

1 **Social learning of fear and safety is determined by the demonstrator's racial group**

2

3 Armita Golkar*, Vasco Castro & Andreas Olsson

4 Karolinska Institutet, Department of Clinical Neuroscience, Division of Psychology

5

6 *Corresponding author

7 Email: armita.golkar@ki.se

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

[Skriv här]

24 Abstract

25 Social learning offers an efficient route through which humans and other animals learn about
26 potential dangers in the environment. Such learning inherently relies on the transmission of
27 social information and should imply selectivity in *what* to learn from *whom*. Here, we conducted
28 two observational learning experiments to assess how humans learn about danger and safety
29 from members ('demonstrators') of another social group than their own. We show that both
30 fear and safety learning from a racial in-group demonstrator was more potent than learning from
31 a racial out-group demonstrator.

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

[Skriv här]

48 **Introduction**

49 Learning to predict potential dangers in the environment is critical for survival. Of equal
50 importance is the ability to flexibly adapt to changing environmental contingencies, such as
51 learning that a past threat is now safe. Whereas previous research has focused on how such
52 predictions are formed and changed through own, direct experiences, much of what humans
53 and other animals learn about the environment comes from social forms of learning, such as
54 observing others [1]. Here, we assessed whether learning about danger and safety was
55 determined by whether the demonstrator belonged to one's own (in-group) or another (out-
56 group) racial group.

57 A substantial body of research on social learning supports that animals are selective with
58 regards to *whom* they learn from. For example, although both primates [2-4] and rodents [5]
59 can observationally acquire fears from unrelated conspecifics, the strength of such learning is
60 enhanced by relatedness [5, 6], familiarity, and social status [6]. Such social learning biases
61 might have conferred an adaptive advantage to animals, including humans, by facilitating the
62 acquisition of locally relevant knowledge [1]. However, in increasingly diverse modern human
63 societies, such social learning biases can be maladaptive both because much important
64 information is of global (as opposed to local) significance, and because traits, such as racial
65 group membership, are less reliable indicators of whether or not individuals have locally
66 relevant knowledge. Recent studies have shown that humans will more readily learn to fear
67 racial out-group members [e.g., 7], but it is unknown whether group membership affects the
68 extent to which humans learn from others. We hypothesized that both fear and safety learning
69 would be more potent when learned from an in-group demonstrator than from an out-group
70 demonstrator. This was based on the presence of similarity-based learning biases in other
71 animals [5, 6], and a general tendency to display greater empathic and otherwise pro-social
[Skriv här]

72 responses to in-group, as compared to out-group, members [8, 9]. To test these hypotheses, we
73 conducted two separate observational learning experiments based on previously established
74 procedures for social fear and safety learning [4, 10]. The first experiment was designed to
75 measure observational fear learning, i.e. the *acquisition* of fear, and the second was designed to
76 measure observational safety learning, i.e. the *extinction* of acquired fear. To assess learning,
77 we used a standard index of conditioned response (CR) in humans: the skin conductance
78 response (SCR), reflecting the phasic increase in skin conductance that occurs in response to
79 physiologically arousing stimuli.

80

81 **Material & Method**

82 In both Experiment 1 and 2, participants were White Swedish residents of European origin
83 and were recruited on, or nearby, Karolinska Institutet campus. Participants were excluded if
84 they either failed to report the correct contingency between the conditioned stimulus (CS) and
85 the unconditioned stimulus (US) during acquisition [10] or their differential SCR were above
86 or below 2.5 standard deviations (*SD*) from the group mean in the acquisition, extinction, or
87 test stage (see Supplementary Material). For both experiments, one image of a snake and one
88 image of a spider served as CSs and four different male individuals (two White and two Black
89 counterbalanced across participants) served as demonstrators. Throughout all experimental
90 stages, each CS was presented six times for six seconds in a pseudo-randomized order with an
91 inter-trial-interval (ITI) ranging between 12-18 seconds. After both experiments, participants
92 completed a computerized Implicit Association Task [11]; assessing implicit racial bias and a
93 questionnaire assessing explicit racial bias (Modern Racism Scale questionnaire) [12].

