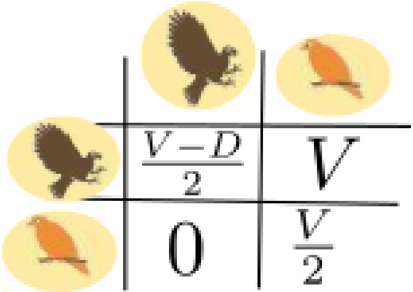
The evolution of badges of status with learners

Andrés Quiñones

The Hawk-Dove game

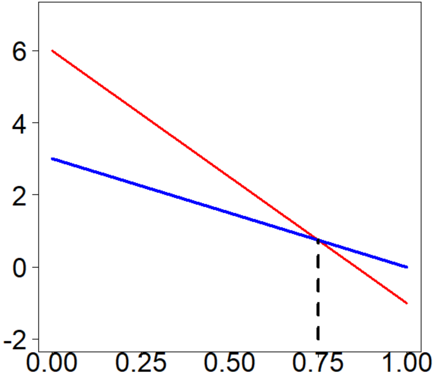
Individuals have one of two genetically determined phenotypic strategies. *Hawks* are willing to start a conflict over resources, while *doves* prefer to stand down in the hope to share the resource without an aggressive contest.



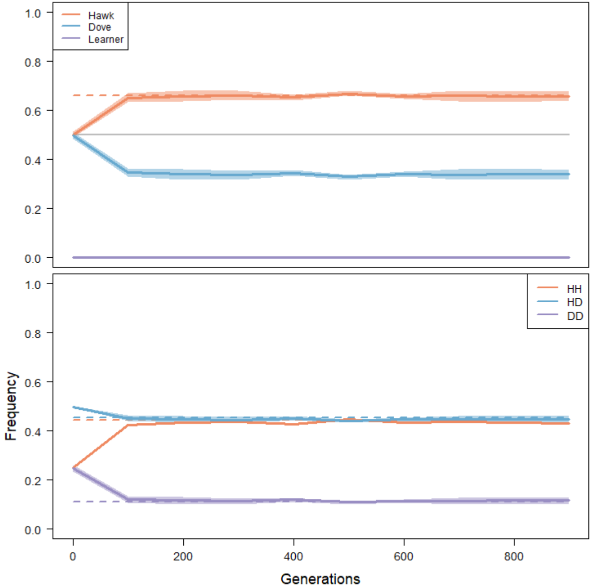
The hawk-dove game

$$w_H = p_H \frac{V - C}{2} + (1 - p_h) V$$

$$w_D = p_H 0 + (1 - p_H) \frac{V}{2}$$



The hawk-dove game



What about signals?

What about signals?

When are signals honest?

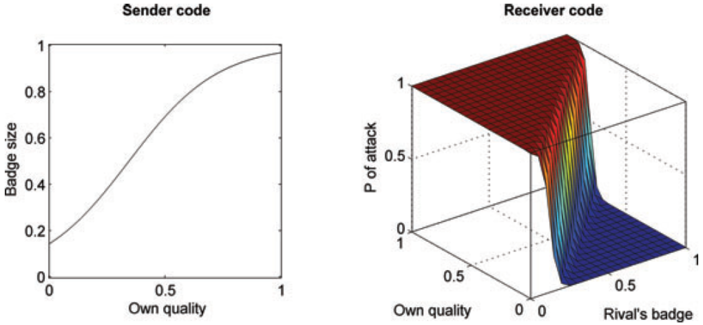
- Impossible to fake
- Individuals have common interests
- Handicap principle (signal´s cost is proportional to quality)
  - Social costs?

Social costs are an underappreciated force for honest signalling in animal aggregations

Michael S. Webster [a, b, *], Russell A. Ligon [a, b], Gavin M. Leighton [a, b]

[a] *Department of Neurobiology and Behavior, Cornell University, Ithaca, NY, USA*
[b] *Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA*

What about signals?



What about learning?

