Avoiding at all costs? An exploration of avoidance costs in a novel Virtual Reality procedure

Anke Lemmens [a],[*], Tom Smeets [a],[b], Tom Beckers [c], Pauline Dibbets [a]

[a] *Faculty of Psychology and Neuroscience, Department of Clinical Psychological Science, Maastricht University, the Netherlands*
[b] *CoRPS - Center of Research on Psychological and Somatic disorders, Department of Medical and Clinical Psychology, Tilburg School of Social and Behavioral Sciences, Tilburg University, the Netherlands*
[c] *Faculty of Psychology and Educational Sciences and Leuven Brain Institute, KU Leuven, Belgium*

ABSTRACT

Approach-avoidance behaviours play a major role in the development and maintenance of anxiety disorders as repeated avoidance behaviours are assumed to prevent fear extinction. Approach-avoidance decisions (Conditioned Stimulus (CS)-avoidance and Unconditioned Stimulus (US)-avoidance) and their effect on fear extinction and renewal were investigated using a novel Virtual Reality fear conditioning procedure that included avoidance costs that are relevant in real-life settings (i.e., temporal delay and physical effort). Participants had to choose between a safe (low approach incentive, no US) and risky stimulus (high approach incentive, US in 75 %). After differential fear acquisition and avoidance learning, participants were randomized to an Avoidance condition or No Avoidance condition during fear extinction. Fear extinction took place in either the original contingency learning context or in a new context and was followed by a renewal test. Furthermore, the influence of trait anxiety, distress tolerance, and intolerance of uncertainty on approach-avoidance decisions was investigated. Exploratively, a second experiment with varying avoidance costs was conducted. Results showed that high (Study 1), but not low (Study 2), avoidance costs resulted in less avoidance behaviour. Even though there were no between-group differences, exploratory comparisons of avoiders and non-avoiders in both studies demonstrated that avoidance behaviours protected from extinction learning, resulting in the maintenance of retrospective US expectancies and a sustained preference for the safe stimulus. Finally, no renewal effect and no robust associations with the individual difference measures were found. Collectively, these findings provide insight in how avoidance behaviours maintain fear and how treatment might be improved by focusing on avoidance costs.

1. Introduction

With an estimated lifetime prevalence of around 30 % and a 12-month prevalence of 18 %, anxiety disorders are among the most prevalent of all mental disorders (Kessler, Berglund et al., 2005; Kessler, Chiu, Demler, & Walters, 2005), imposing a significant economic impact on society (Greenberg et al., 1999; Gustavsson et al., 2011). Anxiety disorders are not only characterized by excessive fear and anxiety, but also by behavioural disturbances such as excessive avoidance of perceived or actual threatening situations, which often interferes with daily-life functioning of the anxious individual (American Psychiatric Association, 2013). For example, avoidance

prevents socially anxious individuals from going to parties and patients with a panic disorder from leaving their houses. While such avoidance behaviours may alleviate anxiety in the short run, they also help sustain anxiety as they preclude patients from learning that the anticipated threat (e.g., getting a heart attack when leaving the house) does not occur.

Pavlovian fear conditioning procedures are widely used in experimental research to investigate elements involved in the development and maintenance of anxiety disorders (Lissek & Grillon, 2010; Lissek et al., 2005; Mineka & Oehlberg, 2008). During fear conditioning, a neutral stimulus is paired with an aversive outcome (i.e., unconditioned stimulus; US). After several pairings, presentation of the neutral, now conditioned stimulus (CS), leads to an increased expectancy of the US and an associated conditioned fear response (CR). Exposure therapy, the treatment of choice for anxiety disorders, originated from the principles of Pavlovian fear conditioning, more specifically from extinction (Hermans, Craske, Mineka, & Lovibond, 2006; Rachman, 2009). In an extinction procedure, the CS is repeatedly presented in the absence of the aversive US, which results in a gradual reduction of acquired CRs (Vervliet, Craske, & Hermans, 2013). Correspondingly, in exposure treatment an anxiety disorder patient is repeatedly exposed to a feared stimulus (e.g., increased heart rate; CS) to learn that the expected catastrophe (e.g., heart attack; US) will not occur, leading to a reduction in fear levels (CR). Unfortunately, the success of current psychotherapeutic treatments such as exposure therapy is rather low and return of fear is pervasive, especially when returning to the original fear learning context (i.e., renewal; Craske, Hermans, & Vervliet, 2018; Hermans et al., 2006).

Craske, Treanor, Conway, Zbozinek, and Vervliet (2014) proposed several strategies to improve the efficacy of exposure therapy. One of them is to encourage patients to inhibit avoidance behaviours during exposure as they are known to interfere with extinction (i. e., *protection from extinction*, see Lovibond, Davis, & O'Flaherty, 2000; Lovibond, Mitchell, Minard, Brady, & Menzies, 2009; Rattel, Miedl, Blechert, & Wilhelm, 2017; van Uijen, Dalmaijer, van den Hout, & Engelhard, 2018; Xia, Dymond, Lloyd, & Vervliet, 2017). It is important to note that in the daily lives of anxiety patients, avoidance behaviours come with significant costs. For example, a panic disorder patient might risk his/her job because he/she may be frequently late or absent at work as he/she has to take time-consuming detours in order to avoid crowded places. The significant costs associated with avoidance behaviours versus the short-term relief in anxiety that they offer, result in a competition between the motivation to obtain a certain goal or reward and the motivation to avoid contact with the aversive stimulus or event (i.e., an approach-avoidance conflict; Forsyth, Eifert, & Barrios, 2006). Several studies included avoidance costs (or approach incentives) in order to investigate behavioural avoidance decisions (Aupperle, Sullivan, Melrose, Paulus, & Stein, 2011; Bublatzky, Alpers, & Pittig, 2017; Dibbets & Fonteyne, 2015; Pittig, 2019; Pittig, Brand, Pawlikowski, & Alpers, 2014; Pittig & Dehler, 2019; Pittig, Schulz, Craske, & Alpers, 2014; Rattel et al., 2017; Sierra-Mercado et al., 2015; Talmi, Dayan, Kiebel, Frith, & Dolan, 2009). In general, the presently available evidence suggests that although anxious individuals tend to avoid a feared stimulus, associating the feared stimulus with higher rewards or avoidance costs may foster approach motivation (Pittig & Dehler, 2019; Rattel et al., 2017).

Inducing a relevant approach-avoidance conflict poses a challenge to researchers interested in avoidance decisions. Studies often use hypothetical monetary rewards (Bublatzky et al., 2017; Pittig & Dehler, 2019; Pittig, Brand et al., 2014, Pittig, Schulz et al., 2014). However, hypothetical monetary rewards might not generalize to true monetary rewards and the loss of virtual money might not be a good match for the avoidance costs anxious individuals are typically confronted with (Pittig, Schulz et al., 2014). Contemporary Virtual Reality (VR) techniques may be ideally suited to manipulate avoidance costs in a way that optimizes ecological validity, for example by introducing personally relevant costs. In the current experiment, the effect of avoidance costs on approach-avoidance decisions was examined in a newly developed VR fear conditioning task in which avoidance did not only lead to a longer duration of the experiment (i.e., temporal delay-based costs; Rattel et al., 2017), but also required more physical effort to complete the experiment, mimicking the real avoidance costs of for instance taking a lengthy detour. Furthermore, VR is suited to investigate fear conditioning as it creates a high sense of immersion and has been shown capable of inducing behavioural and physiological fear (Dibbets & Fonteyne, 2015; Grillon, Baas, Cornwell, & Johnson, 2006).

We investigated whether participants preferred a risky stimulus (high approach incentive, 75 % chance of encountering the US) over a safe alternative (low approach incentive, no US). Note that this resulted in a trade-off between approach (high incentive) and avoidance (more time and effort needed to complete the task). Because the variability in avoidance responses between participants was low, we conducted a second experiment with a larger approach-avoidance conflict (i.e., smaller difference in costs between approaching and avoiding the risky stimulus) and exploratively compared the results of the two studies. In study 1, participants who avoided the risky stimulus missed out on five coins and received one coin when selecting the safe stimulus (high avoidance cost), requiring more time and physical effort to complete the experiment. In study 2, participants who avoided the risky stimulus missed out on five coins and received three coins when selecting the safe stimulus (low avoidance cost). Furthermore, there was one group that could use an avoidance button during extinction, whereas for the rest of the participants the avoidance button was not functional, providing participants in the avoidance group with another option to avoid the US next to stimulus selection. Note that in the current study we operationalised avoidance decisions as both CS-avoidance (i.e., stimulus choice) and US-avoidance (i.e., avoidance button), which have different effects and relevance in clinical settings (Krypotos, Vervliet, & Engelhard, 2018). The rationale behind providing participants with an escape option and a time-consuming safe alternative is because in daily life, anxious individuals also often have a choice between, on the one hand, escaping a situation whenever fear becomes too high or, on the other hand, avoiding the feared situation and choosing an alternative option in advance. We were therefore interested in the effect of avoidance costs on both types of avoidance decisions. It was also investigated whether avoidance behaviours during the extinction procedure impeded extinction learning.

Furthermore, we aimed to investigate renewal of avoidance responses. It is generally acknowledged that during extinction the original CS-US association is not erased, but instead a new inhibitory CS-noUS association is formed (Bouton, 1993; Craske et al., 2014; Rescorla, 1996), remaining the original association existent. Assuch, the original CS-US association may become reactivated by a

switch in context and cause a return of fear (Bouton, 1993, 2002; Vervliet, Craske et al., 2013). Although research on the renewal of fear has demonstrated a robust effect, studies investigating the renewal of avoidance behaviours are sparse. To our knowledge, only two studies investigated approach-avoidance decisions after a context switch (Schlund et al., 2019; Urcelay, Symmons, & Prével, 2019). Both studies demonstrated that extinguished avoidance responses reoccurred when shifting from the extinction context to the acquisition context. Finally, we explored relationships between approach-avoidance decisions and individual difference measures that have been previously shown to be relevant in fear and avoidance conditioning studies (Intolerance of uncertainty, Trait anxiety, Distress tolerance: Morriss, Christakou, & Van Reekum, 2015; Pittig, Schulz et al., 2014; Vervliet, Lange, & Milad, 2017).

It was hypothesized that the use of avoidance behaviours during extinction would result in protection from extinction. Therefore, we expected highest US-expectancy ratings and more avoidance decisions at test in response to the extinguished CS if the avoidance button was available during extinction or if participants avoided the anticipated US by choosing the safe option. Second, only two studies investigated renewal in the context of behavioural avoidance (Schlund et al., 2019; Urcelay et al., 2019). Based on their findings, we expected that participants who were presented with the extinguished CS in a different context as where extinction took place would show highest US-expectancy ratings and most avoidance responses. Finally, we expected that higher avoidance costs would lead to less avoidance behaviours (choosing the safe alternative or pressing the avoidance button) when comparing the two studies.

2. Method study 1

2.1. Participants

A total of 90 individuals (aged 17–58, $M = 20.07$, $SD = 4.38$), 14 males and 76 females, participated in the current study. Participants were recruited through advertisements at the university billboards and an online research participation environment. Exclusion criteria were self-reports of psychological treatment for an anxiety disorder at the time of or within six months before participation and diagnosis of epilepsy. Before the start of the experiment participants signed a written informed consent and were pseudo-randomly allocated to one of three experimental groups (with equal male-to-female ratio in each group). After completion of the experiment participants were compensated with 1.5 course credits. The study was approved by the Ethics Review Committee Psychology and Neuroscience at Maastricht University (ERCPN-187_10_01_2018) in accordance with the Declaration of Helsinki. The design and analyses of the study were preregistered on the Open Science Framework (VRavoidance1, osf.io/xe9b5). Some aspects of the design and analyses have been adjusted based on progressive insight. All deviations from the preregistration are described in the Supplementary material.

