Potentiation of the startle reflex is in line with contingency reversal instructions rather than the conditioning history

Gaëtan Mertens & Jan De Houwer

Department of Experimental Clinical and Health Psychology, Ghent University

Correspondence concerning this article should be addressed to Gaëtan Mertens, Department of Experimental Clinical and Health Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium.

E-mail: Gaetan.Mertens@UGent.be

Tel: +32 9 264 86 13

Fax: +32 9 264 64 89

Abstract

In the context of fear conditioning, different psychophysiological measures have been related to different learning processes. Specifically, skin conductance responses (SCRs) have been related to cognitive expectancy learning, while fear potentiated startle (FPS) has been proposed to reflect affective learning that operates according to simple associative learning principles. On the basis of this two level account of fear conditioning we predicted that FPS should be less affected by verbal instructions and more affected by direct experience than SCRs. We tested this hypothesis by informing participants that contingencies would be reversed after a differential conditioning phase. Our results indicate that contingency reversal instructions led to an immediate and complete reversal of FPS regardless of the previous conditioning history. This change was accompanied by similar changes on US expectancy ratings and SCRs. These results conform with an expectancy model of fear conditioning but argue against a two level account of fear conditioning.

Keywords: Fear conditioning; Instructions; Fear Potentiated Startle; Skin Conductance Response; Reversal learning; Expectancy learning; Affective learning

Potentiation of the startle reflex is in line with contingency reversal instructions rather than the conditioning history

Fear conditioning is an adaptive process through which organisms learn to fear and avoid a conditioned stimulus (CS) that has been paired with an aversive event (unconditioned stimulus, US). This is usually modeled in the lab by pairing neutral stimuli (lights, geometric shapes) with an unpleasant but harmless electric stimulus. For humans, fear conditioning is often believed to be mediated by the generation of cognitive expectancies about the occurrence of the US in the presence of the CS (e.g., Lovibond & Shanks, 2002; Mitchell et al., 2009; Reiss, 1980). However, according to the two level account of human fear conditioning (e.g., Hamm & Weike, 2005; Öhman & Mineka, 2001; Sevenster, Beckers, & Kindt, 2012a), this cognitive contingency learning between the CS and the US is supplemented with affective learning. Affective learning is proposed to be a highly automatic process, taking place independent of cognitive contingency learning (Baeyens, Eelen, & Crombez, 1995; Hamm & Weike, 2005; Mineka & Öhman, 2002; Öhman & Mineka, 2001) and mediated by a specifically dedicated neural system centered on the amygdala (Mineka & Öhman, 2002; Öhman & Mineka, 2001).

These two forms of learning have been mapped onto different physiological responses (Hamm & Weike, 2005). Conditioned skin conductance responses (SCRs) are usually considered to reflect cognitive contingency learning about the occurrence of the aversive US in the presence of the CS (e.g., Dawson & Furedy, 1976; Lovibond & Shanks, 2002). This hypothesis is supported by studies showing that conditioning of the SCRs only occurs when participants are aware of the CS-US contingencies (e.g., Dawson, 1970; Dawson & Furedy, 1976; Sevenster et al., 2014; Singh et al., 2013) and that conditioned SCRs are very sensitive to verbal instructions (Hugdahl, 1978; Luck & Lipp, 2015b; Sevenster et al., 2012a). Conditioned potentiation of the startle reflex (or fear potentiated startle, FPS), on the other hand, is believed to primarily reflect affective learning. Evidence for this idea was provided by studies suggesting that conditioning of the startle reflex does not require awareness of the CS-US contingency (Hamm & Weike, 2005; Hamm & Vaitl, 1996; Sevenster et al., 2014) and that FPS is less sensitive to verbal instructions (Dawson, Rissling, Schell, & Wilcox, 2007; Sevenster et al., 2012). Furthermore, in a number of recent psychopharmacological studies, FPS was abolished by the administration of propranolol during a reconsolidation period while leaving expectancy of the US and SCRs intact, demonstrating a strong dissociation between FPS and cognitive measures of conditioned fear (Kindt, Soeter, & Vervliet, 2009; Soeter & Kindt, 2010). However, the evidence is not unequivocal. For instance, in a number of other studies, conditioning of the startle reflex was obtained only for participants who became aware of the CS-US contingencies (Dawson, Rissling, et al., 2007; Grillon, 2002; Jovanovic et al., 2006; Purkis & Lipp, 2001).

In the current study, we tested a different prediction that follows from the proposal that FPS reflect automatic affective learning. That is, if FPS primarily reflects simple associative learning, it should primarily be a function of the past stimulus pairings (i.e., conditioning history; Lipp & Purkis, 2005) and should be relatively insensitive to verbal instructions about future stimulus pairings (Mineka & Öhman, 2002; Sevenster et al., 2012a). To test this hypothesis, we made use of the contingency reversal procedure (Grings, Schell, & Carey, 1973; McNally, 1981; Wilson, 1968). In this procedure, participants are informed after a differential conditioning phase that the contingencies of the first phase will be reversed in a second phase. Consequently, in this second phase, cognitive contingency information as provided by the verbal instructions is directly opposed to what has been learned through CS-US pairings in the first phase. If learning is a function of experienced stimulus pairings, conditioned responses in the second phase should be in line with the conditioning history of the first phase. However, if learning is the result of cognitive beliefs regarding the CS-US contingency, conditioned responses should be in line with the verbal instructions, regardless of the conditioning history. In previous studies employing this procedure with SCRs as the measure of conditioning (Grings et al., 1973; McNally, 1981; Wilson, 1968), conditioning in the second phase of the experiment was in line with the verbal instructions and no evidence for effects of past stimulus pairings was obtained.

In a recent study by Costa, Bradley, and Lang (2015), fear was installed in a first phase by providing threat information to participants. In a second phase, one of the threatened CSs was instructed to be safe, while the other threatened CS remained a threatening stimulus. Similarly, for the initially safe CSs, one of these was threatened, while the other CS remained safe. This adapted reversal procedure allowed them to compare reversed and non-reversed CSs after the reversal instructions and thus controlled for time-related changes (e.g., habituation, sensitization) that could explain the reversal effect. Costa et al. (2015) found that fear reactions, including FPS, completely reversed on the basis of verbal contingency instructions, which demonstrates that FPS is very sensitive to cognitive information (see also: Grillon, Ameli, Woods, Merikangas, & Davis, 1991). However, conditioned responses in their study were instantiated only via verbal threat instructions and not by direct conditioning. Therefore, pairings of the CS in close proximity to the US were absent in the study of Costa et al. (2015), possibly excluding simple associative learning as the result of actual stimulus pairings (Blair, Schafe, Bauer, Rodrigues, & LeDoux, 2001). Hence, it is possible that affective learning did not take place in the study of Costa et al. (2015) due to the absence of CS-US pairings (see also: Olsson & Phelps, 2007, 2004). Therefore, in the current study, we set out to investigate whether FPS to a threatened CS can be reversed on the basis of verbal instructions, even when this CS has actually been paired with the US. Furthermore, we included threatened, but not actually conditioned CSs in our experiment to conceptually replicate the results of Costa et al. (2015) and to compare reversal on these CSs to reversal of threatened CSs that have been actually paired with the US. In line with the hypothesis that FPS reflects affective learning, we predicted that reversal of conditioned responses would be less pronounced for FPS than for SCR and ratings of US expectancy when threat instructions are combined with direct CS-US pairings.

Method

Participants

Thirty-six right-handed students (11 men, 25 women) at Ghent University participated in the experiment in exchange for €8. Age ranged between 18 and 32 years (M = 21.44, SD = 2.66). Psychophysiological data from one participant was lost due to a recording error. All participants completed an informed consent form and were instructed that they could discontinue the experiment at any point without any negative consequences. This study was approved by the ethics committee of the Faculty of Psychology and Educational Sciences of Ghent University.