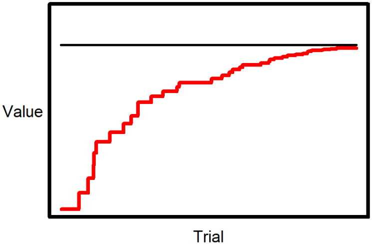
Associative learning

Reinforcement learning theory

$$\Delta V_{t(s)} = \alpha \underbrace{(R_t - V_t)}_{\text{prediction error}}$$

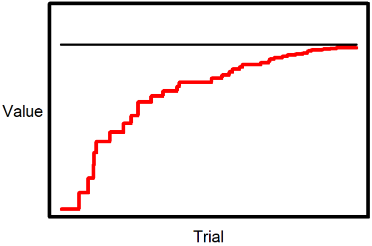
Reinforcement learning theory

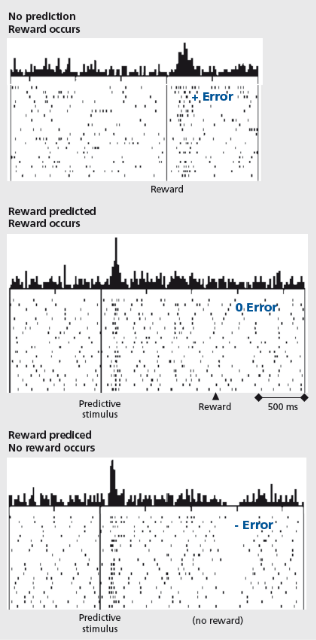
$$\Delta V_{t(s)} = \alpha \underbrace{(R_t - V_t)}_{\text{prediction error}}$$



Reinforcement learning theory

$$\Delta V_{t(s)} = \alpha \underbrace{(R_t - V_t)}_{\text{prediction error}}$$





Reinforcement learning theory

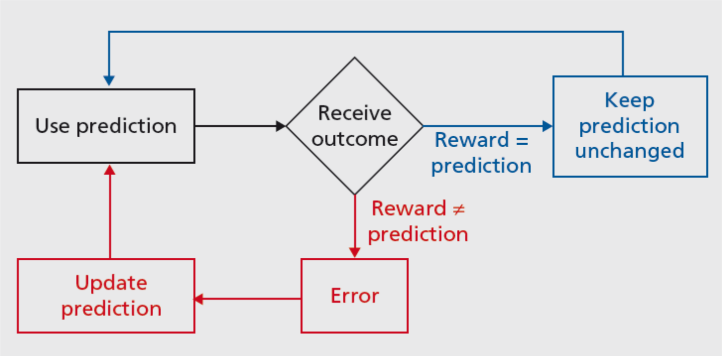
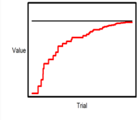
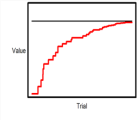


Figure 1. Scheme of learning by prediction error. Red: a prediction error exists when the reward differs from its prediction. Blue: no error exists when the outcome matches the prediction, and the behavior remains unchanged.

Environmental states

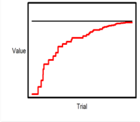
Discrete states

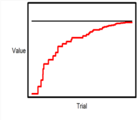




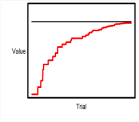
Environmental states

Discrete states



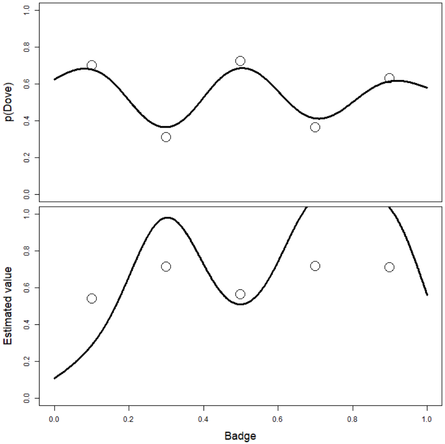


Continuos states



Continuos environmental states

$$\Delta V_{t(s)} = \alpha \underbrace{(R_t - V_t)}_{\text{prediction error}}$$

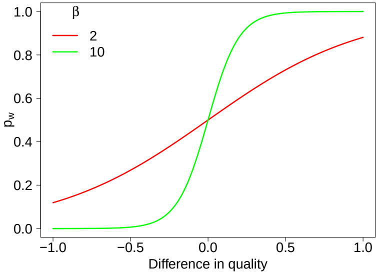


Our model

The hawk-dove game 2.0

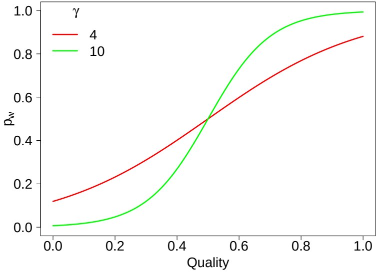
	H	D
H	$p_w V \frac{-C}{2} + (1 - p_w)\frac{-C}{2}$	V
V	0	$\frac{V}{2}$

