Draft: Learned badges vs. Social recognition

Eleanor Brush[1] and Elizabeth A. Hobson[2,3,4]

[1] *University of Maryland; email: xxxxx*
[2] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501 USA*
[3] *Center for Biosocial Complex Systems, Arizona State University*
[4] *National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, TN, USA; email: ehobson@santafe.edu*

2016-09-08

Target journal: Behavioral Ecology

Target submission date: October?

Abstract

to be added..

Keywords: xxxxxxxx

Overview of previous research, outstanding questions, goals

Species have different ways of assessing the quality of others. Quality here is used broadly, and can indicate an individual's body condition or immune status, overall reproductive potential, or fitness. Individuals can gain long-term benefits from assessing each others' quality. In the context of conflicts when individuals differ in fighting ability, resource-holding potential, or dominance rank, it may be beneficial for individuals to be able to identify and order individuals by quality.

...Lots of investigation into evolution of 'honest' signaling...

Species have different ways of assessing the quality of others. Here, we focus on one type of signal – the badge of status – and compare it to a system based on individual recognition. A badge is an arbitrary signal, not directly indicative of fighting ability, where the intensity of the signal correlates with the underlying quality of the individual (here, rank/power) (CITE). Another quality assessment system is a combination recognition/relationship method, where individuals that are able to recognize individuals track their actions and event outcomes to infer quality in the absence of a cue or badge of status (CITE). In both cases, individuals must learn about their opponents in order to assess their quality.

At the two extremes, individuals in a pure badge quality signal system would be completely unable to recognize individuals, and would treat all individuals with similar badge intensities as equivalent; individuals in a pure recognition/relationships based quality system would be completely reliant on remembering an individual's actions, and have no other cues to infer that individual's quality or rank. Binning individuals with similar badge intensities into the same category effectively decreases the functional group size to the number of bins, but also increases noise about quality of each individual. An individual recognition-based system is one in which the number of bins equals the number of individuals in the group.

Justification and Goals

[this is a bit rough] Many factors likely affect the accuracy and speed of quality assessment in these two systems. Recent work ([2]) provides several predictions but focused on predicting differences in signaling systems comparing an innate quality signal to a social recognition system across different species, not within species ([1]). Here, we use a somewhat different approach: we are interested in two hypothetical systems, where individuals learn to assess quality to others in different ways.

Our goal is to better understand the conditions under which a badge signal or individual recognition would be favored. We simulated quality assessment through fights among individuals. Each time individuals fight, they are able to assess each other's quality and update their previous opinion about quality. We use two quality assessment methods: (1) Badge signal systems: individuals group others into quality categories based on the intensity of badge signals and (2) Recognition systems: individuals use individual recognition to remember the outcome of events with particular individuals to estimate quality.

Methods

Model

Each animal in the social group has an inherent quality value. This can be thought of, for example, as fighting ability [], resource holding potential [], or body size. An animal can learn about another's quality value by interacting with it. The animals can also learn by observing the interactions between other pairs. We consider two different ways of learning: individual recognition and badges. An animal using individual recognition can identify each of its group mates. An animal using the badge system only recognizes animals based on the signals they display. It is costly to the animals in the group to learn about each other inaccurately. Since animals can use information about their group mates to decide how to interact with them in the future, having the wrong information can lead to an inappropriate choice. It is also costly to learn slowly. The time spent interacting and learning about one's group mates could be spent in other ways. Additionally, if the interaction is actually a fight between animals, having more interactions than necessary can lead to injury and perhaps even death. Finally, having improved memory and perceptive ability can also be costly because of the associated energetic costs and or because improving these properties involves a tradeoff that negatively affects other traits. We combine the costs associated with inaccurate learning, the time required to learn, and cognitive abilities to assess the performance of animals using each of the two systems.

Interactions and learning

Specifically, each animal has a quality value, q_i. Each animal also displays a signal, s_i, which is always visible. We define the parameter ρ as the correlation between the signals, $\{s_i\}$, and the quality values, $\{q_i\}$. In a group with N animals, we draw quality values $\{q_1, \ldots, q_N\}$ from a normal distribution with mean 0 and standard deviation σ_{q}. We then generate N signal values such that $\min_i s_i \approx -1$, $\max_i s_i \approx 1$, and the correlation between $\{q_i\}$ and $\{s_i\}$ is precisely ρ. (A list of the variables used in the text is provided in Table 1.)