Material

Conditioned Stimuli. CSs were six white geometric shapes (circle, square, triangle, pentagon, trapezium and diamond) with a maximal radius, longitude and/or latitude of 300 pixels presented in the middle of a 17 inch Dell computer screen (resolution: 1024 by 768 pixels) with a black background. Assignment of these shapes to the different CS types (see Table 1) was randomized over participants.

Table 1. Overview of the Different CS Types.

CS	Relationship with the US	Contingency reversal?
CS+T/P	Threatened + paired	No
R-CS+T/P	Threatened + paired	Yes
CS+T	Threatened	No
R-CS+T	Threatened	Yes
CS-	Safe	No
R-CS-	Safe	Yes

Unconditioned Stimulus. The US was an electric stimulus that consisted of 10 rectangular pulses of 2 ms with and inter pulse interval of 8 ms, creating an electric stimulus of 100 ms. This stimulus was administered by two lubricated Fukuda standard Ag/AgCl electrodes (1-cm diameter; inter-electrode distance: ~2-cm) to the left leg over the retromalleolar course of the sural nerve. The stimulus was generated by a constant current stimulator (DS7A, Digitimer, Hertfordshire, UK). The intensity of the electric stimulus was determined for each participant individually to be unpleasant but not painful using a stepwise work-up procedure (see the Procedure section for details concerning this work-up procedure).

Psychophysiology

Skin Conductance Responses (SCRs). SCRs were collected using a Coulbourn V71-23 skin conductance coupler (Coulbourn Instruments, Allentown, PA) and disposable Ag/AgCl electrodes (3M Red Dot 2259-50, 17 mm diameter) attached to the thenar and hypothenar eminences of the non-dominant hand. The signal was measured using a constant voltage coupler

(0.5 V) and digitized at 10 Hz. The collected data were smoothed and further analyzed offline with Psychophysiological Analysis (PSPHA) (De Clercq, Verschuere, De Vlieger, & Crombez, 2006). SCRs were calculated by subtracting the mean value of a baseline period (2 seconds before CS onset) from the highest amplitude within a 1 to 7 seconds interval after CS onset (Milad, Orr, Pitman, & Rauch, 2005; Pineles, Orr, & Orr, 2009; Raes, De Houwer, De Schryver, Brass, & Kalisch, 2014; Soeter & Kindt, 2012). In this scoring method, negative values and values smaller than 0.02µS were recoded to zero. Finally, collected SCRs were range corrected with the highest recorded amplitude for that participant to account for individual differences in responsivity (Lykken & Venables, 1971) and square root transformed to normalize the data (Dawson, Schell, Filion, & Berntson, 2007).

Fear Potentiated Startle (FPS). FPS was measured using two miniature Ag/AgCl electrodes (0.5 cm diameter) filled with conductive gel. One electrode was placed just below the pupil of the left eye and the other electrode was placed approximately 1 cm laterally. A ground electrode was placed in the middle of the forehead (Blumenthal et al., 2005). Electrode sites were first gently cleaned with scrub gel and water. The raw electromyographic signal was amplified 50,000 times, filtered online (band pass: 13 – 1000 Hz) and digitally stored at 1000 Hz using a Coulbourn V75-01 bioamplifier (Coulbourn Instruments, Allentown, PA). The acquired data were rectified and smoothed in the area of interest (0 – 150 ms after probe onset) with a FIR filter (Nitschke, Miller, & Cook, 1998) using PSPHA. The startle probe was a 50 ms white noise burst (104 dB) generated using a V85-05C Coulbourn audio module and administered via Sennheiser headphones.

The acquired signal was scored semi-automatically using PSPHA by subtracting the mean baseline value (0 - 20 ms after probe onset) from the peak value in the 21 - 150 ms window after probe onset. All startle responses were visually inspected and scored as missing values if a voluntary blink occurred just before, during or after probe onset, or if there were any other artifacts obscuring the startle response. On average, 4.25% of the trials were scored as missing for each participant (*SD* = 3.38; *Range* = 0% – 11.11%). The scores were subsequently T-transformed to control for inter-individual differences in responsivity.

US expectancy ratings

US expectancy ratings were collected on each trial using a 9-point Likert scale presented below the CSs with 5 anchor points: 1 = “not at all”, 3 = “probably not”, 5 = “uncertain”, 7 = “to some extent”, 9 = “certainly”. Above the CSs, the question “To what extent do you expect the electric stimulus?” was presented. Participants indicated their answer by clicking one of the response options of the Likert scale with the computer mouse using their dominant hand.

Questionnaire

The trait version of the State-Trait Anxiety Inventory (STAI-T; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983; Dutch translation: van der Ploeg, Defares, & Spielberger, 2000) was used to determine the general anxiety level of the participants.

Procedure

Work-up procedure. After filling in the informed consent form and the STAI-T questionnaire (Spielberger et al., 1983; van der Ploeg et al., 2000), participants first went through a work-up procedure to determine the intensity level of the electric stimulus. During this procedure, participants were exposed to gradually increasing stimulus intensity levels and were asked to report on their experience. Specifically, participants were asked after each intensity level to verbally rate the electric stimulus on a painfulness scale ranging from zero (not painful at all) to ten (maximally tolerable pain). A minimal painfulness threshold for the electric stimulus was set at seven. The procedure was stopped when participants indicated that they felt uncomfortable experiencing higher intensities of the electric stimulus. If a participant gave a rating of less than seven and indicated that he or she did not want to experience a more intense electric stimulus, the work-up was also stopped and the stimulus with the highest tolerated intensity was selected[1]. The final selected electric stimulus intensity levels ranged between 1.6 and 14 mA (*M* = 5.00, *SD* = 2.62) and pain ratings ranged between 6 and 9.5 (*M* = 7.88, *SD* = 0.81). After the work-up procedure, psychophysiology recording electrodes were applied as described above. Finally, headphones for the startle probe administration were put on. Participants were verbally informed that these headphones served to present loud but harmless noises to them throughout the experiment.

Contingency instructions and memory test. After the work-up procedure, further instructions regarding the experiment were provided on the computer screen in the absence of the experimenter. Participants were asked to read the instructions carefully. The instructions started with an overview of the different geometric shapes together with the names of these shapes to make sure participants would understand the instructions regarding these shapes. Next, participants were told that some of the shapes would be followed by the electric stimulus and that their task was to indicate to what extent they expect that stimulus after the shape by clicking one

[1] The results remained similar when the data of four participants who did not reach the painfulness threshold were excluded.

of the options on the scale below the shape. Additionally, participants were told that even CSs that could be followed by an electric stimulus would often not be followed by an electric stimulus. This instruction was added to keep the instructions about CS-US relations credible for the instructed CS+s that were never actually paired with a US (see Table 1; see the supplementary materials for a translation of the instructions in the experiment).

Next, participants were instructed about which four geometrical shapes would be predictive of the electric stimulus and which other two shapes would not be followed by the electric stimulus. Participants were told to remember these instructions well because they would afterwards be tested to ascertain that they had memorized the instructions. During this test, participants were shown all the different geometrical shapes twice in a random order. They were asked to indicate for each shape whether it could be followed by the electric stimulus by clicking one of three response buttons projected on the computer screen below the shapes (yes, no, uncertain). There was no response deadline during the test. After responding, participants received feedback for 400 ms indicating whether they were correct. If they made an error on one of the twelve trials or indicated that they were unsure about the correct response, they received the contingency instructions again and had to redo the test, until they passed it (*average number of memory tests until pass* = 1.31, *SD* = 0.58, *Range* = 1-3).

