This is the accepted version of the following article, Bramwell, S. and Mallan, K. and Lipp, O. 2014. Are two threats worse than one?: The effects of face race and emotional expression on fear conditioning. Psychophysiology: an international journal. 51 (2): pp. 152-158, which has been published in final form at http://doi.org/10.1111/psyp.12155

Are two threats worse than one? The effects of face race and emotional expression on fear conditioning

Shannon Bramwell[1], Kimberley M. Mallan[2], & Ottmar V. Lipp[1],

[1]School of Psychology, University of Queensland, Australia

[2]Institute of Health and Biomedical Innovation, Queensland University of Technology, Australia

Running Head: Are two threats worse than one?

Author Notes

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

Acknowledgements

This work was supported by grant number DP110100490 from the Australian Research Council.

Word count: 4588

Abstract

Facial cues of racial outgroup or anger mediate fear learning that is resistant to extinction. Whether this resistance is potentiated if fear is conditioned to angry, other race faces has not been established. Two groups of Caucasian participants were conditioned with two happy and two angry face conditional stimuli (CSs). During acquisition, one happy and one angry face were paired with an aversive unconditional stimulus whereas the second happy and angry faces were presented alone. CS face race (Caucasian, African American) was varied between groups. During habituation, electrodermal responses were larger to angry faces regardless of race and declined less to other race faces. Extinction was immediate for Caucasian happy faces, delayed for angry faces regardless of race, and slowest for happy racial outgroup faces. Combining the facial cues of other race and anger does not enhance resistance to extinction of fear.

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

The human face is a rich source of information that influences the manner in which we interact with others. It communicates aspects about us which are invariant across time (e.g., sex and ethnic group membership) and also aspects that are variable and subject to change almost from moment to moment (e.g., expression; Bruce & Young, 1986; Young & Bruce, 2011). We process this information very efficiently and it determines our behaviour from the minute and transient (will we smile or frown in a given moment), to the more overt and complex (will we maintain the interaction and approach our fellow person or terminate to avoid future encounters). Much effort has been expended to understand the impact of specific facial cues, such as a person's race or their emotional expression, on behaviour, but less is known about how these facial cues interact.

Prior research on the acquisition of fear in response to facial stimuli has pointed to the special role of particular emotional expression such as anger and fear, and to the ethnicity communicated on another person's face. Following up on earlier research that was concerned with the acquisition of fear of dangerous animals, Öhman and colleagues showed that fear conditioned to male angry faces, like fear conditioned to pictures of snakes and spiders, was acquired fast and was resistant to extinction (Öhman & Dimberg, 1978; see Öhman, 2009). This was interpreted to suggest that interpersonal signals of threat, like angry or fearful expressions, can become triggers of fear in a manner that resembles the fear acquired in evolutionary prepared associations (Seligman, 1971; Öhman & Mineka, 2001).

More recent work by Olsson and colleagues (Olsson, Ebert, Banaji, & Phelps, 2005) broadened this perspective by showing that fear conditioned to faces of racial outgroup members also showed resistance to extinction. Using a within participant procedure, Olsson et al. presented two groups of participants, one African American and one Caucasian American, with four faces, two African American and two Caucasian American. During fear conditioning one picture from each group was followed by an aversive unconditional stimulus (electric shock) whereas the second was presented alone. Fear conditioned to outgroup faces, Caucasian Americans for African American participants and African Americans for Caucasian American participants, was resistant to extinction in comparison to fear conditioned to ingroup faces. Mallan, Sax, and Lipp (2009) replicated this finding in Caucasian Australian participants conditioned with Chinese faces as the outgroup stimuli. However, little is known about the mechanism that mediates this type of preferential social fear learning.

Although the findings that fear conditioned to faces expressing anger or faces of a racial outgroup is resistant to extinction are robust, this social fear learning seems distinct from the so-called prepared fear learning seen with pictures of snakes and spiders in that it fails one of the criteria for ‘evolutionarily’ prepared fear learning: it is not impervious to cognition. Mallan et al. (2009) and Rowles, Lipp, and Mallan (2012) have shown that verbal instructions and removal of the unconditional stimulus electrode will abolish the resistance to extinction found for fear conditioned to other race faces or angry faces respectively (for a review see Mallan, Lipp, & Cochrane, 2013). This is not the case for the resistance to extinction of fear conditioned to snakes and spiders which is unaffected by such manipulations (Hugdahl & Öhman, 1977; Lipp & Edwards, 2002). Moreover, it seems that the resistance to extinction of fear conditioned to facial stimuli is moderated by other facial cues. Navarette et al. (2009) showed that resistance to extinction of fear conditioned to other race faces was gender specific, it emerged for male other race faces, but not for female other race faces. Mazurski, Bond, Siddle, and Lovibond (1996) indicated that fear conditioned to angry faces was resistant to extinction if the expressions were posed by adult males, but not if they were posed by pre-adolescent males. Thus, other information that can be decoded from a face, such as sex or age, can influence the characteristics of the fear that is acquired to these social stimuli.

