On Classical Conditioning

Again a recap of Yael Niv's Cosyne talk

A Tea Talk

by Kristin Völk, 18.03.2014

Which papers?

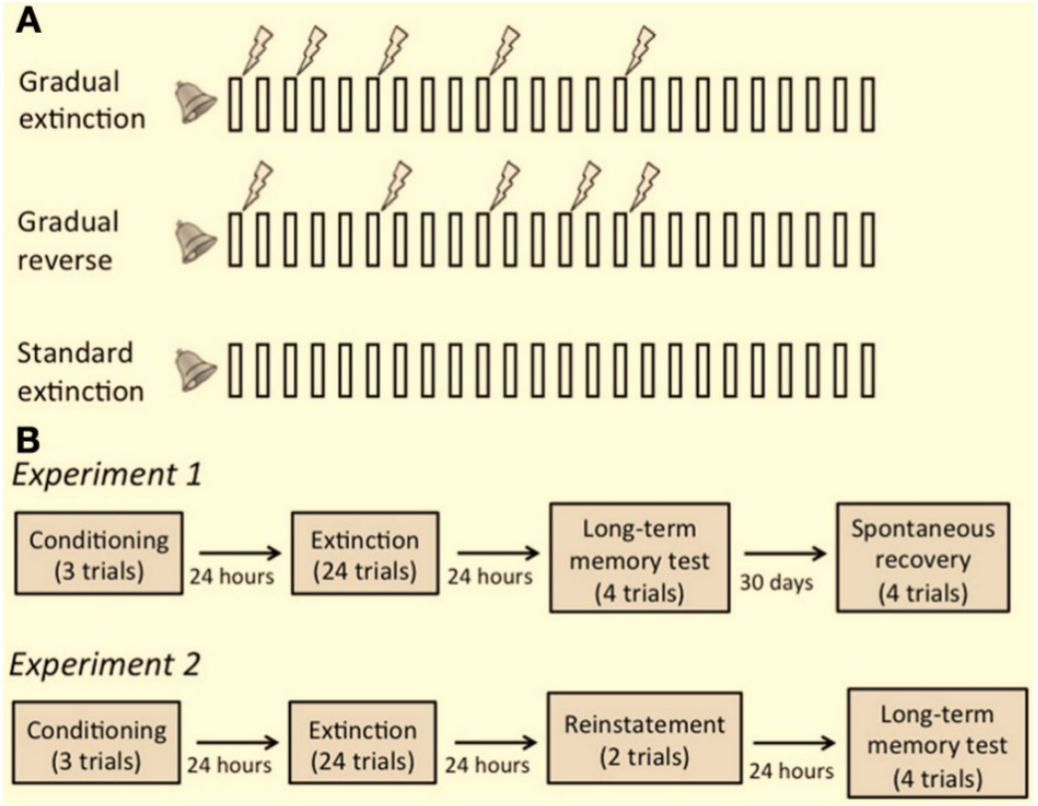
Gradual extinction prevents the return of fear: implications for the discovery of state

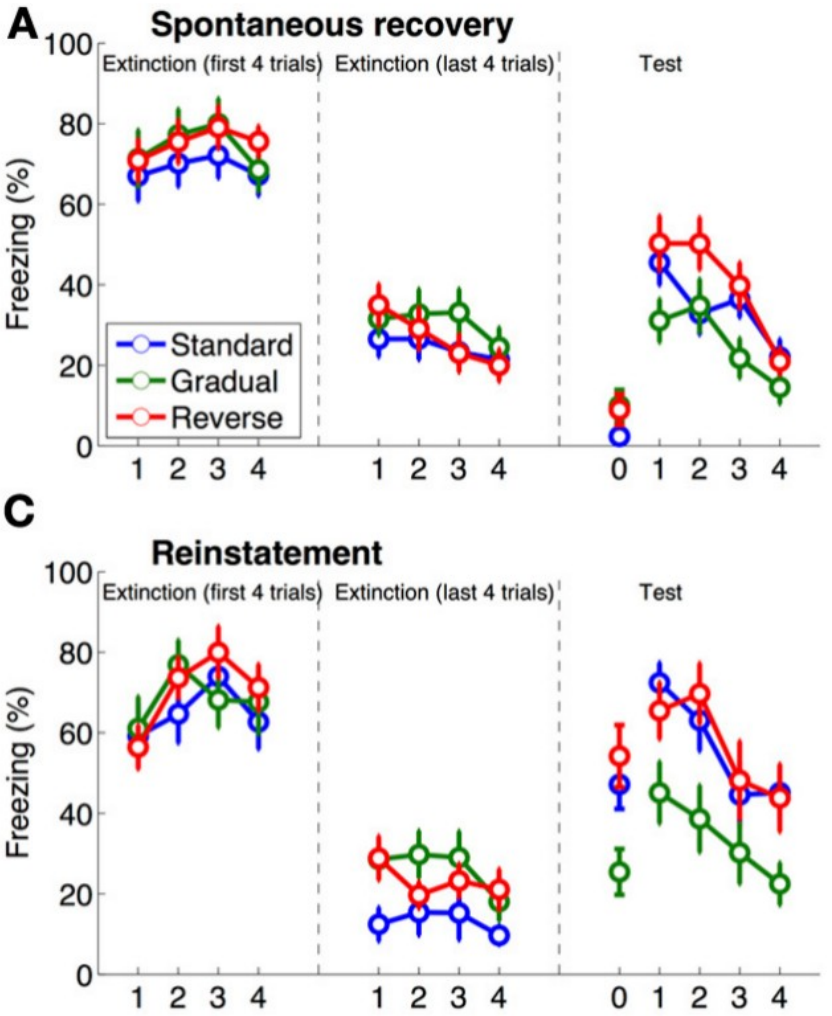
Samuel J. Gershman, Carolyn E. Jones, Kenneth A. Norman, Marie-H. Monfils and Yael Niv