Conditioning phase. Following the contingency instructions, participants continued to the first phase of the experiment. This phase started with six startle probe habituation trials (ITI: 7 s). After these habituation trials, all six different CSs were presented six times (36 trials total). CSs were presented on the screen for eight seconds and were preceded by a fixation cross presented for one second. Startle probes were administered on each trial seven seconds after CS onset

(Sevenster et al., 2012a). The ITI was 13, 15 or 17 seconds, randomly selected. Trial order was randomized in small blocks containing two presentations of each CS (limiting the number of consecutive identical trials to maximally four). Two of the four CSs that had been instructed to be contingent with the electric stimulus were followed by the stimulus immediately at offset on three trials (50% reinforcement rate; CS+T/P's). The other two threatened shapes were never reinforced (CS+T's, see Table 1). The reinforcement rate of the CS+T/P's was chosen to be 50% because this is low enough to maintain the credibility of the threat instructions for the threatened but not paired CS+T's, but also high enough to allow for a sufficient number of CS-US pairings for the CS+T/P's. Reinforcement of the CS+T/P's with the US was distributed equally over the course of the conditioning phase (one reinforcement on the first or the second trial, the third or the fourth trial and the fifth or the sixth trial, randomly determined). SCRs, FPS and US expectancy ratings were collected on every trial as described above.

Reversal instructions and memory test. Following the conditioning phase, participants received new instructions that informed them that in the next phase of the experiment, other shapes would predict the electric stimulus. Three shapes were instructed to be predictive of the electric stimulus during this second phase, of which one was previously reinforced (CS+T/P), one was previously threatened (CS+T) and one was previously safe (R-CS-; see Table 1). Furthermore, three shapes were instructed to not be followed by electric stimulus during the second phase, of which one was previously paired with the stimulus (R-CS+T/P), one was previously threatened (R-CS+T) and one was previously safe (CS-; see Table 1). As for the previous contingency instructions, participants again had to complete a short test to make sure that they remembered these new instructions. The procedure of this memory test was the same as for the previous test. Participants could again continue to the next part of the experiment only if they completed the test without making errors (*average number of memory tests until pass* = 1.39, *SD* = 0.64, *Range* = 1-3).

Reversal phase and believability rating. The procedure of the reversal phase was identical to the procedure of the conditioning phase except that none of the CSs were reinforced during this phase. The reversal phase was followed by a final question that asked participants to indicate to what extent they found the instructions of the experiment believable at the moment they received them. They could select one option of a dropdown list: “not believable”, “not very believable”, “believable” and “very believable”.

Data analysis

In order to present the data of this relatively complex experiment in a concise and clear way, we averaged our collected measures for each of the CSs over trials within the two phases, thus ignoring the factor trial. Exclusion of this factor did not alter our conclusions because the results for the different CSs were consistent over trials. Results and graphs including the factor trial are provided in the supplementary materials. The averaged data were first analyzed with the within-subject factors CS (CS+T/P, CS+T, CS-), reversal (yes, no) and phase (conditioning, reversal). The crucial effect in this analysis is, for our purposes, the three-way interaction which indicates whether reversal instructions resulted in a reduction of conditioned fear for the reversed relative to the consistent CS+T/P and CS+T in the reversal phase, but an increase in conditioned fear for the reversed relative to the consistent CS-. Furthermore, results from the reversal phase were analyzed separately employing omnibus ANOVAs. Specifically, in a first ANOVA we compared the three CSs that, according to the instructions for the second phase, could be followed by the US (i.e., CS+T/P, CS+T, R-CS-; see Table 1). If the prior conditioning history carried over to the reversal phase, conditioned fear reactions in this phase to the CS+T/P should be higher than to the R-CS- and CS+T. A similar ANOVA was carried out to compare the different CSs that, according to the instructions for the second phase, would not be followed by a US (CS-, R-CS+T/P, R-CS+T; see Table 1). Again, if the conditioning history of the conditioning phase carried over to the reversal phase, conditioned fear reactions to the R-CS+T/P should be higher than to the R-CS+T and the CS-. Degrees of freedom of these ANOVAs were corrected with Greenhouse-Geisser corrections when the sphericity assumption was violated. Finally, we calculated Bayes Factors (BF) using JASP (version 0.6; Love et al., 2015) for our different ANOVAs to complement the results of these traditional analyses. As discussed extensively elsewhere, there are several important limitations to classical null hypothesis significance testing (NHST) (e.g., Wagenmakers, 2007). A problem that is of particular relevance for our own research question is that a non-significant result in NHST does not provide evidence for the null hypothesis (and hence, does not provide evidence for the absence of an effect). Thus, the absence of a significant effect in the traditional ANOVAs does not inform us whether there was a genuine absence of an effect of the prior conditioning phase or of verbal instructions, or whether our data was inconclusive in this regard. However, Bayesian hypothesis testing does allow to quantify the evidence in favor of the null hypothesis (reflected by the BF) and thus allows to evaluate whether effects were genuinely absent or whether our data was inconclusive (e.g., Dienes, 2014; Rouder, Speckman, Sun, Morey, & Iverson, 2009). In line with Jeffreys (1961; see also: Andraszewicz et al., 2015) we considered BFs between 1/3 and 1 as anecdotal evidence for the absence of an effect. BFs smaller than 1/3 or smaller than 1/10 were considered substantial and strong evidence, respectively, for the absence of an effect. Similarly, BFs between 1 and 3 were considered anecdotal evidence for the presence of an effect, while BFs greater than 3 or 10 were considered to be substantial and strong evidence for the presence of an effect, respectively.

Results

Believability of the instructions.

Thirty-one of the participants indicated that the instructions in the experiment were believable or very believable. Three participants indicated that the instructions were not very believable. Ratings from two participants were missing. Overall, these results indicate that our instructions were considered believable by the participants. The results remained similar regardless of whether we included or excluded participants who indicated that the instructions were not very believable. Below, we report only the results for the full sample.

US expectancy ratings.

The crucial three-way interaction between CS, reversal and phase reached significance for the US expectancy ratings, $F(1.27, 44.36) = 339.08$, $p < .001$, $\eta_p^2 = .91$. This interaction was driven by a significant increase in US expectancy ratings for the R-CS- from the conditioning phase to the reversal phase, while US expectancy ratings for the R-CS+T/P and R-CS+T decreased (see Figure 1, all t-values > 18, p-values < .001, *Cohen's d*'s > 3.8). US expectancy ratings for the consistent CS-, CS+T/P and CS+T did not differ significantly across the two phases (all t-values < 1, p-values > .3, *Cohen's d*'s < 0.14; see Figure 1). Hence, US expectancy ratings were very sensitive to the contingency reversal instructions. The Bayesian analysis confirmed that this three-way interaction was a very robust result (BF = $\sim\infty$).

The omnibus ANOVA comparing the CS-, the R-CS+T/P and the R-CS+T, did not reach significance, $F(1.69, 59.17) = 2.06$, $p = .143$, $\eta_p^2 = .06$. This results suggests that there is little difference in US expectancy between a consistent CS- and a previously conditioned CS+ or threatened CS+ after contingency reversal instructions (see Figure 1). However, the result of the Bayesian analysis shows that there is only anecdotal evidence for an absence of a difference between these CSs (BF = .445). The omnibus ANOVA comparing CS+T/P, CS+T and R-CS- did reach significance, $F(1.73, 60.61) = 3.63$, $p = .039$, $\eta_p^2 = .09$. This overall effect was due to significantly higher US expectancy ratings for CS+T/P than for R-CS-, $t(35) = 2.24$, $p = .031$, *Cohen's d* = 0.17, and CS+T, $t(35) = 1.93$, $p = .062$, *Cohen's d* = 0.13 (see Figure 1). There was no significant difference between R-CS- and CS+T, $t(35) < 1$, $p = .449$, *Cohen's d* = 0.04. Hence, this result demonstrates that US expectancy was slightly elevated for a consistent CS+ that was previously paired with the electric stimulus compared to a previously threatened CS+ or a reversed CS-. However, there is only anecdotal evidence for this effect according to the corresponding Bayesian analysis (BF = 1.49).