If gender or age cues on an outgroup race or angry face can determine whether fear conditioned to it is resistant to extinction, the question arises as to whether social group and emotional expression can also interact to facilitate their effect on fear learning. Results of research in the area of face processing suggest that this may indeed be the case. The ‘happy face advantage’ refers to the observation that happy faces are categorised faster as ‘happy’ than are angry faces as ‘angry’ (Leppänen & Hietanen, 2003). This advantage is absent, however, if the faces expressing the emotions are of a social outgroup (Hugenberg, 2005; for further elaboration of this finding, see Craig, Mallan, & Lipp, 2012). In implicit evaluation tasks such as the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) or affective priming (Fazio & Olsen, 2003), happy faces are usually evaluated as more positive than are fearful faces. Weisbuch and Ambadi (2008) demonstrated that this does not hold for racial outgroup faces for which a happy expression is evaluated as more negative than a fearful one. Conversely, anger is detected earlier on other race faces than on same race faces (Hugenberg & Bodenhausen, 2003) suggesting an interaction between cues of emotional and social outgroup threat.

The present research was designed to investigate whether, like implicit evaluation or emotion detection, fear conditioned to a human face is affected by a stimulus person’s racial group and their expressed emotion. In particular, we were interested to see whether cues of anger and racial outgroup would combine to yield a ‘super’ fear-relevant stimulus and whether the differences seen in fear conditioning with happy and angry faces would also be evident if the emotional expressions were posed by members of another race. Thus, we compared fear conditioning to angry and happy faces drawn either from a racial ingroup or from a racial outgroup.

Method

Participants

Fifty Caucasian undergraduate students (mean age of 20.7 years; range 17-40; 15 male) volunteered participation and provided informed consent. Participants were assigned upon arrival at the laboratory to two groups, one trained with African American faces and one with Caucasian faces. The rating data from one participant in Group Caucasian were excluded due to failure to enter all ratings and the electrodermal data from another participant in Group Caucasian were excluded due to failure to display phasic electrodermal responses.

Apparatus and stimuli

Colour pictures of four angry and four happy male African American and of four angry and four happy male Caucasian faces (NimStim database: images AN_O and HA_O: models 20, 21, 31, 32, 38, 40, 42, and 43; Tottenham et al., 2009) served as conditional and control stimuli and were displayed for six seconds on a 17” color LCD screen at a size of 506 x 650 pixels. According to the group allocation, a participant was presented with either only African American or with Caucasian faces. Four of the faces, the angry or happy expressions of two individuals were used as conditional stimuli, CSs+ and CSs-, during conditioning training whereas the remaining four faces were used as primes in a subsequent affective priming task. We used expressions posed by two rather than four different individuals to ensure that any difference in extinction did not reflect on the particular poser, but on the facial expression of the poser. Use of images as CSs+ and CSs- and as conditional or as control stimuli was counterbalanced across participants. The 200ms electrotactile unconditional stimulus (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) which also recorded ratings and reaction times in the affective priming task. 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 a sampling frequency of 1000 Hz.

Procedure

Upon arrival at the laboratory, participants were informed about the general procedure and provided informed consent. After washing their hands, they were seated in an experimental 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 to accustom participants to the laboratory environment and to record their levels of electrodermal activity. After the baseline, participants were presented with the four face stimuli, the happy and angry expressions of two different African American or Caucasian posers, that were to be used as CSs+ and CSs- 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, acquisition, and extinction phases presented without interruption. During habituation, the four CS+ and CS- faces were shown four times each for six seconds. Acquisition consisted of six presentations of each of the four faces. The CS+ faces, a happy and an angry expression of the same individual were followed by the electrotactile US whereas the CS- faces, a happy and an angry expression of the second individual were presented alone. Presentation of the US coincided with the offset of the CS+, resulting in a six second delay conditioning procedure. Extinction consisted of six presentations of each of the four CSs, however, no shock stimuli were presented. In each phase, CSs were presented in a pseudo random sequence with no more than two consecutive trials being the same. Use of stimulus set as CSs or controls, use of stimuli within a set as CSs+ and CSs-, and the nature of the first stimulus presented within each phase, whether it was a CS+ or a CS-, whether it was of a happy or an angry expression, were counterbalanced across participants. Thus a total of 16 different trial sequences were used (controlling: CSs vs. control stimuli; CS+ vs. CS-; first trial a CS+ or a CS-; first trial an angry or a happy face). The intertrial interval was 11, 13, or 15 s in all phases of the experiment.

Extinction was followed by a second pleasantness rating, an affective priming task, and the completion of a set of questionnaire measures. In the affective priming task, participants were presented with six pleasant (Appealing, Charming, Desirable, Favourable, Nice, Enjoyable) and six unpleasant target words (Annoying, Disturbing, Inferior, Nasty, Repulsive, Terrifying) and asked to evaluate them as quickly as possible as either pleasant or unpleasant. Target words were preceded by face primes presented for 200 ms at a stimulus onset asynchrony of 293 ms. After indicating whether the target word was pleasant or unpleasant by pressing the right or left ‘Shift’ key, participants were asked to say aloud whether the face prime depicted an angry or a happy expression. The priming task comprised 96 trials, two pairings of each prime face with each target word. The post experimental questionnaire comprised a recognition questionnaire which depicted the eight faces used in each group and required the identification of the angry and happy face paired with the US, a rating of US unpleasantness on a 7 point Likert scale, and an answer to the question: ‘How many black people do you know?’.

