On the resistance to extinction of fear conditioned to angry faces

Monique E. Rowles, Ottmar V. Lipp, & Kimberley M. Mallan

School of Psychology, University of Queensland, Australia

Running Head: On the resistance to extinction of fear conditioned to angry faces

Address for correspondence:

Ottmar V. Lipp, PhD

School of Psychology, University of Queensland, QLD, 4072, Australia.

Email: o.lipp@uq.edu.au

Phone: +61 7 3365 6385

Fax: + 61 7 3365 4466

Word count:

Acknowledgements: This work was supported by grant number DP0770844 from the Australian Research Council.

Abstract

The present study investigated whether, like fear conditioned to pictures of snakes and spiders, fear conditioned to angry faces resists extinction even after verbal instruction and removal of the shock electrode. Participants were trained in a differential Pavlovian fear conditioning procedure with angry face or happy face conditional stimuli (CSs). Prior to extinction, half the participants in each group were informed that no more unconditional stimuli would be presented and the shock electrodes removed. In absence of this manipulation, participants showed resistance to extinction after training with angry, but not after training with happy face CSs. Instructed extinction and electrode removal abolished fear conditioning regardless of the emotion expressed by the CS faces. This finding suggests that fear conditioned to angry faces, like fear conditioned to racial out-group faces, is more malleable than fear conditioned to snakes and spiders.

Key words: Preparedness, fear learning, electrodermal responses, facial expressions, fear relevance.

Since Seligman's seminal 1971 paper that proposed preparedness theory to reconcile the apparent inconsistencies between then current learning theory and the phenomenology of phobic fear, much work has been done to confirm the importance of associative learning in psychopathology (for a current review see Mineka & Zinbarg, 2006). Fears of snakes and spiders, in particular, seem to be the result of prepared learning which is proposed to be a) selective to associations with aversive events, b) resistant to extinction, c) observed after only one trial of CS-US pairing, and d) irrational, such that information about the non-occurrence of the US will not diminish it (Seligman, 1971). Öhman, Mineka and colleagues have confirmed this across an extended series of studies (for a review see Öhman & Mineka, 2001). Most of the studies with human participants have used electrodermal responses as an index of fear learning, although cardiovascular indices have been used as well (e.g., Hamm, Vaitl, & Lang, 1989). Using electrodermal responses only, Öhman, Fredrikson, Hugdahl, and Rimmö (1976) demonstrated that, relative to flower and mushroom controls, fear conditioned to snake or spider pictures is resistant to extinction, a resistance that is selective to associations with an aversive event, an electrotactile US, but absent for a non aversive US. Öhman, Eriksson, and Olofsson (1975) demonstrated that this resistance to extinction can be acquired in one CS-US pairing. Of most relevance to the present study, there is ample evidence that the verbal instruction that the US will no longer be presented in combination with removal of the shock electrode does not accelerate extinction of fear, as indexed by electrodermal responses, conditioned to snake or spider pictures, whereas it accelerates extinction of differential electrodermal responses conditioned to flowers and mushrooms (Öhman, Erixon, & Löfberg, 1975; Hugdahl & Öhman, 1977; Hugdahl, 1978; Soares & Öhman, 1993; Lipp & Edwards, 2002).

Prepared learning has also been proposed in the context of social fear. Öhman and

Dimberg (1978) showed that differential electrodermal responses conditioned to pictures of angry faces display resistance to extinction relative to electrodermal responses conditioned to happy or neutral faces (see also Dimberg, 1983). More recently, Olsson, Ebert, Banaji, and Phelps (2005) reported enhanced resistance to extinction for electrodermal responses conditioned to racial out-group faces, relative to electrodermal responses conditioned to faces of racial in-group members. Differential electrodermal responding to racial out-group faces, but not to racial in-group faces was maintained across 12 extinction trials. This led to the suggestion that social fears, like fears conditioned to snakes and spiders, may reflect on prepared learning (Öhman, 2009). However, to be regarded an instance of prepared learning an association has to fulfill not only the criterion of resistance to extinction, but also the criteria of selectivity, one trial learning, and encapsulation from cognition (Seligman, 1971).

