Biological Psychology 129 (2017) 195–206

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Fear acquisition and liking of out-group and in-group members: Learning bias or attention?

Stephan Koenig*,[1], Peter Nauroth[1], Sara Lucke, Harald Lachnit, Mario Gollwitzer, Metin Uengoer

Department of Psychology, Philipps-Universität Marburg, Germany

ARTICLE INFO

Keywords:
Intergroup bias
Preparedness
Fear conditioning
Attention
Skin conductance response
Pupil response
Evaluative learning

ABSTRACT

The present study explores the notion of an out-group fear learning bias that is characterized by facilitated fear acquisition toward harm-doing out-group members. Participants were conditioned with two in-group and two out-group faces as conditioned stimuli. During acquisition, one in-group and one out-group face was paired with an aversive shock whereas the other in-group and out-group face was presented without shock. Psychophysiological measures of fear conditioning (skin conductance and pupil size) and explicit and implicit liking exhibited increased differential responding to out-group faces compared to in-group faces. However, the results did not clearly indicate that harm-doing out-group members were more readily associated with fear than harm-doing in-group members. In contrast, the out-group face not paired with shock decreased conditioned fear and disliking at least to the same extent that the shock-associated out-group face increased these measures. Based on these results, we suggest an account of the out-group fear learning bias that relates to an attentional bias to process in-group information.

1. Introduction

In many intergroup conflicts, members of the conflicting parties hold negative attitudes about each other. The dislike for out-group members is a major obstacle in resolving these conflicts. For instance, the Israeli-Palestinian conflict is marked by a mutual antipathy between Israelis and Palestinians that severely impedes political solutions (Maoz & McCauley, 2005). Understanding the formation of such antipathy requires studying how individuals learn to dislike out-group members. In this regard, the hypothesis of an *out-group fear learning bias* where out-group members are more readily associated with negative attributes has gained much attention (Mallan, Lipp, & Cochrane, 2013; Ohman, 2005). Empirically, this hypothesis is supported by enhanced differential fear conditioning to out-group faces (Navarrete et al., 2012; Olsson, Ebert, Banaji, & Phelps, 2005). The purpose of the present study is to critically evaluate this notion. As illustrated with reference to standard associative learning theory, greater differential responding to out-group faces is also predicted by an *attentional bias for the in-group*. The introduction is structured in two sections :The first section introduces the differential conditioning paradigm that we used in our experiment and that has previously been reported to yield evidence in favor of an out-group fear learning bias. The second section derives two contrasting hypotheses of (a) enhanced out-group fear learning and (b) enhanced attention to the in-group. Both hypotheses are in accord with greater differential fear conditioning to out-group members. However the hypotheses differ with respect to the underlying pattern of acquired CS-US associations.

1.1. Empirical evidence for an out-group fear learning bias

Olsson et al. (2005 also see Ohman, 2005) investigated whether racial out-group members were more readily associated with conditioned fear than in-group members. They used a within-subject conditioning procedure that presented participants with two racial out-group faces (i.e., African-American faces for Caucasian-American participants and Caucasian-American faces for African-American participants) and two racial in-group faces (having the same race as the participant) as conditioned stimuli (CS). In the acquisition stage of a differential conditioning procedure, one out-group face (Out+) and one in-group face (In+) were paired with an aversive electric shock (unconditioned stimulus; US) while the shock was omitted after the other two faces (Out−, In−). In the following extinction phase all stimuli were presented without the US. The authors measured the skin conductance response (SCR) elicited by the faces and reported that the *differential* SCR to the out-group faces was more pronounced than the *differential* SCR to the in-group faces, i. e. the difference "Out+ minus

* Corresponding author at: Philipps-Universität Marburg, Department of Psychology, Gutenbergstrasse 18, 35032 Marburg, Germany.
E-mail address: koenigst@uni-marburg.de (S. Koenig).
[1] Stephan Koenig and Peter Nauroth have equally contributed to the manuscript and share first authorship.

http://dx.doi.org/10.1016/j.biopsycho.2017.08.060
Received 3 June 2016; Received in revised form 30 August 2017; Accepted 30 August 2017
Available online 01 September 2017

Out − ” was greater than the difference “In + minus In − ” (in fact the authors only reported these difference scores).

Olsson et al. interpreted their finding as evidence of a human “preparedness” to readily associate out-group members with harm. This interpretation was based on similar findings for enhanced fear conditioning with phylogenetically fear-relevant stimuli such as snakes (Ohman, Fredrikson, Hugdahl, & Rimmö, 1976) or angry faces (Ohman & Dimberg, 1978; Ohman & Soares, 1993, 1998; for review see Mallan et al., 2013; Ohman & Mineka, 2001). The conceptual link to these previous studies also justified the focus on differential responding to control for “preexisting differences in the emotional salience of stimulus categories as a confounding variable.” (Olsson et al., 2005, p. 785). If a racial out-group face (much like an angry face) elicits a more pronounced SCR than an in-group face (happy face) *prior* to conditioning, focusing on the differences “CS + minus CS − ” of course controls for these baseline differences. However, the problem with this rational is that the CS − might very well change its association with shock during conditioning. In particular, the non-reinforced CS − is established as a safety signal that is supposed to exhibit an inhibitory association with the US in standard theories of associative learning (Pearce and Hall, 1980; Pearce, 1987, 1994, 2002; Rescorla and Wagner, 1972; Wagner and Rescorla, 1972; Wagner, 1981, 2003; see Lovibond, Siddle, & Bond, 1993 for an explicit reference to inhibitory learning about fear-relevant CS). From this perspective, the difference “Out + minus Out − ” could not only increase because of *increased* responding to the reinforced out-group face (with Out + > In +) but also could increase because of *decreased* responding to the non-reinforced out-group face (with Out − < In −). It should be clear that the latter effect stands in stark contrast to the assumption of an out-group fear learning bias because it would mean that out-group faces are also readily established as a cue that signals safety from harm.