Scoring and response definition

Any discernible electrodermal response during baseline was counted to provide a measure of spontaneous electrodermal activity. Electrodermal recordings were inspected for respiration induced artifacts and electrodermal responses to the CSs were quantified as the largest response that started within 1-4s after picture onset. Unconditional electrodermal responses were quantified as the largest response that started within 1-4s after the onset of the electrotactile stimulus during acquisition (Prokasy & Kumpfer, 1973). It should be noted that the responses are defined by the timing of response onset and that the response peak does not have to fall within this latency interval, however, these responses typically peak within 5 s after response onset. Prior to analysis, electrodermal responses were square root transformed to reduce the positive skew of the distribution, 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, which in the majority of the cases was 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 (Group [African American, Caucasian] x Emotion [Angry, Happy] x CS [CS+, CS-] x Block [2 in habituation, 3 in acquisition, 3 in extinction]) factorial ANOVAs and unconditional electrodermal responses were analysed in a 2 x 2 x 3 (Group [African American, Caucasian] x Emotion [Angry, Happy] x Block) factorial ANOVA.

Ratings were subjected to a 2 x 2 x 2 x 2 (Group [African American, Caucasian] x Emotion [Angry, Happy] x CS [CS+, CS-] x Phase [Before, After]) factorial ANOVA. Evaluation times from the affective priming task shorter than 100 ms or three standard deviations above and below the participants' mean were removed and coded as errors. Response times and average error percentages were subjected to separate 2 x 2 x 2 (Group [African American, Caucasian] x Emotion [Angry, Happy] x Target Valence [Positive, Negative]) factorial ANOVAs. 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 F-tests protected against the accumulation of α-error using IBM SPSS 21. The level of significance was set at .05 for all statistical analyses.

Results

The groups did not differ in age ($t(48) < 1.10$, $p = .321$), sex distribution (African American: 7:18; Caucasian: 8:17), the number of spontaneous electrodermal responses during the 3 minute baseline (African American: $M = 23.14$, $SD = 8.74$; Caucasian: $M = 19.78$, $SD = 10.97$; $t(48) = 1.13$, $p = .264$), or the number of black people they reported knowing (African American: $M = 4.36$, $SD = 6.05$; Caucasian: $M = 4.20$, $SD = 9.93$; $t(48) = .069$, $p = .945$). Both groups set the US to similar levels of intensity (African American: $M = 27.4$ V, $SD = 7.38$; Caucasian: $M = 31.8$ V, $SD = 8.77$; $t(48) = 1.92$, $p = .061$) and rated it as similarly unpleasant (African American: $M = 5.57$, $SD = 0.84$; Caucasian: $M = 5.36$, $SD = 1.38$; $t(48) = 0.615$, $p = .542$). Two participants in group African American reported a third face as having been paired with the US. Running the analyses excluding these participants did not alter the results and hence, the results from the complete sample are reported.

Figure 1 summarises the pleasantness ratings of the CSs collected before habituation and after extinction training. Participants did not differ in their evaluation of the CSs+ and CSs- prior to habituation training, but rated the CSs+ as less pleasant than the CSs- after extinction. Moreover, angry faces were overall rated as less pleasant than happy faces. These impressions were confirmed by the analysis which yielded main effects for CS, $F(1, 47) = 11.86$, $p = .001$, □ η_p^{2}□ = .202, and Emotion, $F(1, 47) = 175.14$, $p < .001$, □ η_p^{2}□ = .789, as well as Group x

Emotion,

F(1, 47) = 5.70, *p* = .021, η_p^2 = .108, and CS x Phase interactions, *F*(1, 47) = 14.25, *p* < .001, η_p^2 = .233. Participants evaluated African American happy faces as more pleasant than Caucasian happy faces, *M* = 7.40, *SD* = 1.74 vs. *M* = 6.29, *SD* = 2.03, *F*(1, 47) = 7.53, *p* = .009, η_p^2 = .138, whereas there was no difference in the rating of angry faces, *M* = 2.81, *SD* = 1.91, vs. *M* = 3.10, *SD* = 2.15, *F*(1, 47) = 0.67, *p* = .417, η_p^2= .014. The CS x Phase interaction reflects that there was no difference in the evaluation of CSs+ and CSs- before Habituation, *M* = 4.86, *SD* = 2.16 vs. *M* = 4.90, *SD* = 2.25, *F*(1, 47) = 0.02, *p* = .891, η_p^2 = < .001, but that the CSs+ were evaluated as more unpleasant than the CSs- after Extinction, *M* = 4.25, *SD* = 2.11 vs. *M* = 5.59, *SD* = 1.57, *F*(1, 47) = 26.84, *p* < .001, η_p^2 = .363. The change in evaluation was significant for the CS+ and the CS-, both *F*(1, 47) > 10.0, *p* = .002, η_p^2 > .180.

Insert Figure 1 about here

Electrodermal responses during habituation were larger to angry than to happy faces, *F*(1, 47) = 7.44, *p* = .009, η_p^2 = .137 (see Figure 2, left sections), and declined from block 1 to block 2, *F*(1, 47) = 71.40, *p* < .001, η_p^2 = .603. This decline was larger in participants presented with Caucasian faces *F*(1, 47) = 54.51, *p* < .001, η_p^2 = .537, than in participants presented with African American faces, *F*(1, 47) = 20.60, *p* < .001, η_p^2 = .305, as indicated by a Group x Block interaction, *F*(1, 47) = 4.40, *p* = .041, η_p^2 = .086.

