about the absence of the aversive outcome from the demonstrator's behavior. As the video did not explicitly convey this information, participants were to some degree uncertain about the actual contingencies. Combined with the moderate level of avoidance, this uncertainty may have resulted in a stronger increase of conditioned fear when avoidance became unavailable. However, the absolute level of threat expectancy during test was low and did not differ compared to incentives and instruction, thereby suggesting that the increase of conditioned fear in the Observation group was limited.

The present findings are in line with clinical interventions to reduce avoidance and safety behavior during exposure therapy. Novel approaches to exposure, such as the inhibitory learning model, highlight the elimination of avoidance and safety behaviors to maximize the violation of a patient's threat belief (Craske et al., 2014; Pittig et al., 2016). In this regard, it has been argued that threat expectancies should not be reduced by other interventions before exposure (e.g., cognitively challenging threat beliefs) to allow for maximization of expectancy violation (Craske et al., 2014). Our findings support the beneficial effects of this rationale on the reduction of avoidance and fear. Our result on incentive-based extinction of avoidance also support the use of incentive-based strategies to highlight positive consequences of exposure. Importantly, although instructional and incentive-based interventions are somewhat used in behavioral treatments (Heinig et al., 2017; Neudeck & Wittchen, 2012), their use is not well understood and can be optimized by proper understanding of the underlying mechanisms and moderators. For example, our findings indicate that incentives may have the potential to reduce avoidance motivation for subsequent encounters with a feared stimulus or situation. Such reduced avoidance motivation is especially important as avoidance tends to persist despite successful fear extinction and may trigger a return of fear (see Pittig et al., 2020). Incentives may thus be beneficial for long-term success of exposure. The present paradigm provides a controlled, laboratory-based proxy of these interventions and may also be useful to examine moderators and mediators of their avoidance reducing effect under controlled conditions. Future research may, for example, investigate the combination of instructional and incentive-based interventions and their long-term effects on the return of avoidance to further inform clinical implications. The present study used a one-day paradigm. Previous research highlighted the use of multiple day designs allow for consolidation between learning and intervention (Lonsdorf et al., 2017), which may be more suitable for clinical interventions to reduce fear and avoidance in the long-run. Future research on social observation of non-avoidance may help to disentangle the role of demonstrator characteristics on the degree of avoidance extinction, such as behaving anxiously or calmly or the demonstrator's trustworthiness. This research could better manipulate the specific information conveyed by the demonstrator to better understand the contribution of observed non-avoidance versus observed absence of the aversive outcome on the individual's own avoidance and fear responses.

In conclusion, the present study provides evidence for the extinction of avoidance by means of incentives, instruction, and social observation. Incentives and instructions evoked fear-opposite actions, thereby initiated fear extinction learning, preventing the increase of fear when avoidance was unavailable (i.e., prevented protection from extinction). Social observation of non-avoidance may convey additional information about the non-occurrence of the aversive event and thereby more directly reduce threat expectancy. However, it may also result in a

stronger return of fear when avoidance becomes unavailable. The effects of social observation thus depend more on the nuances of the demonstrator's behavior and characteristics.

CRediT authorship contribution statement

Andre Pittig: Conceptualization, Funding acquisition, Methodology, Preregistration, Formal analysis, Writing - original draft, Writing - review & editing, Visualization, Software, Supervision. **Alex H.K. Wong:** Conceptualization, Methodology, Writing - review & editing.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) grant PI1269/2-1 - 389569971 to AP. The authors have no conflicts of interest regarding this manuscript. The authors would like to thank Johanna Braun, Paulina Eckel, Charlotte Erlinghagen, and Martin Lorenz for their help with data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2020.103797.

References

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*(2), 455–463. https://doi.org/10.1016/j.bbr.2011.08.003.

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

Beierlein, C., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2014). *Eine Single-Item-Skala zur Erfassung von Risikobereitschaft: Die Kurzskala Risikobereitschaft-1 (R-1). GESIS-Working Papers*.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., et al. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology, 49*(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x.

Bublatzky, F., Gerdes, A. B. M., & Alpers, G. W. (2014). The persistence of socially instructed threat: Two threat-of-shock studies. *Psychophysiology, 51*(10), 1005–1014. https://doi.org/10.1111/psyp.12251.

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. https://doi.org/10.1016/j.jbtep.2015.06.003.

