Potency of Instructed Fear Reversal on Fear Responses: The Underlying Psychological Mechanisms

Twinkle Darlene Correa

S6538770

MSc Clinical Psychology, Utrecht University

Supervisor: Dr. Gaëtan Mertens

March 13th, 2020

Acknowledgements

The completion of this thesis may not have been possible without the love and support of many people. I would like to express my utmost gratitude and deep appreciation to the following:

Dr. Gaëtan Mertens for your time and patience in assisting me with my research. Thank you for your useful critiques and insightful comments.

My family for your endless love and guidance. Thank you for being my main pillars in my life. I am so grateful to have each and every one of you.

Abstract

Fear conditioning is a study of much interest because of its contribution to understanding fear-related disorders. In fear conditioning, a conditioned stimulus (CS) is repeatedly paired with an aversive unconditioned stimulus (US) resulting in conditioned fear responses to the CS. Of importance, learned contingencies are shown to be reversible by the provision of contingency reversal instructions. Indeed, fear reversal allows one to flexibly readjust reactions to altered circumstances. It seems crucial then to clarify the psychological mechanisms behind fear reversal for further understanding of appropriate and inappropriate control of fear. Research has identified two competing models of fear learning, the single process model and the dual process model, that differently explain the influence of reversal instructions on fear responses. While some studies show support to a single process (e.g., Mertens & De Houwer, 2016), other findings are consistent with dual processes (e.g., Soeter & Kindt, 2010), thus resulting in an inconclusive gap in the fear conditioning literature. Therefore, the study aimed to examine the malleability of skin conductance response and fear potentiated startle to contingency reversal instructions. Using multiple repeated measures ANOVAs, findings revealed that both fear responses immediately reversed, therefore, supporting the single process model in explaining conditioned fear. The findings of the study provide insight to appropriate emotional control and the role of cognitive expectancies which could be of clinical relevance in understanding the control of fear in anxiety disorders and facilitating its treatment. Limitations and important future recommendations are also discussed.

Keywords: instructions, cognitive expectancy, fear responses, fear reversal

Potency of Instructed Fear Reversal on Fear Responses: The Underlying Psychological Mechanisms

Fear learning is an adaptive ability where species use environmental cues to predict potential dangers ahead of time and form long-lasting fear memories even after a single learning experience (Schiller & Delgado, 2010). Translated from animal behavior research, fear conditioning in humans has been studied for a century (Watson & Rayner, 1920) and is often used in various research areas such as psychopharmacology and behavioral neuroscience (Lonsdorf et al., 2017). Indeed, studying human fear conditioning has contributed considerably to the understanding of how a fear memory is acquired, expressed, and maintained (Sevenster, Beckers & Kindt, 2012). In the laboratory, fear learning is studied through the use of the fear conditioning paradigm. During the acquisition phase, one of the stimuli (i.e., the conditioned stimulus, CS+) is followed by an aversive unconditional stimulus (US) whilst the other stimulus (CS-) is not. Therefore, the CS+ becomes predictive of the US and acquires the capacity to elicit conditional fear responses (CR) due to the anticipation of the US (Luck & Lipp, 2015). An electrical shock or loud sounds are commonly used as USs and neutral stimuli such as colored lights or geometrical shapes are typically used as CSs (Öhman & Mineka, 2001). Common physiological responses involved with conditioned fear are an increase in skin conductance response (SCR) and potentiation of the startle reflex (FPS).

With regard to the underlying psychological mechanism for fear conditioning, there is still ongoing debate. According to the expectancy model of Davey (1992), the mediating factor of cognitive expectancies explain the conditioning of fear which bare evidence to one of the basic models of fear learning known as the single process expectancy model (Humphreys, 1939). According to this model, learning can only take place when an individual is consciously aware of the contingency. This model also emphasizes the potency of contingency knowledge for eliciting all types of fear responses including highly automatic defensive responses such as the startle reflex (Lovibond, 2011). Accordingly, individuals cognitively learn the expectancies in a slow and controlled way, which is in contrast to another competing model of fear learning. According to the dual process model, learning takes place not only through a slow and controlled way but also through a fast and uncontrolled automatic process, known to be affective in nature (Grillon, 2009; Mitchell, De Houwer & Lovibond, 2009). Within this model, the different learning processes are mapped to different CRs (Hamm & Weike, 2005). Cognitive learning is reflected by SCRs and is dependent on contingency awareness (Soeter & Kindt, 2010). The unconscious learning process called affective learning affects highly automatic defensive reflexes, such as FPS, that does not require contingency awareness but relies more on actual CS-US pairings (Sevenster et al., 2012).

One way to investigate this premise is by looking at effects of instructions as instructed knowledge is considered as one of the major pathways to the acquisition of fear (Davey, 1992). Surprisingly, the potency of verbally transmitted information has been often overlooked in fear conditioning research. Although it is well established in psychology that humans possess higher language abilities than animals, fear learning studies have often turned a blind eye to its contribution (Rachman, 1977). Nevertheless, instructed knowledge is important to consider as fear conditioning experiments require humans to be instructed for reasons of informed consent and task instruction. Furthermore, instructions are considered to be a powerful manipulation for installing defensive reactions even before actual stimulus pairings (Davey, 1992; Mertens & De Houwer, 2017; Rachman, 1977). However, with regard to the single and dual process models, different explanations have been put forward for the effect of instructions. The single process model emphasizes the potency of instructed knowledge for eliciting all types of fear responses including highly automatic defensive responses such as FPS (Lovibond, 2011). In contrast, according to the dual process model, only SCR is sensitive to verbal information

(Soeter & Kindt, 2010). FPS is considered less susceptible to verbal information and more affected by actual CS-US pairings (Sevenster et al., 2012).