[Skriv här]

94 In Experiment 1, 46 White European participants (7 of which were excluded, see
95 Supplementary material) first underwent an *observational acquisition* stage. During this stage,
96 participants viewed a video depicting either the in- (white) or the out-group (black)
97 demonstrator displaying discomfort when receiving an electrical shock (the US) paired with the
98 presentation of an image of a snake (the CS+), but never when paired with a spider image (the
99 CS-). During a subsequent direct test, participants were re-exposed to the CSs, without being
100 shocked, and in the absence of the demonstrator. Accordingly, any fear response to the CS+
101 was due to social learning taking place during the observational acquisition stage.

102 In Experiment 2, 55 White European participants (10 of which were excluded; see
103 Supplementary Material) underwent an observational extinction paradigm developed by our
104 group [10]. Briefly, participants first underwent standard fear conditioning during which they
105 received a mild electrical shock (the US) when presented with an image of a snake (CS+), but
106 never when presented with an image of a spider (CS-). Then, during *observational extinction*,
107 participants watched a video depicting either the in- (white) or the out-group (black)
108 demonstrator acting calmly when exposed to presentations of both CSs. Finally, to assess fear
109 recovery, participants underwent a standard reinstatement test during which they received
110 three reminder shocks before they were re-exposed to the CSs during the *direct test*. This
111 procedure has been shown to reinstate the expression of the original fear memory in both
112 animals [13] and humans [14].

113 In both experiments, physiological fear responses were based on the mean SCR to the
114 CS+ and CS- separately for each stage and analyzed using separate mixed analysis of variance
115 (ANOVA) with CS (CS+, CS-) as a within-subject factor and Group (In-group demonstrator,
116 Out-group demonstrator) as a between-subject factor. As CRs extinguish rapidly in the absence
117 of shocks [15], the *direct test* data in both Experiment 1 and 2 was divided into an early and
[Skriv här]

118 late stage, defined by the mean first three responses versus the mean subsequent three responses
119 and analyzed in a CS (CS+, CS-) x Time (early, late) x Group (In-group demonstrator, Out-
120 group demonstrator) mixed ANOVA.

121

122 **Results**

123 The CR was indexed as the differential SCR to the snake (CS+) and spider (CS-) image.
124 The mean CR from both experiments are presented in Figure 1.

125 In Experiment 1, we observed a significant group difference during *observational*
126 *acquisition*, (CS x Group: $F(1,37) = 7.57, p = .009, \eta^2 = .17$). Participants that were exposed to
127 the in-group demonstrator acquired a CR, $t(18) = 4.04, p = .001$, whereas this effect was not
128 significant in the group of participants exposed to the out-group demonstrator, ($t(19) = .28, p =$
129 $.79$). During the *direct test*, participants still differed during the early stage (CS x Time x Group:
130 $F(1,37) = 4.21, p = .04, \eta^2 = .10$), but both groups of participants nevertheless expressed a
131 significant CR (in-group demonstrator: $t(18) = 6.68, p < .001$; out-group- demonstrator: $t(19) =$
132 $2.95, p = .008$). Interestingly, correlation analysis demonstrated that explicit, but not implicit,
133 negative racial attitudes predicted CRs during the early test stage ($r = -.5, p = .026$), so that
134 participants with more negative racial attitudes expressed less learning from a racial out-group
135 demonstrator (see Supplementary material).

136 In Experiment 2, participants showed equivalent level of CR during *direct acquisition*
137 (Main effect of CS: $F(1,43) = 13.10, p = .001, \eta^2 = .23$) and there was a small remaining effect
138 of CS during *observational extinction* ($F(1,43) = 4.11, p = .049, \eta^2 = .09$). Importantly,
139 participants differed in CR during the subsequent *direct test*, (CS x Group: $F(1,43) = 4.85, p =$
140 $.033, \eta^2 = .10$, CS x Time x Group: $F(1,43) = .15, p = .70$). As predicted, the recovery of fear
141 was stronger in the group of participants exposed to an out-group demonstrator (in-group: $t(22)$

[Skriv här]

142 = 1.74, $p = .095$; out-group: $t(21) = 5.07, p < .001$). Finally, neither implicit nor explicit racial
143 attitudes moderated the out-group learning bias in Experiment 2 (see Supplementary material).