Exploring a latent cause theory of classical conditioning

Samuel J. Gershman & Yael Niv

Experiments

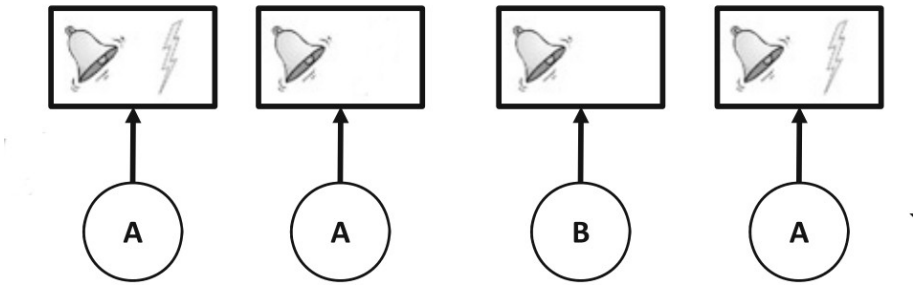




How to explain these results?

- Classical theories of conditioning (e.g. Rescola-Wagner) fail: prediction error → decrease in associability between CS and US → no reinstatement

- Latent cause theory: Animal infers latent cause of observations → constantly large prediction errors are signals for cause switching



The Theory

Assumptions on generative process:

1. Each trial is caused by one latent cause.
2. Each latent cause has some characteristic probability of emitting observed features (CS, US, etc.).
3. All else being equal, a prolific latent cause (i.e., one that has caused many trials) is more likely to cause another trial.
4. There is some small probability that the current trial results from a completely new latent cause (i.e., onethat has not yet generated any observations).

A short formalization

- Observations: $\mathbf{f}_t = \{f_{t,1}, \ldots, f_{t,D}\}$
- Mixture model:
  - First sample cause c_t from P(c)
  - Sample observations: P($\mathbf{f}$|c_t)
- Generate cause *k* on trial *t* with probability:

$$P(c_t = k|\mathbf{c}_{1:t-1}) = \begin{cases} \frac{N_k}{t-1+a} & \text{if } k \text{ is an old cause} \\ \frac{a}{t-1+a} & \text{if } k \text{ is an new cause} \end{cases}$$

- with $\mathbf{c}_{1:t} = \{c_1, \ldots, c_t\}$

A short formalization

- Animals belief about latent causes:

$$P(\mathbf{c}_{1:t}|\mathbf{F}_{1:t}) = \frac{P(\mathbf{F}_{1:t}|\mathbf{c}_{1:t})P(\mathbf{c}_{1:t})}{\sum_{\mathbf{c}1:t} P(\mathbf{F}_{1:t}|\mathbf{c}_{1:t})P(\mathbf{c}_{1:t})} \qquad \mathbf{F}_{1:t} = \{\mathbf{f}_1, \ldots, \mathbf{f}_t\}$$

- Predicted probability of US (CR)

$$\begin{aligned} V_t &= P\left(f_{t,1} = \mathrm{US}|\mathbf{f}_{t,2:D}, \mathbf{F}_{1:t-1}\right) \\ &= \sum_{\mathbf{c}_{1:t}} P\left(f_{t,1} = \mathrm{US}|c_t, \mathbf{c}_{1:t-1}, \mathbf{f}_{1:t-1,1}\right) P\left(c_t|\mathbf{f}_{t,2:D}, \mathbf{F}_{1:t-1,2:D}, \mathbf{c}_{1:t-1}\right) \\ &\quad \times P\left(\mathbf{c}_{1:t-1}|\mathbf{F}_{1:t-1}\right). \end{aligned}$$

Take-Home-Message

- Classical theories cannot account for experimental results
- Latent cause theory offers an alternative explanation, but needs to be expanded (no exchangeability of observation order)
- Ideas on how to eliminate context dependent anxieties