Mallan, Sax, and Lipp (2009) provided evidence that verbal instruction in conjunction with removal of the shock electrode eliminated the resistance to extinction of fear conditioned to racial out-group faces. Using pictures of Caucasian and Chinese persons as CSs, Mallan et al. replicated the resistance to extinction of fear conditioned to racial out-group faces documented by Olsson et al. (2005). Moreover, this resistance to extinction was evident in electrodermal responses and in fear potentiated startle. However, verbal instruction that no more USs would be forthcoming and removal of the shock electrode was sufficient to eliminate this resistance in both indices. This result can be taken to suggest that fear conditioned to social CSs is more malleable than is fear conditioned to animal CSs. The present experiment is, to the best of our knowledge, the first to assess whether verbal instruction and removal of the shock electrode will eliminate the resistance to extinction of differential electrodermal conditioning with angry face conditional stimuli. This manipulation has, in previous studies, failed to affect resistance to extinction of fear conditioned to animal fear-relevant stimuli, but did accelerate the extinction of fear conditioned to racial out-group faces. Thus, the current experiment addresses an important gap in the literature.

Method

Participants

Fifty eight undergraduate students (mean age of 20.6 years; range 17-39; 22 male) volunteered participation and provided informed consent. Data from an additional 6 participants were excluded due to computer error (1), failure to verbalise the contingencies (2), and failure to believe the instructions (3).

Apparatus and stimuli

Colour pictures of four angry and four happy male Caucasian faces (NimStim database: images AN_O and HA_O: models 20, 23, 34 and 37; Tottenham et al., 2009) served as conditional and control stimuli and were displayed on a 17" color CRT screen at a size of 506 x 650 pixels. Participants were presented either with four angry or four happy faces during initial and final rating tasks, but only with two images during conditioning. Use of images as CS+ and CS- and as conditional or as control stimuli was counterbalanced across participants. The 200ms electrotactile US was generated by a Grass SD9 Stimulator pulsed at 50Hz and presented via a concentric electrode to the participants' preferred forearm. Stimulus presentation was controlled with DMDX (Forster & Forster, 2003). Electrodermal activity was monitored with Ag/AgCl electrodes filled with an isotonic electrolyte and attached to the thenar and hypothenar prominences of the participant's non-preferred hand. Respiration was monitored with an elasticized chest gauge. Physiological responses were recorded with a Biopac MP150 system at 1000 Hz.

Procedure

Upon arrival at the laboratory, participants were informed about the general procedure and provided informed consent. They were seated in a recording room, adjacent to the control room, in front of the monitor and the measurement devices were attached. The experiment commenced with a shock work-up during which the intensity of the US was set individually to be ‘unpleasant, but not painful’. This was followed by a three minute baseline recording to determine participant’s level of electrodermal responsiveness. After the baseline, participants were presented with the four angry or happy face stimuli and asked to rate their pleasantness on a 9-point Likert scale using the instruction "Please rate on a scale of 1 to 9 where 1=unpleasant and 9=pleasant”. This initial rating was followed by habituation and acquisition phases presented without interruption. During habituation, the CS+ and CS- faces were shown four times each for 6 s with an intertrial interval of 11, 13 or 15 s. Acquisition consisted of 8 trials of the CS+ face followed by the electrotactile US and 8 trials of CS- alone. Prior to the commencement of extinction training, the experimenter entered the testing room pretending to inspect the shock electrode. In groups No instruction, the experimenter re-attached the electrode and informed the participant that the experiment was to continue. In groups Instruction, the experimenter told the participant that no more electrotactile stimuli would be presented and did not re-attach the shock electrode before informing the participant that the experiment was to continue. The experimenter then left the room and extinction commenced with 10 presentations each of CS+ and CS-. This procedure which includes removal of the shock electrode is a direct replication of the methodology used in previous studies and is standard in research on instructed extinction. It has been used in all studies that failed to find an effect of verbal instruction on the extinction of fear conditioned to snake and spider stimuli (Öhman, Erixon, & Löfberg, 1975; Hugdahl & Öhman, 1977; Hugdahl, 1978; Soares & Öhman, 1993; Lipp & Edwards, 2002) and in the study that documented an effect of verbal instruction on the resistance to extinction of fear conditioned to racial out-group faces (Mallan et al., 2009).

Trials in all phases were presented in a pseudo random sequence with no more than two consecutive trials being the same. Use of stimuli as CS+ and CS- and the nature of the first stimulus presented were counterbalanced across participants, thus a total of eight different trial sequences were used (controlling: CSs vs. control stimuli; CS+ vs. CS-; nature of the first trial). Extinction was followed by a second pleasantness rating and the completion of a recognition questionnaire. This required the identification of the two angry or happy faces shown during conditioning, the identification of the angry or happy face paired with the US, and, in groups Instruction, the question whether the participant believed the instruction. Participants were then thanked and debriefed.