144 --Insert Figure 1 --

145 Discussion

146 Our results show that both fear and safety learning are more potently learned from an in-
147 group as compared to an out-group demonstrator. These findings concur with research in non-
148 human animals showing that fears are readily learned from both social in- and out-group
149 demonstrators [2, 5], but that fear learning can be enhanced depending on the social relationship
150 between the observer and demonstrator [5, 6]. Experiment 2 demonstrated that the efficacy of
151 social safety learning was restricted to conditions in which safety information was transmitted
152 from an in-group demonstrator.

153 Taken together with previous research on social learning in non-human animals, the
154 evolutionary interpretation of our findings suggests that humans are predisposed to learn about
155 potential danger and safety from individuals belonging to their own social group. Because social
156 categorization based on racial difference among humans occurred relatively recently in human
157 evolutionary history [16], it is unlikely that responses to racial differences are under strong
158 genetic influence [17], and therefore race *per se* is unlikely to explain our findings. Even if a
159 genetic predisposition to preferentially learn from in-group individuals more generally exists,
160 this is likely to be influenced by socially acquired attitudes concerning racial groups [18]. In
161 support of this, we found a significant relation between negative racial attitudes and less
162 expressed fear learning from a racial out-group member, stressing the importance of assessing
163 how stereotypic beliefs about members of out-groups can bias both whom we learn from and
164 what we learn from whom. Indeed, children show superior learning from similar others [19]
165 and develop negative beliefs about racial out-group members early in life [20, 21], and similar

[Skriv här]

166 racial-biases have been demonstrated in related social situations that include inter-individual
167 encounters, such as cooperation and trust [22], empathy [9, 23], and altruism [24].

168 Theoretical and empirical work highlights the tendency in animals to be selective with
169 regards to *when* and *whom* to learn from, and that natural selection will favor the use of adaptive
170 social strategies to guide the individual's reliance on social information [1]. Here, we
171 demonstrate that human social learning is determined by the demonstrator's racial group. Given
172 the possibility that using participants and demonstrators of other racial groups could produce
173 different results, future work should address whether these effects generalize to other racial and
174 social groups.

175

176 **Acknowledgements**

177 We thank Elizabeth A. Phelps and Carlos D. Navarrete for helpful comments on an earlier draft.

178

179 **Author contributions**

180 A. Golkar and A. Olsson developed the study concept and designed the experiments. Testing
181 and data collection were performed by V. Castro. A. Golkar performed the data analysis and
182 drafted the manuscript, and A. Olsson provided critical revisions. All authors approved the
183 final version of the manuscript for submission.

184

185 **Competing interests**

186 The authors declare that they had no conflicts of interest with respect to their authorship or the
187 publication of this article.

188

[Skriv här]

189 **Data accessibility**

190 Data associated with the manuscript are accessible through Dryad; Provisional DOI:
191 doi:10.5061/dryad.n9v18.

192

193 **Ethics approval and consent to participate procedures.**

194 The study was approved by the local ethics committee at Karolinska Institutet and all
195 participants gave written consent prior to participation

196

197 **Funding**

198 This research was supported by an Independent Starting Grant (284366; ELSI) from the
199 European Research Council to A. Olsson.

200

201

202

203

204

205

206

207

208

209

210

211 **References**

- 212 1. Laland K.N. 2004 Social learning strategies. *Learning & Behavior* **32**(1), 4-14.
- 213 2. Cook M., Mineka S., Wolkenstein B., Laitsch K. 1985 Observational conditioning of snake
214 fear in unrelated rhesus-monkeys. *Journal of Abnormal Psychology* **94**(4), 591-610.
- 215 3. Hygge S., Ohman A. 1978 MODELING PROCESSES IN ACQUISITION OF FEARS -
216 VICARIOUS ELECTRODERMAL CONDITIONING TO FEAR-RELEVANT STIMULI.
217 *Journal of Personality and Social Psychology* **36**(3), 271-279. (doi:10.1037//0022-
218 3514.36.3.271).
- 219 4. Olsson A., Phelps E.A. 2004 Learned fear of "unseen" faces after Pavlovian, observational,
220 and instructed fear. *Psychological Science* **15**(12), 822-828. (doi:10.1111/j.0956-
221 7976.2004.00762.x).
- 222 5. Jeon D., Kim S., Chetana M., Jo D., Ruley H.E., Lin S.Y., Rabah D., Kinet J.P., Shin H.S.
223 2010 Observational fear learning involves affective pain system and Ca(v)1.2 Ca²⁺
224 channels in ACC. *Nature Neuroscience* **13**(4), 482-U105. (doi:10.1038/nn.2504).
- 225 6. Kavaliers M., Colwell D.D., Choleris E. 2005 Kinship, familiarity and social status
226 modulate social learning about "micropredators" (biting flies) in deer mice. *Behavioral*
227 *Ecology and Sociobiology* **58**(1), 60-71. (doi:10.1007/s00265-004-0896-0).
- 228 7. Olsson A., Ebert J.P., Banaji M.R., Phelps E.A. 2005 The role of social groups in the
229 persistence of learned fear. *Science* **309**(5735), 785-787. (doi:10.1126/science.1113551).
- 230 8. Hein G., Silani G., Preuschoff K., Batson C.D., Singer T. 2010 Neural Responses to
231 Ingroup and Outgroup Members' Suffering Predict Individual Differences in Costly
232 Helping. *Neuron* **68**(1), 149-160. (doi:10.1016/j.neuron.2010.09.003).