2.2. Materials

2.2.1. Virtual Reality (VR)

The study was conducted in the VR lab of Maastricht University, a room of 6 by 4 m. Participants could freely walk wearing

Fig. 1. Screenshots from the VR fear conditioning task. Top row (from left to right): context of the practice phase, and the two experimental contexts. Bottom row (from left to right): trial with avoidance button present; trial with avoidance button used; example of retrospective US-expectancy rating.

headphones, a head mounted display (HMD, HTC Vive®, Seattle, USA) that provides a 3D stereoscopic view, and a controller (HTC Vive controller®, Seattle, USA). The virtual environment automatically adjusted to the motions of the participants. The task was programmed using Unity®.

For the purpose of our renewal paradigm two different virtual reality environments (A and B) were created in which it would be natural to encounter computer screens. One environment represented an aisle in a consumer electronics store, whereas the other represented a computer lab. In both environments one screen was placed on one side of the room and three other screens were placed next to each other on the opposite side of the room. The VR environments were counterbalanced across participants. Fig. 1 provides images of the fear conditioning task in VR.

2.2.2. Stimuli fear conditioning

Three coloured screens served as the CS during our experimental task: a yellow, blue, and orange screen. The role of the CS was counterbalanced across participants. The US was a 9-s film fragment showing a screaming female cook who accidentally poured a pan with boiling water over herself. The film clip consisted of a 4-s neutral part (no accident, no scream) to allow for avoidance decisions and a 5-s aversive part. The peak volume was set at 95 dB to increase the aversiveness of the US. In one of our previous studies, this film clip was experienced as highly unpleasant and was capable of inducing conditioned fear responses (Dibbets, Lemmens, & Voncken, 2018). Both the CSs (colours) and the US (film clip) were presented on the computer screens in VR.

2.2.3. Fear conditioning task

The current study adopted a new free-choice fear conditioning procedure in VR in which avoidance costs were manipulated. We informed participants that they had to collect a prefixed amount of coins in each phase and were allowed to leave as soon as the task was finished. The sole purpose of introducing the coins was to implement delay-based costs and physical effort costs in the task. By receiving different amounts of coins for the different stimuli, participants were faced with a trade-off between avoidance of the aversive stimulus and duration of the experiment (and the corresponding physical effort required to finish the task). As such the coins were a marker of the participant's progress in terms of reaching the prefixed amount of coins and might have motivated participants to approach the risky stimulus. Both the delay costs and the movement effort contribute to the ecological validity of the task, because both
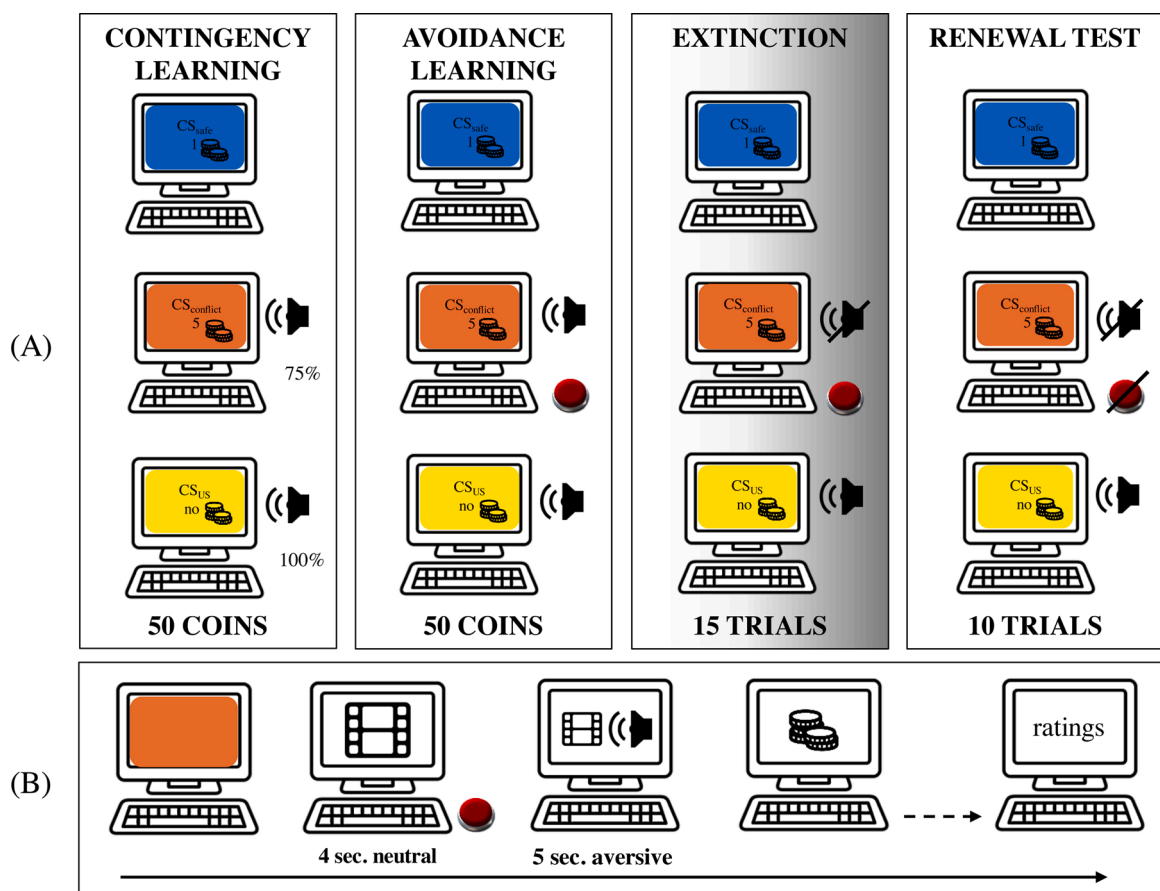

Fig. 2. Overview of the VR fear conditioning paradigm (Study 1). (A) Overview of stimuli and experimental phases. (B) Overview of a single trial during the avoidance learning phase. Not pressing the avoidance button in the first 4 s of the film clip results in the US and 5 coins.

costs are also involved in the trade-off anxiety patients make.

The experimental task consisted of a practice phase, contingency learning phase, avoidance phase, extinction phase, and a renewal phase. The conditioning task was carried out in one day, similar to studies with comparable designs (Rattel et al., 2017; Vervliet and Indekeu, 2015). Fig. 2 presents an overview of the fear conditioning task. The detailed task instructions are provided in the Supplementary material.

2.2.3.1. Practice phase. The practice phase took place in a blank room with a main screen and three black-and-white coloured screens (dark grey, light grey, white). In order to familiarize participants with acting in a VR environment, they completed four unreinforced trials and ratings with the controller.

2.2.3.2. Contingency learning phase. The contingency learning phase took place in context A. During this phase participants were free to approach one of the three coloured screens. Selection of one colour resulted in one coin and absence of the US (CS_{safe}), selection of another colour resulted in presentation of the US on 75 % of the trials and five coins on all trials ($CS_{conflict}$), and selection of a third colour always resulted in presentation of the US and no coins (CS_{US}). As it would be highly disadvantageous to choose the latter stimulus, it served as a control stimulus that would inform us about whether or not participants understood the task. To make sure that participants would explore the $CS_{conflict}$ option, the first two $CS_{conflict}$ trials ended in receiving five coins but no US. Selection of a screen with the controller led to CS offset and US onset (in case of a reinforced trial). Location of the CSs varied between trials, in all phases. Participants were instructed to obtain 50 coins, therefore the amount of trials depended on the decisions of the participants.

2.2.3.3. Avoidance learning phase. During this phase (in context A), participants were again free to approach either of the three coloured screens. In case they selected the $CS_{conflict}$ a red avoidance button appeared next to the screen that allowed participants to terminate the film clip during the neutral first 4 s of the film clip. Pressing the avoidance button changed the colour of the button into green to indicate its functionality. Use of the button resulted in withholding of the coins. The amount of trials of this phase was dependent on participants' avoidance decisions. Note that from this phase onwards, participants were able to choose between two types of avoidance responses: CS-avoidance (stimulus choice) or US-avoidance (avoidance button).

2.2.3.4. Extinction phase. The extinction context varied depending on group assignment. Participants in one group (AAAA) underwent extinction in the contingency learning and avoidance learning context, whereas for the other two groups extinction took place in a new environment (AABA). In all groups, selecting the $CS_{conflict}$ resulted in the neutral part of the film clip to allow for avoidance decisions, but the aversive part never followed. Moreover, groups differed in the availability of the avoidance button: In one group the avoidance button was available (AABA-Avoidance), whereas for the other groups the button was also present, but not functional (AAAA-No Avoidance and AABA-No Avoidance). Use of the avoidance button in the avoidance group changed the colour of the button into green and resulted in withholding of coins. In the No Avoidance groups, pressing the button did not result in a green activated button and a withholding of coins. To guarantee sufficient extinction trials, the extinction phase consisted of 15 trials for all participants. However, participants were not informed about this change and were instructed to "*collect coins*".

2.2.3.5. Renewal phase. The renewal phase took place in context A in all groups. On the first five trials the avoidance button was present (but not functional) when selecting the $CS_{conflict}$ for all participants. On the next five trials the avoidance button was not present. We included two separate phases in the renewal phase for the following reasons: First, trial 1–5 (with the avoidance button present) were included in order to be able to measure button use. Second, we included trial 6–10 in order to measure retrospective US expectancies and behavioural responses when the button was absent, because we expected that the button would function as a discriminative stimulus that would inform participants that they could avoid the US. As trials without the avoidance button were most similar to trials during the contingency learning phase, an increase in retrospective US expectancies and CS_{safe} selections was anticipated. Similar to the extinction phase, only the neutral part of the film clip was displayed. The order of trials with and without avoidance button was not counterbalanced, because trials without the avoidance button could be considered as additional extinction trials, potentially influencing behaviour on consecutive trials with the avoidance button.

2.3. Questionnaires

The current study explored associations between approach-avoidance decisions and individual difference measures that have previously been shown to be relevant in fear and avoidance conditioning studies: the State-trait anxiety inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970), the Distress Tolerance Scale (DTS; Simons & Gaher, 2005), and the Intolerance of Uncertainty Scale (IUS; Freeston, Rhéaume, Letarte, Dugas, & Ladouceur, 1994). For more details about the questionnaires see Supplementary material.

2.3.1. Stimulus ratings

2.3.1.1. US ratings. During the practice phase and after the retention phase, participants were asked how they felt about the film clip in order to check for habituation effects. Participants rated the pleasantness (unpleasant – pleasant), valence (negative – positive), and evoked arousal (not startled at all – very startled) on a 7-point scale ranging from −3 to 3.

2.3.1.2. Button effectiveness rating.. As a manipulation check, participants rated whether they thought the avoidance button was effective in terminating the film clip on a 7-point scale ranging from −3 to 3 (never – always) after the avoidance phase and the renewal phase.

2.3.1.3. Retrospective US expectancy.. Participants rated US expectancy for each CS retrospectively at the end of all experimental phases on a 7-point scale ranging from -3 to 3 (never – always).

2.4. Procedure

After reading information about the procedure of the study and signing informed consent, participants filled in the STAI, DTS and the IUS. Next, the loud scream that was part of the US was presented via headphones (volume 95 dB). If the sound was experienced as painful, the volume level was adjusted. Subsequently, participants were equipped with the head mounted display and controller and the practice phase started. After the four practice trials and one presentation of the US, participants completed the practice ratings and the first US rating.

Participants were informed via the screen that the goal of the contingency learning phase was to obtain 50 coins. They were told that every colour had its own reward scheme (combination of coins and presence or absence of the film clip) and it was their task to discover the relationships. Every trial, participants had to walk towards one of the screens and click on it. The chance that participants avoided watching the US by walking away from the screen was minimized by the fact that coins had to be collected by clicking on the screen after the US was displayed. At the end of every trial participants had to walk back to the main screen to start the next trial. The contingency learning phase concluded with retrospective US-expectancy ratings.