Scoring and response definition

Electrodermal recordings were inspected for respiration induced artifacts and the number of spontaneous responses was counted during the three minute baseline. Spontaneous responses were defined as any response that exceeded .05 µS during baseline. Electrodermal responses to the CSs were quantified as the largest response that started within 1-4s after picture onset and unconditional electrodermal responses were quantified as the largest response that started within 1-4s after the onset of the electrotactile stimulus (Prokasy & Kumpfer, 1973). Prior to analysis, electrodermal responses were square root transformed, to reduce the positive skew of the distribution, and range corrected, to give even weight to all participants' responses, and averaged into blocks of two trials. The reference for the range correction was the largest response displayed by a participant, in the majority of the cases the response elicited by the first or second

US presented during acquisition. Electrodermal responses to the CSs were subjected to separate 2 x 2 x 2 x n (Emotion [Angry, Happy] x Instruction [Instruction, No instruction] x CS [CS+, CS-] x Block [2 in habituation, 4 in acquisition, 4 in extinction]) factorial ANOVAs and unconditional electrodermal responses were analysed in a 2 x 2 x 4 (Emotion [Angry, Happy] x Instruction [Instruction, No instruction] x Block) factorial ANOVA. Electrodermal responses elicited during the first block of extinction training were excluded from the analyses as these responses were confounded by considerable sensitization effects in the two not instructed groups apparently due to the intervention by the experimenter to check the shock electrode – see Figure 2. Sensitization (Groves & Thompson, 1970) is a stimulus non-specific effect that can result in increased autonomic responding as a result of increased activation. A similar, if less pronounced, increase in electrodermal responding at the beginning of extinction training in the groups receiving no instructions was observed by Lipp and Edwards (2002). As the emphasis of the current analysis was on the extent of resistance to extinction which should be evident later, omission of the first extinction block was not seen to limit the conclusions that can be drawn. Please note that although the factors and groups are named ‘Instruction’ the experimental manipulation was a combination of verbal instruction and electrode removal.

Multivariate F values (Pillai’s trace) and partial eta-squares are reported for main effects and interactions involving repeated measures. Significant interactions were subjected to follow-up analyses using two-tailed t-tests. To protect against violation of the assumption of sphericity, Greenhouse-Geisser corrected mean square error values and degrees of freedom were used for the follow-up t-tests and the critical values for these t-tests were derived from Sidak’s tables to protect against the accumulation of α-error (Rohlf & Sokal, 1995). The level of significance was set at .05 for all statistical analyses.

Results

The four groups were of comparable gender distribution (Angry-No Instruction: M:F = 10:6; Angry-Instruction: 10:4; Happy-No Instruction: 7:8; Happy-Instruction: 5:8; Pearson Chi^2 (3) = 1.01, $p > .790$). Separate 2 x 2 (Emotion [Angry, Happy] x Instruction [Instruction, No instruction]) factorial ANOVAs were conducted to assess for differences in age, number of spontaneous electrodermal responses during baseline, and intensity to which the US was set. The groups differed in average age (Angry-No Instruction: Mean=18.19 years, SD=1.05; Angry-Instruction: M=18.36 years, SD=1.28; Happy-No Instruction: Mean=23.40 years, SD = 4.61; Happy-Instruction: M = 20.74 years, SD=2.99) with participants trained with happy faces older than those trained with angry faces, $F(1,54)=42.95$, $p<.001$, $p\eta^2=.443$. The four groups did not differ in the number of spontaneous electrodermal responses during baseline (Angry-No Instruction: Mean=17.19, SD=7.12; Angry-Instruction: M=12.0, SD=7.10, Happy-No Instruction: Mean=14.8, SD=10.5; Happy-Instruction: M=10.5, SD=5.48, all $F < 1.5$, $p > .280$) or the intensity to which the US was set (Angry-No Instruction: Mean=42.5 V, SD=10.8; Angry-Instruction: M=42.9 V, SD=10.7, Happy-No Instruction: Mean=39.0 V, SD=9.9; Happy-Instruction: M=38.1 V, SD=7.2, all $F < 2.60$, $p > .115$).