Researchers have only recently started considering the possibility of testing the flexibility of these fear memories known to be resistant to change for survival purposes (e.g., Mertens & De Houwer, 2016; Sevenster et al., 2012). Indeed, it seems more advantageous to be able to quickly and flexibly readjust fear responses in the face of changing circumstances as it has been shown that having inappropriate response control to altered predictions are related to the development and maintenance of anxiety disorders (Schiller, Levy, Niv, LeDoux & Phelps, 2008; Sehlmeyer et al., 2009). However, little is known about the extent to which instructed knowledge can alter the subsequent expression of fear. Researchers have used reversal instructions within the fear conditioning paradigm to modify conditioned fear and test their malleability to the instructed reversal knowledge (Mertens & De Houwer, 2016). Here, after the acquisition phase, participants are informed that previous shock contingencies are reversed in that the previous CS+ is now the safe cue (i.e., new CS-) and the previous CS- will now be followed by a shock (i.e., new CS+). Considered to be a more demanding process than extinction, fear reversal requires simultaneously updating CRs with different stimuli as the fear is not diminished but rather switched to the other CS (Schiller et al., 2008). Most prior studies have only investigated and found reversal of SCR reactions (Grings, Schell, & Carey, 1973; McNally, 1981; Wilson, 1968). Recent studies by Luck and Lipp (2016) and Mertens and De Houwer (2016) have confirmed these findings. Moreover, Mertens and De Houwer (2016) extended these findings to include FPS reactions which was found to be reversible upon contingency reversal instructions. However, much of the past research on FPS has focused on its effect to other types of instructions, such as extinction. A study by Sevenster et al. (2012) found FPS reactions were still maintained following instructed extinction, therefore showing support that FPS is less susceptible to change and contingency learning (Soeter & Kindt, 2010).

This was in contrast to a more recent study by Costa, Bradley and Lang (2015) who found an immediate decrease of various defensive reactions, including FPS when threat cues were signaled as safe. Furthermore, no influence of initial learning associations has been supported upon reversal reactions (Atlas, 2019). Considering that past research on FPS has revealed mixed findings in addition to its lack of research to reversal instructions, the effects on FPS to instructions still continues to be a more debated issue.

Therefore, the present experiment investigated whether fear responses, namely SCR and FPS, conditioned to geometrical shapes are sensitive to reversal instructions using a fear conditioning paradigm. Neutral stimuli were used because certain physiological responses conditioned to fear-relevant stimuli (e.g., pictures of snakes and spiders) are known to be resistant to change and therefore might form a bias in subsequent conditioning (Öhman, Erixon, & Lofberg, 1975). SCR and FPS were included as the two main outcome measures because they have been argued to be differently sensitive to contingency reversal instructions (see above). It is worth mentioning that this experiment included a second objective to investigate the effects of contingency instructions (i.e., no instructions vs. general instructions vs. precise contingency instructions) on the acquisition of fear. However, for the purpose of this research paper, only contingency aware participants were selected to erase any systematic differences that may have arose from the instruction manipulation.

In sum, two competing hypotheses will be tested. According to single process model, SCR and FPS will reverse in the reversal phase. According to the dual process model, SCR will reverse and FPS will not reverse in the reversal phase.

Method

Participants

Participants were Dutch and International students at Utrecht University fluent in English, and did not participate in prior studies involving electrical stimulation. A total of 108 participants took part in the study and were recruited through convenience sampling by the use of advertisement posters around Utrecht Science Park and social media (e.g. Facebook groups). Six participants were excluded due to equipment and experimenter error. The sample was selected on voluntarily basis and consisted of 69% females and 31% males aged between 18 and 35 (*M* = 23.25; *SD* = 3.57).

For this research paper, only contingency aware participants were selected because fear conditioning can only take place when individuals are aware of the CS-US contingencies (see Mertens & De Houwer, 2016) and only those that are contingency aware are thought to be able to reverse responses (see Luck and Lipp, 2016). Participants that selected the correct CS as being paired with the US and selected “very sure” and “quite sure” for the contingency awareness rating (see Procedure) were therefore included in this investigation. The final sample included 72 participants consisting of 72% females and 28% males aged between 18 and 35 (*M* = 23.13; *SD* = 3.47) out of which 8 were left handed.

Participation reward consisted of 8€/hour or one participant credit point. Exclusion criteria consisted of various conditions such as pregnancy or current psychiatric problems/diagnosis (see Appendix A for more information). Table 1 depicts various demographic information and the average scores of the final selected participants.

Table 1.

Information about the participants in the study

Information	Contingency aware
Mean shock pain rating	5.6
Mean shock intensity	6.8
STAI mean score	41.18(9.12)

Note. STAI = State-Trait Anxiety Inventory

Material

Conditioned Stimuli (CS). CSs were two grey geometric shapes (circle, square) of 300 by 300 pixels presented on a white background of a HP EliteDisplay E231 screen with resolution of 1920 by 1080 pixels. Both CSs were presented during the acquisition and the reversal phase.