The avoidance learning phase started with the instruction that participants had to collect 50 coins again. In every phase of the experimental task it was emphasized to use the information they had learned during previous phases. Specifically, for this phase of the experiment, participants were informed about the red button. The avoidance learning phase ended with retrospective US-expectancy ratings and the button effectiveness rating.

The extinction phase and renewal phase followed a similar procedure. However, during these phases no specific amount of coins that had to be obtained was stated, as the amount of trials was fixed. The extinction phase concluded with retrospective US-expectancy ratings. At the end of the renewal phase, participants completed retrospective US-expectancy ratings and another US and avoidance effectivity rating. After finishing the experimental task in VR, participants returned the equipment and were compensated for their participation. They were debriefed after completion of the study.

2.5. Data analysis

Participants who were not aware of the contingencies between the CSs and the US during and after the contingency learning phase were excluded from the analyses as extinction and renewal are redundant when no contingency learning has occurred. Participants were excluded if they selected both the CS_{safe} and CS_{US} on more than 15 % of the trials yet rated the retrospective US expectancy of the CS_{safe} higher than that of the CS_{US}. Likewise, participants were excluded if they selected both the CS_{safe} and $CS_{conflict}$ on more than 15 % of the trials yet rated the retrospective US expectancy of the CS_{safe} higher than that of the $CS_{conflict}$. Note that the 15 % selection condition was added after exclusion criteria were preregistered (see Supplementary material), because we realized that retrospective US-expectancy ratings were only reliable if participants had enough opportunity to learn about the contingencies. As there were no previous studies with a similar design, this condition was based on the assumption that if participants selected both stimuli at least two times (> 15 %), they should be able to distinguish the unreinforced stimulus (CS_{safe}) from the reinforced stimuli ($CSc_{conflict}$ and CS_{US}). Furthermore, we also added the exclusion criterion that participants were excluded from the analyses when they selected the CS_{US} more than once during the final five trials of the contingency learning phase, because this would indicate that participants either did not understand the task or did not participate in a serious manner, as there were only negative consequences of choosing this stimulus.

After excluding participants that were contingency unaware, a randomization check was performed, comparing demographic variables, STAI-T, DTS, and IUS scores, and the first US ratings using ANOVAs and χ^2-tests. GLM repeated measures ANOVAs were used to check for habituation to the US. After the checks, retrospective US-expectancy data, stimulus choice and use of the avoidance button were analysed for each phase separately. Univariate ANOVAs and GLM repeated measures ANOVAs with Group (AAAA-No Avoidance; AABA-No Avoidance; AABA-Avoidance) as between-subjects factor and Stimulus (CS_{safe}, $CS_{conflict}$, CS_{US}) and Time (Phase or Block) as within-subjects factors were used. The standard rejection criterion was set at $p < .05$ throughout. In case of multiple comparisons, Bonferroni corrections were applied. If Mauchly's test indicated that the assumption of sphericity had been violated, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. In case of an interaction effect, main effects are not reported. Significant (interaction) effects were followed up with pairwise comparisons or post-hoc analyses (e.g., separate GLM repeated measures per stimulus). In order to investigate renewal, responses (i.e., stimulus choice and use of avoidance button) on the final extinction trial and the first renewal trial were compared between groups using Fisher's exact tests.

3. Results study 1

3.1. Included sample

Two female participants dropped out during the experiment, one due to equipment failure and one that indicated that the US was too aversive. Thirteen participants (12 female) were labelled as contingency unaware based on the abovementioned criteria. Hence, the total sample consisted of 75 participants. Demographic information of the included sample and questionnaire data are presented in Table 1. Groups did not differ in sex ratio (χ^2 ($N = 75$) = 1.72, $p = .42$), age, STAI-T, DTS and IUS scores ($Fs < 2.70$, $ps > .07$). Participants did not differ in their evaluations of the US presented once before the contingency learning phase ($Fs < 1.76$, $ps > .18$). They evaluated the US as unpleasant ($M = -2.07$; $t(74) = -22.07$, $p < .001$), negative ($M = -2.21$; $t(74) = -22.32$, $p < .001$) and reported to feel moderately startled by the film clip ($M = .75$; $t(74) = 3.93$, $p < .001$). Three repeated measures ANOVAs revealed that the US ratings did not significantly differ before the contingency learning phase and at the end of the experiment ($Fs < 3.23$, $ps > .08$), indicating that participants did not habituate to the US after repeated presentations.

3.2. Fear conditioning

3.2.1. Retrospective US expectancy
The pattern of retrospective US expectancies over the different phases of the experiment is presented in Fig. 3.

3.2.1.1. Contingency learning. The GLM repeated measures detected a significant Stimulus effect ($F(1.57,112.87) = 98.13$, $p < .001$, $\eta_p^2 = .58$). Pairwise comparisons showed that after contingency learning the CS_{safe} received lower retrospective US-expectancy ratings than the $CS_{conflict}$ and CS_{US} ($ps < .001$), indicating successful differential fear acquisition. There was no significant difference between $CS_{conflict}$ and CS_{US} ($p = .74$). If anything, in contrast to the reinforcement schedule, participants expected the US more after $CS_{conflict}$ ($M = 1.50$) than after CS_{US} ($M = 1.17$). This might be explained by the fact that the $CS_{conflict}$ was selected most often (54 %) and the CS_{US} in only 13 % of the trials, making differential learning for the $CS_{conflict}$ and the CS_{US} harder. As expected, the main effect of Group and the Group*Stimulus interaction were not significant ($Fs < .74$, $ps > .54$).

3.2.1.2. Avoidance learning. The GLM revealed a significant Stimulus effect ($F(1.46,105.44) = 153.99$, $p < .001$, $\eta_p^2 = .68$). The main effect of Group and the Group*Stimulus interaction were not significant ($Fs < 2.40$, $ps > .07$). Pairwise comparisons showed that retrospective US-expectancy scores of all stimuli differed significantly ($ps < .03$) and followed the pattern of the contingency learning phase, $CS_{safe} < CS_{US} < CS_{conflict}$.

US-expectancy ratings after the avoidance phase were compared with US-expectancy ratings after contingency learning. There was a significant Phase*Stimulus interaction ($F(1.71,123.39) = 3.95$, $p = .03$, $\eta_p^2 = .05$), caused by a significant decrease in retrospective US expectancy of CS_{safe} ($F(1,74) = 11.10$, $p = .001$, $\eta_p^2 = .13$) and an absence of change for the remaining stimuli ($Fs(1,74) < 2.32$, $ps > .13$). The Group effect and all possible interactions with Group were not significant ($Fs < 1.97$, $ps > .13$).

3.2.1.3. Extinction. The GLM revealed a significant Stimulus effect ($F(1.86,133.84) =$

77.11, $p < .001$, $\eta_p^2 = .52$). The main effect of Group and the Group*Stimulus interaction were not significant ($Fs < 1.23$, $ps > .30$). Pairwise comparisons showed that retrospective US-expectancy scores of all stimuli differed significantly ($ps < .009$), $CS_{safe} < CS_{conflict} < CS_{US}$.

US-expectancy ratings after the extinction phase were compared with US-expectancy ratings after the avoidance phase. The GLM showed a significant Phase*Stimulus interaction ($F(1.78,127.81) = 70.18$, $p < .001$, $\eta_p^2 = .49$), caused by a significant decrease in retrospective US expectancy of $CS_{conflict}$ ($F(1,74) = 152.49$, $p < .001$, $\eta_p^2 = .67$) and an absence of such decrease for the remaining

Table 1
Demographic information and mean scores (+SD) for the questionnaires, US and button ratings per condition (Study 1).

|	AAAA - No Avoidance ($n = 24$)	AABA - No Avoidance ($n = 27$)	AABA - Avoidance ($n = 24$)
Age	19.54 (1.82)	19.67 (1.44)	19.87 (1.73)
Male / Female	6 / 18	3 / 24	4/20
STAI-T	41.54 (7.06)	45.93 (8.74)	41.71 (9.22)
DTS	55.96 (9.42)	49.41(10.99)	51.75 (9.77)
IUS	54.63 (12.82)	59.59 (16.53)	55.83 (16.59)
US ratings			
#1, pleasant	−1.96 (0.86)	−2.15 (0.82)	−2.08 (0.78)
#1, positive	−2.29 (0.81)	−2.22 (0.85)	−2.12 (0.95)
#1, startle	0.29 (1.63)	1.15 (1.41)	0.75 (1.85)
#2, pleasant	−1.96 (1.23)	−2.30 (0.78)	−2.25 (0.85)
#2, positive	−2.33 (0.76)	−2.48 (0.75)	−2.33 (0.76)
#2, startle	0.13 (1.87)	0.96 (1.63)	1.08 (1.64)
Button effectiveness			
#1	0.21 (1.64)	0.37 (1.50)	0.92 (1.38)
#2	−0.17 (1.17)	0.41 (1.50)	0.33 (1.55)
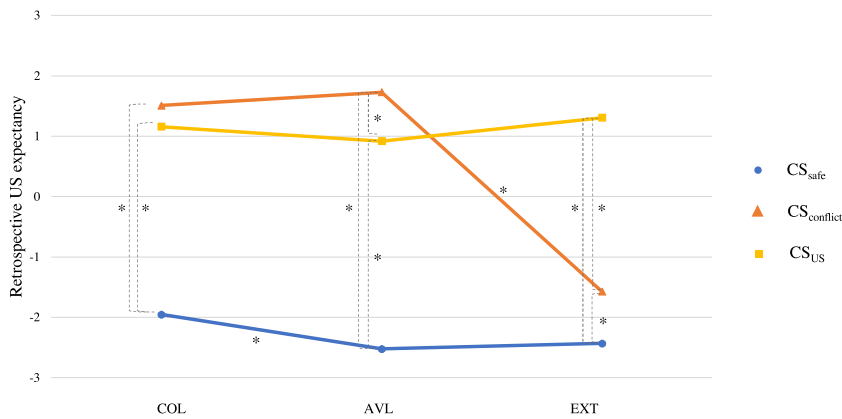
Fig. 3. Retrospective US expectancies after the contingency learning, avoidance learning, and extinction phase (Study 1). Results of the total sample are displayed as there were no significant main or interaction effects of Group. Significant Stimulus and Phase effects are marked, * $p < .05$.

stimuli ($Fs(1,74) < 1.66$, $ps > .20$). The strong decrease in retrospective US expectancy of $CS_{conflict}$ (i.e., from 1.73 to −1.57) is suggestive of fear extinction. The Group effect and all possible interactions with Group were not significant ($Fs < 2.20$, $ps > .08$).