During acquisition differential electrodermal conditioning was evident in both groups for happy and angry faces (see middle sections of Figure 2). The analysis yielded main effects for CS, *F*(1, 47) = 29.28, *p* < .001, η_p^2 = .384, and Block, *F*(2, 46) = 6.81, *p* = .003, η_p^2 = .228, and a CS x Block interaction, *F*(2, 46) = 7.24, *p* = .002, η_p^2 = .239. Responses to CS+ exceeded those to CS- in blocks 2, *F*(1, 47) = 37.77, *p* < .001, η_p^2 = .446, and 3, *F*(1, 47) =

14.70, $p < .001$, □ η_p^2 = .238, but not in block 1, $F(1, 47) = 2.989$, $p = .090$, □ η_p^2 = .060. No interaction involving the group or emotion factors was significant, largest $F(2, 46) = 1.37$, $p = .264$, □η_p^2 = .056. Analysis of electrodermal unconditional responses yielded no significant results with the largest effect being a main effect for Block, $F(2, 46)=2.89$, $p = .065$, η_p^2 = .112. No interaction involving the group or emotion factors was significant, largest $F(2, 46) = 2.37$, $p = .105$, η_p^2 = .093.

Insert Figure 2 about here

The right sections of Figure 2 display the electrodermal responses during extinction. As can be seen, differential electrodermal responding was present in the first Extinction block for African American faces regardless of emotional expression and for angry Caucasian faces. In the second extinction block, differential electrodermal responses were evident only for happy African American faces. The analysis confirmed this impression yielding main effects for CS, $F(1, 47) = 16.10$, $p < .001$, □ η_p^2 = .255, and Block, $F(2, 46) = 4.46$, $p = .017$, □ η_p^2 = .162, as well as Group x CS, $F(1, 47) = 5.19$, $p = .027$, □ η_p^2 = .099, and Group x Emotion x CS x Block interactions, $F(2, 46) = 4.35$, $p = .019$, □ η_p^2 = .159. Follow up analyses confirmed larger responses to CS+ than to CS- for happy African American faces on Block 1, $F(1, 47) = 7.44$, $p = .009$, □η_p^2= .137, and Block 2, $F(1, 47) = 23.14$, $p < .001$, □η_p^2 = .330, for angry African American faces on Block 1, $F(1, 47) = 7.28$, $p = .010$, □η_p^2= .134, and for angry Caucasian faces on Block 1, $F(1, 47) = 5.48$, $p = .023$, □η_p^2 = .104. No such difference emerged for happy Caucasian faces on either block, both $F < 1$. To confirm the difference in resistance to extinction of fear conditioned to angry and happy African American faces, we subjected the electrodermal data from extinction block 2 to a 2 x 2 (Emotion [Angry, Happy] x CS [CS+, CS-]) factorial ANOVA for Group African American only[1]. This analysis yielded a main effect for CS, $F(1, 24)$

= 8.87, p = .007, η_p^2 = .270, and an Emotion x CS interaction, F(1, 24) = 4.36, p = .048, η_p^2 = .154, confirming the difference in differential responding between emotions.

Figure 3 summarises the results of the affective priming task. Participants committed more errors and were slower on trials where prime face emotion and target word were of different valence than when they were valence matched. The analysis of response times yielded a main effect for Emotion, F(1, 48) = 12.98, p = .001, η_p^2 = .213, and a Emotion x Target valence interaction, F(1, 48) = 45.70, p < .001, η_p^2 = .488, indicating that participants were faster to evaluate positive targets after happy faces than after angry faces, F(1, 48) = 39.10, p < .001, η_p^2 = .449, and faster to evaluate negative targets after angry faces than after happy faces, F(1, 48) = 7.23, p = .010, η_p^2 = .131. The analysis of error percentages yielded an Emotion x Target valence interaction, F(1, 48) = 38.35, p < .001, η_p^2 = .444, indicating that participants made fewer errors evaluating positive targets after happy faces than after angry faces, F(1, 48) = 30.46, p < .001, η_p^2 = .388, and fewer errors evaluating negative targets after angry faces than after happy faces, F(1, 48) = 23.80, p < .001, η_p^2 = .331.

Insert Figure 3 about here

The results of the affective priming task confirm the differential evaluation of angry and happy faces that was evident in the explicit ratings as well. They do not provide support for a more positive evaluation of happy African American than of happy Caucasian faces. The only interaction involving the factor Group with an F larger than 1, the Group x Emotion x Target valence, F(1, 48) = 3.94, p = .053, η_p^2 = .076, found for error data reflects on a difference in response to angry primes. Moreover, post hoc analyses of between group differences in the effect of happy primes yielded no significant outcomes, t < 1, for error percentages and evaluation times.

Discussion

Angry racial outgroup faces combine two facial cues (emotion and race) that have been shown to mediate fear learning that is resistant to extinction. The current experiment was designed to assess whether fear conditioned to angry racial outgroup faces shows superior resistance to extinction. Resistance to extinction was not enhanced for fear conditioned to angry outgroup faces in comparison to either happy outgroup faces or angry ingroup faces. Rather, conditioning to happy outgroup faces was the most robust. This occurred although happy outgroup faces were rated as more pleasant than happy ingroup faces and did not differ from the evaluation of happy ingroup faces in an implicit evaluation task.