Figure 1 summarises the pleasantness ratings of the CSs collected before habituation and after extinction training. Participants did not differ in their evaluation of CS+ and CS- prior to habituation training, but rated the CS+ as less pleasant than the CS- after extinction. Moreover, angry faces were overall rated as less pleasant than happy faces. These impressions were confirmed by a 2 x 2 x 2 x 2 (Emotion [Angry, Happy] x Instruction [Instruction, No instruction] x CS [CS+, CS-] x Phase [Pre, post]) factorial ANOVA which yielded main effects for Emotion, $F(1,54)=8.10$, $p<.01$, $p\eta^2=.130$, and CS, $F(1,54)=10.69$, $p<.01$, $p\eta^2=.165$, as well as a CS x Phase interaction, $F(1,54)=14.71$, $p<.001$, $p\eta^2=.214$. Participants evaluated angry faces as less pleasant than happy faces, M = 4.20, SD=2.78 vs. M=5.69, SD=2.88, and evaluated CS+ faces as less pleasant than CS- faces after extinction, M = 4.10, SD=2.67 vs. M=5.42, SD=1.93, $t(54) = 5.71$, $p<.01$, but not before habituation training M = 5.10, SD=2.34 vs. M=5.16, SD=2.46, $t(54) < 1$, $p > .30$. There was no evidence for an interaction of this differential evaluation observed after extinction training with CS emotion or the instruction manipulation in either the overall analysis, all $F(54) < 1.60$, $p > .22$, or in an analysis of the post extinction ratings only, all $F(54) < 2.20$, $p > .14$.

Insert Figure 1 about here

Electrodermal responses to the CSs during habituation were larger in Groups No instruction than in Groups Instruction, $F(1,54)=4.57$, $p<.05$, $p\eta^2=.078$ (see Figure 2, left sections), and declined from block 1 to block 2, $F(1,54)=50.41$, $p<.001$, $p\eta^2=.483$. During acquisition differential electrodermal conditioning was evident in all groups (see middle sections of Figure 2). The analysis yielded main effects for CS, $F(1,54)=64.23$, $p<.001$, $p\eta^2=.543$, and Block, $F(3,52)=5.68$, $p<.01$, $p\eta^2=.247$, and a CS x Block interaction, $F(3,52)=9.24$, $p<.001$, $p\eta^2=.348$. Responses to CS+ exceeded those to CS- in blocks 2-4, all $t(52) > 6.70$, $p<.01$, but not in block 1, $t(53)< 1$, $p > .30$. No interaction involving the between group factors was significant, largest $F(1,54) = 2.19$, $p > .140$, and conditioning was significant in each of the four groups, smallest $t(14) = 3.44$, $p = .004$.. Differential responding in the last block of acquisition did not differ across groups, $F < 1.80$, $p > .18$ for all interactions involving the between group factors. Electrodermal unconditional responses in block 1 exceeded those in later blocks, all $t(52)>5.10$, $p < .01$, which did not differ from each other, all $t(52) < 2.0$, $p > .06$, main effect for Block, $F(3,52)=15.90$, $p<.001$, $p\eta^2=.478$. No interaction involving the between group factors was significant, largest $F(1,54) = 1.98$, $p = .128$.

Insert Figure 2 about here

The right sections of Figure 2 display the electrodermal responses during extinction. As sensitization effects affected electrodermal responding in both No Instruction groups at the outset of extinction training, responses from the first block of extinction were omitted from the analysis. As can be seen, differential electrodermal responding was present in Group Angry-No Instruction throughout the later part of extinction training, but not in the remaining three groups. The analysis confirmed this impression a main effect for Instruction, $F(1,54)=6.95$, $p<.05$, $p\eta^2=.114$, and a Instruction x Emotion x CS interaction, $F(1,54)=5.17$, $p<.05$, $p\eta^2=.087$. Follow up t-tests confirmed that responding to CS+ exceeded responding to CS- in Group Angry-No Instruction, $t(54) = 2.50$, $p < .05$, but not in the other three groups, all $t(54) < 1$, $p > .30$.