$$p_w = \frac{1}{1 + e^{-\beta(Q_i - Q_j)}}$$



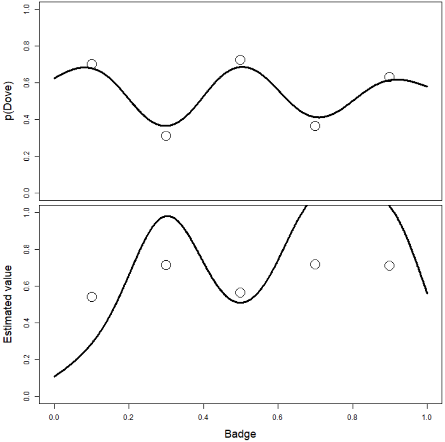
Sender Code

$$s_i = \frac{1}{1 + e^{-(\epsilon_i + \gamma_i Q_i)}}$$



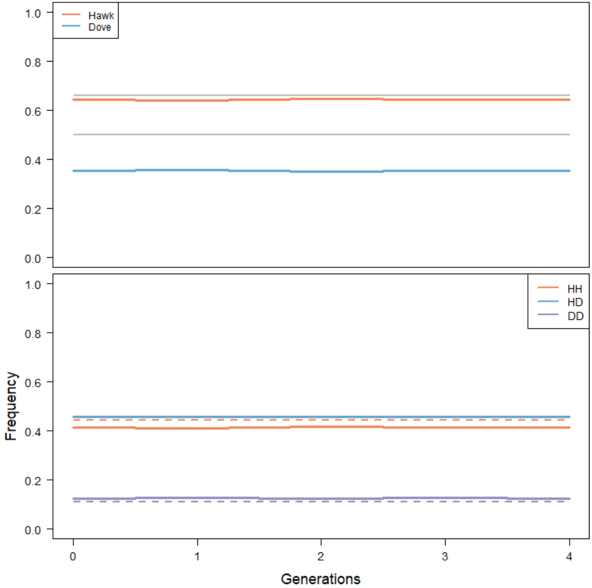
Receiver code

$$\Delta V_{t(s)} = \alpha(R_t - V_t)$$

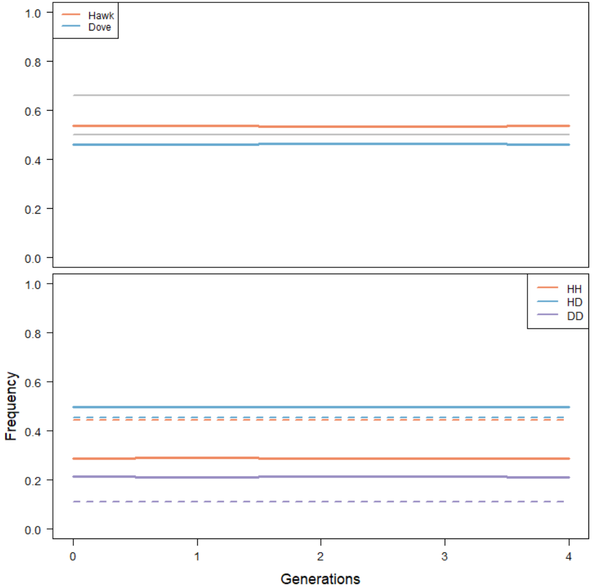


Results

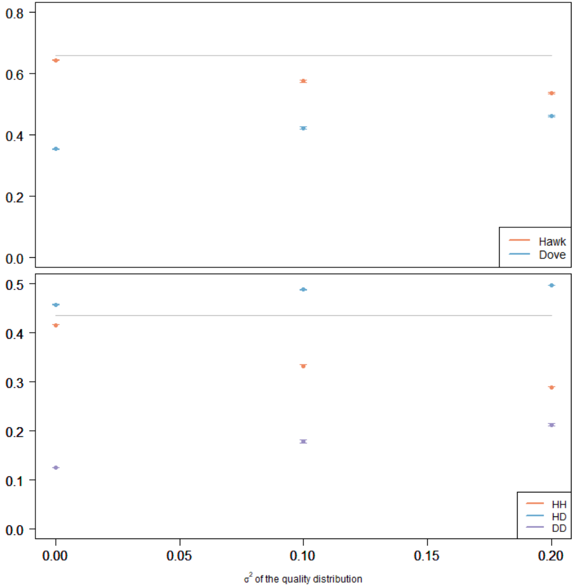
When individuals do *NOT* vary in quality