The present experiment replicates a number of findings from research on face processing (Hart et al., 2000; Hess, Sabourin & Kleck, 2007) and from the human conditioning literature (Öhman, & Dimberg, 1978; Olsson et al., 2005). Angry faces were rated and evaluated as less pleasant than happy faces and, during habituation, elicited larger electrodermal orienting responses than did happy faces. Electrodermal orienting responses to racial outgroup faces were slower to habituate than orienting responses to ingroup faces, a finding consistent with the observation that amygdala activation to ingroup, but not to outgroup faces declines across repeated stimulus presentations (Hart et al., 2000). Fear conditioning to angry or happy ingroup or outgroup faces did not differ during acquisition, however, clear differences across face categories emerged during extinction. Extinction of fear conditioned to racial ingroup faces was retarded if these faces displayed anger in comparison to happiness. This is to the best of our knowledge the first demonstration of this difference in a within subject conditioning design. It should be noted that contrary to prior evidence (see Öhman, & Dimberg, 1978; Rowles et al., 2012), fear conditioned to angry ingroup faces did show extinction within the three blocks of extinction training used and that extinction of fear conditioned to happy ingroup faces was evident very early in extinction. This may reflect on the use of a within subject design in which two different CS+ were presented without the US during extinction training. Thus, for half of the participants the first presentation of, for instance, the happy same race CS+ during extinction was preceded by a presentation of the angry same race CS+ without the US. This may reduce responding to the CS+ although it was reliably paired with the US during acquisition or enhance responding to the corresponding CS- as some participants may expect a contingency reversal during the initial trials of extinction. The experience that two different CS+ are no longer paired with the unconditional stimulus may facilitate extinction of fear conditioning that has been shown to be subject to cognitive control (Mallan et al., 2009; Rowles et al., 2012) and not encapsulated from cognition (Seligman, 1971; Öhman & Mineka, 2001).

Fear conditioning to outgroup race faces was resistant to extinction in comparison to fear conditioned to happy ingroup faces (Olsson et al., 2005; Mallan et al. 2009). This held regardless of the emotion displayed by the outgroup faces indicating that, contrary to gender information (Navarrete et al., 2009), expressions of positive emotion do not ameliorate the effect of ‘other race’ on fear conditioning. The finding that fear conditioned to happy outgroup faces was more robust than fear conditioned to angry outgroup faces can be seen as consistent with prior reports that happy outgroup faces are not regarded as positive, are evaluated as more negative than fearful outgroup faces, and may be associated with fear rather than happiness by members of the ingroup (Weisbuch & Ambadi, 2008). However, this interpretation is not consistent with the explicit and implicit evaluations obtained in the present study. Participants rated happy outgroup faces as more pleasant than angry outgroup faces – and as more pleasant than happy ingroup faces. It should be noted, however, that the latter difference was not evident in the implicit evaluations obtained in the affective priming task. Here, happy faces were evaluated as more pleasant than angry faces regardless of race. It seems likely that this difference between explicit and implicit evaluations reflects on the participant's desire to appear unbiased against persons from a different racial background. This may have led them to evaluate the happy outgroup faces as more positive than the happy ingroup faces. Nevertheless, both findings are inconsistent with Weisbuch and Ambadi's report that happy outgroup faces are seen as fear provoking rather than as affiliative. It should be noted, however, that Weisbuch and Ambadi's findings were obtained in a context in which fear, not anger, was the negative expression with which happiness was contrasted and in which participants saw both ingroup and outgroup faces. Given that fear conditioned to fearful faces has also been suggested to be subject to prepared learning (Orr & Lanzetta, 1980), future studies should contrast fear conditioning to happy and to fearful in- and outgroup faces to assess whether Weisbuch and Ambadi's finding of affective divergence extends to fear conditioning.

Differential electrodermal responding diminished across extinction in all conditions to the extent that it was no longer significant in Block 3. Nevertheless, post experimental ratings of conditional stimulus pleasantness were significantly different for CS+ and CS-. This recovery of differential stimulus evaluation from the end of extinction to post experimental assessments is well documented (Lipp, Oughton, & LeLievre, 2003). Interestingly, it does seem to occur rather non-selectively and independently of any of the between condition differences that are observed during extinction training (Lipp & Edwards, 2002; Rowles et al., 2012). Thus, the recovery seems to reflect the differential evaluations that were acquired during acquisition which were uniform across the experimental conditions. This re-emergence of differential stimulus evaluations has been interpreted as reflecting renewal, the re-emergence of conditioned responding after successful extinction due to context change (Bouton, 2002). It is thought to reflect the fact that extinction learning is context specific whereas the initial learning that occurred during acquisition is not. Thus, when assessed in a different context, responses to a CS are more likely to reflect behavior shown during acquisition than during extinction. This raises a number of interesting issues in reference, for instance, to the assessment of the efficacy of exposure based behavioural treatments which are thought to be mediated mainly by extinction processes.

In summary, the current study assessed whether resistance to extinction of fear conditioned to facial cues of race and emotional expression is enhanced if fear is conditioned to angry other race faces. It replicated previous reports of resistance to extinction of fear conditioned to angry or other race faces, regardless of whether these expressed happiness or anger. It did not, however, find evidence for an enhanced resistance to extinction of fear conditioned to angry other race faces. Thus, at least in the current context, two threats seem no worse than one.

References

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry, 52*, 976-986. doi: 10.1016/S0006-3223(02)01546-9

Bruce, V., & Young, A. (1986). Understanding face perception. *British Journal of Psychology, 77*, 305-327.