- 233 9. Xu X.J., Zuo X.Y., Wang X.Y., Han S.H. 2009 Do You Feel My Pain? Racial Group
234 Membership Modulates Empathic Neural Responses. *Journal of Neuroscience* **29**(26),
235 8525-8529. (doi:10.1523/jneurosci.2418-09.2009).
- 236 10. Golkar A., Selbing I., Flygare O., Ohman A., Olsson A. 2013 Other People as Means to a
237 Safe End: Vicarious Extinction Blocks the Return of Learned Fear. *Psychological Science*
238 **24**(11), 2182-2190. (doi:10.1177/0956797613489890).
- 239 11. Greenwald A.G., mcghee D.E., Schwartz J.L.K. 1998 Measuring individual differences in
240 implicit cognition: The implicit association test. *Journal of Personality and Social*
241 *Psychology* **74**(6), 1464-1480. (doi:10.1037/0022-3514.74.6.1464).
- 242 12. Akrami N., Ekehammar B., Araya T. 2000 Classical and modern racial prejudice: a study
243 of attitudes toward immigrants in Sweden. *European Journal of Social Psychology* **30**(4),
244 521-532. (doi:10.1002/1099-0992(200007/08)30:4<521::aid-ejsp5>3.0.co;2-n).
- 245 13. Bouton M.E. 2002 Context, ambiguity, and unlearning: Sources of relapse after behavioral
246 extinction. *Biological Psychiatry* **52**(10), 976-986. (doi:10.1016/s0006-3223(02)01546-9).
- 247 14. Haaker J., Golkar A., Hermans D., Lonsdorf T.B. 2014. A review on human reinstatement
248 studies: An overview and methodological challenges. *Learning and Memory* **15**;
249 21(9):424-40. doi: 10.1101/lm.036053.114.
- 250 15. Schiller D., Monfils M.H., Raio C.M., Johnson D.C., ledoux J.E., Phelps E.A. 2010
251 Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*
252 **463**(7277), 49-U51. (doi:10.1038/nature08637).
- 253 16. Molnar S. 1998 *Human Variation: Races, Types, and Ethnic Groups* Prentice Hall, ed. 4,
254 Upper Saddle River, NJ.
- 255 17. Cosmides L., Tooby J., Kurzban R. 2003 Perceptions of race. *Trends in Cognitive Sciences*
256 **7**(4), 173-179. (doi:10.1016/s1364-6613(03)00057-3).

- 257 18. Fiske S., Taylor S. 2008 *Social Cognition: From Brains to Culture* New York, mcgraw-
258 Hill.
- 259 19. Kinzler K.D., Corriveau K.H., Harris P.L. 2011 Children's selective trust in native-
260 accented speakers. *Developmental Science* **14**(1), 106-111. (doi:10.1111/j.1467-
261 7687.2010.00965.x).
- 262 20. Aboud F.E. 1988 *Children and prejudice*. New York, NY Blackwell.
- 263 21. Stanley D.A., Sokol-Hessner P., Banaji M.R., Phelps E.A. 2011 Implicit race attitudes
264 predict trustworthiness judgments and economic trust decisions. *Proceedings of the*
265 *National Academy of Sciences of the United States of America* **108**(19), 7710-7715.
266 (doi:10.1073/pnas.1014345108).
- 267 22. Knauff B.M. 1991 VIOLENCE AND SOCIALITY IN HUMAN-EVOLUTION. *Current*
268 *Anthropology* **32**(4), 391-428. (doi:10.1086/203975).
- 269 23. Tajfel H., Billig M.G., Bundy R.P., Flament C. 1971 Social categorization and intergroup
270 behavior. *European Journal of Social Psychology* **1**(2), 149-177.
271 (doi:10.1002/ejsp.2420010202).
- 272 24. De Waal F.B.M. 2008 Putting the altruism back into altruism: The evolution of empathy.
273 In *Annual Review of Psychology* (pp. 279-300).