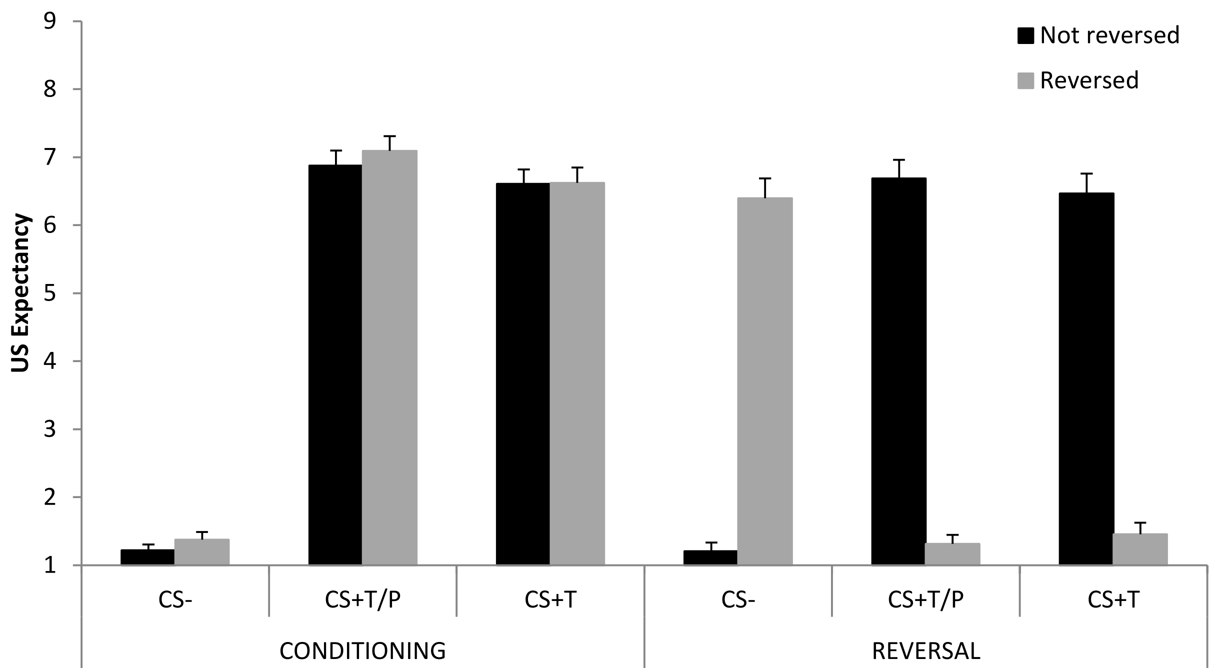

Figure 1. Mean US expectancy rating for the different types of CSs in the two phases of the experiment. Error bars represent standard errors.

SCR.

The crucial interaction between CS, reversal and phase reached significance for SCRs as well, $F(1.74, 59.20) = 5.29$, $p = .010$, $\eta_p^2 = .10$. SCRs were lower for all CSs in the reversal phase compared to the conditioning phase (t-values > 1.9, p-values < .07, *Cohen's d*'s > 0.39; see Figure 2), except for R-CS-, $t(34) = -1.17$, $p = .250$, *Cohen's d* = -0.24. More importantly, SCRs were larger for R-CS- than for CS- in the reversal phase, while SCRs for these two CSs were comparable in the conditioning phase (see Figure 2), interaction between reversal and phase for the two types of CS-'s, $F(1, 34) = 5.65$, $p = .023$, $\eta_p^2 = .14$. The reversed pattern was found for CS+T. That is, smaller SCRs were found for the R-CS+T compared to the CS+T in the reversal phase, again despite these two CSs being comparable in the conditioning phase (see Figure 2), interaction between reversal and phase for the CS+T's: $F(1, 34) = 4.50$, $p = .041$, $\eta_p^2 = .12$. However, there was no indication for an effect of verbal instructions on the CS+T/P's, interaction between reversal and phase for CS+T/P's, $F(1, 34) < 1$. That is, SCRs were comparable for CS+T/P and R-CS+T/P both in the reversal and conditioning phase (see Figure 2). These results demonstrate that our reversal instructions successfully induced larger SCRs for a reversed CS- and reduced SCRs for a reversed threatened CS+. Interestingly, however, our reversal instructions did not significantly influence conditioned SCRs to a threatened CS+ that has actually been paired with the US (i.e., CS+T/P). In fact, a Bayesian analysis showed that there was substantial evidence for an absence of an effect of verbal instructions on the CS+T/P's (BF interaction reversal and phase = 0.254). Furthermore, the Bayesian analysis of the three-way interaction between CS, reversal and phase showed that there was only anecdotal evidence for an effect of verbal instructions on SCRs (BF= 2.527).

The results from the reversal phase were again further explored using an ANOVA that compared responses to CS-, R-CS+T and R-CS+T/P. This ANOVA did not reveal a significant effect, $F(2, 68) = 2.05$, $p = .136$, $\eta_p^2 = .06$, thus failing to provide evidence for transfer effects of the conditioning phase to the reversal phase for these CSs. The corresponding Bayesian analysis showed that there was only anecdotal evidence for an absence of difference between these CSs (BF = 0.454). Likewise, the ANOVA comparing R-CS-, CS+T and CS+T/P did not reach significance, $F(2, 68) = 1.17$, $p = .318$, $\eta_p^2 = .03$, thus also failing to provide evidence for transfer effects of the conditioning history to the verbally established CS+s in the reversal phase on

SCRs. In fact, the corresponding Bayesian analysis showed that our data provided substantial evidence for an absence of transfer effects (BF = 0.221).

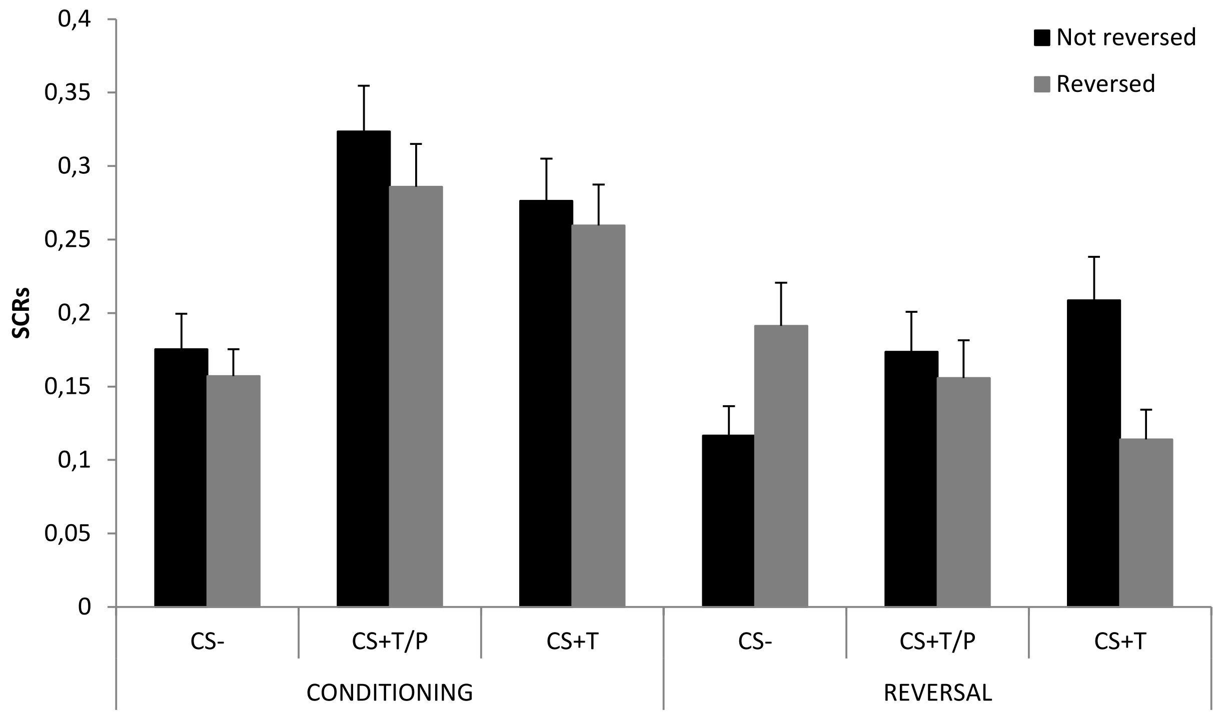

Figure 2. Mean range corrected and square root transformed SCRs (measured in µS) for the different types of CSs in the two phases of the experiment. Error bars represent standard errors.