Discussion

The current data provide evidence that verbal instruction combined with the removal of the shock electrodes eliminates fear conditioned to angry face conditional stimuli. This finding in group Angry-Instruction is in contrast to the strong resistance to extinction observed in group Angry-No Instruction. The present pattern of results resembles that reported for fear conditioned to racial out-group faces by Mallan et al. (2009). This, like the present study, replicated the finding of resistance to extinction in the groups receiving standard extinction training, but failed to find evidence for differential conditioning after instruction that no more shocks were forthcoming and removal of the shock electrodes. The current pattern of results is in contrast to the failures to find an effect of this intervention on the extinction of fear conditioned to snakes or spiders (Öhman, Erixon, & Löfberg, 1975; Hugdahl & Öhman, 1977; Hugdahl, 1978; Soares & Öhman, 1993; Lipp & Edwards, 2002). The current research does not permit the dissociation of the effects of verbal instruction and of the removal of the shock electrode on extinction. However, this should not detract from the overall contribution of the present result. All previous studies of the effects of instructed extinction on prepared learning with animal fear-relevant CSs employed the same methodology which included removal of the shock electrode. Nevertheless, the five studies that employed animal fear-relevant stimuli failed to find an effect of this intervention (Öhman, Erixon, & Löfberg, 1975; Hugdahl & Öhman, 1977; Hugdahl, 1978; Soares & Öhman, 1993; Lipp & Edwards, 2002) whereas the two studies that employed social fear-relevant stimuli did, Mallan et al. (2009) and the present study.

The finding that fear conditioned to angry faces, like fear conditioned to racial out-group faces, is affected by the instruction/removal of the electrode procedure whereas fear conditioned to snakes and spiders is not suggests a difference in the learning processes that underlie animal and socially related fears. The notion that fear conditioned to angry faces may be more similar to fear of racial out-group faces than to animal related fear is also supported by suggestions that both seem affected by the gender of the models used. Navarrete et al. (2009) replicated Olsson et al's (2005) finding of resistance to extinction of fear conditioned to racial out-group faces, but only for male faces, not for female faces. The results of Öhman and Dimberg's (1978) second experiment suggest a similar pattern in that resistance to extinction was evident for fear conditioned to angry male faces, but not to angry female faces. The latter observation may require further study, however, as Öhman and Dimberg used faces of different gender as CS+ and CS-; conditioning to an angry female face CS+ was compared to an angry male CS- and vice versa.

One may argue that the relevance of the current – and much previous – work on prepared learning to our understanding of the acquisition of fear is limited as electrodermal responses were the only dependent physiological measure employed. Previous research has shown that electrodermal responses are sensitive to stimulus arousal, not valence (Lang, Greenwald, Bradley, & Hamm, 1993), and will emerge during aversive and during non-aversive conditioning (Lipp, Siddle, & Dall, 2003). Other indices, such as fear potentiated startle have been suggested as more specific indices of fear learning (Hamm & Weike, 2005) although startle facilitation has been observed during non-aversive conditioning as well (Lipp et al., 2003). Moreover, it seems unlikely that inclusion of fear potentiated startle would have yielded a different pattern of results. Mallan et al. (2009) employed electrodermal responses and fear potentiated startle in their study of the effects of verbal instruction and removal of the electrodes on extinction of fear conditioned to racial out-group faces. Both variables yielded similar results in that the resistance to extinction of fear conditioned to racial out-group faces that was evident in both measures during normal extinction was equally eliminated in both measures after verbal instruction and removal of the electrodes.

The finding that extinction of fear conditioned to angry faces was facilitated after instruction and removal of the shock electrode seems inconsistent with the view that animal related and socially related fears are mediated by the same evolutionarily traditioned learning system (Öhman, 1986, 2009). Rather they suggest that the learning which underlies fears of threatening or differently looking conspecifics is more malleable by cognition than is the learning that leads to animal phobias. However, the learning processes in question also share common characteristics such as resistance to extinction or sensitivity to severely degraded stimulus input (Soares & Öhman, 1993; Esteves, Parra, Dimberg, & Öhman, 1994). Further research that delineates the similarities and differences between learning about animal and social fear-relevant stimuli is required to gain a complete understanding of the learning mechanisms involved. This should include a replication of the differential effect of verbal instruction and removal of the shock electrode on fear conditioned to social and animal fear-relevant stimuli in a single experiment as well as a delineation of the separate effects of verbal instruction and of removal of the shock electrode.

The current results, together with findings that learning about stimuli such as pointed guns satisfies some of the criteria for prepared learning (Hugdahl & Johnsen, 1989; Flykt, Esteves, & Öhman, 2007), suggest that resistance to extinction of fear learning is not limited to evolutionary prepared associations. Social learning that reflects on negative stereotypes and prejudice, as discussed in the context of fear conditioned to racial out-group faces (Olsson et al., 2005), may also enable stimuli to enter into associations that are resistant to extinction. Whereas the boundary conditions of these sources of preferential learning remain to be mapped, the current data suggest that a single account cannot accommodate fear learning with animal and social fear-relevant stimuli (Öhman, 2009).