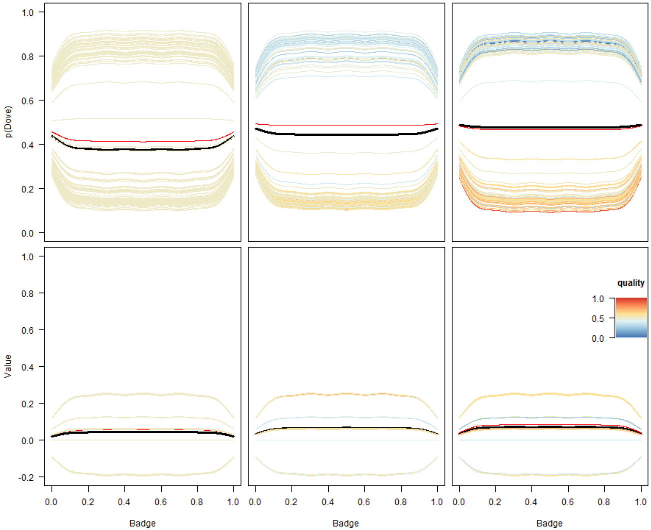
When individuals *DO* vary in quality



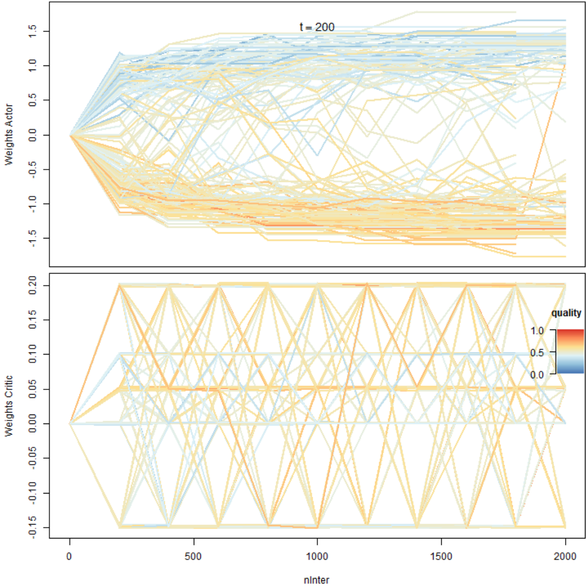
Overal effect of quality variation



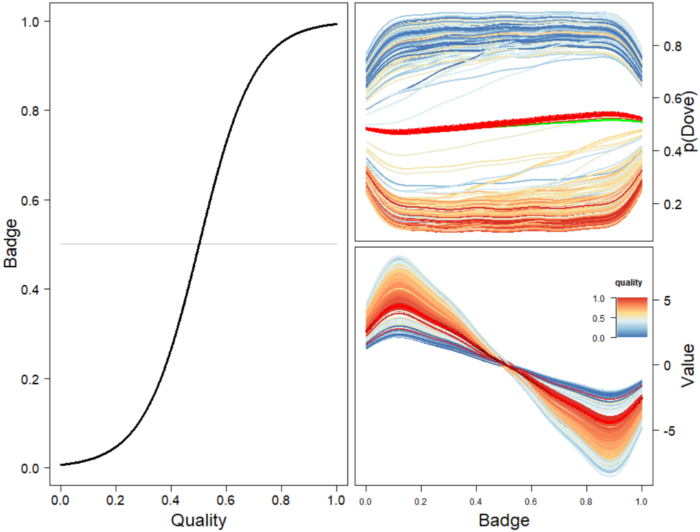
How does variation among *learners* look like?



How do the learning dynamics look like?

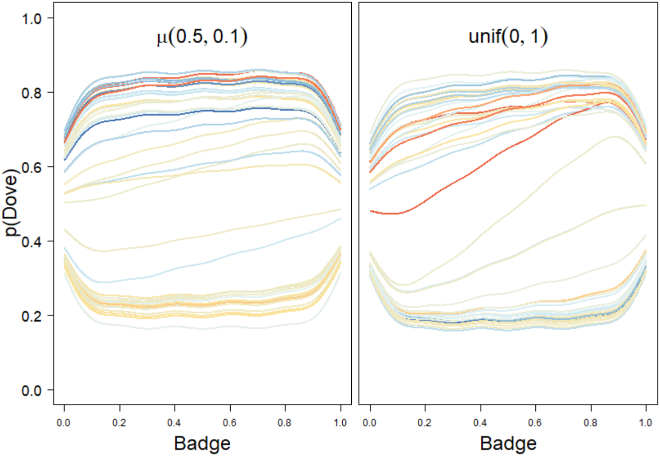


How do learners behave when signals are honest?