Craig, B. M., Mallan, K. M., & Lipp, O. V. (2012). The effect of poser race on the happy categorisation advantage depends on stimulus type, set size, and presentation duration. *Emotion, 12*(6), 1303-1314. doi: 10.1037/a0028622

Fazio, R. H., & Olsen, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*, 297-327. doi: 10.1146/annurev.psych.54.101601.145225

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

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology, 74*, 1464-1480. doi: 10.1037/0022-3514.74.6.1464

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *NeuroReport, 11*(11), 2351-2355.

Hess, U., Sabourin, G., & Kleck, R. E. (2007). Postauricular and eye-blink startle responses to facial expressions. *Psychophysiology, 44*, 431-435. doi: 10.1111/j.1469-8986.2007.00516.x

Hugenberg, K. (2005). Social categorization and the perception of facial affect: Target race moderates the response latency advantage for happy faces. *Emotion, 5*(3), 267-276. doi: 10.1037/1528-3542.5.3.267

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, *14*, 640-643.

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

Leppänen, J. M., & Hietanen, J. K. (2003). Affect and face perception: Odors modulate the recognition advantage of happy faces. *Emotion, 3*(4), 315-326. doi: 10.1037/1528-3542.3.4.315

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., Oughton, N., & LeLievre, J. (2003). Evaluative learning in human Pavlovian conditioning: Extinct, but still there? *Learning and Motivation, 34*, 219-239. doi: doi:10.1016/S0023-9690(03)00011-0

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

Mallan, K. M., Lipp, O. V., & Cochrane, B. (2013). Slithering snakes, angry men and out-group members: What and whom are we evolved to fear? *Cognition & Emotion*. doi: 10.1080/02699931.2013.778195

Mazurski, E. J., Bond, N. W., Siddle, D. A. T., & Lovibond, P. F. (1996). Conditioning with facial expressions of emotion: Effects of CS sex and age. *Psychophysiology 33*, 416-425. doi: 10.1111/j.1469-8986.1996.tb01067.x

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. (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

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

Orr, S.P., & Lanzetta, J.T. (1980). Facial expressions of emotion as conditioned stimuli for human autonomic responses. *Journal of Personality and Social Psychology, 38*, 278-282. doi: 10.1037/0022-3514.38.2.278

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.

Rowles, M. E., Lipp, O. V., & Mallan, K. M. (2012). On the resistance to extinction of fear conditioned to angry faces. *Psychophysiology, 49*, 375-380. doi: 10.1111/j.1469-8986.2011.01308.x

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

Tottenham, N., Tanaka, J.W., Leon, A.C., McCarry, T., Nurse, M., Hare, T. A., ….Nelson, C. (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

Weisbuch, M., & Ambady, N. (2008). Affective divergence: Automatic responses to others' emotions depend on group membership. *Journal of Personality and Social Psychology, 95*, 1063–1079. doi: 10.1037/a0011993

Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology, 102*, 959-974. doi: 10.1111/j.2044-8295.1986.tb02199.x

Footnotes

[1] We would like to thank the anonymous reviewer who suggested this analysis.