References

Dimberg, U., & Öhman, A. (1983). The effects of directional facial cues on electrodermal conditioning to facial stimuli. *Psychophysiology, 20*, 160-167. doi:10.1111/j.1469-8986.1983.tb03282.x

Esteves, F., Parra, C., Dimberg, U., & Öhman, A. (1994). Nonconscious associative learning: Pavlovian conditioning of skin conductance responses to masked fear-relevant facial stimuli. *Psychophysiology*, *31*, 375-385. doi:10.1111/j.1469-8986.1994.tb02446.x

Flykt, A., Esteves, F., & Öhman, A. (2007). Skin conductance responses to masked conditioned stimuli: phylogenetic/ontogenetic factors versus direction of threat? *Biological Psychology, 74*, 328-336. doi:10.1016/j.biopsycho.2006.08.004

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers, 35*, 116-124.

Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychological Review, 70*, 419-450.

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

Hamm, A.O., Vaitl, D., & Lang, P.J. (1989). Fear conditioning, meaning, and belongingness: A selective association analysis. *Journal of Abnormal Psychology, 98*, 395-406. doi: 10.1037/0021-843X.98.4.395

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

Hugdahl, K., & Öhman, A. (1977). Effects of instruction on acquisition and extinction of electrodermal responses to fear-relevant stimuli. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 608-618. doi: 10.1037/0278-7393.3.5.608

Hugdahl, K., & Johnsen, B. H. (1989). Preparedness and electrodermal fear-conditioning: Ontogenetic vs phylogenetic explanations. *Behaviour, Research and Therapy, 27*, 269-278. doi: 10.1016/0005-7967(89)90046-6

Lang, P.J., Greenwald, M.K., Bradley, M.M., & Hamm, A.O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*, 261-273. doi:10.1111/j.1469-8986.1993.tb03352.x

Lipp, O.V., & Edwards, M.S. (2002). Effect of instructed extinction on verbal and autonomic indices of Pavlovian learning with fear-relevant and fear-irrelevant conditional stimuli. *Journal of Psychophysiology, 16*, 176-186. doi: 10.1027//0269-8803.16.3.176

Lipp, O.V., Siddle, D.A.T., & Dall, P.J. (2003). The effects of unconditional stimulus valence and conditioning paradigm on verbal, skeleto-motor, and autonomic indices of human Pavlovian conditioning. *Learning and Motivation, 34*, 32-51. doi:10.1016/S0023-9690(02)00507-6

Mallan, K.M., Sax, J., & Lipp, O.V. (2009). Verbal instruction abolishes fear conditioned to racial out-group faces. *Journal of Experimental Social Psychology, 45*, 1303-1307. doi:10.1016/j.jesp.2009.08.001

Mineka, S., & Zinbarg, R. (2006). A Contemporary Learning Theory Perspective on the Etiology of Anxiety Disorders: It's Not What You Thought It Was. *American Psychologist, 61*, 10-26. doi: 10.1037/0003-066X.61.1.10

Navarrete, C.D., Olsson, A., Ho, A.K., Mendes, W.B., Thomsen, L., & Sidanius, J. (2009). Fear extinction to an out-group face: The role of target gender. *Psychological Science, 20*, 155-158. doi: 10.1111/j.1467-9280.2009.02273.x

Öhman, A. (1983). The orienting response during Pavlovian conditioning. In D. A. T. Siddle (Ed.), *Orienting and habituation: Perspectives in human research* (pp. 315-370). New York: Wiley.

Öhman, A. (1986). Face the beast and fear the face: Animal and social fears as prototypes for evolutionary analyses of emotion. *Psychophysiology, 23,* 123-145. doi: 10.1111/j.1469-8986.1986.tb00608.x

Öhman, A. (2009). Of snakes and faces: An evolutionary perspective on the psychology of fear. *Scandinavian Journal of Psychology, 50*, 543-552. doi: 10.1111/j.1467-9450.2009.00784.x

Öhman, A., & Dimberg, U. (1978). Facial Expressions as Conditioned Stimuli for Electrodermal Responses: A Case of "Preparedness"? *Journal of Personality and Social Psychology, 36*, 1251-1258. doi: 10.1037/0022-3514.36.11.1251

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

Öhman, A., Eriksson, A., & Olofsson, C. (1975). One-trial learning and superior resistance to extinction of autonomic responses conditioned to potentially phobic stimuli. *Journal of Comparative and Physiological Psychology, 88*, 619-627. doi: 10.1037/h0078388

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*, 41-45. doi:10.1037/h0076255