Unconditioned Stimulus (UCS). The UCS was an electric stimulus that was presented six times during the acquisition phase and once during the reversal phase. Shocks were delivered through two lubricated Fukuda standard Ag/AgCl electrodes (1-cm diameter, inter-electrode distance: ~2cm). A wristband with the electrodes administered the shocks by use of a constant current stimulator (DS7A, Digitimer, Hertfordshire, UK). Each participant determined the intensity of the electric stimuli individually in a stepwise work-up procedure starting with the lowest intensity of 0.5 (see Procedure). An unpleasant, but not painful intensity of the stimulus was selected for each participant.

Psychophysiology

Fear potentiated startle (FPS). FPS was measured using two BioSemi EMG electrodes (0.4 cm diameter) filled with conductive gel (signal gel by Parker). One electrode was placed below the pupil of the left eye and the other one approximately 1cm laterally on the side. Two ground electrodes were placed in the middle of the participants forehead 1 inch below the hairline (Blumenthal et al., 2005). Additionally, an auditory stimulus in the form of a loud noise (50 ms duration; ~85 dB) delivered by headphones was used to elicit the startle response, which is known to increase in anticipation of a shock (Lonsdorf et al., 2017). Startle responses were scored automatically by subtracting the mean baseline value (0-20 ms) from the highest peak value in the 20–120 ms time frame following the startle probe onset. T-transformations were then applied to these values using each participants' individual mean and standard deviation (Blumenthal et al., 2005).

Skin Conductance Response (SCR). SCR was collected using two BioSemi GSR electrodes (0.8 cm diameter) with conductive gel that were attached to the thenar and hypothenar eminences of the left palm and measured using the BioSemi system. SCRs were calculated by subtracting a mean baseline value (2s preceding CS onset) from the highest response value within a 1- to 7-s interval after CS onset (Pineles, Orr, & Orr, 2009). A minimum criterion of 0.02 μS was applied for the SCRs. SCR values were also range corrected to account for individual differences in responsivity. In order to normalize the data, square root transformation was applied to all SCR responses (Dawson, Schell, Filion, & Berntson, 2007).

Questionnaires

The trait version of the State-Trait Anxiety Inventory (STAI-T; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) was used to assess the participants general trait anxiety level.

The STAI consists of 20 items (e.g.: "I feel secure") and participants were asked to use the rating scale to rate how much the item describes themselves on a 4-point Likert scale ranging from "almost never" to "almost always".

Two additional questionnaires (The Context Sensitivity Index and the Intolerance of Uncertainty Scale) were also used for an unrelated research question and will not be included in this analysis.

Procedure

General information and work-up procedure. The research study has been approved by Faculty Ethical Review Board of the Faculty of Social Sciences of Utrecht University. Upon arrival, participants washed their hands and were then given an information sheet, and a declaration consent form to sign. Following this, participants were asked to complete the three questionnaires. Next, the SCR and FPS electrodes were attached after the participants skin was cleaned with a scrub gel. This was followed by the work-up procedure to determine the intensity of the electric shock. Participants were reminded to select an intensity that they found unpleasant but not painful and were asked to rate their discomfort verbally using a scale ranging from 0 (not painful) to 10 (extremely painful). The shock intensity was gradually increased and was stopped when participants rated the intensity 6 or higher. Finally, for the startle probe administration, headphones were put on. Participants were then asked to provide their age, their dominant hand and their gender on the screen. Further instructions were provided on the screen that participants were asked to read carefully. Participants received different instructions depending on which group they were assigned to. The no instructions group received general information about the experiment. The general instructions group was given the general information with general contingency instructions stating that one of the shapes will sometimes be followed by the shock. Participants in the precise instructions group received the general information with precise contingency instructions of the two shapes and the shock. Upon completion of the experiment, participants were given a debriefing sheet and thanked for their participation and time.

Conditioning phase. Following the contingency instructions, the conditioning phase started with six startle probe trials. The two geometric shapes were presented eight times each (16 in total) in a pseudo-random order (i.e., no more two consecutive trials of each CS type; see Figure 1 for an overview). Additionally, the CS associated with the shock (CS+) was followed by the electric shock six times on a partial reinforcement schedule (75% reinforcement rate). The other geometric shape (CS-) was never reinforced during the conditioning phase. Startle probes were delivered during CS presentations (i.e., 1s before CS offset).

Following the conditioning phase, participants were questioned to determine their contingency awareness between each shape and the electric shock (i.e.: *Did you think that the circle[/square] would be followed by the electric shock?*). An additional question (i.e.: *How sure are you about your answer?*) was presented to which the participants had to answer using a scale with four forced-choice options ("very sure", "quite sure", "quite unsure" and "very unsure").

Reversal phase. After rating their contingency awareness, the experiment continued to the reversal phase. Irrespective of the contingency instructions received in the acquisition phase, participants in the reversal phase were informed that relationships between the CSs and the UCS were reversed in that the CS+ will not be followed by the electric shock, and that the CS- will sometimes be followed by the electric shock. The CSs were presented five times each and the new CS+ was reinforced by an electric shock once (see Figure 1). Startle probes were again delivered during CS presentations. After the reversal phase, participants were asked to rate their contingency awareness again.