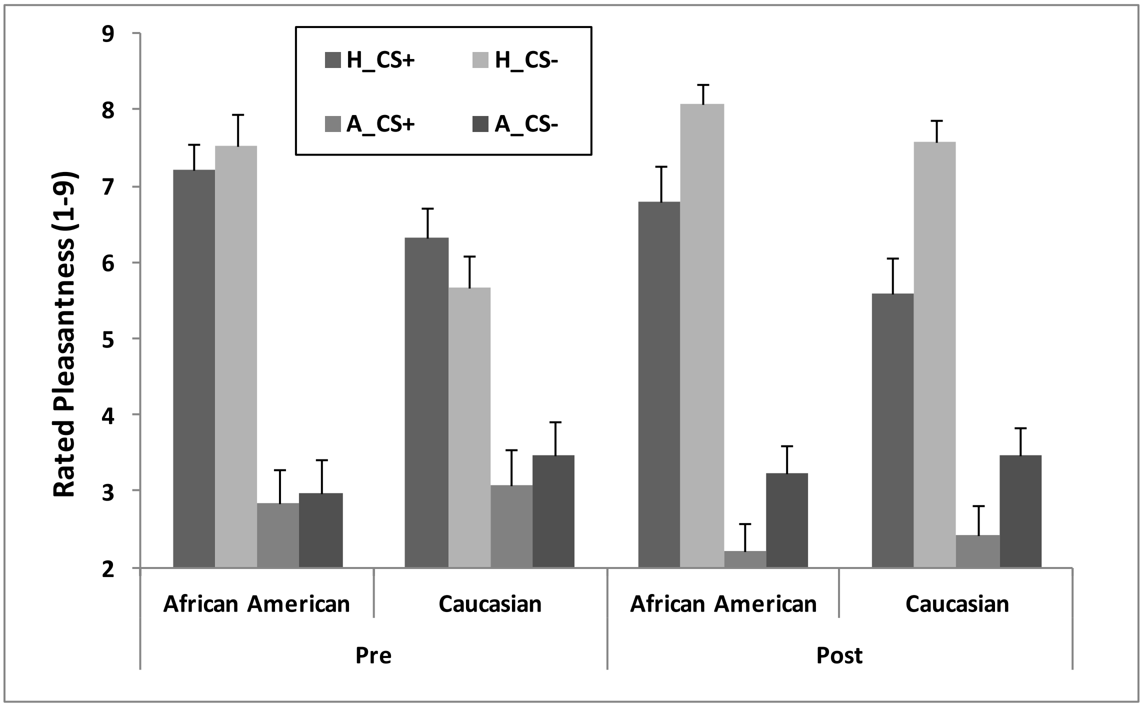
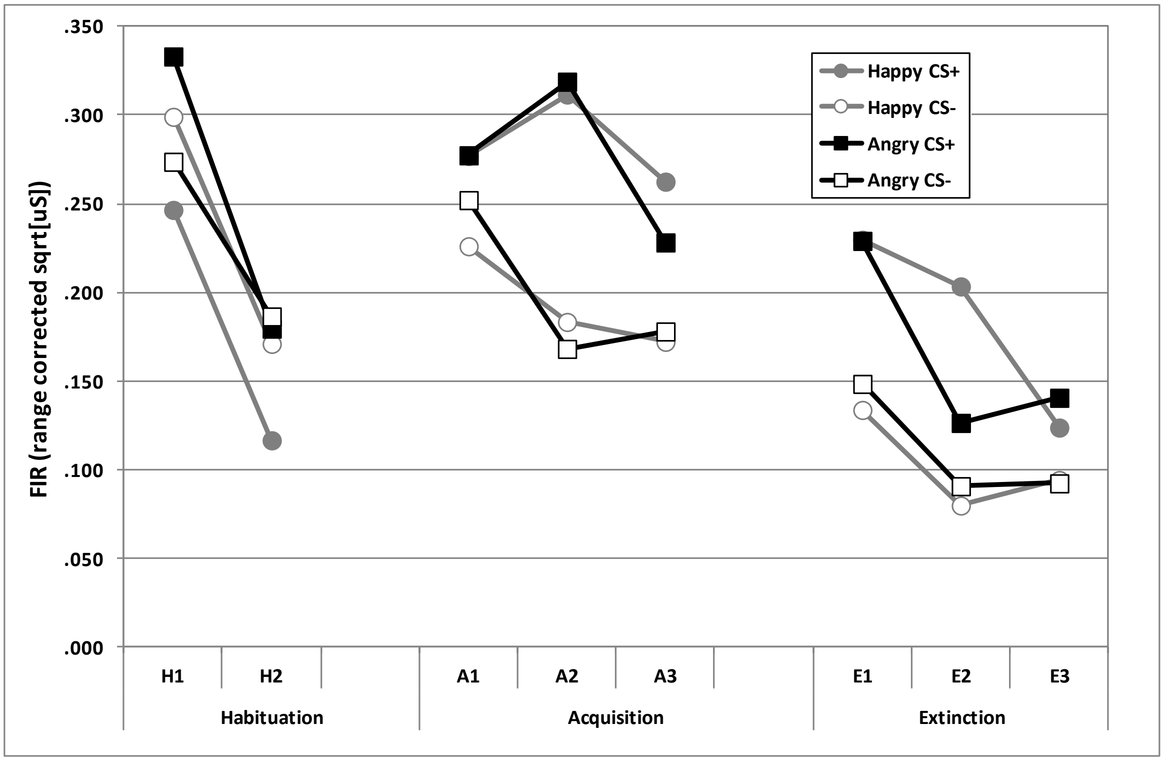


Figure 1. Rated CS pleasantness before habituation (Pre) and after extinction (Post) in groups trained with African American or Caucasian faces as a function of emotion (H: Happy; A: Angry) and CS condition (CS+, CS-; error bars represent standard errors of the means).