Öhman, A., Fredrikson, M., Hugdahl, K., & Rimmö, P.A. (1976). The premise of equipotentiality in human classical conditioning: Conditioned electrodermal responses to potentially phobic stimuli. *Journal of Experimental Psychology: General, 105*, 313-337. doi:10.1037/0096-3445.105.4.313

Olsson, A., Ebert, J.P., Banaji, M.R., & Phelps, E.A. (2005). The role of social groups in the persistence of learned fear. *Science, 309*, 785-787. doi: 10.1126/science.1113551

Prokasy, W.F., & Kumpfer, K.L. (1973). Classical conditioning. In W. F. Prokasy & D. C. Raskin (Eds.), *Electrodermal activity in psychological research* (pp. 157-202). San Diego: Academic Press.

Rohlf, F.J., & Sokal, R.R. (1995). *Statistical tables (3rd ed.).* New York: Freeman.

Soares, J.J., & Öhman, A. (1993). Preattentive processing, preparedness and phobias: Effects of instruction on conditioned electrodermal reponses to masked and non-masked fear-relevant stimuli. *Behaviour Research and Therapy, 31*, 87-95. doi:10.1016/0005-7967(93)90046-W

Seligman, M.E.P. (1971). Phobias and preparedness. *Behavior Therapy, 2*, 307-320. doi:10.1016/S0005-7894(71)80064-3

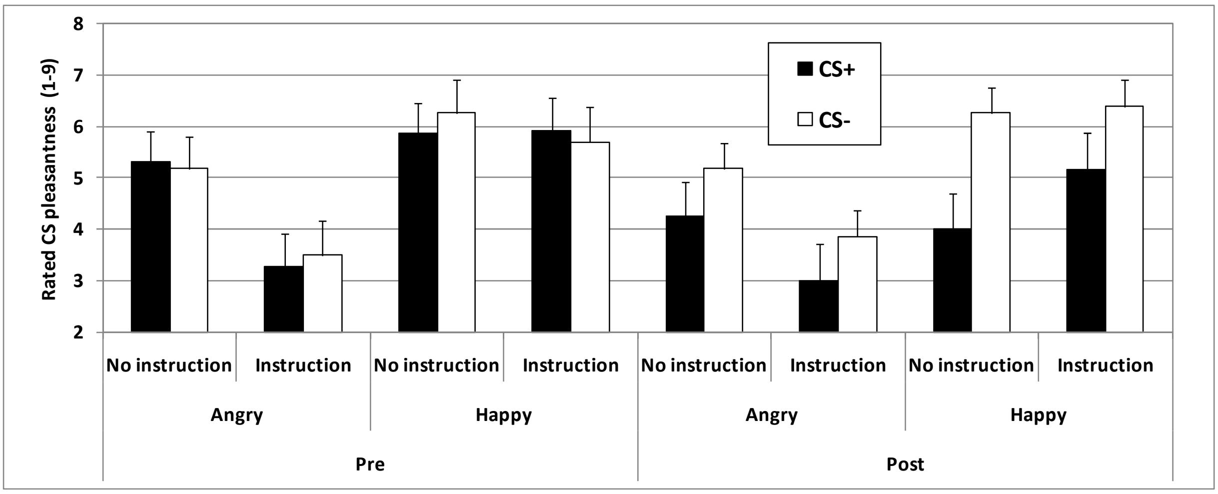
Siddle, D.A.T. (1985). Effects of stimulus omission and stimulus change on dishabituation of the skin conductance response. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11*, 206-216. doi: 10.1037/0278-7393.11.2.206

Tottenham, N., Tanaka, J.W., Leon, A.C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*, 242–249. doi:10.1016/j.psychres.2008.05.006