3.2.1.4. Renewal test.. Results of the GLMs were not indicative of renewal. For details see Supplementary material.

3.2.2. Stimulus choice
Stimulus choice patterns during the different phases of the experiment are presented in Fig. 4.

3.2.2.1. Contingency learning.. We analysed change in stimulus choice from the first five contingency learning trials to the final five trials. The GLM revealed a significant interaction between Block, Stimulus, and Group ($F(2.52,90.60) = 3.19$, $p = .035$, $\eta_p^2 = .08$). Post-hoc analyses per group showed that the Block*Stimulus interaction was significant in all groups ($Fs > 4.37$, $ps < .04$). The significant interactions were followed up with GLM repeated measures ANOVAs per stimulus per group. With respect to the AAAA-No Avoidance group it was demonstrated that there was a significant increase from the first five to the final five contingency learning trials in choice of $CS_{conflict}$ ($F(1,23) = 4.40$, $p = .047$, $\eta_p^2 = .16$) and decrease in choice of CS_{US} ($F(1,23) = 47.77$, $p < .001$, $\eta_p^2 = .68$), but no change in choice of CS_{safe} ($F(1,23) = 1.60$, $p < .22$). In the AABA-No Avoidance group there was also a significant increase in choice of $CS_{conflict}$ (F
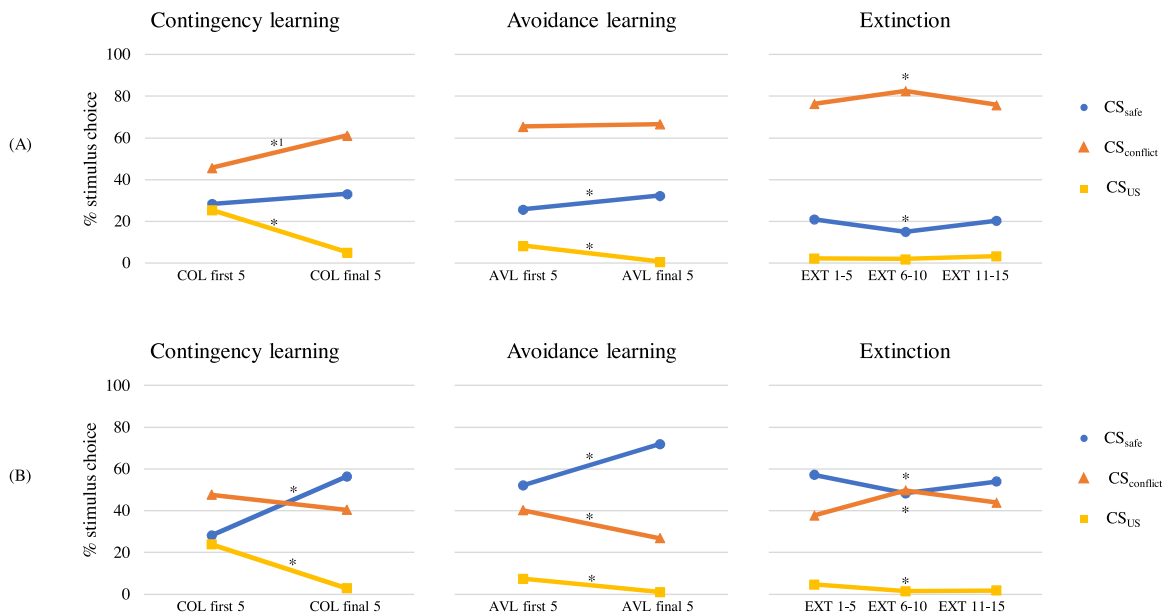

Fig. 4. Percentages of CS_{safe}, $CS_{conflict}$, and CS_{US} selections during the contingency learning, avoidance learning, and extinction, phase in Study 1 (A) and Study 2 (B). Results of the total sample are displayed as there were no significant main or interaction effects of Group. Significant Block effects are marked, * $p < .05$.
[1] The Block effect was not significant in the AABA-avoidance group.

$(1,26) = 22.24, p < .001, \eta_p^2 = .46$) and decrease in choice of CS_{US} ($F(1,26) = 108.43, p < .001, \eta_p^2 = .81$), but no change in choice of the CS_{safe} ($F(1,26) = 1.55, p < .23$). In the AABA-Avoidance group there was only a significant decrease in choice of CS_{US} ($F(1,23) = 33.48$, $p < .001, \eta_p^2 = .59$) and no change for the remaining stimuli ($Fs(1,23) < 3.60, ps > .07$). Note that during the final five trials of the contingency learning phase there were no group differences anymore ($F(2.22,80) = 2, p = .14$).

3.2.2.2. Avoidance learning.. We analysed change in stimulus choice from the first five avoidance trials to the final five trials. The GLM detected a significant interaction between Block and Stimulus ($F(1.26,90.93) = 4.41, p = .03, \eta_p^2 = .06$). Post-hoc analyses per stimulus showed that stimulus choice increased during the avoidance phase from the first five to the final five avoidance learning trials for CS_{safe} ($F(1,74) = 4.38, p = .04, \eta_p^2 = .06$) and decreased for CS_{US} ($F(1,74) = 29.86, p < .001, \eta_p^2 = .29$), but did not change for $CS_{conflict}$ ($F(1,74) = .10, p < .75$). The main effect of Group and interaction effects with Group were all non-significant ($Fs < .82, ps > .47$).

3.2.2.3. Extinction.. Analysis of the change in stimulus choice during the extinction phase in the different groups did not show main or interaction effects of Group ($Fs < .81, ps > .52$). Unexpectedly, the GLM did not show a Block*Stimulus interaction ($F(2.02,145.23) = 2.03, p = .14$). Only the Stimulus effect was significant ($F(1.03,73.84) = 180.34, p < .001, \eta_p^2 = .72$). Pairwise comparisons showed significant differences between all stimuli, $CS_{conflict} > CS_{safe} > CS_{US}$ ($ps < .001$)

3.2.2.4. Renewal test.. Results of the GLMs and Fisher's exact tests were not indicative of renewal. For details see Supplementary material.

3.2.3. Avoidance button

3.2.3.1. Avoidance learning.. An ANOVA was conducted to test whether the groups differed in use of the avoidance button during the avoidance phase, which was not the case ($F(2,74) = 1.54, p = .22$). During the avoidance learning phase only 22 out of 75 participants used the avoidance button once or more often, and those who did pushed it only in 27 % (range: 9–100 %) of the $CS_{conflict}$ trials. Based on the results, it is questionable whether participants learned to effectively use the avoidance button.

We checked whether participants who used the avoidance button ($n = 22$) thought the avoidance button was effective in terminating the film clip. The majority (82 %) correctly indicated that the button was effective.

3.2.3.2. Extinction.. Because only one group was able to use the avoidance button during extinction, it is not possible to directly compare the groups on use of the avoidance button. In the avoidance condition, only five out of 24 participants used the avoidance button. Participants who pushed the avoidance button did this on average on 49 % (range: 7–100 %) of the trials. We also assessed unsuccessful (i.e., pushed the button, even though it was not functional) use of the avoidance button during the experiment: In the AAAA-No Avoidance group one participant used the avoidance button in 44 % of the trials and in the AABA-No Avoidance group the button was used by four out of 27 participants in 30 % of the trials on average (range: 7–87 %), even though participants were instructed that the button was not functional.

3.2.3.3. Renewal test.. Results were not indicative of renewal. For details see Supplementary material.

3.3. Correlational analyses

No consistent pattern of correlations with individual difference measures was found and the measures did not change the overall pattern of the main results. For details see Supplementary material.

4. Discussion study 1

The current study used a new paradigm to investigate the effect of avoidance costs on approach-avoidance decisions and whether avoidance behaviours during the extinction procedure impeded extinction learning. The new VR-paradigm succeeded in inducing differential fear conditioning and extinction learning as can be concluded from the pattern of retrospective US-expectancy ratings. Also, the stimulus choice patterns suggested differential fear acquisition: During contingency learning participants learned that the CS_{US} predicted danger and the frequency of selected CS_{US} trials decreased. However, the non-significant Block*Stimulus interaction during extinction seems to indicate a lack of extinction. Because participants could learn that the $CS_{conflict}$ did not predict danger anymore, increases in the choice of the $CS_{conflict}$ were expected. The lack of this increase could be a result of ceiling effects, because this stimulus was already chosen in the majority of the trials (i.e., 77 %) by the majority of the participants ($n = 61$) during the first extinction block.

The study also aimed to investigate renewal effects. Contrary to our expectations, the groups did not differ in retrospective US-expectancy ratings, use of the avoidance button or choice of CS_{safe} during the renewal test phase (see Supplementary material). The results of the current study therefore suggest that context in our VR study had no influence on response patterns. Possible explanations for the lack of renewal are discussed in the General Discussion.

We cannot draw firm conclusions about whether avoidance behaviours during extinction impede extinction learning. In contrast to earlier studies (Lovibond et al., 2000, 2009; Rattel et al., 2017; van Uijen et al., 2018; Xia et al., 2017), we did not observe the

protection from extinction effect. However, this might be explained by a lack of variability in responses between groups. Only 20 % of the participants in the avoidance condition actually used the avoidance button, and when so, only used it in less than half of the trials. Furthermore, frequency of CS$_{conflict}$ selection was high in all groups, ranging from 73 % to 81 % between conditions. As a result, the majority of the participants was repeatedly exposed to the unreinforced conflict stimulus during extinction, allowing them to learn that the CS$_{conflict}$ now predicted safety.

Findings of the current study suggest that individuals can overcome their fears and that extinction can generalize over contexts. As expected, we observed that when avoidance costs are high, people are motivated to approach a feared stimulus. In the current study, avoiding the US led to a loss of 80 % (choosing CS$_{safe}$) or 100 % (using the avoidance button) of the reward compared to facing the feared stimulus. Participants frequently made remarks during the experiment that they rather watched the film clip a couple of times than having to walk back and forth on more trials. We therefore conducted a second experiment, in which we investigated avoidance behaviour and its effect on extinction learning when avoidance costs are lower, resulting in a more balanced approach-avoidance conflict. We expect that lower avoidance costs will lead to more variability in avoidance behaviours, allowing us to investigate protection from extinction.

5. Method study 2

5.1. Participants

A total of 90 individuals (aged 18–30, $M = 20.48$, $SD = 1.83$), 38 males and 52 females, participated in the study. Participant recruitment and allocation followed the same procedure as in Study 1.

5.2. Materials, questionnaires, procedure and data analyses

The materials, questionnaires, procedure and data analyses were identical to Study 1, except from one adaptation. To induce a larger conflict between approach and avoidance motivations, the reward of the CS$_{safe}$ was increased from one coin in Study 1 to three coins in Study 2. The amount of coins was based on a pilot study in $N = 20$ participants, aimed at determining the optimal reward conflict in terms of variability in responses between participants. When questioned about their response tendencies if they would be rewarded with three coins for the CS$_{safe}$ versus five coins for the CS$_{conflict}$, 12 participants indicated they would be inclined to choose the CS$_{conflict}$, six participants reported they would choose the CS$_{safe}$, and two participants reported they would use both stimuli.

6. Results study 2

6.1. Included sample

One female participant dropped out during the experiment due to equipment failure. Fourteen participants (10 female) were labelled as contingency unaware based on the aforementioned criteria. Hence, the total sample included 75 participants. Demographic information of the included sample and questionnaire data are presented in Table 2. Groups did not differ in sex ratio (χ^2 ($N = 75$) = .05, $p = .98$), age, STAI-T, DTS, and IUS scores ($Fs < 1.51$, $ps > .23$). Participants did not differ in their evaluations of the US presented once before the contingency learning phase ($Fs < 1.1$, $ps > .34$). They evaluated the US as unpleasant ($M = -1.81$; $t(74) = -14.79$, $p < .001$), negative ($M = -2.01$; $t(74) = -16.88$, $p < .001$) and reported to feel moderately startled by the film clip ($M = .61$; $t(74) = 3.43$, $p = .001$). Three repeated measures ANOVAs were conducted to check for habituation to the US. We found a significant difference between the pleasantness ratings before the contingency learning phase and at the end of the experiment ($F(1,74) = 17.60$, $p < .001$, $\eta_p^2 = .19$), but not for valence and feeling startled ($Fs < 2.87$, $ps > .09$). Of note, rather than increase, the pleasantness of the US

Table 2
Demographic information and mean scores (+SD) for the questionnaires, US and button ratings per condition (Study 2).