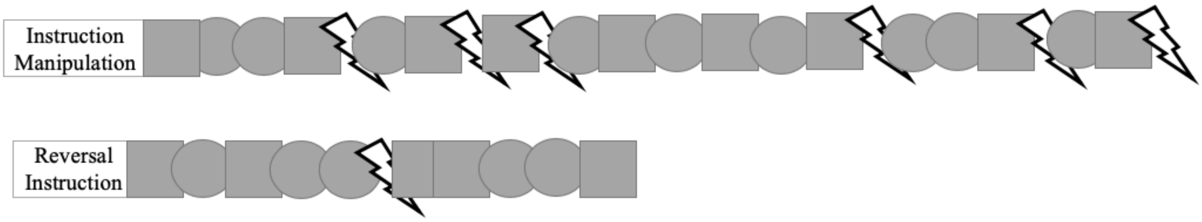

Figure 1. Schematic overview of the CS and US presentations in the conditioning and reversal phase. Note that the exact order of the CSs was pseudo-randomized (see Procedure).

Data analysis

The different measures of fear (SCR and FPS) were analyzed separately using repeated measures ANOVAs. Specifically, the conditioning and reversal phase were first analyzed with two within-subjects factors namely, CS type (CS+, CS-) and trail number (conditioning phase: 8 trials each; reversal phase: 5 trials each). Additionally, for direct comparisons between the acquisition and reversal phase, each measurement of fear was analyzed using two within-subjects factors namely, CS type (CS+, CS-) and phase (acquisition, reversal) where average responses of the CS+ and CS- were used across the two phases. Violations of the sphericity assumption were corrected using Greenhouse-Geisser corrections. For all statistical analyses, alpha level of .05 was applied.

A correlation between STAI scores and reversal efficiency was included in a follow-up analysis. Reversal efficiency was calculated by dividing the difference between the CS- and CS+ in the reversal phase by the difference between the CS+ and the CS- in the conditioning phase (hence, a larger value indicates greater reversal efficiency).

Results

The aim of the experiment was to assess whether fear responses, namely SCR and FPS are sensitive to reversal instructions using a fear conditioning paradigm. Only contingency aware participants were selected for the analyses. Prior to being analyzed separately using repeated measures ANOVA, a response criterium (> .02 microS), range correction, and a square root transformation were applied to the SCR recordings and startle responses were T-transformed. This was done to obtain a normal distribution and account for inter-individual differences in responsivity.

SCR

Conditioning. The RM-ANOVA of CS type and trial number revealed a significant main effect for CS type ($F(1, 71) = 76.649$, $p < .001$, $\eta_p^2 = .519$), trial number ($F(5.792, 411.246) = 11.716$, $p < .001$, $\eta_p^2 = .142$) as well as an interaction ($F(7, 497) = 2.779$, $p = .008$, $\eta_p^2 = .038$). These results show that there was successful conditioning of the CSs. Figure 2 presents the interaction graph.

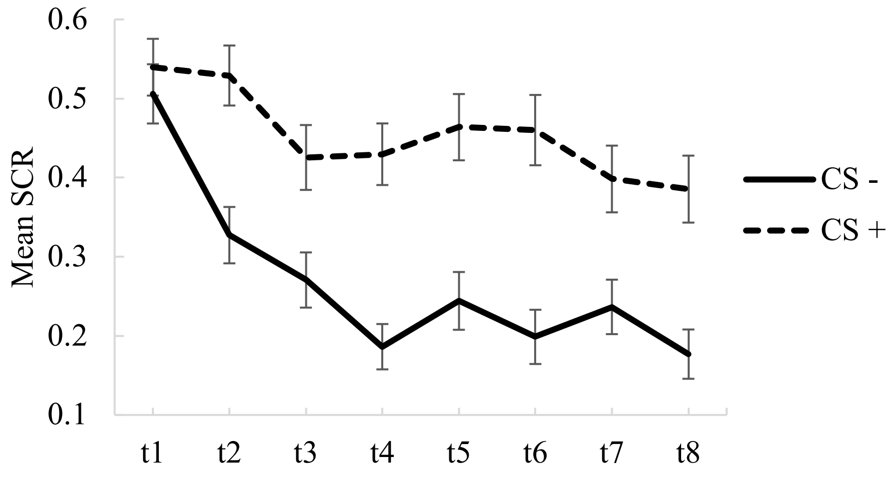

Figure 2. SCR in the conditioning phase consisting of 8 trials.

Reversal. Results using the reversal phase showed a significant main effect for CS type ($F(1, 71) = 32.461$, $p < .001$, $\eta_p^2 = .314$) and trial number ($F(4, 284) = 6.189$, $p < .001$, $\eta_p^2 = .080$). However, the interaction was non-significant ($F(3.497, 248.313) = 1.011$, $p = .396$, $\eta_p^2 = .014$). Crucially, as predicted, the effect of CS type on SCR in this phase was in the opposite direction compared to the conditioning phase. Figure 3 presents the observed mean responses.