FPS.

The crucial three-way interaction between CS, reversal and phase was significant for FPS as well, $F(1.64, 55.91) = 13.35$, $p < .001$, $\eta_p^2 = .28$. As for SCRs, a general reduction of startle magnitude was observed in the reversal phase for all CSs (see Figure 3; t-values > 3.9, p-values

< .001, *Cohen's d*'s > 1.00) except for R-CS-, $t(34) < 1$, $p = .925$, *Cohen's d* = 0.03. Importantly, reversal instructions resulted in larger FPS for R-CS- than for CS- in the reversal phase, while FPS was comparable for both these CSs in the conditioning phase (see Figure 3), interaction between phase and reversal for CS-'s: $F(1, 34) = 11.83$, $p = .002$, $\eta_p^2 = .26$. This pattern was reversed for the CS+T's, with smaller FPS for R-CS+T than for CS+T in the reversal phase, even though FPS was comparable for these two CSs in the conditioning phase (see Figure 3), interaction between phase and reversal for CS+T's: $F(1, 34) = 11.86$, $p = .002$, $\eta_p^2 = .26$. A similar pattern was obtained for CS+T/P and R-CS+T/P. That is, FPS was also smaller for R-CS+T/P than for CS+T/P in the reversal phase, while it was comparable for these two CSs in the conditioning phase (see Figure 3), interaction between phase and reversal for CS+T/P's, $F(1, 34) = 9.26$, $p = .004$, $\eta_p^2 = .21$. Combined, these results demonstrate that our verbal instructions were successful in influencing FPS both for previously safe and threatened CSs, including CSs that were actually followed by US. That is, reversal instructions resulted in an increase of FPS for the reversed CS- while it decreased FPS for a reversed threatened CS+, regardless of whether this CS+ was actually paired with the US. The Bayesian analysis confirmed the robustness of this effect of verbal instructions (BF three-way interaction CS, reversal and phase = 567 304).

The ANOVAs comparing the different CS-'s (CS-, R-CS+T/P, R-CS+T) and CS+'s (R-CS-, CS+T/P, CS+T) in the reversal phase did not reach significance, *F*-values < 1. The corresponding BFs for these respective ANOVAs were 0.160 and 0.111. Hence, our data provide substantial evidence for an absence of transfer effects of the previous conditioning history to the reversal phase for FPS.

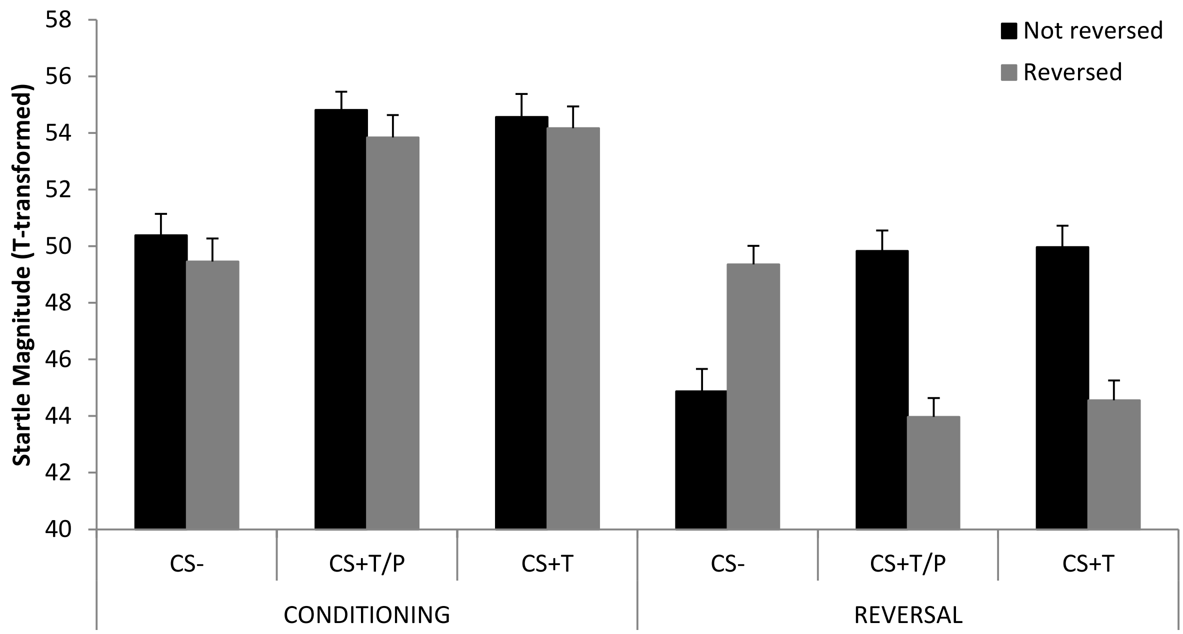

Figure 3. Mean T-transformed startle response (measured in μV) for the different types of CSs in the two phases of the experiment. Error bars represent standard errors.

STAI scores

STAI-T scores ranged between 21 and 59 (M = 37.31, SD = 9.72). For exploratory purposes, participants were divided in a low and high anxiety group by means of a median-split. When this factor was added to the analyses as a between-subjects variable, it did not interact with the crucial three-way interaction between CS, reversal and phase for any of our measures of fear, F-values < 1, BFs < 0.1. Furthermore, the three-way interaction between CS, reversal and phase remained present for both the low and high anxiety groups for all measures, F-values > 4, p-values < .05, η_p^2's > .20, except for the low anxiety group on SCRs, $F(2, 28) = 1.27$, $p = .297$, $\eta_p^2 = .08$. However, this latter result could be due to the limited size of this sample (N = 15).

Furthermore, our results should be interpreted with caution because of the small samples in the low and high anxiety groups and because there was only a limited range in anxiety scores. Nevertheless, these results indicate that even in a sample with relatively elevated anxiety scores (*Mean STAI-T* = 45.29, *SD* = 7.56, *Range* = 37 – 59), contingency reversal instructions seemed to be successful.

Discussion

In the current study, we investigated whether verbal instructions can reverse conditioned fear responses. In the two level account of human fear conditioning (Hamm & Weike, 2005; Sevenster et al., 2012), FPS is considered to be a measure of automatic affective learning that operates according to simple associative learning principles whereas SCR is assumed to capture cognitive expectancies. We therefore predicted that the effect of reversal instructions on conditioned fear reactions would be less pronounced for FPS than for SCR, especially when CSs have been repeatedly paired with the US. Our results demonstrated that all measures of conditioned fear were sensitive to the contingency reversal instructions. Interestingly, we also obtained suggestive evidence for effects of CS-US pairings for US expectancy ratings and SCRs, but not FPS.