	AAAA-No Avoidance ($n = 26$)	AABA-No Avoidance ($n = 26$)	AABA-Avoidance ($n = 23$)
Age	21.00 (2.47)	20.54 (1.50)	20.35 (1.47)
Male / female	12/14	12/14	10/13
STAI-T	41.19 (5.96)	38.35 (4.79)	40.17 (7.11)
DTS	53.19 (7.49)	52.35 (6.58)	53.35 (6.87)
IUS	53.58 (12.46)	52.00 (10.21)	51.91 (14.05)
US ratings			
#1, pleasant	−2.04 (0.82)	−1.65 (1.29)	−1.74 (1.01)
#1, positive	−2.23 (1.07)	−1.81 (1.17)	−2.00 (0.80)
#1, startle	0.42 (1.63)	0.85 (1.49)	0.57 (1.56)
#2, pleasant	−2.23 (1.07)	−2.19 (1.23)	−2.43 (0.66)
#2, positive	−2.15 (1.22)	−2.23 (1.21)	−2.22 (1.09)
#2, startle	0.46 (1.73)	1.23 (1.45)	1.09 (1.54)
Button effectiveness			
#1	0.35 (2.02)	0.77 (2.01)	0.70 (2.01)
#2	0.23 (2.03)	0.15 (1.97)	0.91 (1.86)

decreased over time. Taken together, it can be assumed that participants did not habituate to the US after repeated presentations.

6.2. Fear conditioning

6.2.1. Retrospective US expectancy
The pattern of retrospective US expectancies over the different phases of the experiment is presented in Fig. 5.

6.2.1.1. Contingency learning.. The GLM detected a significant Stimulus effect ($F(1.64,117.87) = 188.91, p < .001, \eta_p^2 = .72$). Pairwise comparisons showed that after contingency learning the CS_{safe} received lower retrospective US-expectancy ratings than the $CS_{conflict}$ and CS_{US} ($ps < .001$), indicating successful differential fear acquisition. However, there was no significant difference between $CS_{conflict}$ and CS_{US} ($p = .13$). The pattern of retrospective US-expectancy ratings was in accordance with the reinforcement schedule (i.e., CS_{safe}: $M = -2.12 < CS_{conflict}$: $M = 1.45 < CS_{US}$: $M = 1.94$). As expected, the main effect of Group and the Group*Stimulus interaction were not significant ($Fs < .28, ps > .86$).

6.2.1.2. Avoidance learning.. The GLM revealed a significant Stimulus effect ($F(1.64,117.76) = 208.39, p < .001, \eta_p^2 = .74$). The main effect of Group and the Group*Stimulus interaction were not significant ($Fs < 1.28, ps > .29$). Pairwise comparisons showed that after avoidance learning the CS_{safe} received lower retrospective US-expectancy ratings than the $CS_{conflict}$ and CS_{US} ($ps < .001$), but retrospective US-expectancy ratings of $CS_{conflict}$ and CS_{US} did not differ significantly ($p = .50$).

US-expectancy ratings after the avoidance phase were compared with US-expectancy ratings after contingency learning. There was a significant Phase*Stimulus interaction ($F(2,144) = 11.15, p < .001, \eta_p^2 = .13$), caused by a significant decrease in retrospective US expectancy of CS_{safe} ($F(1,74) = 13.77, p < .001, \eta_p^2 = .16$) and increase in retrospective US expectancy of $CS_{conflict}$ ($F(1,74) = 12.90, p = .001, \eta_p^2 = .15$), and no change in retrospective US expectancy of CS_{US} ($F(1,74) = 2.60, p = .11$). The Group effect and all possible interactions with Group were not significant ($Fs < 1.10, ps > .34$).

6.2.1.3. Extinction.. The GLM revealed a significant Stimulus effect ($F(1.81,129.97) = 111.70, p < .001, \eta_p^2 = .61$). The main effect of Group and the Group*Stimulus interaction were not significant ($Fs < 1.82, ps > .17$). Pairwise comparisons showed that retrospective US-expectancy scores of all stimuli differed significantly ($ps < .001$), $CS_{safe} < CS_{conflict} < CS_{US}$.

US-expectancy ratings after the extinction phase were compared with US-expectancy ratings after the avoidance phase. The analysis revealed a significant Phase*Stimulus interaction ($F(1.82,131.16) = 48.21, p < .001, \eta_p^2 = .40$). The Group effect and all possible interactions with Group were not significant ($Fs < 2.67, ps > .08$). GLMs per stimulus indicated that there was a significant decrease in retrospective US expectancy of $CS_{conflict}$ ($F(1,74) = 83.16, p < .001, \eta_p^2 = .53$), but not of CS_{safe} and CS_{US} ($Fs < 3.57, ps > .06$). The strong decrease in retrospective US expectancy (i.e., from 1.99 to – 0.87) of the $CS_{conflict}$ is suggestive of fear extinction.

6.2.1.4. Renewal test.. Results of the GLMs were not indicative of renewal. For details see Supplementary material.

6.2.2. Stimulus choice
Stimulus choice patterns during the different phases of the experiment are presented in Fig. 4.

6.2.2.1. Contingency learning.. We analysed change in stimulus choice from the first five contingency learning trials to the final five trials. The GLM revealed a significant interaction between Block and Stimulus ($F(1.16,83.71) = 36.21, p < .001, \eta_p^2 = .34$) and was followed up with separate repeated measures ANOVAs per stimulus. There was a significant increase in choice of CS_{safe} ($F(1,74) = 48.16, p < .001, \eta_p^2 = .39$) and decrease in choice of CS_{US} ($F(1,74) = 238.80, p < .001, \eta_p^2 = .76$), but no change in $CS_{conflict}$ (F
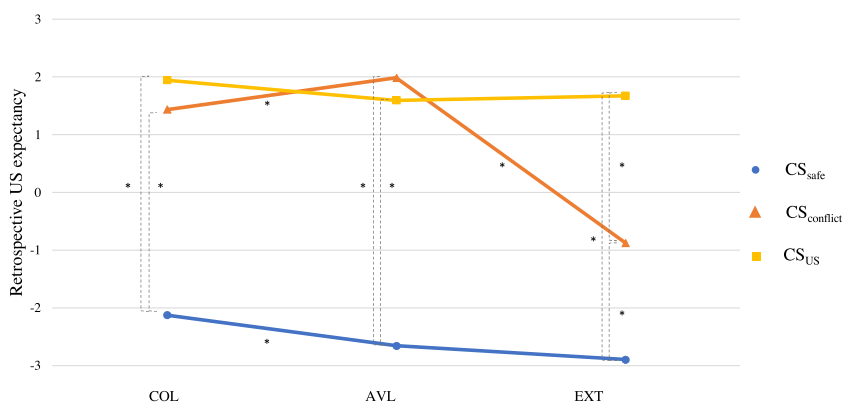

Fig. 5. Retrospective US expectancies after the contingency learning, avoidance learning, and extinction phase (Study 2). Results of the total sample are displayed as there were no significant main or interaction effects of Group. Significant Stimulus and Phase effects are marked, * $p < .05$.

$(1,74) = 2.96$, $p = .09$). The main effect of Group and interaction effects with Group were all non-significant ($Fs < 1.00$, $ps > .38$).

6.2.2.2. Avoidance learning.. We analysed change in stimulus choice from the first five avoidance trials to the final five trials. The GLM showed a significant interaction between Block and Stimulus ($F(1.21,87.28) = 29.86$, $p < .001$, $\eta_p^2 = .29$). Post-hoc analyses per stimulus showed that stimulus choice increased during the avoidance phase for CS_{safe} ($F(1,74) = 37.27$, $p < .001$, $\eta_p^2 = .34$) and decreased for $CS_{conflict}$ ($F(1,74) = 21.51$, $p < .001$, $\eta_p^2 = .23$) and CS_{US} ($F(1,74) = 27.97$, $p < .001$, $\eta_p^2 = .27$). The main effect of Group and interaction effects with Group were all non-significant ($Fs < .57$, $ps > .58$).

6.2.2.3. Extinction.. Analysis of the change in stimulus choice during the extinction phase in the different groups showed no main or interaction effects of Group ($Fs < 1.18$, $ps > .31$). There was a significant interaction between Block and Stimulus ($F(2.14,154.01) = 4.09$, $p = .02$, $\eta_p^2 = .05$). Separate GLMs per stimulus showed that the block effect for CS_{safe} ($F(1.85,137.14) = 2.98$, $p = .058$) did not reach statistical significance. Analyses revealed significant Block effects for $CS_{conflict}$ ($F(2,148) = 5.48$, $p = .005$, $\eta_p^2 = .07$) and CS_{US} ($F(1.75,129.83) = 5.00$, $p = .01$, $\eta_p^2 = .06$). The significant effects were caused by a significant increase in choice of $CS_{conflict}$ from the first five trials to the five intermediate trials ($p = .002$) and a significant decrease in choice of CS_{US} from the first five trials to the five intermediate trials ($p = .03$).

6.2.2.4. Renewal test.. Findings of the GLMs and Fisher's exact tests were not indicative of renewal. For details see Supplementary material.

6.2.3. Avoidance button

6.2.3.1. Avoidance learning.. During the avoidance learning phase there were no group differences in use of the avoidance button ($F(2,74) = .25$, $p = .78$). Only 28 out of 75 participants used the avoidance button once or more often. Participants who used the button pushed it only in 53 % (range: 9–100 %) of the trials on average. Based on the results, it is questionable whether participants learned to use the avoidance button.

We checked whether participants who used the button ($n = 28$) thought the avoidance button was effective in terminating the film clip. Ninety-three percent of the participants correctly indicated that the button was effective.

6.2.3.2. Extinction.. Because only one group was able to use the avoidance button during extinction, it was not possible to directly compare the groups on use of the avoidance button. In the avoidance condition, 11 out of 24 participants used the avoidance button. Participants who pushed the avoidance button did this on average on 56 % (range: 7–100 %) of the trials. We also assessed unsuccessful (i.e., pushed the button, but button was not functional) use of the avoidance button during the experiment: In the AAAA-No Avoidance group two participants used the avoidance button in 9% of the trials on average (range: 7–11 %) and in the AABA-No Avoidance group the button was used by two out of 27 participants in 11 % of the trials on average (range: 9–13 %), even though participants were instructed that the button was not functional.

6.2.3.3. Renewal test.. Results were not indicative of renewal. For details see Supplementary material.

6.3. Correlational analyses

No consistent pattern of correlations with individual difference measures was found and the measures did not change the overall pattern of the main results. For details see Supplementary material.

7. Discussion study 2

Study 2 investigated the effect of avoidance costs on approach-avoidance decisions and whether avoidance behaviours during the extinction procedure impeded extinction learning. Similar to Study 1, the pattern of retrospective US-expectancy ratings showed that the newly developed VR-paradigm succeeded in inducing differential fear conditioning and extinction learning. In contrast to Study 1, patterns of stimulus choice were not only indicative of differential fear learning, but now also of extinction learning. During the first 10 extinction trials, participants' preference for the $CS_{conflict}$ increased, which suggests that participants learned that the $CS_{conflict}$ was now completely safe. Regarding renewal effects (see Supplementary material), Study 2 again showed no group differences during the renewal test in retrospective US-expectancy ratings, use of the avoidance button or choice of CS_{safe}. As mentioned earlier, possible explanations for the lack of renewal are discussed in the General Discussion.

The changes in Study 2 resulted in more avoidance behaviour. Forty-six percent of the participants used the button at least once during extinction in comparison to 20 % in Study 1 and the CS_{safe} was selected in 53 % of the extinction trials in Study 2 in comparison to 19 % in Study 1. Especially choosing the CS_{safe} option to avoid the US was persistent throughout the experiment. The fact that the design of Study 2 resulted in a wider range of avoidance responses, might be explained by the lower difference in cost between approaching and avoiding the conflict stimulus (5 coins versus 3 coins) in Study 2 in comparison to Study 1 (5 coins versus 1 coins). This is supported by the finding of Schlund et al. (2016), who found that most diversity in approach-avoidance responses was found at levels where the value differences between the approach and avoidance option were smallest.