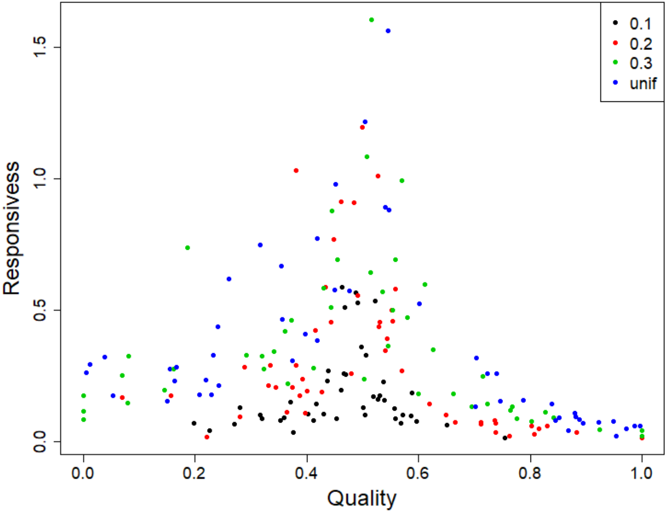
What parameters influence responsiveness?

Distribution of quality



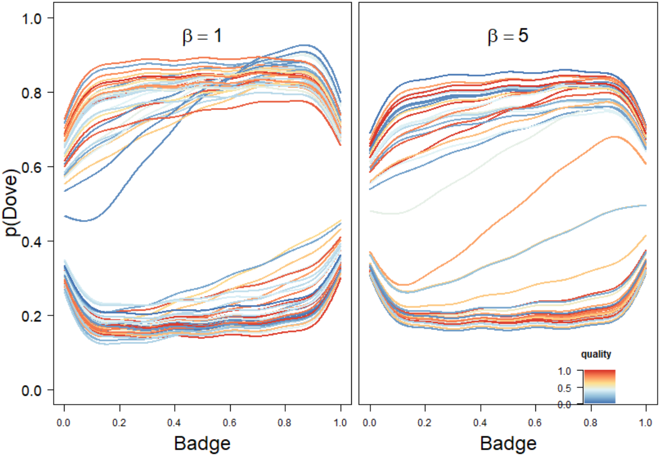
What parameters influence responsiveness?

Distribution of quality



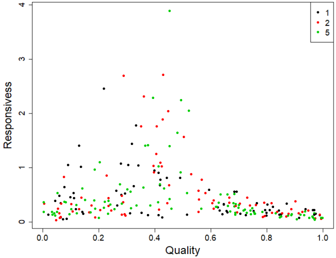
What parameters influence responsiveness?

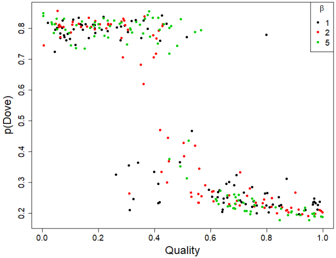
Strength of quality influence on competition outcome



What parameters influence responsiveness?

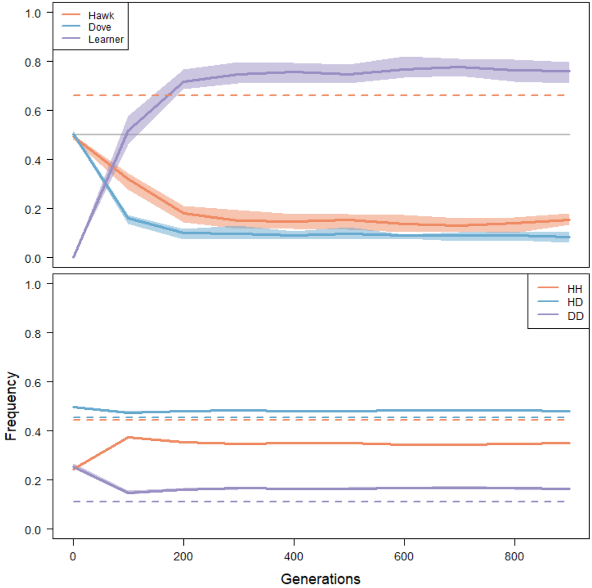
Strength of quality influence on competition outcome





Why the hell is responsiveness so low?

How do learners fare against pure strategies?



Concluding

- Learning is good, even if you do not learn about your opponent. You still learn about you
- Learning reduces aggressive interactions
- Learning splits up individuals into *types*

What's next?

- Let reaction norm evolve, under different initial conditions.
- Let learning parmeters evolve
- Let the communication system co-evolve