FPS reactions in the second phase of the experiment were completely in line with the verbal instructions, and no evidence for effects of the prior CS-US pairings were obtained for FPS. These results extend the findings of Costa et al. (2015) by showing that reversal of FPS can take place even when threat instructions are combined with actual CS-US pairings. Hence, even though there was opportunity for simple associative learning to take place in the current study (Olsson & Phelps, 2004), verbal instructions still primarily determined FPS. This finding is even more striking in light of the significant impact of actual CS-US pairings on other measures that are typically considered to be more cognitive in nature (i.e., SCRs, US expectancy ratings). Hence, our results do not fit well with the two level account of fear conditioning that propose that FPS is a measure of automatic affective learning that operates according to the principles of simple associative learning (Blair et al., 2001; Lipp & Purkis, 2005) and that is independent from cognitive contingency learning as measured by SCRs and expectancy ratings (Hamm & Weike, 2005; Hamm & Vaitl, 1996; Sevenster et al., 2012a). Rather, the results in the current study suggest that FPS is very sensitive to verbal instructions and is not influenced by previous CS-US pairings.

Less surprisingly, US expectancy ratings were also very sensitive to verbal reversal instructions as illustrated by an increase in US expectancy ratings for R-CS- and a decrease for the R-CS+T/P and R-CS+T after the contingency reversal instructions (see Figure 1). Furthermore, also a small but reliable effect of the previous conditioning history was obtained for US expectancy ratings. That is, US expectancy ratings were slightly higher in the second phase for a threatened CS that was previously paired with the US (CS+T/P), compared to a threatened CS that was not previously paired with the US (CS+T) and a threatened CS that was previously safe (R-CS-). This latter result is in line with a prior study from our lab showing that US expectancy ratings were slightly elevated for a threatened CS when it was previously paired with the US (Mertens et al., 2015). However, the results of our Bayesian analysis showed that there was only anecdotal evidence for this effect in our data. Combined, these results show that participants adapted their expectancies about receiving an electric stimulus in accordance with the instructions, demonstrating that the instructions were clear. Interestingly, our data suggests that participants also took previous CS-US pairings into account when providing US expectancy ratings.

Finally, results obtained for SCRs were also in line with the verbal instructions, except for the CSs that had been actually paired with the US (CS+T/P's). For these latter CSs, SCRs were comparable in the second phase of the experiment, regardless of the reversal instructions (see Figure 2). However, results in the reversal phase for the CS+T/P's were not completely in line with the prior conditioning history either. That is, SCRs to R-CS+T/P were not significantly higher than to CS- in the reversal phase. Hence, SCRs seem to have been influenced by two opposing influences, that is, instructions on the one hand and actual CS-US pairings on the other hand. This result is in contrast with previous studies employing the contingency reversal procedure that found that SCRs were completely in line with the reversal instructions (Grings et al., 1973; McNally, 1981; Wilson, 1968). One potential reason for this discrepancy between our own results and the results from these earlier studies is the inclusion of threatened, but not actually conditioned, CSs. The fact that participants noticed that there were threatened and actually conditioned CS+s, as illustrated by the US expectancy ratings, may have prompted them to be more cautious about the instructions concerning the actually conditioned CS+s. However, this interpretation does not explain why we did not obtain a similar pattern for FPS. Alternatively, this result could suggest that SCRs reflect both simple associative learning and cognitive contingency learning. This interpretation is not in line with the findings of the studies outlined in the introduction, but does fit with the results of other studies that have found that SCRs can be dissociated under certain conditions from cognitive contingency learning (Bechara et al., 1995; McAndrew, Jones, McLaren, & McLaren, 2012). Regardless of the exact interpretation of the SCRs results, our results illustrate that SCRs were insensitive to verbal instructions when a CS had been paired with the USs, while such an effect was not observed for FPS. This finding demonstrates that the classification of FPS and SCRs as affective and cognitive measures of fear conditioning, respectively, does not correspond with our data.

Our conclusion runs counter the results of a number of studies that we mentioned in the introduction. We will discuss these studies in more detail here. First, some studies have found that conditioning of the startle reflex can occur in the absence of CS-US contingency awareness while such unaware conditioning was not obtained for SCRs (Hamm & Weike, 2005; Hamm & Vaitl, 1996; Sevenster et al., 2014). However, as mentioned before, a number of other studies have found conditioning of the startle reflex only for participants that became aware of the CS-US contingencies (Dawson, Rissling, et al., 2007; Grillon, 2002; Jovanovic et al., 2006; Purkis & Lipp, 2001). Whether fear conditioning, or learning in general, can occur without contingency awareness is a question that has been proven to be difficult to answer and that requires appropriate measurement of contingency awareness (e.g., Shanks & St. John, 1994) and careful experimental control of other contingencies that could explain learning (e.g., Singh et al., 2013). A recent study by Sevenster et al. (2014) seems to meet these two criteria and nevertheless provide evidence for unaware conditioning of FPS but not of SCRs. While these results are certainly promising, it seems premature to us to conclude that conditioning of the startle reflex can occur in the absence of awareness given the conflicting evidence. Further replication of the result of Sevenster et al. (2014) will clarify whether this claim can be upheld.

Second, two studies have shown that verbal instructions that the US will no longer be applied results in a complete reduction of SCRs but not of FPS, suggesting a dissociation between SCRs and FPS in their sensitivity to verbal instructions (Dawson, Rissling, et al., 2007; Sevenster et al., 2012a). However, in a recent study by Luck and Lipp (2015a), in which instructed extinction was combined with removal of the shock electrodes, a complete reduction of both SCRs and FPS was observed. As argued by Luck and Lipp (2015b), the incomplete reduction of FPS in the study of Sevenster et al. (2012a) can perhaps be explained by a subset of participants in their experiment that did not find the extinction instructions believable (because the shock electrodes remained attached in the study of Sevenster et al., 2012a). Furthermore, this incomplete reduction was perhaps not observed for SCRs due to increased SCRs to the CS- in the extinction phase for the instructed extinction group in the study of Sevenster et al. (2012a). A similar reasoning can also be applied to the results of Dawson, Rissling, et al. (2007) because their experiment employed a picture-picture evaluative conditioning procedure. Hence, participants could not be sure that the USs would no longer be applied (because the computer screen was not removed), resulting in unconvincing extinction instructions. However, it remains unclear from the data of Dawson, Rissling, et al. (2007) why such an effect was not obtained for SCRs because pre and post extinction SCR data are not provided in their article. Hence, taken together, the studies investigating effects of instructed extinction on SCRs and FPS do not provide definitive evidence for a dissociation between SCRs and FPS either.

Finally, a number of recent studies have demonstrated that behavioral or pharmacological manipulations during a reconsolidation phase specifically reduce FPS but leave expectancy ratings and conditioned SCRs intact (Kindt et al., 2009; Sevenster, Beckers, & Kindt, 2012b, 2013; Soeter & Kindt, 2010, 2012). These studies provide persuasive evidence that FPS can be dissociated from cognitive measures of conditioned fear. However, other studies have found a reduction of conditioned SCRs after disruption of reconsolidation as well (Oyarzún et al., 2012; Schiller et al., 2010), while still others did not find the disruption of reconsolidation effect either for FPS or SCRs (Bos, Beckers, & Kindt, 2014; Golkar, Bellander, Olsson, & Ohman, 2012).

Furthermore, erasure of fear memories through reconsolidation has been shown to depend on prediction error as captured by US expectancy ratings (Sevenster et al., 2013) and some evidence could even suggest that these disruption of reconsolidation effects are more pronounced with concurrent US expectancy ratings (Warren et al., 2014). Thus, reduction of FPS through disruption of reconsolidation may not be as independent of expectancies and SCRs as some studies suggest. Hence, considering all these different studies, currently the data available with regard to unaware learning, instructed extinction and disruption of reconsolidation do not provide definitive evidence that FPS reflects automatic affective learning. The results of the current study provide further evidence that FPS may reflect cognitive contingency learning rather than automatic affective learning.