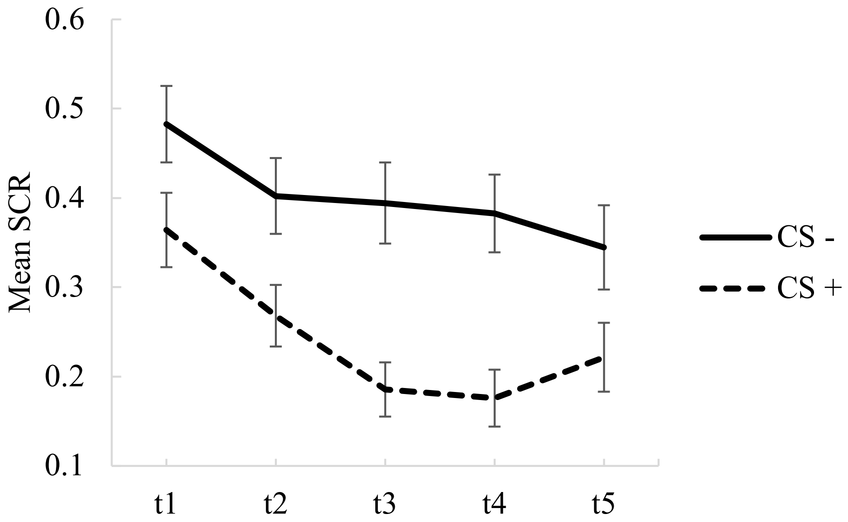

Figure 3. SCR in the reversal phase consisting of 5 trials.

Direct comparison between the acquisition and reversal phase. The 2x2 RM-ANOVA of CS type and phase revealed a non-significant main effect for CS type ($F(1, 71) = .928$, $p = .339$, $\eta_p^2 = .013$) and a significant main effect for phase ($F(1, 71) = 5.623$, $p = 0.20$, $\eta_p^2 = .073$). More importantly, the interaction of this analysis was significant ($F(1, 71) = 72.379$, $p < .001$, $\eta_p^2 = .505$) therefore showing successful reversal of SCR in the reversal phase. Figure 4 displays this interaction.

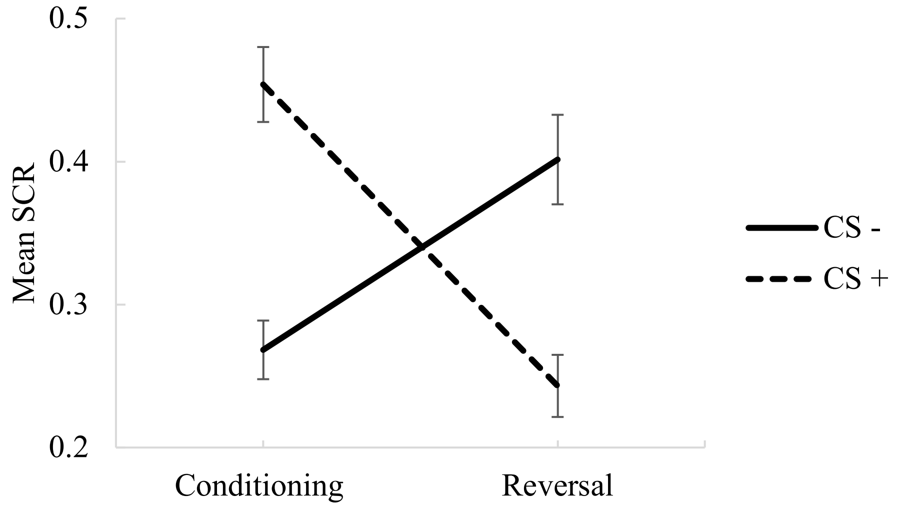

Figure 4. SCR in conditioning and reversal phase.

FPS

Conditioning. The analysis for CS type and trial number revealed a significant main effect for CS type ($F(1, 71) = 33.438$, $p < .001$, $\eta_p^2 = .320$) and trial number ($F(5.678, 403.120) = 12.384$, $p < .001$, $\eta_p^2 = .149$). However, the interaction was non-significant ($F(7, 497) = 1.401$, $p = .202$, $\eta_p^2 = .019$). Figure 5 presents the mean FPS values.

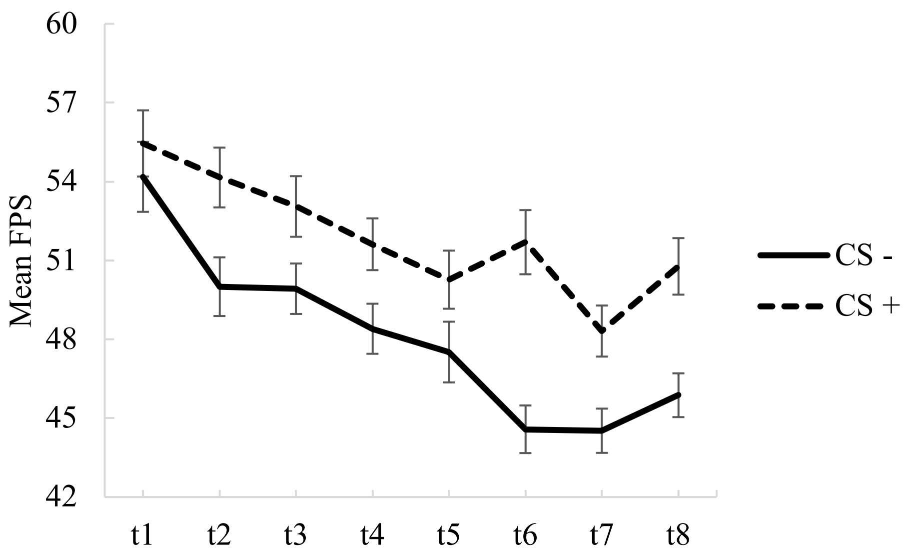

Figure 5. FPS in the conditioning phase consisting of 8 trials.