Each animal assesses the quality of each other animal: $a_{ij}(t)$ is the opinion of animal i about animal j at time t. At first, the group consists entirely of naive animals: at $t = 0$, no animal has an opinion of any other. Every animal has a memory window of length w. At each point in time, first, if animal k has not updated its opinion of animal ℓ within the last w timesteps, it forgets its opinion of ℓ. Next, two animals, i and j, are chosen randomly to interact. If i already has an opinion of j, its updated opinion is

$$a_{ij}(t) = (1 - \ell_{\mathsf{i}})a_{ij}(t-1) + \ell_{\mathsf{i}} q_j + \xi,$$

where ℓ_{i} is a parameter describing how much i changes its opinion based on the interaction and ξ is drawn from a normal distribution with mean 0 and standard deviation σ_{i}. If i and j have not previously interacted or i has forgotten its opinion of j, then after interacting

$$a_{ij}(t+1) = (1 - r)b_i(t) + r q_j + \xi,$$

where $b_i(t)$ is a baseline opinion of any animal it encounters. Specifically, $b_i(t)$ is drawn from a normal distribution with mean 0 (the expected quality value of any animal) and standard deviation σ_{b}. The other animals in the group observe the interaction with probability p_{o}. If animal k observes the interaction, it uses the same rules to update its opinion of both i and j, with parameters ℓ_{o} and σ_{o}. We use $\ell_{\mathsf{o}} \leq \ell_{\mathsf{i}}$ and $\sigma_{\mathsf{o}} \geq \sigma_{\mathsf{i}}$ so that observational learning is noisier. This completes the description of how learning occurs for animals using individual recognition.

Badge system

An animal that uses the badge system to learn cannot perceive the difference between animals with similar signals and is forced to make a noisy estimate of the quality values of all such animals. Specifically, an animal divides the rest of its group into categories before any interactions take place. It does so by picking another animal, j, at random. All animals whose signals are within $\delta/2$ of s_j are put in the same category. Then the focal animal picks an uncategorized animal at random, k, forming a second category of all uncategorized animals whose signals are within $\delta/2$ of s_k. The process continues until every animal in the group has been assigned a category. Category width δ depends on the animals' perceptive abilities. For example, if δ is large, an animal might only be able to identify animals with small, medium, and large signals, whereas if $\delta = 0$ an animal can identify every individual in its group and using the badge is no different than using individual recognition. Different animals will categorize the group differently and may perceive different numbers of categories. Figure 6 shows one example of this process. The average number of categories animals perceive decreases with group size and category width, δ (Figure 7). When an animal using the badge system updates its opinion of another animal based on direct interaction or on observation, it simultaneously updates its opinions of all the other animals in the same category. If an animal observes an interaction between two animals in the same category, it does not updates its opinion of that category.

Measures of performance

The animals interact and learn about each other for T timesteps. To measure each animal's learning ability, we average the errors it makes in its opinions about the rest of the group:

$$\epsilon_i(T) = \frac{1}{|\mathscr{O}_i(T)|} \sum_{j \in \mathscr{O}_i(T)} |a_{ij}(T) - q_j|,$$

where $\mathscr{O}_i(T)$ is the set of animals about whom i has an opinion at time T. We also calculate the average time it takes for an animal's error about each other animal to drop below a threshold:

$$\tau_i = \frac{1}{|\mathscr{O}_i|} \sum_{j \in \mathscr{O}_i} \min_{t=1,\ldots,T} \{t : \epsilon_{ij}(t) \leq 0.2\}.$$

(If ϵ_{ij} is always greater than the threshold, then i's learning time about j is taken to be T.) Figure 8 shows one example of this calculation. We model 25 groups of animals using each of the two learning systems and calculate the average error $\bar{\epsilon}$ and average learning time $\bar{\tau}$ of all animals in all groups. In Figure 1 we show how the average error of animals using each of the two systems changes over time.

Learning errors and learning time can both be costly. Improving perceptive ability (by decreasing δ) and improving memory (by increasing w) can also be costly. We describe the cost of the category width with the function

$$c_\delta = \left(\frac{2-\delta}{2}\right)^\alpha$$

and that of memory with the function

$$c_w = \left(\frac{w}{2000}\right)^\alpha,$$

where α is a parameter determining the concavity of the cost function. The animals pay no costs for having no cognitive ability ($\delta = 2$ and $w = 0$), and 1 unit of cost for optimal cognitive ability ($\delta = 0$ and $w = 2000$, the highest w we consider). (Figure 9 shows these cost functions.) We then combine the four sources of cost—error, time, memory window, and category width—into the total cost

$$C = 2\bar{\epsilon} + 10^{-4}\bar{\tau} + c_w + c_\delta,$$

where the terms in front of $\bar{\epsilon}$ and $\bar{\tau}$ ensure that the four factors, $\bar{\epsilon}$, $\bar{\tau}$, c_w and c_δ, are on the same scale.