However, there are several limitations to this study that should be acknowledged. First, as described by Öhman and Mineka (2001), the affective learning module is only selectively activated by biologically relevant or highly aversive stimuli. Therefore, mild electric stimuli as USs and geometric shapes as CSs might not be sufficient to recruit this affective learning module in the learning process. It would be interesting to conduct a follow-up study looking at whether similar results would be obtained with more ecologically valid CSs and USs. A second limitation is that the CS+T/P's and the US were paired on only three occasions throughout the experiment, which might not be a sufficient number of pairings for simple associative learning to take place. On the other hand, if affective learning is an evolutionary adaptive process, it is unlikely that a high number of CS-US pairings is required for simple associative learning to take place. Third, we gave explicit verbal instructions about all the contingencies and asked participants on every trial to provide ratings about the extent to which they expected the US. There is evidence that including online US expectancy ratings maintains fear conditioning in patients with damage in the amygdala (Coppens, Spruyt, Vandenbulcke, Van Paesschen, & Vansteenwegen, 2009). Furthermore, some studies have compared participants who were instructed about the stimulus contingencies with participants who were either unaware of the contingencies or who learned the contingencies spontaneously. These studies revealed increased activation in brain areas that have been related to decision making and cognitive control in the instructed group (e.g., rostral dmPFC, lateral OFC; Mechias, Etkin, & Kalisch, 2010; Tabbert et al., 2011). Combined, these studies suggest that online US expectancy ratings and contingency instructions may induce a more cognitive way of learning about the CS-US pairings and consequently limited the contribution of affective learning (Coppens et al., 2009). Therefore, it is possible that if CS-US contingencies are learned in a spontaneous manner, stronger effects of the previous conditioning history could be obtained.

Taking into account these reservations, we conclude that FPS should not by default be regarded as a measure of affective learning that is independent of SCRs and expectancy ratings. The results of our experiment demonstrate that conditioning of the startle reflex can depend on verbal instructions and expectancies about the occurrence of the US and does not necessarily follow simple associative learning rules.

Acknowledgements

The research reported in this paper was funded by the Interuniversity Attraction Poles Program initiated by the Belgian Science Policy Office (IUAPVII/33) and by Ghent University Methusalem Grant BOF09/01M00209.

References

Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., & Wagenmakers, E.-J. (2015). An Introduction to Bayesian Hypothesis Testing for Management Research. *Journal of Management*, *41*(2), 521–543. doi:10.1177/0149206314560412

Baeyens, F., Eelen, P., & Crombez, G. (1995). Pavlovian associations are forever: On classical conditioning and extinction. *Journal of Psychophysiology*, *9*(2), 127–141. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/0005796792901365

Bechara, a., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science*, *269*(5227), 1115–1118. doi:10.1126/science.7652558

Blair, H. T., Schafe, G. E., Bauer, E. P., Rodrigues, S. M., & LeDoux, J. E. (2001). Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. *Learning & Memory*, *8*(5), 229–242. doi:10.1101/lm.30901

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V, & van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, *42*(1), 1–15. doi:10.1111/j.1469-8986.2005.00271.x

Bos, M. G. N., Beckers, T., & Kindt, M. (2014). Noradrenergic Blockade of Memory Reconsolidation: A Failure to Reduce Conditioned Fear Responding. *Frontiers in Behavioral Neuroscience*, *8*(November), 1–8. doi:10.3389/fnbeh.2014.00412

Coppens, E., Spruyt, A., Vandenbulcke, M., Van Paesschen, W., & Vansteenwegen, D. (2009). Classically conditioned fear responses are preserved following unilateral temporal lobectomy in humans when concurrent US-expectancy ratings are used. *Neuropsychologia*, *47*(12), 2496–2503. doi:10.1016/j.neuropsychologia.2009.04.021

Costa, V. D., Bradley, M. M., & Lang, P. J. (2015). From threat to safety: Instructed reversal of defensive reactions. *Psychophysiology*, *52*(3), 325–332. doi:10.1111/psyp.12359

Dawson, M. E. (1970). Cognition and conditioning: Effects of masking the CS-UCS contingency on human GSR classical conditioning. *Journal of Experimental Psychology*, *85*(3), 389–396. doi:10.1037/h0029715

Dawson, M. E., & Furedy, J. J. (1976). The Role of Awareness in Human Differential Autonomic Classical Conditioning: The Necessary-Gate Hypothesis. *Psychophysiology*, *13*(1), 50–53. doi:10.1111/j.1469-8986.1976.tb03336.x

Dawson, M. E., Rissling, A. J., Schell, A. M., & Wilcox, R. (2007). Under what conditions can human affective conditioning occur without contingency awareness? Test of the evaluative conditioning paradigm. *Emotion*, *7*(4), 755–66. doi:10.1037/1528-3542.7.4.755

Dawson, M. E., Schell, A. M., Filion, D. L., & Berntson, G. G. (2007). The Electrodermal System. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of Psychophysiology* (3rd ed., pp. 157–181). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511546396

De Clercq, A., Verschuere, B., De Vlieger, P., & Crombez, G. (2006). Psychophysiological Analysis (PSPHA): A modular script-based program for analyzing psychophysiological data. *Behavior Research Methods*, *38*(3), 504–510. doi:10.3758/BF03192805

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*(July), 1–17. doi:10.3389/fpsyg.2014.00781

Golkar, A., Bellander, M., Olsson, A., & Ohman, A. (2012). Are fear memories erasable?-reconsolidation of learned fear with fear-relevant and fear-irrelevant stimuli. *Frontiers in Behavioral Neuroscience*, *6*(November), 80. doi:10.3389/fnbeh.2012.00080

Grillon, C. (2002). Associative learning deficits increase symptoms of anxiety in humans. *Biological Psychiatry*, *51*(11), 851–858. doi:10.1016/S0006-3223(01)01370-1

Grillon, C., Ameli, R., Woods, S. W., Merikangas, K., & Davis, M. (1991). Fear-Potentiated Startle in Humans: Effects of Anticipatory Anxiety on the Acoustic Blink Reflex. *Psychophysiology*, *28*(5), 588–595. doi:10.1111/j.1469-8986.1991.tb01999.x

Grings, W. W., Schell, A. M., & Carey, C. a. (1973). Verbal control of an autonomic response in a cue reversal situation. *Journal of Experimental Psychology*, *99*(2), 215–221. doi:10.1037/h0034653

Hamm, A. O., & Weike, A. I. (2005). The neuropsychology of fear learning and fear regulation. *International Journal of Psychophysiology*, *57*(1), 5–14. doi:10.1016/j.ijpsycho.2005.01.006

Hamm, A., & Vaitl, D. (1996). Affective learning: Awareness and aversion. *Psychophysiology*, *33*, 698–710. doi:10.1111/j.1469-8986.1996.tb02 366.x

Hugdahl, K. (1978). Electrodermal conditioning to potentially phobic stimuli: effects of instructed extinction. *Behaviour Research and Therapy*, *16*(5), 315–321. doi:10.1016/0005-7967(78)90001-3

Jeffreys, H. (1961). *Theory of Probability*. Oxford: Oxford University Press.