Reversal. The RM-ANOVA revealed a significant main effect for CS type ($F(1, 70) = 37.447, p < .001, \eta_p^2 = .349$) and trial number ($F(3.420, 239.382) = 6.979, p < .001, \eta_p^2 = .091$). However, non-significant interaction was found ($F(4, 280) = 1.447, p = .219, \eta_p^2 = .020$). Crucially, the results revealed the effect of CS type on FPS was in the opposite direction compared to the conditioning phase (see Figure 6).

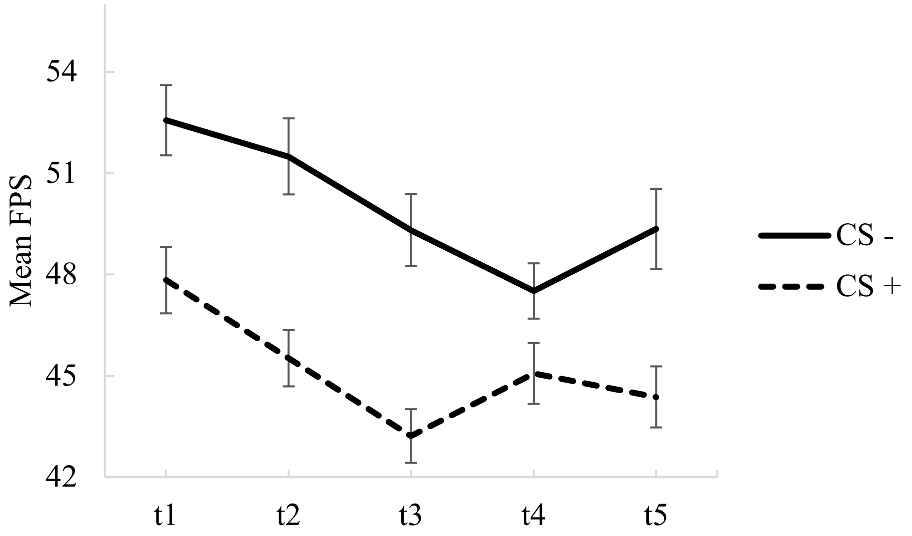

Figure 6. FPS in the reversal phase consisting of 5 trials.

Direct comparison between the acquisition and reversal phase. The 2x2 RM-ANOVA of CS type and phase revealed a nonsignificant main effect for CS type ($F(1, 71) = 1.300, p = .258, \eta_p^2 = .018$) and a significant main effect for phase ($F(1, 71) = 16.859, p < .001, \eta_p^2 = .192$). Crucially, the CS type and phase interaction was significant ($F(1, 71) = 63.196, p < .001, \eta_p^2 = .471$) therefore showing successful reversal of FPS in the reversal phase. Figure 7 presents these findings.

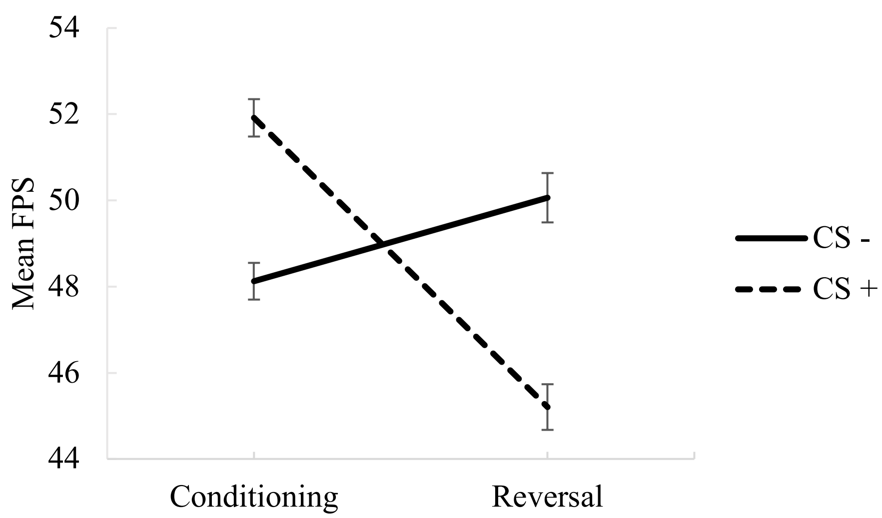

Figure 7. FPS in conditioning and reversal phase.

Follow-up analysis

Follow-up analysis revealed no significant correlation between STAI scores and reversal efficiency in SCR, $r(70) = .07$, $p = .58$ and no correlation between STAI scores and reversal efficiency in FPS, $r(70) = .19$, $p = .11$.

Discussion

The present study investigated the effect of reversal contingency instructions on two fear responses, SCR and FPS, using a fear conditioning paradigm. Two competing models of fear learning were used to derive the research hypotheses: according to the single process model it was predicted that both FPS and SCR would reverse following the reversal instructions. Alternatively, according to the dual process model it was hypothesized that only SCR, and not FPS, would reverse.