Table 1: Description of variables

Variables	Description of variables
C	total cost of learning
δ	category width
ϵ_i	average error of animal i about all other animals
$\bar{\epsilon}$	average error of all animals in all groups
ℓ_i	learning rate in direct interaction
ℓ_o	learning rate in indirect observation
N	number of animals in group
p_o	probability of observing interactions between other pairs
q_i	quality of animal i
ρ	correlation between quality and signal
s_i	signal of quality / badge of status of animal i
σ_b	standard deviation of baseline opinion
σ_i	standard deviation of noise in opinion updating during interaction
σ_o	standard deviation of noise in opinion updating during observation
σ_q	standard deviation of quality values
T	total number of fights
τ_i	average learning time of animal i
$\bar{\tau}$	average learning time of all animals in all groups
w	memory window

Results

The effects of group size and cognitive abilities on learning.

We first analyze the model without any observational learning ($p_\mathrm{o} = 0$). The effects of group size N, memory window w, and signal-quality correlation ρ on error and learning time are intuitive and confirm that our model is a reasonable description of a real social system. For animals using both learning systems, smaller groups and longer memory windows make it easier to learn. As group size decreases and memory window increases, both error and learning time decrease (Figure 2). Animals using the badge system learn more accurately and more quickly when the correlation between the signal and quality increases (Figure 2).

While changing any of these first three parameters (group size, memory window, and correlation) will either increasing or decreasing both error and learning time, changing category width leads to a tradeoff between error and learning time, increasing one while decreasing the other. Increasing category width results in low learning time and high error (Figure 2). When a focal animal lumps many of its peers together, it updates its opinion about each of those animals more frequently and by chance may at any given point in time have an accurate opinion about one of the animals in the category, leading to a lower average learning time. On the other hand, by lumping many animals into the same category, it loses its ability to accurately learn about their individual quality values, leading to a higher average error. When the signal-quality correlation is poor, decreasing category width decreases error, by allowing animals to distinguish between their peers more easily. However, when the signal is highly correlated with quality, an intermediate category width minimizes error (Figure 2). When the signal is a good indicator of quality, it is advantageous to lump animals with similar signals together because it makes sense to transfer information about one of their quality values to other animals in the category. This effect is weakened when the memory window is very long because in that case one can simply learn about all the other animals without a chance of forgetting what has been learned (Figure 10).

The effect of observational learning.

Observational learning helps animals using both systems to learn more accurately and more quickly, but it helps animals using individual recognition much more (Figure 3). Observational learning increases the number of animals about which a focal animal can learn at any given time. In fact, for animals using individual recognition, the same level of error can be achieved by increasing the probability of observing (p_{o}), by increasing memory window (w), or by increasing both to a lesser degree. The same is true of learning time. This can be seen in the strong interaction in how error and learning time change as a function of the probability of observing (p_{o}) and memory window (w) (Figure 3).

Cost of learning.

Since increasing group size increases both error and learning time (Figure 2), it increases the overall cost of using both learning systems (Figure 4). The effects of memory window (w) and category width (δ) on overall cost depend on the cost function, specifically on the parameter α (Figure 4). When $\alpha < 1$, the costs of memory window (c_w) and category width (c_δ) increase most quickly at low cognitive ability; when $\alpha > 1$, the costs of memory window and category width increase most quickly at high cognitive ability (Figure 9).

When the costs of the memory window increase quickly at low w ($\alpha < 1$), the rapid increase in the cost of using a larger memory window is almost exactly offset by the decrease in error and learning time, so the costs of using both systems are almost a constant function of memory window (Figure 4). However, when the cost of memory window increases quickly at high w ($\alpha > 1$), overall costs initially decrease as w increases because of the improvement in error and learning time and only eventually increase because of the cost of using a higher memory window, so an intermediate memory window is the least costly (Figure 4).

Using a category width δ equal to 0 is always costly because it leads to high error and learning time, as we saw above (Figure 2). When the costs of category width increase quickly as δ decreases from 2 ($\alpha < 1$), overall costs tend to increase as δ decreases, but there is a large range of intermediate δ at which overall costs are nearly constant: in this range, the increase in cost from decreasing δ is offset by a decrease in cost from reduced error and learning time. When the costs of category width increase only at low δ ($\alpha > 1$), at first decreasing δ decreases overall cost by improving error and learning time and only at low δ does overall cost increase because of the higher cost of this cognitive ability: an intermediate category width is favored (Figure 4).