Costa, P. T., & McCrae, R. R. (1992). *Neo personality inventory-revised (NEO PI-R). Odessa, FL: Psychological Assessment Resources*.

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

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

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, G. G. Berntson, & Hrsg (Eds.), *Handbook of psychophysiology (3rd aufl., S. 159–181)*. Cambridge University Press.

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Hinne, M., et al. (2020). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon. Bull. Rev.*, 1–14.

Dugas, M. J., Buhr, K., & Ladouceur, R. (2004). *The role of intolerance of uncertainty in etiology and maintenance*.

Dymond, S. (2019). Overcoming avoidance in anxiety disorders: The contributions of Pavlovian and operant avoidance extinction methods. *Neuroscience & Biobehavioral Reviews, 98*, 61–70. https://doi.org/10.1016/j.neubiorev.2019.01.007. January.

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*(10). https://doi.org/10.1371/journal.pone.0047539.

Gerlach, A. L., Andor, T., & Patzelt, J. (2008). Die Bedeutung von Unsicherheitsintoleranz für die Generalisierte Angststörung Modellüberlegungen und Entwicklung einer deutschen Version der Unsicherheitsintoleranz-Skala. *Zeitschrift für Klinische Psychologie und Psychotherapie, 37*(3), 190–199. https://doi.org/10.1026/1616-3443.37.3.190.

Golkar, A., Selbing, I., Flygare, O., Öhman, A., & Olsson, A. (2013). Other people as means to a safe end: Vicarious extinction blocks the return of learned fear. *Psychological Science, 24*(11), 2182–2190. https://doi.org/10.1177/0956797613489890.

Heinig, I., Pittig, A., Richter, J., Hummel, K., Alt, I., Dickhöver, K., et al. (2017). Optimizing exposure-based CBT for anxiety disorders via enhanced extinction: Design and methods of a multicentre randomized clinical trial. *International Journal of Methods in Psychiatric Research, 26*(2), Article e1560. https://doi.org/10.1002/mpr.1560.

Jasp Team. (2020). *Jasp*.

Javanbakht, A., Duval, E. R., Cisneros, M. E., Taylor, S. F., Kessler, D., & Liberzon, I. (2017). Instructed fear learning, extinction, and recall: Additive effects of cognitive information on emotional learning of fear. *Cognition & Emotion, 31*(5), 980–987. https://doi.org/10.1080/02699931.2016.1169997.

Kirlic, N., Young, J., & Aupperle, R. L. (2017). Animal to human translational paradigms relevant for approach avoidance conflict decision making. *Behaviour Research and Therapy, 96*, 14–29. https://doi.org/10.1016/j.brat.2017.04.010.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine, 16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

Krypotos, A.-M., Blanken, T., Arnaudova, I., Matzke, D., & Beckers, T. (2017). A primer on Bayesian analysis for experimental psychopathologists. *Journal of Experimental Psychopathology, 8*(2), 140–157. https://doi.org/10.5127/jep.057316.

Krypotos, A.-M., Vervliet, B., & Engelhard, I. M. (2018). The validity of human avoidance paradigms. *Behaviour Research and Therapy, 111*, 99–105. https://doi.org/10.1016/j.brat.2018.10.011. May.

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., Heitland, I., Hermann, A., Kuhn, M., Kruse, O., Meir Drexler, S., Meulders, A., Nees, F., Pittig, A., Richter, J., Römer, S., Shiban, Y., Schmitz, A., Straube, B., … Merz, C. J. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience & Biobehavioral Reviews, 77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026.

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*(8), 716–720. https://doi.org/10.1016/j.brat.2009.04.013.

Lovibond, P. F., Saunders, J. C., Weidemann, G., & Mitchell, C. J. (2008). Evidence for expectancy as a mediator of avoidance and anxiety in a laboratory model of human avoidance learning. *The Quarterly Journal of Experimental Psychology, 61*(8), 1199–1216. https://doi.org/10.1080/17470210701503229.

Mertens, G., Boddez, Y., Krypotos, A.-M., & Engelhard, I. (2020). *Human fear conditioning depends on stimulus contingency instructions* [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/by964.

Navarrete, C. D., Olsson, A., Ho, A. K., Mendes, W. B., Thomsen, L., & Sidanius, J. (2009). Fear extinction to an out-group face: The role of target gender. *Psychological Science, 20*(2), 155–158.