Findings revealed that both SCR and FPS immediately and completely reversed following reversal instructions. These findings show support to the single process expectancy model in explaining that both fear responses behave in a similar way and place much importance to the role of cognitive expectancies in influencing these conditioned fear responses. That is, the appearance of the CS, retrieved the CS-US contingency, and thus lead to a US expectancy which influenced various defensive fear responses similarly (Humphreys, 1939; Lovibond & Shanks, 2002). In this study, for example, an increase in *both* SCR and FPS to the CS+ in the conditioning phase and a decrease in *both* SCR and FPS to the CS+ in the reversal phase. This mediating effect of cognitive expectancies is also supported by Davey (1992). Moreover, it may also be noted that the provision of reversal instructions immediately reversed the fear responses on the first reversal trial while the actual shock was only administered on the third trial. This is in line with prior research about how fear reactions can be influenced by instructions prior to actual stimulus pairings (Davey, 1992; Rachman, 1977). The current study also supports the findings of earlier research such as Wilson (1968) and Luck and Lipp (2016) regarding the reversal of SCR. Crucially, however, the FPS responses also reversed. This result replicates the findings of Mertens and De Houwer (2016) who found that reversal instructions reversed FPS reactions, including SCR, therefore, supporting conscious cognitive learning as a causal role for CR production. To an extent, Costa et al. (2015) findings were also supported as the study found that participants decreased their FPS responses when threat cues were signaled as safe (instructed extinction).

With regard to the implications, the study gives support to the notion that healthy participants have the ability to engage in appropriate emotional control following changes in predictions via instructions. That is, to update their expectancies by learning in a slow and controlled way and therefore readjust their conditioned fear reactions in response to decreased fear (threat to safe) and increased fear (safe to threat), simultaneously. This cognitive flexibility could be assumed to be less pronounced in individuals with anxiety disorders as research has shown that these individuals are less capable of updating their conditioned responses (Izquierdo, Brigman, Radke, Rudebeck & Holmes, 2017; Sehlmeyer et al., 2009). Furthermore, as the follow-up analysis revealed no correlation between STAI scores and reversal efficiency in both fear responses, it can be concluded that updating fear responses based on reversal instructions is not predicable of a healthy participant's trait anxiety score. Moreover, the study also highlights the potency of instructions to *alone* reverse fear responses even before actual reinforcement. Specifically, the finding of FPS reversal further demonstrates the power of instructed knowledge as even responses that are believed to be automatic and unconscious are susceptible, therefore, adding to the inconclusive research on FPS reversal.

Nonetheless, there are several potential limitations to the present study that should be addressed. First, as the research was carried out using only healthy participants, the findings of the study may be difficult to generalize to a clinical sample especially to those individuals who have anxiety-related disorders or cognitive impairments. Further, although the study has placed a high emphasis on cognitive learning, the insufficient contribution of the affective learning process could set limitations for the study, for example, participants were reinforced partially, therefore reducing the opportunity for experiential affective learning to take place. Similarly, using fear-irrelevant stimuli as CSs and mild electric shocks as USs provides a further restriction in assessing affective learning. Third, although few studies have investigated sex differences in fear conditioning, some have reported males to be more resistant to extinction (Dalla & Shors, 2009). Thus, considering that extinction (and conditioning) is a part of reversal, the majority of females in the study could have, to some extent, skewed the results. Finally, as the study had another research objective, the selection criteria for contingency awareness might be argued as an unreliable measure of contingency awareness and could, therefore, impact examining the results with full confidence.

Therefore, future studies should consider assessing reversal contingency instructions on fear responses using a clinically-relevant sample with an important focus to those that have cognitive deficits and/or anxiety-related disorders. Although the follow-up analysis revealed no correlation between trait anxiety and reversal efficiency for the participants in the study, research has shown that anxiety disorders are maintained and developed by an inappropriate response control to altered circumstances (Schiller et al., 2008; Sehlmeyer et al., 2009). Thus, studying reversal instructions in these populations might assist in understanding the biased maintenance and treatment of fear. In terms of reinforcement, it would be interesting to investigate whether reversal instructions are as powerful for an anxiety population as they are for the healthy participants in the study by focusing on whether anxious individuals require less (or more) reinforcement and its impact on updating their responses. Therefore, the inclusion of varying amounts of reinforcement (e.g.., delayed, continuous, no reinforcement) is deemed appropriate to understand the extent of reversal instructions on defensive responses. Moreover, replication of the study with the inclusion of fear-relevant stimuli might offer a more ecologically appropriate CS to measure the affective learning component of fear conditioning. Future studies may also consider testing the long-term effect of reversal instructions on fear responses as this could deepen the understanding of the maintenance of reversal conditioning.

In conclusion, findings of the current study provide additional insight to the existing literature on the malleability of fear responses to reversal contingency instructions. In support of the single process model of fear learning, the mediating factor of conscious cognitive expectancies has shown to reverse both SCR and FPS responses of the participants in the study. Therefore, this selective and accurate responding to stimuli via instructions may provide a better understanding of appropriate control of fear and be of clinical interest for important considerations for possible treatments of fear in anxiety-related disorders.