The overall costs of using the badge system are higher than the costs of using individual recognition when group size is low and memory window is high (Figure 5A). In these cases, animals using individual recognition are capable of learning accurately about their peers and there is no benefit from grouping animals together into categories. When the signal-quality correlation is high, individual recognition is only better for very small groups and very high memory windows (Figure 5C).

At an intermediate group size ($N = 50$), animals using individual recognition perform similarly to or better than animals using the badge system when category width is low and memory window is high (Figure 5B). In these cases, animals using the badge system are paying high costs for their cognitive abilities without receiving much benefit from either improved error or improved learning time. When the signal-quality correlation is high, individual recognition is more costly except when category width is equal to 0 and the two systems are equivalent (Figure 5D).

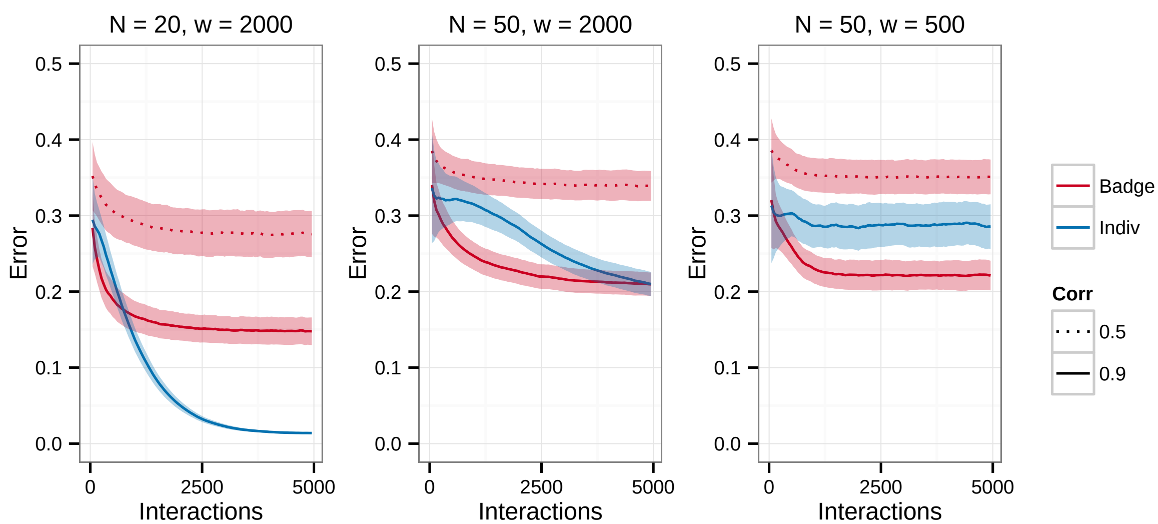

Figure 1: **Examples of how average error decreases over time.** In each panel, we show time in number of interactions on the x-axis and average error, $\bar{\epsilon}$, on the y-axis. The line shows the average error across all animals in all groups and the shaded area shows this average ± 0.5 the standard deviation of error across all animals in all groups. In each panel, blue lines show animals using individual recognition and red lines show animals using the badge system. The solid lines correspond to groups where signal-quality correlation $\rho = 0.9$ and the dotted lines correspond to groups where signal-quality correlation $\rho = 0.5$. As groups size increases, from A to B, the error of animals using both systems increases, but it affects the error of animals using individual recognition more. As memory window decreases, from B to C, the error of animals using the badge system is not strongly affected, whereas the error of animals using individual recognition becomes higher. Parameters: in all panels $\delta = 0.5$, $\ell_{\mathrm{i}} = 0.2$, $p_{\mathrm{o}} = 0$, $\sigma_{\mathrm{b}} = 0.2$, $\sigma_{\mathrm{i}} = 0.01$, $\sigma_{\mathrm{q}} = 0.5$, $T = 10000$; in A $N = 20$, $w = 2000$; in B $N = 50$, $w = 2000$; in C $N = 50$, $w = 500$.