Even though more avoidance behaviour was observed in Study 2, we did not observe the protection from extinction effect (Lovibond et al., 2000, 2009; Rattel et al., 2017; van Uijen et al., 2018; Xia et al., 2017). This could be explained by the fact that all groups could choose the CS_{safe} to avoid the US. We expected a difference between the AABA-No Avoidance and the AABA-Avoidance group, confirming the protection from extinction effect, however, participants rather chose the CS_{safe} (in all groups) than using the avoidance button as it led to an extra loss of coins. Therefore, the expected group difference was not observed.

Similar to the conclusions of Study 1, Study 2 suggests that individuals can overcome their fears and that extinction can generalize over contexts. The study also indicates that if a safe, substantially rewarding alternative is available, the use of avoidance behaviours increases. Besides investigating the effect of avoidance costs on approach-avoidance decisions and protection from extinction within the two studies, we can also compare the results of Study 1 and 2. In the section Joint analysis the two studies, and hence two different approach-avoidance conflicts, will be compared in order to explore how different avoidance costs affect approach-avoidance decisions.

8. Joint analysis

Study 1 (high avoidance costs) and Study 2 (low avoidance costs) were compared post-hoc to examine the effect of avoidance costs on approach-avoidance decisions and retrospective US expectancy. To this end, we first compared the two samples regarding age, sex, and scores on the questionnaires. If the samples differed on some of the factors, those factors were included as covariates to control for their influence. Because we checked for the influence of the DTS in our main analyses in the two separate studies, we also included the DTS as a covariate in the joint analyses. Only in case the covariates changed the results of the main analyses, these results are reported. Second, to investigate the influence of avoidance costs on response patterns, we compared frequencies of 'avoiders' for the separate phases between studies using Fisher's exact tests. Avoiders were operationalized as participants who either used the avoidance button in more than 75 % of the trials with the avoidance button present or selected the CS_{safe} in more than 75 % of the trials. Using ANOVAs, we compared retrospective US-expectancy ratings and stimulus choice during renewal of avoiders and non-avoiders in order to examine the protection from extinction effect. Third, use of the avoidance button was compared between studies using ANOVAs. Fourth, ANOVAs were used to compare choices for the CS_{safe} and $CS_{conflict}$ between studies. Fifth, we investigated patterns of stimulus choice within the separate phases and compared them between studies using GLMs.

8.1. Samples

Since the procedures of Study 1 and 2 were identical, we compared the effects of high and low avoidance costs on behavioural avoidance patterns. As no renewal effects were observed in either study, findings of the renewal test phase are discussed in the Supplementary material. The two study samples differed in age ($F(1,149) = 10.78$, $p = .001$), STAI-T scores ($F(1,149) = 7.38$, $p = .007$), and sex (χ^2 ($N = 150$) = 13.67, $p < .001$). The difference between the samples in IUS scores ($F(1,149) = 3.58$, $p = .06$) almost reached significance, but no differences in DTS scores ($F(1,149) = .23$, $p = .63$) were observed. Mean age in Study 1 ($M = 19.69$) was somewhat lower than in Study 2 ($M = 20.64$). In Study 1 participants had slightly higher STAI-T ($M = 43.17$) and IUS scores ($M = 56.80$) than in Study 2 ($M = 39.89$ and $M = 52.52$ respectively). The largest difference between samples was found in sex: 83 % was female in Study 1, whereas this was 55 % in Study 2. After contingency learning, the study samples did not differ in retrospective US expectancy of the CS_{safe} and $CS_{conflict}$ ($ps > .46$), but they did differ in retrospective US expectancy of the CS_{US} ($F(1,149) = 6.56$, $p = .01$; Study 1: $M = 1.16$ and Study 2: $M = 1.95$). As we were interested in the conflict between the CS_{safe} and $CS_{conflict}$, results of the joint analysis only focus on those two stimuli.

To control for the influence of the factors that differed between the samples, we ran separate ANCOVAs with age, STAI-T and IUS scores as covariates. Sex was included as a between-subjects factor to check for main or interaction effects of this variable. Results of the analyses of covariance are discussed in the Supplementary material.

8.2. Fear conditioning

8.2.1. Avoiders

The frequencies of avoiders and results of the Fishers's exact tests are presented in Table 3. ANOVAs demonstrated protection from extinction: Avoiders (CS_{safe}: $M = 1.26$; button: $M = 2.14$) had significantly higher retrospective US-expectancy ratings of $CS_{conflict}$ after extinction compared to non-avoiders (CS_{safe}: $M = -1.87$; button: $M = -1.38$; $ps < .001$). Although we did not find renewal effects (see Supplementary material), we did find evidence for the protection from extinction effect during the renewal test phase. Avoiders during extinction (CS_{safe}: $M = .74$; button: $M = .74$) selected CS_{safe} more frequently during the renewal test than non-avoiders (CS_{safe}: $M = .14$;

Table 3
Comparison of frequencies of avoiders in Study 1 and 2.

	COL Avoiders CS_{safe}	AVL Avoiders CS_{safe}	EXT Avoiders CS_{safe}	AVL Avoiders button	EXT Avoiders button	EXT Avoiders uns. button
Study 1	2 / 75	3 / 75	5 / 75	1 / 75	2 / 75	1 / 75
Study 2	1 / 75	29 / 75	26 / 75	8 / 75	5 / 75	0 / 75
	$p > .99$	$p < .001$	$p < .001$	$p = .03$	$p = .44$	$p > .99$

button: $M = .24$; $ps < .001$). There was no difference between avoiders and non-avoiders during extinction in use of the avoidance button during the renewal test ($ps > .12$).

8.2.2. Avoidance button

The percentages of button presses during the different phases of the experiment and the results of the ANOVAs are presented in Table 4. Only during the avoidance learning phase was the avoidance button used by significantly more participants in Study 2 than in Study 1 ($p = .006$).

8.2.3. Stimulus choice

The percentages of CS$_{safe}$ and CS$_{conflict}$ chosen during the different phases of the experiment and the results of the ANOVAs are presented in Table 5. During the contingency learning, avoidance learning and extinction phase, the CS$_{safe}$ was more often selected by participants in Study 2 and the CS$_{conflict}$ was more often selected by participants in Study 1 (p's $< .006$).

Fourth, the patterns of stimulus choice within the separate phases are presented in Fig. 4. During the contingency learning and avoidance learning phase, there was a significant interaction between Study, Stimulus, and Block ($Fs > 8.70$, $ps < .002$). We examined the effect of Study per Stimulus per Block, using six separate ANOVAs during the contingency learning and avoidance learning phase. To correct for multiple testing the alpha level was set at $p = .008$. Six ANOVAs revealed that during the final five trials of the contingency learning phase, participants in Study 2 selected the CS$_{safe}$ more frequently ($F(1,148) = 18.79$, $p < .001$, $\eta_p^2 = .11$) and the CS$_{conflict}$ less frequently ($F(1,148) = 15.77$, $p < .001$, $\eta_p^2 = .10$) than participants in Study 1. There were no other significant differences between studies during the contingency learning phase ($Fs < 3.32$, $ps > .07$). During the avoidance learning phase participants in Study 2 again selected the CS$_{safe}$ more frequently during both the first and final five trials ($Fs > 42.17$, $ps < .001$) and the CS$_{conflict}$ less frequently during the first and final five trials ($Fs > 36.78$, $ps < .001$) in comparison to participants in Study 1. Studies did not differ in selection of the CS$_{US}$ during the avoidance learning phase ($Fs < .36$, $ps > .55$).

During the extinction phase, there was no three-way interaction ($F(2.12, 314.13) = .78$, $p = .47$). The Block*Stimulus ($F(2.12,314.13) = 5.72$, $p = .003$, $\eta_p^2 = .04$) and Study*Stimulus ($F(1.02,150.99) = 47.09$, $p < .001$, $\eta_p^2 = .24$) interactions were significant. Post-hoc analyses of the Block effect per stimulus showed that the effect was significant for CS$_{safe}$ ($F(1.84,273.89) = 5.06$ $p = .01$, $\eta_p^2 = .03$) and CS$_{conflict}$ ($F(1.82,271.47) = 6.69$ $p = .002$, $\eta_p^2 = .04$), but not for CS$_{US}$ ($F(2,298) = 2.2$ $p = .11$). Choice of CS$_{safe}$ significantly decreased from the first five trials to the five intermediate trials ($p = .003$) and choice of CS$_{conflict}$ increased from the first five trials to the five intermediate trials ($p < .001$) and significantly decreased from the five intermediate trials to the five final trials ($p = .03$). The significant Study*Stimulus effect was followed up with separate ANOVAs of Study per Stimulus. To correct for multiple testing the alpha level was set at $p = .017$. The analyses revealed that participants in Study 2 selected the CS$_{safe}$ more frequently ($F(1,148) = 46.85$, $p < .001$, $\eta_p^2 = .24$) and CS$_{conflict}$ less frequently ($F(1,148) = 47.98$, $p < .001$, $\eta_p^2 = .25$) than participants in Study 1. There was no difference between studies in selection of the CS$_{US}$ ($F(1,148) = .02$, $p = .88$).

8.3. Correlational analyses

No consistent pattern of correlations with individual difference measures was found. For details see Supplementary material.

9. General discussion

The current studies investigated whether different avoidance costs led to different approach-avoidance decisions and whether avoidance behaviours during the extinction procedure impeded extinction learning using a novel VR-paradigm. Comparing the two studies, we observed that when avoidance costs were high (Study 1), less participants avoided the aversive stimulus than when they were low (Study 2). This finding is in line with our hypothesis that higher avoidance costs would lead to less avoidance behaviours. Our findings also coincide with earlier studies (e.g., Aupperle et al., 2011; Dibbets & Fonteyne, 2015; Pittig, 2019; Pittig, Brand et al., 2014; Pittig & Dehler, 2019; Pittig, Schulz et al., 2014; Rattel et al., 2017; Sierra-Mercado et al., 2015; Talmi et al., 2009). In a study by Rattel et al. (2017) participants could choose between a risky shortcut or a costly detour. Manipulating the length of the detour (i.e., avoidance cost) between subjects, it was revealed that elevated avoidance costs went along with less avoidance behaviours. Findings of Dibbets and Fonteyne (2015) were also in line with the current findings, showing that spider-fearful individuals were motivated to approach a stimulus that predicted the risk of encountering a virtual spider when approach was rewarded with a large amount of virtual coins even if a low-rewarding, safe alternative was available. Schlund et al. (2016) investigated the tipping point of approach and avoidance behaviour. They concluded that as value differences between approach and avoidance options decrease, there is increasing conflict, choice difficulty, threat responses and time needed to make a decision. They also observed that when the

Table 4
Percentages of use of the avoidance button compared between Study 1 and 2.

	AVL	EXT	EXT unsuccessful
Study 1	8%	3%	2%
Study 2	20 %	9%	0.5 %
	$p = .006$	$p = .11$	$p = .22$

Table 5
Percentages of CS_{safe} and $CS_{conflict}$ selections compared between Study 1 and 2.

	COL CS_{safe}	COL $CS_{conflict}$	AVL CS_{safe}	AVL $CS_{conflict}$	EXT CS_{safe}	EXT $CS_{conflict}$
Study 1	33 %	54 %	29 %	67 %	19 %	78 %
Study 2	42 %	45 %	64 %	33 %	53 %	44 %
	$p = .008$	$p = .006$	$p < .001$	$p < .001$	$p < .001$	$p < .001$

probability of losing money was lowest (hence the avoidance costs were highest), participants showed least avoidance responses. In conclusion, it can be assumed that when benefits of approaching a feared stimulus or situation outweigh the costs, approach rather than avoidance behaviour will be observed even when a safe alternative is available.