References

Atlas, LY (2019). How instructions shape aversive learning: higher order knowledge, reversal learning, and the role of the amygdala. *Current Opinion in Behavioral Sciences*, *26*, 121-129.

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.

Cook, S. W., & Harris, R. E. (1937). The verbal conditioning of the galvanic skin reflex. *Journal of Experimental Psychology*, *21*(2), 202–210.

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

Dalla, C., & Shors, T. J. (2009). Sex differences in learning processes or classical and operant conditioning. *Physiology and Behavior*, *97*(2), 229-238.

Davey, G. C. L. (1992). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behaviour Research and Therapy*, *14*(1), 29–66.

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.

Grillon, C. (2009). D-Cycloserine Facilitation of Fear Extinction and Exposure-Based Therapy Might Rely on Lower-Level, Automatic Mechanisms. *Biological Psychiatry, 66*(7), 636– 641.

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.

Hamm, A. O., & Weike, A. I. (2005). The neuropsychology of fear learning and fear regulation. *International Journal of Psychophysiology*, *57*(1), 5–14.

Humphreys, L, G. (1939). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology*, *25*(3), 294-301.

Izquierdo, A., Brigman, JL, Radke, A. K., Rudebeck, PH, & Holmes, A. (2017). The neural basis of reversal learning: an updated perspective. *Neuroscience*, *345*, 12-26.

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., ... & Drexler, S. M. (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 and Biobehavioral Reviews, 77,* 247-285.

Lovibond, P. F. (2011). Learning and anxiety: A cognitive perspective. In T. R. Schachtman & S. Reilly (Eds.), *Associative Learning and Conditioning: Human and Non-human Applications* (pp. 104–120). New York: Oxford University Press.

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

Luck, C. and Lipp, O. (2015). To remove or not to remove? Removal of the unconditional stimulus electrode does not mediate instructed extinction effects. *Psychophysiology*, *52*(9), 1248-1256.

Luck, C., & Lipp, O. (2016). The influence of contingency reversal instructions on electrodermal responding and conditional stimulus valence evaluations during differential fear conditioning. *Learning and Motivation*, *54*, 1-11.

Mertens, G., & De Houwer, J. (2016). Potentiation of the startle reflex is in line with contingency reversal instructions rather than the conditioning history. *Biological Psychology*, *113*, 91-99.

Mertens, G., & De Houwer, J. (2017). Can threat information bias fear learning? Some tentative results and methodological considerations. *Journal of Experimental Psychopathology*, *8*(4), 390-412.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *The Behavioral and Brain Sciences*, *32*(2), 183-198.

Öhman, A., Erixon, G., & Löfberg, I. (1975). Phobias and preparedness: Phobic versus neutral pictures as conditioned stimuli for human autonomic responses. *Journal of Abnormal Psychology*, *84*(1), 41–45.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522.

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), 984–995.

Rachman, S. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour Research and Therapy*, *15*(5), 375-387.

Schiller, D., & Delgado, MR (2010). Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends in Cognitive Sciences*, *14*(6), 268-276.

Schiller, D., Levy, I., Niv, Y., LeDoux, J. E., & Phelps, E. A. (2008). From fear to safety and back: reversal of fear in the human brain. *Journal of Neuroscience*, *28*(45), 11517-11525.

Sehlmeyer, C., Schöning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V., & Konrad, C. (2009). Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS One*, *4*(6), e5865.

Sevenster, D., Beckers, T., & Kindt, M. (2012). Instructed extinction differentially affects the emotional and cognitive expression of associative fear memory. *Psychophysiology*, *49*(10), 1426-1435.

Soeter, M., & Kindt, M. (2010). Dissociating response systems: Erasing fear from memory. *Neurobiology of Learning and Memory*, *30*(1), 4990– 4998.

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

Watson, J.B., Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, *3*(1), 1-14.

Appendix A

Demographic Survey

About you

1. What is your email address?

   ____

2. How old are you

   ____

3. What is your gender
   - o Female
   - o Male
   - o Other
4. Are you right-handed?
   - o Yes
   - o No
5. What is the highest level of education you have completed?
   - o Less than a high school diploma
   - o High school or equivalent (e.g., GED)
   - o Some college, no degree
   - o Associate degree
   - o Bachelor's degree (e.g. BA, BS)
   - o Master's degree (e.g. MA, MS, Med)
   - o Professional degree (e.g. MD, DDS, DVM)
   - o Doctorate (e.g. PhD, EdD)
6. Are you a student? If so: I am enrolled in an education at the following level:
   - o Bachelor
   - o Master

- o Doctorate
- o Anders: __________________

7. Can you see clearly? (good vision or adjusted to correct vision?)
   - o No
   - o Yes
8. Do you have trouble hearing? If so, could you tell us a bit more about these problems?

   ____

9. Do you use medication that could influence your attention, responsiveness, memory or concentration?
   - o Yes
   - o No
10. Are you pregnant?
    - o Yes
    - o No
11. Do you suffer or have you suffered in the past from a sever neurological or medical condition (like epilepsy or heart disease)?
    - o Yes
    - o No
12. Do you have an electronic implant (for example a pacemaker)?
    - o Yes
    - o No
13. To the extent of my knowledge, I am healthy and I have no medical conditions like the ones asked in this screening. I have answered all questions truthfully.
    - o Yes
    - o No