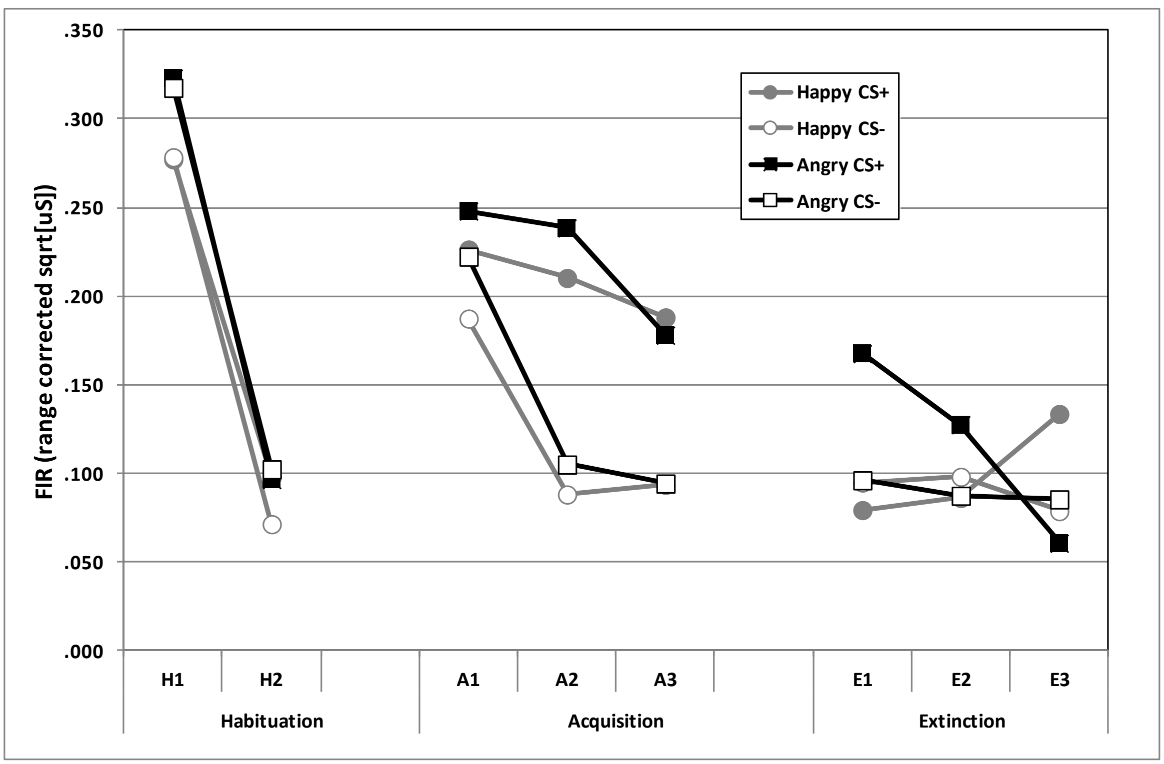


Figure 2. Electrodermal responses during Habituation, Acquisition, and Extinction in groups trained with African American (upper panel) or Caucasian faces (lower panel) as a function of emotion (Happy; Angry), CS condition (CS+, CS-), and trial blocks.

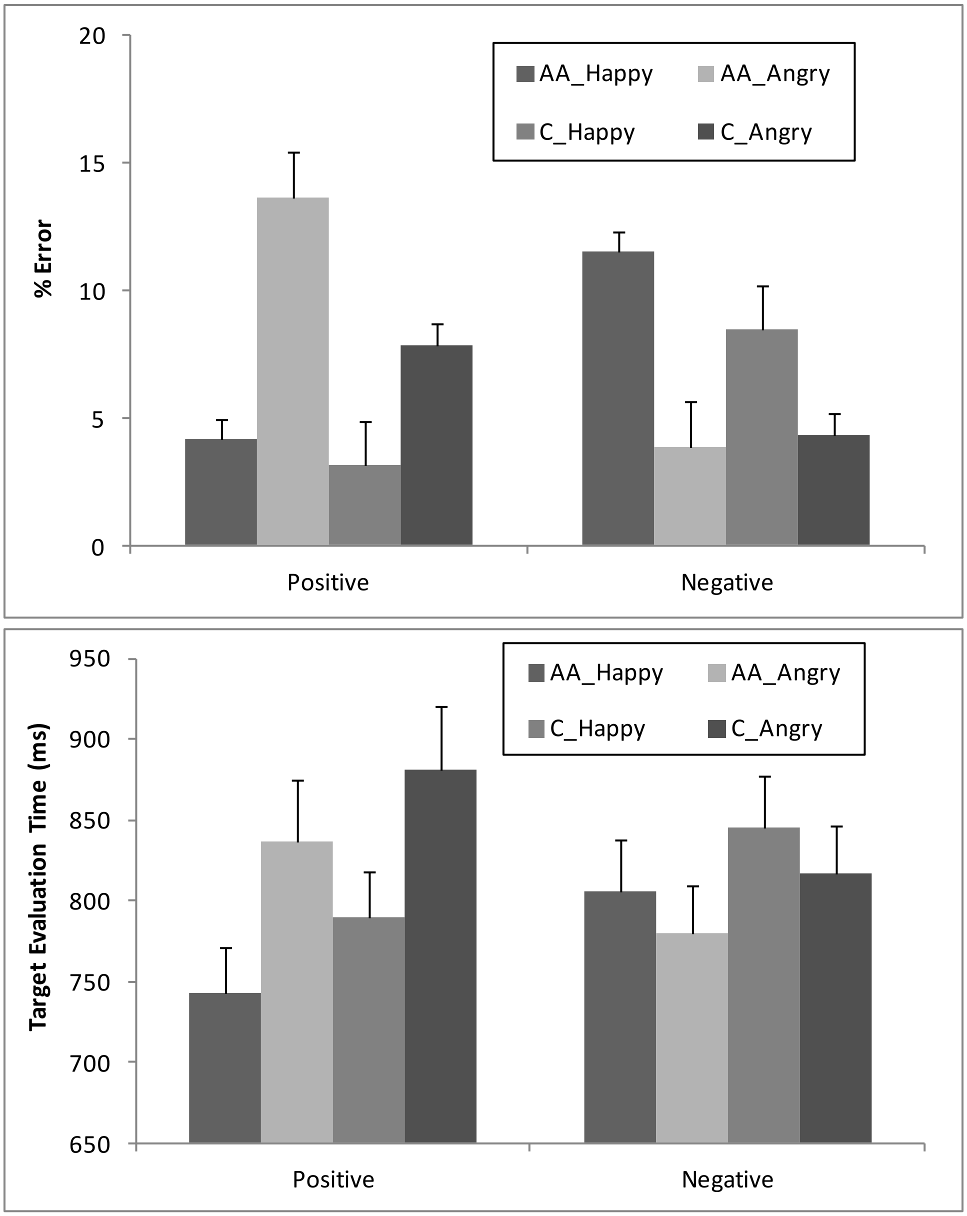


Figure 3. Percent error (upper panel) and target evaluation time (lower panel) in the affective priming task as a function of prime face ethnicity (AA: African American; C: Caucasian) and emotion (Happy; Angry) and target word valence (Positive, Negative; error bars represent standard errors of the means).