Secondly, it was hypothesized that avoidance behaviours during extinction would result in protection from extinction. We did not find differences in retrospective US expectancies, stimulus choice and use of the avoidance button between participants in the Avoidance and No Avoidance conditions. This finding is in contrast to a recent study that directly tested our two main hypotheses (Pittig, 2019). Although both our studies as well as the study by Pittig (2019) concluded that avoidance costs (or approach incentives) reduce avoidance behaviours, Pittig (2019) did observe that, in contrast to a group that received a reward when they did not avoid the US, participants in the no reward-group executed safety behaviours persistently (also in absence of the US), which resulted in protection from extinction. However, the lack of group differences in the present study need not come as a surprise. As stated earlier, even though groups differed in the availability of the avoidance button, all groups had the possibility to select the safe stimulus on every trial. Selecting the CS_{safe} frequently was preferred over using the avoidance button, as this led to a small reward instead of the omission of a larger reward. Unfortunately, allowing avoidance responding by selecting the CS_{safe} conflated the contingency ratings and amount of button presses, which made the avoidance group differences difficult to interpret. One might also argue that the lack of group differences and protection from extinction could be explained by the US being not aversive enough or the reinforcement rate being too high. However, the fact that avoidance behaviours significantly increased in Study 2 (when avoidance costs were lower) seems to suggest that participants did experience the US as unpleasant. Also, reinforcement rates in other studies investigating protection from extinction were higher (i.e., 100 %), and nonetheless demonstrated this phenomenon, even when avoidance was costly (e.g., Rattel et al., 2017).

Although we did not find group differences, results comparing avoiders to non-avoiders were suggestive of protection from extinction. Avoiders during extinction had higher retrospective US-expectancy ratings of the extinguished stimulus and selected the safe stimulus more often during the renewal test than non-avoiders. The fact that avoiders did not learn that the $CS_{conflict}$ predicted safety shows that avoiding the anticipated US protected participants from undergoing extinction. Several theoretical underpinnings for this phenomenon have been proposed. For example, it has been suggested that avoidance behaviours function as a negative occasion setter. That is, the use of the avoidance behaviours signals that the stimulus is not followed by the US and therefore prevents ordinary inhibitory learning (i.e., CS-noUS; De Houwer, Crombez, & Baeyens, 2005; Krypotos, Effting, Kindt, & Beckers, 2015). By this account, participants in the current studies learned during the avoidance learning phase that the $CS_{conflict}$ was followed by the US in the absence of an avoidance response and was not followed by the US when they made a correct avoidance response (Declercq & De Houwer, 2011). Hence, avoiders had high retrospective US-expectancy ratings after extinction as their use of avoidance behaviours signalled the non-occurrence of the US. In contrast, non-avoiders learned new information that was not consistent with the expected outcome signalled by the absence of the occasion setter (i.e., CS-noUS), allowing for inhibitory learning during the extinction phase.

Another explanation for the protection from extinction effect lies in the *violation of expectancy* approach. This approach postulates that inhibitory learning is enhanced by increasing the discrepancy between the expectancy (i.e., feared outcome) and the actual outcome (i.e., typically the nonoccurrence of the feared outcome; Craske et al., 2014; Rescorla & Wagner, 1972). Previous studies that assessed US expectancy after an avoidance response was made or not made (e.g., Lovibond et al., 2009; Rattel et al., 2017) have shown that avoidance behaviours reduce the expectancy of the US, which is expected to have a negative impact on extinction learning (Craske et al., 2014). However, the current studies assessed US expectancies retrospectively and asked for the expectancy when no avoidance response would have been made. As the current paradigm did not allow us to determine whether avoidance behaviours reduced US expectancy, results of the current studies cannot be related to the violation of expectancy explanation of protection from extinction.

Finally, in a treatment context, the *misattribution of safety* hypothesis (Salkovskis, 1991) posits that avoidance behaviours prevent disconfirming learning experiences because anxious patients misattribute the non-occurrence of the feared outcome to their use of avoidance instead of to the innocuousness of the situation. Translating this to the results of the current studies, avoiders attributed the non-occurrence of the US on $CS_{conflict}$ trials to the use of the avoidance button. This resulted in higher retrospective US-expectancy ratings of $CS_{conflict}$ and more CS_{safe} selections during the renewal test when the avoidance button was dysfunctional in comparison to non-avoiders, who had no alternative explanation for the absence of the US during the extinction phase. What follows from these theoretical explanations is that treatment should not only focus on CS-US contingencies, but also on patients' beliefs about avoidance behaviours (Declercq & De Houwer, 2011).

A third aim of the studies was to investigate the influence of trait anxiety, distress tolerance, and intolerance of uncertainty on approach-avoidance decisions exploratory. We found a positive association between STAI-T scores and use of the avoidance button during the avoidance learning phase in Study 2. Significant associations were also detected between DTS scores and selections of CS_{safe} during the contingency learning and extinction phase in Study 1 and between DTS scores and use of the avoidance button during the avoidance learning phase and selections of CS_{safe} during the renewal test phase in Study 2. IUS scores correlated negatively with use of

the avoidance button during the renewal test phase when the two samples were combined. Note that no robust pattern of correlations over studies and groups was detected. Considering multiple testing, the significant correlations should be interpreted with caution. Although other studies found significant correlations between individual difference measures and fear and avoidance conditioning, these correlations were also weak to moderate (Pittig, Schulz et al., 2014), and not specifically related to avoidance behaviours (Morriss et al., 2015; Vervliet et al., 2017). More support for the lack of robust effects of individual difference measures on avoidance behaviours comes from including the measures as covariates in our analyses. Only in Study 1, controlling for STAI-T scores changed a significant result in a non-significant Block*Stimulus*Group interaction, which was in line with our expectations. The covariates did not change other results of the main analyses. Covariate analyses did reveal significant effects of STAI-T (positive effect) and DTS scores (negative effect) on use of the avoidance button in Study 2 and effects of DTS scores on CS_{safe} (negative effect) and $CS_{conflict}$ selections (positive effect). We also found that avoiders had lower DTS scores than non-avoiders. However, considering multiple testing these effects should again be interpreted with caution.

Finally, we investigated renewal in the context of behavioural avoidance. It was expected that participants who were presented with the extinguished CS in a different context as where extinction took place would show highest retrospective US-expectancy ratings and most avoidance responses. Unexpectedly, no group differences were found during the renewal test in retrospective US-expectancy ratings, use of the avoidance button or selection of the safe stimulus. However, renewal is typically measured as the change in fear responses from the last extinction trial(s) to the first renewal test trial(s) (Lonsdorf et al., 2017; Vervliet, Baeyens, Van den Bergh, & Hermans, 2013; Vervliet, Craske et al., 2013). In Study 1, the percentage of participants who selected the CS_{safe} during the final extinction trial ranged between 19 % and 29 % between groups, whereas this difference was larger at the first renewal test trial. Even though it did not reach statistical significance, the pattern of percentages suggested renewal: Only 4% in the AAAA condition chose CS_{safe} on the first renewal test trial, in comparison to 25 % and 22 % in the AABA-No Avoidance and AABA-Avoidance group respectively. However, comparing fear responses from the last extinction trial to the first renewal test trial in Study 2 did not provide evidence for renewal. Percentages of participants who selected the CS_{safe} during the final extinction trial ranged between 42 % and 61 % between groups, and between 39 % and 57 % during the first renewal trial, with lowest percentages in the AABA-No Avoidance group.

A possible explanation for the lack of renewal, in contrast to previous studies that did find evidence for the renewal of avoidance responses (Schlund et al., 2019; Urcelay et al., 2019), is that participants did not distinguish well between contexts. Even though the contexts were visually discriminable, they also shared some similarities. For example, the computer screens that displayed the CSs and the setting of the rooms were similar. Moreover, research shows that as arousal increases for negative stimuli or events, individuals are more focused on central details than on peripheral details (Yegiyan & Lang, 2010). As the central details of the contexts (i.e., the computer screens) were similar, it might be the case that our context manipulation was not as strong compared to other computer or VR-paradigms that used background colours or rooms without largely overlapping central items as a context variable (e.g., Kroes, Dunsmoor, Mackey, McClay, & Phelps, 2017; Milad, Orr, Pitman, & Rauch, 2005; Schlund et al., 2019; Urcelay et al., 2019). Second, it is possible that fear for the $CS_{conflict}$ did not extinguish completely. This would be a problem, as Urcelay et al. (2019) found that participants that did not show extinction also did not show renewal. Although retrospective US expectancy decreased after extinction, stimulus choice for $CS_{conflict}$ and CS_{safe} did not change from the first five to the final five extinction trials in both studies. In Study 1, this might be explained by a ceiling effect as participants selected the $CS_{conflict}$ already in 76.5 % of the first five extinction trials and continued selecting the $CS_{conflict}$ during the final five extinction trials (76 %). However, in Study 2 participants still avoided the US in more than half of the trials (54 %) in the final block of the extinction phase.

The current studies tried to establish a new paradigm to investigate avoidance behaviours. One of the major strengths of this paradigm is that we implemented avoidance costs that resemble the restrictions to the daily lives of anxiety patients more closely than previous studies that for example used (hypothetical) monetary rewards (e.g., Bublatzky et al., 2017; Pittig, 2019; Pittig, Brand et al., 2014; Pittig & Dehler, 2019; Pittig, Schulz et al., 2014). Avoiding the anticipated US required more physical effort and resulted in having to stay in the laboratory for a longer time, for instance simulating the costs a panic disorder patient is facing when taking a detour to avoid a place that has been associated with a previous panic attack. Second, providing participants with an escape option and a time-consuming safe alternative also increased the ecological validity of this study. In daily life, anxiety patients also often have a choice whether, on the one hand, to expose themselves to a feared stimulus or situation and escape the situation whenever the fear level becomes too high or, on the other hand, to choose an alternative option (e.g., taking a detour) that takes more time.

9.1. Limitations

Although the paradigm succeeded in inducing fear conditioning and extinction, the studies also comprise some limitations. First, the current paradigm used two different avoidance responses, namely CS-avoidance (stimulus choice) and US-avoidance (avoidance button), which have different effects and relevance in clinical settings (Krypotos et al., 2018). Also, in the current paradigm, choices on CS-avoidance trials are likely driven by US-cancellation as well as CS-termination. Hence, participants' responses not only denote avoidance, but also escape behaviour. This deserves mention because traditionally, signalled avoidance learning involves US-cancellation, but does not terminate the CS. Although intentional, the mix of CS- and US-avoidance made it difficult to establish between-group effects and interpret the behavioural findings. As the free-choice paradigm allowed participants in No Avoidance groups to avoid the US as well, our conclusions regarding protection from extinction could only be based on a comparison between avoiders and non-avoiders and not on between-group differences. Furthermore, those conclusions can only be tentative, because the sample of avoiders in both studies was rather small and combined both types of avoidance behaviours. The interpretation of the behavioural findings is further limited by the fact that stimulus choice was used both as a measure of avoidance (see Joint analysis) and

as a measure of fear conditioning (see Study 1 and 2). Finally, offering a safe alternative seems to have interfered with the use of the avoidance button, which made it difficult to interpret the results regarding contextual renewal. In future studies, we would advise to separate both types of avoidance responses in order to investigate between-group differences in avoidance behaviours, protection from extinction, and context renewal, for example by using a forced-choice paradigm following the design of Lovibond et al. (2009). In such a forced-choice paradigm, it would also be useful to add an extra extinction group receiving extinction in the same context as conditioning with the avoidance button available (AAAA-avoidance) in order to increase the sensitivity to detect a protection from extinction effect.