Neudeck, P., & Wittchen, H.-U. (2012). *Exposure therapy: Rethinking the model—refining the method*. Springer.

Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The role of social groups in the persistence of learned fear. *Science, 309*(5735), 785–787.

Olsson, A., & Phelps, E. A. (2004). Learned fear of „unseen" faces after Pavlovian, observational, and instructed fear. *Psychological Science, 15*(12), 822–828. https://doi.org/10.1111/j.0956-7976.2004.00762.x.

Olsson, A., & Phelps, E. A. (2007). Social learning of fear. *Nature Neuroscience, 10*(9), 1095–1102. https://doi.org/10.1038/nn1968.

Pittig, A. (2019). Incentive-based extinction of safety behaviors: Positive outcomes competing with aversive outcomes trigger fear-opposite action to prevent protection from fear extinction. *Behaviour Research and Therapy, 121*, 103463. https://doi.org/10.1016/j.brat.2019.103463.

Pittig, A., & Dehler, J. (2019). Same fear responses, less avoidance: Rewards competing with aversive outcomes do not buffer fear acquisition, but attenuate avoidance to accelerate subsequent fear extinction. *Behaviour Research and Therapy, 112*, 1–11. https://doi.org/10.1016/j.brat.2018.11.003.

Pittig, A., Hengen, K., Bublatzky, F., & Alpers, G. W. (2018). Social and monetary incentives counteract fear-driven avoidance: Evidence from approach-avoidance decisions. *Journal of Behavior Therapy and Experimental Psychiatry, 60*, 69–77. https://doi.org/10.1016/j.jbtep.2018.04.002.

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

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

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

Rattel, J. A., Miedl, S. F., Blechert, J., & Wilhelm, F. H. (2017). Higher threat avoidance costs reduce avoidance behaviour which in turn promotes fear extinction in humans. *Behaviour Research and Therapy, 96*, 37–46. https://doi.org/10.1016/j.brat.2016.12.010.

Selbing, I., & Olsson, A. (2017). Beliefs about others' abilities alter learning from observation. *Scientific Reports, 7*(1), 16173. https://doi.org/10.1038/s41598-017-16307-3.

Sierra-Mercado, D., Deckersbach, T., Arulpragasam, A. R., Chou, T., Rodman, A. M., Duffy, A., et al. (2015). Decision making in avoidance–reward conflict: A paradigm for non-human primates and humans. *Brain Structure and Function, 220*(5), 2509–2517. https://doi.org/10.1007/s00429-014-0796-7.

Solomon, R. L., Kamin, L. J., & Wynne, L. C. (1953). Traumatic avoidance learning: The outcomes of several extinction procedures with dogs. *Journal of Abnormal and Social Psychology, 48*(2), 291–302. https://doi.org/10.1037/h0058943.

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. *Journal of Neuroscience, 29* (46), 14617–14626. https://doi.org/10.1523/JNEUROSCI.2026-09.2009.

Vervliet, B., & Indekeu, E. (2015). Low-cost avoidance behaviors are resistant to fear extinction in humans. *Frontiers in Behavioral Neuroscience, 9*, 351. https://doi.org/10.3389/fnbeh.2015.00351.

Wittchen, H.-U., Kessler, R. C., Pfister, H., Höfler, M., & Lieb, R. (2000). Why do people with anxiety disorders become depressed? A prospective-longitudinal community study. *Acta Psychiatrica Scandinavica*. https://doi.org/10.1111/j.0065-1591.2000.acp29[dash]03.x. Supplement.

Wittchen, H.-U., Knappe, S., Andersson, G., Araya, R., Banos Rivera, R. M., Barkham, M., Bech, P., Beckers, T., Berger, T., Berking, M., Berrocal, C., Botella, C., Carlbring, P., Chouinard, G., Colom, F., Csillag, C., Cujipers, P., David, D., Emmelkamp, P. M. G., … Schumann, G. (2014). The need for a behavioural science focus in research on mental health and mental disorders. *International Journal of Methods in Psychiatric Research, 23*(S1), 28–40. https://doi.org/10.1002/mpr.1409.

Wong, A. H. K., & Pittig, A. (2020). *A dimensional measure of safety behavior: Validation of a novel, non-dichotomous measure of avoidance in human fear conditioning [preprint]*. PsyArXiv. https://doi.org/10.31234/osf.io/n67h2.