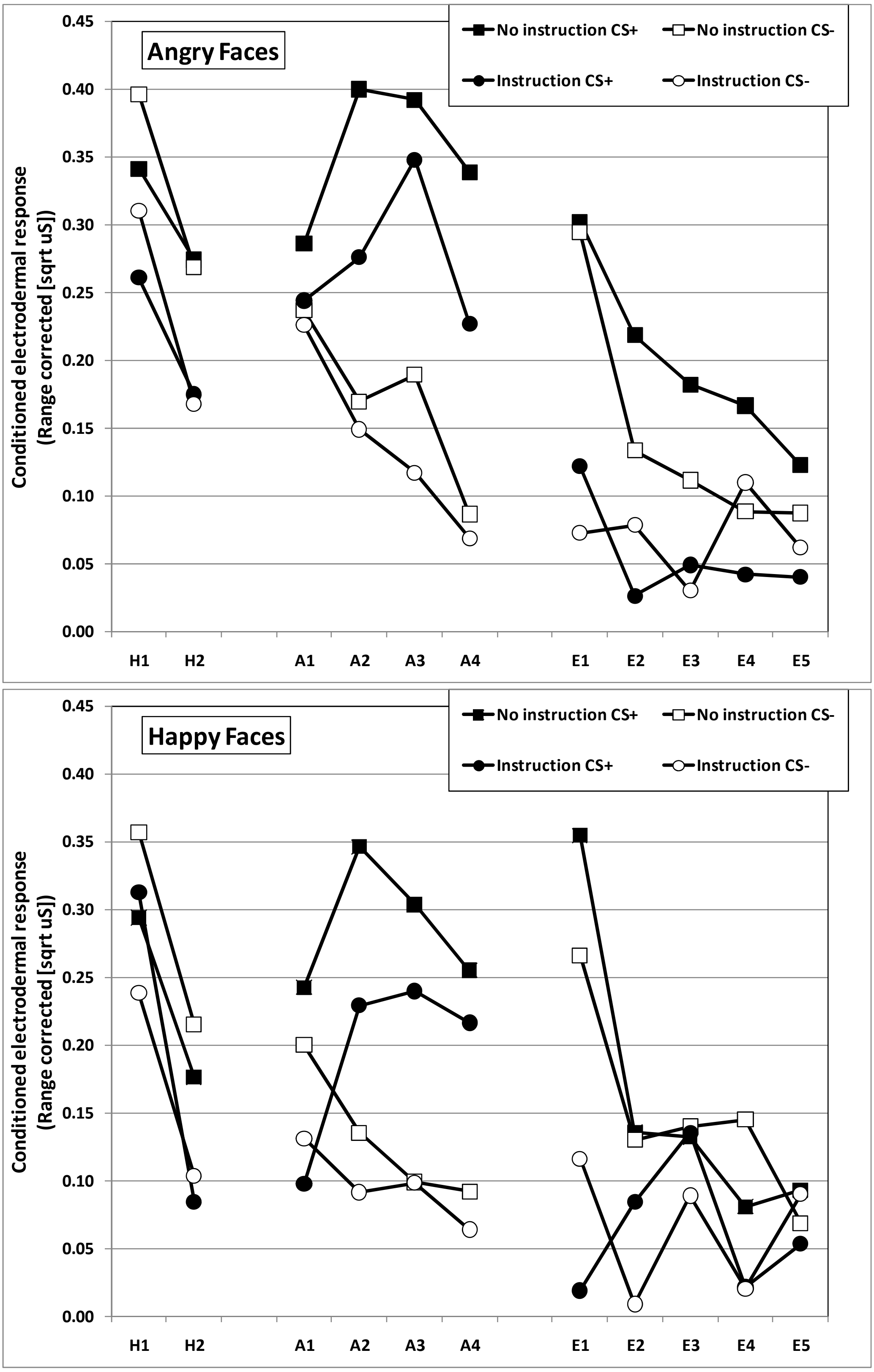
Figure captions

Figure 1: Rated CS pleasantness before habituation (Pre) and after extinction (Post) in groups trained with angry and happy faces as a function of instructions provided before extinction (No Instruction, Instruction) and CS (CS+, CS-; error bars represent standard errors of the means).

Figure 2: Electrodermal responses during Habituation, Acquisition, and Extinction in groups trained with angry (upper panel) and happy faces (lower panel) as a function of instructions provided before extinction (Instruction, No Instruction), CS (CS+, CS-) and trial blocks.



Rated CS pleasantness (1-9)
8
7
6
5
4
3
2
CS+
CS-
No instruction
Instruction
No instruction
Instruction
No instruction
Instruction
No instruction
Instruction
Angry
Happy
Angry
Happy
Pre
Post


Angry Faces
No instruction CS+
No instruction CS-
Instruction CS+
Instruction CS-
Conditioned electrodermal response (Range corrected [sqrt uS])
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
H1 H2 A1 A2 A3 A4 E1 E2 E3 E4 E5
Happy Faces
No instruction CS+
No instruction CS-
Instruction CS+
Instruction CS-
Conditioned electrodermal response (Range corrected [sqrt uS])
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
H1 H2 A1 A2 A3 A4 E1 E2 E3 E4 E5