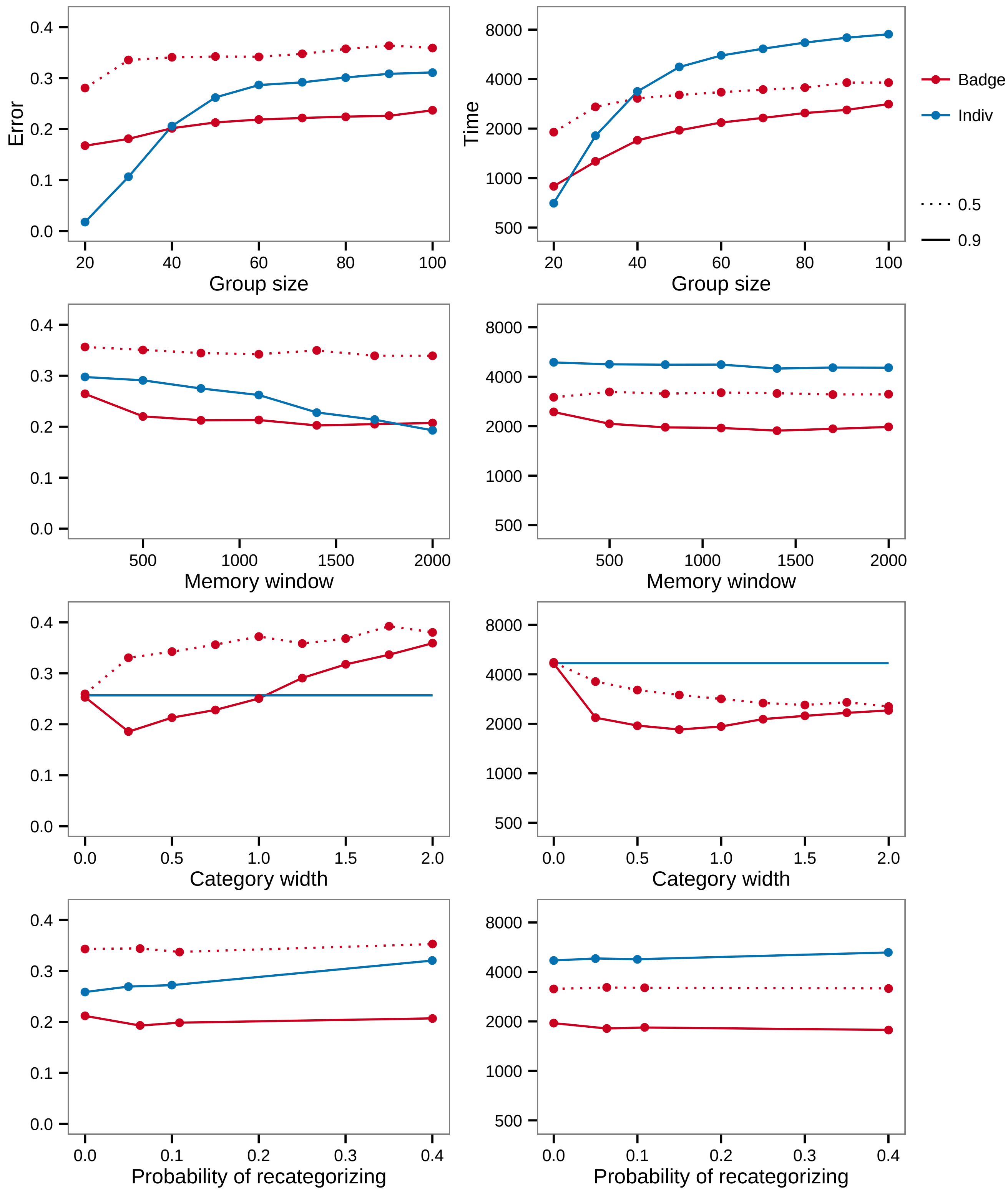

Figure 2: **Increasing group size and decreasing memory window both improve learning.** Error and learning time are minimized at an intermediate category width. In the left column, we show average error $\bar{\epsilon}$ as a function of various parameters, and in the right column, we show average learning time $\bar{\tau}$ as a function of various parameters. In each panel, blue lines show animals using individual recognition and red lines show animals using the badge system. The solid red lines correspond to groups in which the signal-quality correlation $\rho = 0.9$ and the dotted lines correspond to groups in which the signal-quality correlation $\rho = 0.5$. Parameters: unless the parameter is being varied $\delta = 0.5$, $\ell_\mathrm{i} = 0.2$, $N = 50$, $p_\mathrm{o} = 0$, $\sigma_\mathrm{b} = 0.2$, $\sigma_\mathrm{i} = 0.01$, $\sigma_\mathrm{q} = 0.5$, $T = 10000$, $w = 1100$.