Jovanovic, T., Norrholm, S. D., Keyes, M., Fiallos, A., Jovanovic, S., Myers, K. M., … Duncan, E. J. (2006). Contingency awareness and fear inhibition in a human fear-potentiated startle paradigm. *Behavioral Neuroscience*, *120*(5), 995–1004. doi:10.1037/0735-7044.120.5.995

Kindt, M., Soeter, M., & Vervliet, B. (2009). Beyond extinction: erasing human fear responses and preventing the return of fear. *Nature Neuroscience*, *12*(3), 256–258. doi:10.1038/nn.2271

Lipp, O. V, & Purkis, H. M. (2005). No support for dual process accounts of human affective learning in simple Pavlovian conditioning. *Cognition & Emotion*, *19*(2), 269–82. doi:10.1080/02699930441000319

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(1), 3–26. doi:10.1037//0097-7403.28.1.3

Luck, C. C., & Lipp, O. V. (2015a). A potential pathway to the relapse of fear? Conditioned negative stimulus evaluation (but not physiological responding) resists instructed extinction. *Behaviour Research and Therapy*, *66*, 18–31. doi:10.1016/j.brat.2015.01.001

Luck, C. C., & Lipp, O. V. (2015b). To remove or not to remove? Removal of the unconditional stimulus electrode does not mediate instructed extinction effects. *Psychophysiology*, *00*, n/a–n/a. doi:10.1111/psyp.12452

Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: a proposal for standardization. *Psychophysiology*, *8*(5), 656–672. doi:10.1111/j.1469-8986.1971.tb00501.x

McAndrew, a, Jones, F. W., McLaren, R. P., & McLaren, I. P. L. (2012). Dissociating expectancy of shock and changes in skin conductance: an investigation of the Perruchet effect using an electrodermal paradigm. *Journal of Experimental Psychology. Animal Behavior Processes*, *38*(2), 203–8. doi:10.1037/a0026718

McNally, R. J. (1981). Phobias and preparedness: instructional reversal of electrodermal conditioning to fear-relevant stimuli. *Psychological Reports*, *48*(1), 175–180. doi:10.2466/pr0.1981.48.1.175

Mechias, M.-L., Etkin, A., & Kalisch, R. (2010). A meta-analysis of instructed fear studies: implications for conscious appraisal of threat. *NeuroImage*, *49*(2), 1760–8. doi:10.1016/j.neuroimage.2009.09.040

Mertens, G., Kuhn, M., Raes, A. K., Kalisch, R., De Houwer, J., & Lonsdorf, T. B. (2015). Fear expression and return of fear following threat instruction with or without direct contingency experience. *Cognition and Emotion*, (May), 1–17. doi:10.1080/02699931.2015.1038219

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), 456–64. doi:10.1111/j.1469-8986.2005.00302.x

Mineka, S., & Öhman, A. (2002). Phobias and preparedness: the selective, automatic, and encapsulated nature of fear. *Biological Psychiatry*, *52*(10), 927–937. doi:10.1016/S0006-3223(02)01669-4

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *The Behavioral and Brain Sciences*, *32*(2), 183–98; discussion 198–246. doi:10.1017/S0140525X09000855

Nitschke, J. B., Miller, G. a., & Cook, E. W. (1998). Digital filtering in EEG/ERP analysis: Some technical and empirical comparisons. *Behavior Research Methods, Instruments, & Computers*, *30*(1), 54–67. doi:10.3758/BF03209416

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522. doi:10.1037//0033-295X.108.3.483

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

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

Oyarzún, J. P., Lopez-Barroso, D., Fuentemilla, L., Cucurell, D., Pedraza, C., Rodriguez-Fornells, A., & de Diego-Balaguer, R. (2012). Updating fearful memories with extinction training during reconsolidation: A human study using auditory aversive stimuli. *PLoS ONE*, *7*(6). doi:10.1371/journal.pone.0038849

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology*, *46*(5), 984–95. doi:10.1111/j.1469-8986.2009.00852.x

Purkis, H., & Lipp, O. V. (2001). Does Affective Learning Exist in the Absence of Contingency Awareness? *Learning and Motivation*, *32*(1), 84–99. doi:10.1006/lmot.2000.1066

Raes, A. K., De Houwer, J., De Schryver, M., Brass, M., & Kalisch, R. (2014). Do CS-US Pairings Actually Matter? A Within-Subject Comparison of Instructed Fear Conditioning with and without Actual CS-US Pairings. *PloS One*, *9*(1), e84888. doi:10.1371/journal.pone.0084888

Reiss, S. (1980). Pavlovian conditioning and human fear: An expectancy model. *Behavior Therapy*, *11*(3), 380–396. doi:10.1016/S0005-7894(80)80054-2

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. doi:10.3758/PBR.16.2.225

Schiller, D., Monfils, M.-H., Raio, C. M., Johnson, D. C., Ledoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*, *463*, 49–53. doi:10.1038/nature08637

Sevenster, D., Beckers, T., & Kindt, M. (2012a). Instructed extinction differentially affects the emotional and cognitive expression of associative fear memory. *Psychophysiology*, *49*(10), 1426–1435. doi:10.1111/j.1469-8986.2012.01450.x

Sevenster, D., Beckers, T., & Kindt, M. (2012b). Retrieval per se is not sufficient to trigger reconsolidation of human fear memory. *Neurobiology of Learning and Memory*, *97*(3), 338–345. doi:10.1016/j.nlm.2012.01.009

Sevenster, D., Beckers, T., & Kindt, M. (2013). Prediction error governs pharmacologically induced amnesia for learned fear. *Science (New York, N.Y.)*, *339*(6121), 830–3. doi:10.1126/science.1231357

Sevenster, D., Beckers, T., & Kindt, M. (2014). Fear conditioning of SCR but not the startle reflex requires conscious discrimination of threat and safety. *Frontiers in Behavioral Neuroscience*, *8*(32). doi:10.3389/fnbeh.2014.00032

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning-systems. *Behavioral and Bain Sciences*, *17*(3), 367–395. doi:10.1017/S0140525X00035032

Singh, K., Dawson, M. E., Schell, A. M., Courtney, C. G., & Payne, A. F. H. (2013). Can human autonomic classical conditioning occur without contingency awareness? The critical importance of the trial sequence. *Biological Psychology*, *93*(1), 197–205. doi:10.1016/j.biopsycho.2013.02.007

Soeter, M., & Kindt, M. (2010). Dissociating response systems: erasing fear from memory. *Neurobiology of Learning and Memory*, *94*(1), 30–41. doi:10.1016/j.nlm.2010.03.004

Soeter, M., & Kindt, M. (2012). Erasing fear for an imagined threat event. *Psychoneuroendocrinology*, *37*(11), 1769–1779. doi:10.1016/j.psyneuen.2012.03.011

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Tabbert, K., Merz, C. J., Klucken, T., Schweckendiek, J., Vaitl, D., Wolf, O. T., & Stark, R. (2011). Influence of contingency awareness on neural, electrodermal and evaluative responses during fear conditioning. *Social Cognitive and Affective Neuroscience*, *6*(4), 495–506. doi:10.1093/scan/nsq070

Van der Ploeg, H. M., Defares, P. B., & Spielberger, C. D. (2000). *Handleiding bij de Zelfbeoordelings Vragenlijst. Een Nederlandstalige bewerking van de Spielberger State-Trait Anxiety Inventory*. Lisse, The Netherlands.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems ofp values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. doi:10.3758/BF03194105

Warren, V. T., Anderson, K. M., Kwon, C., Bosshardt, L., Jovanovic, T., Bradley, B., & Norrholm, S. D. (2014). Human fear extinction and return of fear using reconsolidation update mechanisms: The contribution of on-line expectancy ratings. *Neurobiology of Learning and Memory*, *113*(OCTOBER), 165–173. doi:10.1016/j.nlm.2013.10.014

Wilson, G. D. (1968). Reversal of differential GSR conditioning by instructions. *Journal of Experimental Psychology*, *76*(3, Pt.1), 491–493. doi:10.1037/h0025540