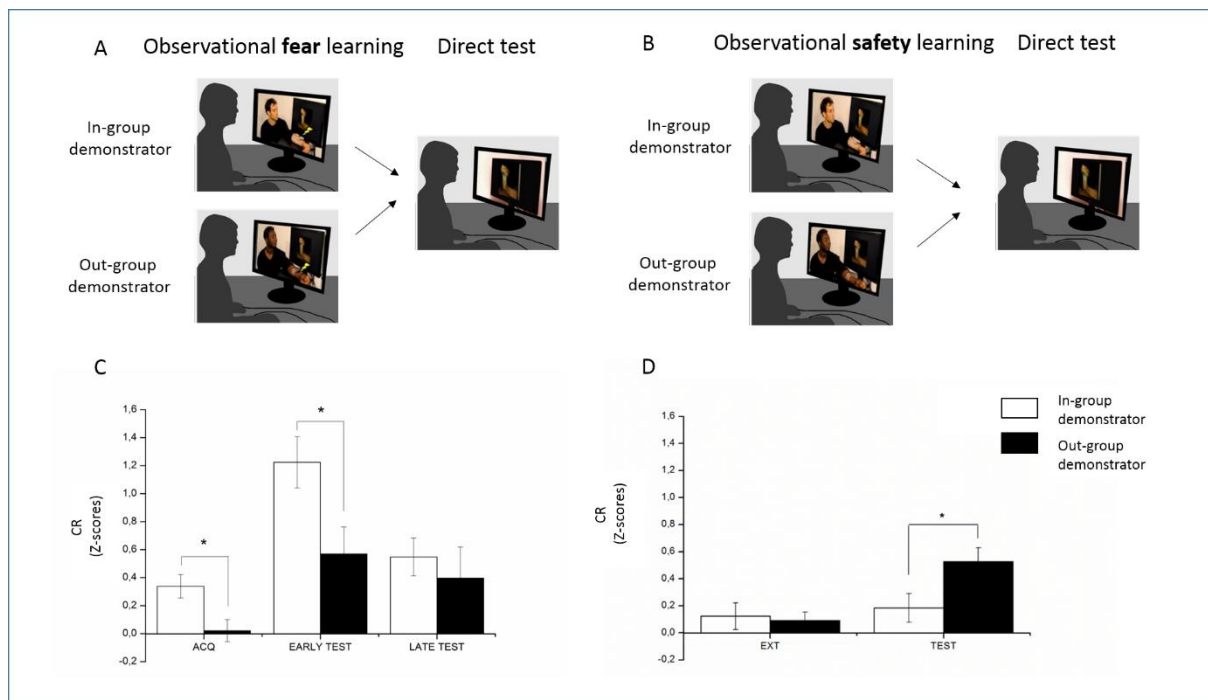
274

275

276

277

278



279

280 Figure. 1. Experimental design and results from Experiment 1 (left panels) and Experiment 2

281 (right panels). **A-B**. Both experiments were divided into an *observational stage* and a *direct test*

282 stage. Note that in Experiment 2, the observational extinction stage was preceded by a direct

283 acquisition stage (see Supplementary Figure 1). **C-D**. Mean CR (expressed as CS+/CS- SCR

284 difference) as a function of in and out-group demonstrator displayed for the observational and

285 direct test stage separately for Experiment 1 (left) and Experiment 2 (right). Given the

286 significant CS x Time x Group interaction during the test stage in Experiment 1 (but not in

287 Experiment 2), data is plotted separately for the early and late stage of the *direct test* in

288 Experiment 1. Error bars indicate standard error of the mean (SEM). Asterisks indicate

289 statistically significant differences ($p < .05$).

290

291

[Skriv här]