Neural and computational underpinnings of social threat learning

Björn Lindström[1,2], Jan Haaker[1], Andreas Olsson[1]

[1] Department of Clinical Neuroscience, Karolinska Institutet, Sweden [2] Department of Economics, University of Zurich, Switzerland

www.emotionlab.se

Introduction

- Learning about danger by observing others is ubiquitous in both humans and non-human animals [1].
- Little is known about the neural and computational underpinnings of such social threat learning.
- Previous research suggests both overlap and divergence between the neural mechanism involved in social and direct threat learning [1].
- We used fMRI and computational modeling to directly compare social and direct threat learning during fear conditioning.

Methods

- Direct and Observational fear conditioning within-subject (n=27).
- 2 conditioned stimuli (CS) per phase.
- 50% reinforcement (direct/observational shock)
- Contingency reversal after 24 (of 48) trials in both Direct and Observational phases
- Binary shock expectancy ratings

Direct Phase   Observational Phase

Reinforcement Learning (RL) Model

A Pearce-Hall/Rescorla-Wagner Hybrid model [2] was used to model learning and generate *associability*, and *prediction error* time-series:

$$Q_{i(t+1)} = Q_{i(t)} + a_{i(t)}\delta_{(t)}$$

Prediction error: $\delta_{(t)} = R_{(t)} - Q_{i(t)}$

Associability: $a_{i(t+1)} = \lambda \mid \delta_{(t)} \mid + (1-\lambda)a_{i(t)}$

fMRI Results: Direct/Observational overlap

Conjunction of shock (Direct/Observational). P(SVC) < .001 in bilateral insula and amygdala

Conjunction of CS+>CS- (Direct/Observational). P(SVC) = .06 left amygdala.

Conjunction of RL model-derived *associability* (Direct/Observational). P(SVC) < .01 in right insula and left amygdala (left insula: p < .06).

Behavioral Results

No behavioral differences between social and direct threat learning. ** = p < .01

fMRI Results: Direct/Observational divergence

(CS+>CS- Observational) > (CS+>CS- Direct). P < .001 uncorrected.

Direct > Social RL prediction error. P(SVC) = .09

Social > Direct RL prediction error. P < .001 uncorrected.

Conclusions

- The behavioral expressions of social and direct threat learning were comparable
- Social and direct threat learning involve overlapping neural and computational processes.
- Both observing and directly experiencing a shock involved the amygdala and insula.
- The amygdala represented the conditioned response in both social and direct threat learning.
- Both social and direct threat learning were supported by an associability signal in the bilateral anterior insula and the amygdala [2].
- Social threat learning involves the medial prefrontal cortex to a higher degree than direct threat learning [1].

References

1. Olsson, A. & Phelps, E. A. Social learning of fear. *Nat. Neurosci.* **10**, 1095–102 (2007).
2. Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A. & Daw, N. D. Differential roles of human striatum and amygdala in associative learning. *Nat. Neurosci.* **14**, 1250–2 (2011).