In another experiment that reported evidence in favor of an out-group fear learning bias, Navarrete et al. (2012) demonstrated that during fear conditioning participants exhibited a larger differential SCR towards arbitrarily-defined social out-group members. Here, groups were established using a variant of the minimal group paradigm (Tajfel, Billig, Bundy, & Flament, 1971). On arrival, participants judged the predominant hue in a bicolored stimulus and were assigned to their “group” depending on their choice of color. They wore a t-shirt with their group color during fear conditioning and the CSs consisted of four images of white man with neutral expression that were digitally manipulated to appear as wearing the same t-shirt color (in-group) or the other color (out-group). Using the same design as Olsson et al. (2005) participants underwent differential fear conditioning Out +, Out −, In +, In −, using electric shock as the US. Again, participants exhibited a larger differential SCR “Out + minus Out − ” and again these difference scores were the only ones reported. In this procedure, even more so, it seems highly questionable whether the random assignment of faces to the minimal out-group was sufficient to establish a baseline difference Out − > In −, and again the observed difference could have been caused by *less* conditioned responding to the Out − face than the In − face (rather than or in addition to Out + > In +).

1.2. Discrimination learning, stimulus generalization and inhibitory conditioning

As outlined above, differential fear conditioning CS +, CS − may result in inhibitory learning about the CS − . The CS − is established as a cue that signals safety from harm (electric shock) in a context in which such harm sometimes occurs (in CS + trials). For further illustration, we will refer to the learning theory of Rescorla and Wagner (1972) as the “most widely accepted description of associative changes during classical conditioning” (Gluck & Bower, 1988, p. 228). Here, conditioned responding to a CS is determined by the sum of associations held by all elements comprising the CS. For example, simple differential fear conditioning formally consists of AC +, BC − where A denotes elements unique to the CS +, B denotes elements unique to the CS −, and C denotes elements common to both CS (usually the experimental context but possibly also perceptual or conceptual elements common to both stimuli). With respect to two angry faces, common elements are common facial features that cause both faces to be perceived as angry. With respect to racial out-group faces, common elements are race-defining characteristics such as a common skin color. With respect to the minimal groups of Navarrete et al. (2012), the element common to out-group members was the t-shirt color (all example similarly refer to the in-group). During differential conditioning AC +, BC − element A is supposed to acquire an excitatory association with the US. Since the common element C is followed by shock on 50% of the trials (in AC + trials) this common element also acquires some excitatory association (but less than A). To correctly predict the absence of shock in BC − trials, element B must acquire an inhibitory association to counteract the excitatory association of element C (for further details see Appendix A). Put less formally, participants learn that out-group members in general exhibit some probability of causing harm, out-group face A is especially dangerous but out-group face B actively prevents harm in an otherwise dangerous environment.

The left panel of Fig. 1 depicts a simulation of the Rescorla-Wagner model where free model parameters have been chosen to be in accord with the hypothesis of an out-group fear learning bias: (a) Out + and Out − exhibit some association with fear prior to conditioning as indicated by their non-zero association in the first block; (b) the aversive US is simulated to be more effective when paired with out-group faces so that the Out + face approaches a higher asymptote than the In + face; (c) the learning rate is Out + > In + promoting faster excitatory fear conditioning for out-group faces, (d) the learning rate is Out − < In − promoting slower inhibitory fear conditioning for out-group faces (for further details see Appendix A). The simulation makes clear that a standard theory of associative learning such as the Rescorla-Wagner theory can predict a pattern of greater differential responding to out-group faces consistent with the hypothesis that out-group faces are more effective CS specifically when paired with an aversive US.

The right panel of Fig. 1 illustrates that the same theory predicts the

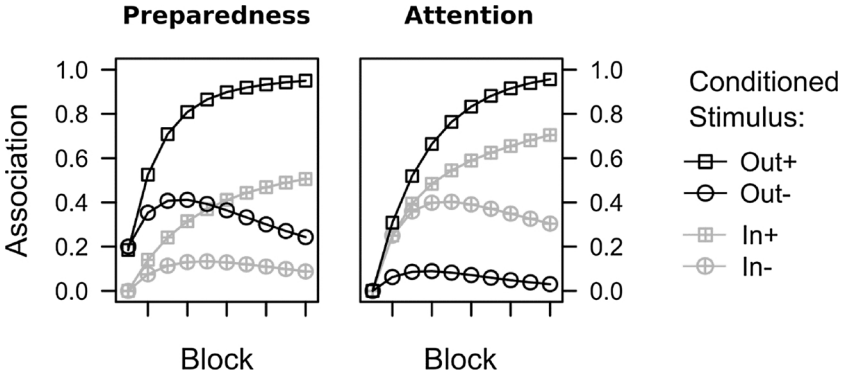

Fig. 1. Illustration of different processes that may lead to an out-group conditioning bias in the learning theory of Rescorla and Wagner (1972). For further explanation see text. For simulation details see Appendix A.