Second, retrospective US-expectancy ratings were used in contrast to trial-by-trial assessments of US expectancy as used in previous studies (e.g., Lovibond et al., 2009; Rattel et al., 2017). Even though we were able to compare retrospective US-expectancy ratings of the separate phases, the current design did not allow us to track the pattern of US-expectancy ratings over the course of conditioning and extinction. It is possible that the lack of online ratings of US expectancy has contributed to the null results regarding protection from extinction and renewal. First, retrospective US-expectancy ratings measure something different than US-expectancy ratings on each trial. Online expectancy ratings represent probability estimates of the US, and hence, might be interpreted as risk-estimates. It has been suggested that as risk over-estimation is an important feature of pathological anxiety, online US-expectancy ratings have high predictive value (Lonsdorf et al., 2017). On the other hand, retrospective US-expectancy ratings only indicate whether participants were able to discriminate between a CS that was followed by the US and a CS that was not. Retrospective ratings fail to measure subtle changes in risk-estimation when learning occurs (Boddez et al., 2013; Lonsdorf et al., 2017). Second, the retrospective ratings might have been biased by the information learned during the renewal test phase that the $CS_{conflict}$ predicted safety. Online US-expectancy ratings are therefore recommended for future studies.

Third, we did not include psychophysiological measures in our study as the VR-setting required participants to make movements that interfere with reliably measuring physiological responses. The necessity of psychophysiological measures might be debated as research suggests a relatively close correspondence between psychophysiological measures (e.g., skin conductance responses or the fear-potentiated startle reflex) and threat expectancy in fear conditioning (Davey, 1987; Lovibond, 1992). However, including physiological outcome measures has some important benefits as they are not prone to self-report biases and provide a link to animal research (Lonsdorf et al., 2017). Future studies should aim to integrate psychophysiological measures in paradigms of high ecological validity.

Fourth, the influence of high and low avoidance costs was investigated in two separate studies. Although both study samples were recruited from the same cohort, we observed some between sample differences, for example with regard to sex. This could bias the results as research suggests sex-related differences in anxiety vulnerability and behavioural avoidance (McLean & Anderson, 2009; Sheynin, Beck, Pang et al., 2014; Sheynin, Beck, Servatius, & Myers, 2014). Nevertheless, our studies came to the same conclusions as studies that varied avoidance costs within one study (Rattel et al., 2017). For future studies, however, it is recommended to include avoidance costs as a between-subjects variable and directly compare groups within one sample.

Fifth, considering three experimental groups were investigated and the complexity of the design and analyses, the sample size of both studies ($n = 75$) could be seen as rather small. Even though based on our power calculation we included enough participants to make statistical inferences with sufficient power, it might be the case that for some of the analyses in particular the small sample size was problematic. For example, avoidance button press was rather rare, which made it difficult to study this behaviour with such a small sample size. Furthermore, a larger sample size might also have increased the sensitivity to detect associations with individual difference measures. Two final limitations of the current studies are that they used simple CS-US associations that might not be fully translatable to complex situations in the daily lives of anxiety patients and that they used a convenience sample, which makes it difficult to translate findings to, for example, other age groups and populations (Lonsdorf & Merz, 2017).

9.2. Conclusions

In sum, the present preliminary findings demonstrate that avoidance costs play an important role in avoidance decisions and fear extinction. Higher avoidance costs foster the motivation to approach a feared stimulus. On the other hand, the findings show that, during extinction, participants who are motivated to avoid are protected from undergoing fear extinction. As the current studies adopted an ecologically valid paradigm including personally relevant avoidance costs, the current studies provide insight in the underlying mechanisms that contribute to the development and maintenance of anxiety disorders. Future research with respect to approach-avoidance conflicts is warranted, as this is a highly relevant yet understudied topic in anxiety disorders.

Funding

CRediT authorship contribution statement

Anke Lemmens: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing - original draft. **Tom Smeets:** Conceptualization, Methodology, Supervision, Writing - review & editing. **Tom Beckers:** Conceptualization, Methodology, Writing - review & editing. **Pauline Dibbets:** Conceptualization, Methodology, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgments

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.lmot.2021.101710.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5®* (5th ed. ed.). Arlington, VA: American Psychiatric Association.

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.

Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., & Beckers, T. (2013). Rating data are underrated: Validity of US expectancy in human fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry, 44*(2), 201–206.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114*(1), 80–99.

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

Bublatzky, F., Alpers, G. W., & Pittig, A. (2017). From avoidance to approach: The influence of threat-of-shock on reward-based decision making. *Behaviour Research and Therapy, 96*, 47–56.

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

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.

Davey, G. C. L. (1987). *Cognitive processes and Pavlovian conditioning in humans*. Chichester: Wiley & Sons, Inc.

De Houwer, J., Crombez, G., & Baeyens, F. (2005). Avoidance behavior can function as a negative occasion setter. *Journal of Experimental Psychology. Animal Behavior Processes, 31*(1), 101–106.

Declercq, M., & De Houwer, J. (2011). Evidence against an occasion setting account of avoidance learning. *Learning and Motivation, 42*(1), 46–52.

Dibbets, P., & Fonteyne, R. (2015). High spider fearfuls can overcome their fear in a virtual approach-avoidance conflict task. *Journal of Depression & Anxiety, 4*(2), 1–7.

Dibbets, P., Lemmens, A., & Voncken, M. (2018). Turning negative memories around: Contingency versus devaluation techniques. *Journal of Behavior Therapy and Experimental Psychiatry, 60*, 5–12.

Forsyth, J. P., Eifert, G. H., & Barrios, V. (2006). Fear conditioning in an emotion regulation context: A fresh perspective on the origins of anxiety disorders. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications* (pp. 133–153). Washington, DC, US: American Psychological Association.

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences, 17*(6), 791–802.

Greenberg, P. E., Sisitsky, T., Kessler, R. C., Finkelstein, S. N., Berndt, E. R., Davidson, J. R., ... Fyer, A. J. (1999). The economic burden of anxiety disorders in the 1990s. *The Journal of Clinical Psychiatry, 60*(7), 427–435.

Grillon, C., Baas, J. M., Cornwell, B., & Johnson, L. (2006). Context conditioning and behavioral avoidance in a virtual reality environment: Effect of predictability. *Biological Psychiatry, 60*(7), 752–759.

Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., ... Fratiglioni, L. (2011). Cost of disorders of the brain in Europe 2010. *European Neuropsychopharmacology, 21*(10), 718–779.

Hermans, D., Craske, M. G., Mineka, S., & Lovibond, P. F. (2006). Extinction in human fear conditioning. *Biological Psychiatry, 60*(4), 361–368.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593–602.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617–627.

Kroes, M. C., Dunsmoor, J. E., Mackey, W. E., McClay, M., & Phelps, E. A. (2017). Context conditioning in humans using commercially available immersive Virtual Reality. *Scientific Reports, 7*, 8640.

Krypotos, A. M., Effting, M., Kindt, M., & Beckers, T. (2015). Avoidance learning: A review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience, 9*, 189.

Krypotos, A. M., Vervliet, B., & Engelhard, I. M. (2018). The validity of human avoidance paradigms. *Behaviour Research and Therapy, 111*, 99–105.

Lissek, S., & Grillon, C. (2010). Overgeneralization of conditioned fear in the anxiety disorders: Putative memorial mechanisms. *The Journal of Psychology, 218*(2), 146–148.

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., ... Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy, 43*(11), 1391–1424.

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

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., ... Kruse, O. (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.

Lovibond, P. F. (1992). Tonic and phasic electrodermal measures of human aversive conditioning with long duration stimuli. *Psychophysiology, 29*(6), 621–632.

Lovibond, P. F., Davis, N. R., & O'Flaherty, A. S. (2000). Protection from extinction in human fear conditioning. *Behaviour Research and Therapy, 38*(10), 967–983.

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.

McLean, C. P., & Anderson, E. R. (2009). Brave men and timid women? A review of the gender differences in fear and anxiety. *Clinical Psychology Review, 29*(6), 496–505.

Milad, M. R., Orr, S. P., Pitman, R. K., & Rauch, S. L. (2005). Context modulation of memory for fear extinction in humans. *Psychophysiology, 42*(4), 456–464.

Mineka, S., & Oehlberg, K. (2008). The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta Psychologica, 127*(3), 567–580.

Morriss, J., Christakou, A., & Van Reekum, C. M. (2015). Intolerance of uncertainty predicts fear extinction in amygdala-ventromedial prefrontal cortical circuitry. *Biology of Mood & Anxiety Disorders, 5*, 4.

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*, Article 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.

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*(3), 326–334.

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*(2), 314–329.

Rachman, S. (2009). Psychological treatment of anxiety: The evolution of behavior therapy and cognitive behavior therapy. *Annual Review of Clinical Psychology, 5*, 97–119.

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.

Rescorla, R. A. (1996). Preservation of Pavlovian associations through extinction. *The Quarterly Journal of Experimental Psychology B, Comparative and Physiological Psychology, 49*(3), 245–258.

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

Salkovskis, P. M. (1991). The importance of behaviour in the maintenance of anxiety and panic: A cognitive account. *Behavioural and Cognitive Psychotherapy, 19*(1), 6–19.

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M., … Dymond, S. (2016). The tipping point: Value differences and parallel dorsal-ventral frontal circuits gating human approach-avoidance behavior. *NeuroImage, 136*, 94–105.

Schlund, M. W., Ludlum, M., Magee, S. K., Tone, E. B., Brewer, A., Richman, D. M., … Dymond, S. (2019). Renewal of fear and avoidance in humans to escalating threat: Implications for translational research on anxiety disorders. *Journal of the Experimental Analysis of Behavior, 113*, 153–171.

Sheynin, J., Beck, K. D., Pang, K. C., Servatius, R. J., Shikari, S., Ostovich, J., … Myers, C. E. (2014). Behaviourally inhibited temperament and female sex, two vulnerability factors for anxiety disorders, facilitate conditioned avoidance (also) in humans. *Behavioural Processes, 103*, 228–235.

Sheynin, J., Beck, K. D., Servatius, R. J., & Myers, C. E. (2014). Acquisition and extinction of human avoidance behavior: Attenuating effect of safety signals and associations with anxiety vulnerabilities. *Frontiers in Behavioral Neuroscience, 8*, 323.

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 & Function, 220*(5), 2509–2517.

Simons, J. S., & Gaher, R. M. (2005). The distress tolerance scale: Development and validation of a self-report measure. *Motivation and Emotion, 29*(2), 83–102.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory (Self-evaluation questionnaire)*. Palo Alto, CA: Consulting Psychologists 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. *Journal of Neuroscience, 29*(46), 14617–14626.

Urcelay, G. P., Symmons, K., & Prével, A. (2019). *Renewal of instrumental avoidance in humans.* https://doi.org/10.31234/osf.io/2nxkh.

van Uijen, S. L., Dalmaijer, E. S., van den Hout, M. A., & Engelhard, I. M. (2018). Do safety behaviors preserve threat expectancy? *Journal of Experimental Psychopathology, 9*(4), 1–14.

Vervliet, B., & Indekeu, E. (2015). Low-cost avoidance behaviors are resistant to fear extinction in humans. *Frontiers in Behavioral Neuroscience, 9*, 351.

Vervliet, B., Baeyens, F., Van den Bergh, O., & Hermans, D. (2013). Extinction, generalization, and return of fear: A critical review of renewal research in humans. *Biological Psychology, 92*(1), 51–58.

Vervliet, B., Lange, I., & Milad, M. R. (2017). Temporal dynamics of relief in avoidance conditioning and fear extinction: Experimental validation and clinical relevance. *Behaviour Research and Therapy, 96*, 66–78.

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: State of the art. *Annual Review of Clinical Psychology, 9*, 215–248.

Xia, W., Dymond, S., Lloyd, K., & Vervliet, B. (2017). Partial reinforcement of avoidance and resistance to extinction in humans. *Behaviour Research and Therapy, 96*, 79–89.

Yegiyan, N. S., & Lang, A. (2010). Processing central and peripheral detail: How content arousal and emotional tone influence encoding. *Media Psychology, 13*(1), 77–99.