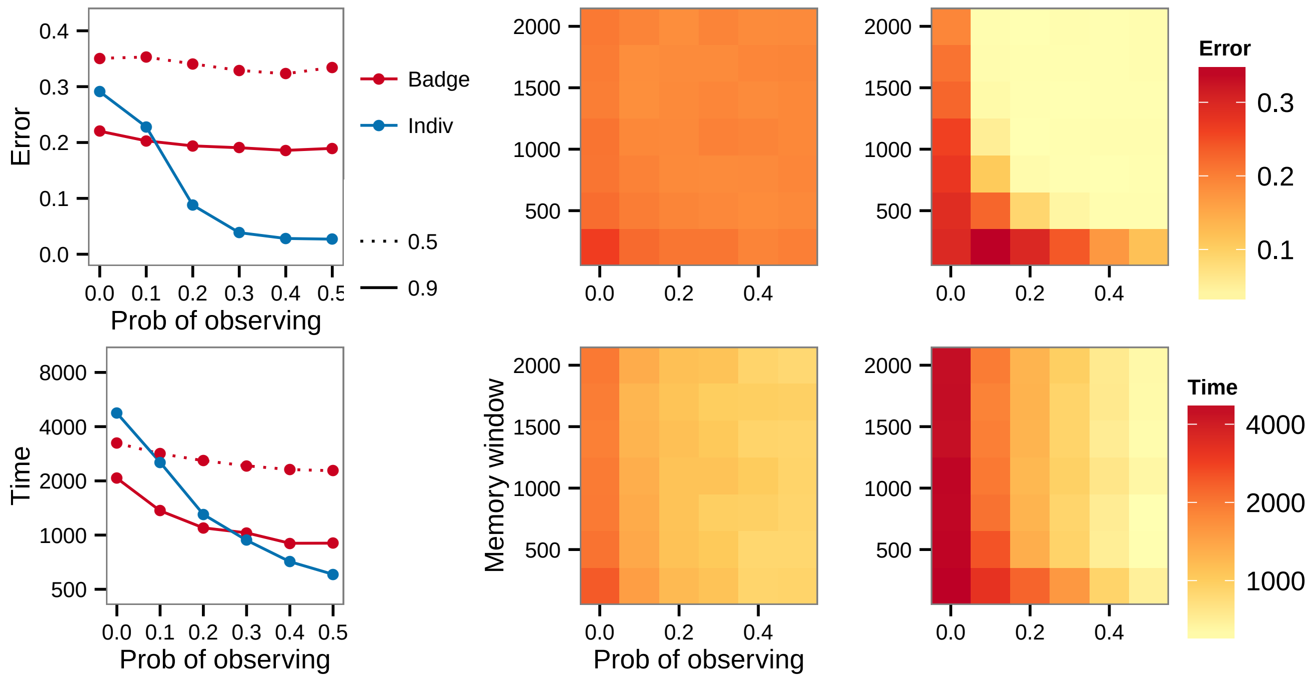

Figure 3: **Observational learning improves both error and learning time and does so much more for animals using individual recognition.** For animals using individual recognition, there is a strong interaction between the probability of observing and memory window. In A and D, we show average error $\bar{\epsilon}$ and average learning time $\bar{\tau}$ respectively as a function of the probability of observing, p_{o}. In each panel, blue lines show animals using individual recognition and red lines show animals using the badge system. The solid red lines correspond to groups in which the signal-quality correlation $\rho = 0.9$ and the dotted lines correspond to groups in which the signal-quality correlation $\rho = 0.5$. In B and C we show average error $\bar{\epsilon}$ of animals using the badge system and individual recognition respectively as a function of both the probability of observing p_{o} on the horizontal axis and memory window w on the vertical axis. In E and F we show average learning time $\bar{\tau}$ of animals using the badge system and individual recognition respectively as a function of both the probability of observing p_{o} on the horizontal axis and memory window w on the vertical axis. Parameters: in all panels $\delta = 0.5$, $\ell_{\text{i}} = 0.2$, $\ell_{\text{o}} = 0.2$, $N = 50$, $\sigma_{\text{b}} = 0.2$, $\sigma_{\text{i}} = 0.01$, $\sigma_{\text{o}} = 0.01$, $\sigma_{\text{q}} = 0.5$, $\rho = 0.9$, $T = 10000$; in A and D $w = 1100$, in B and E $\rho = 0.9$.

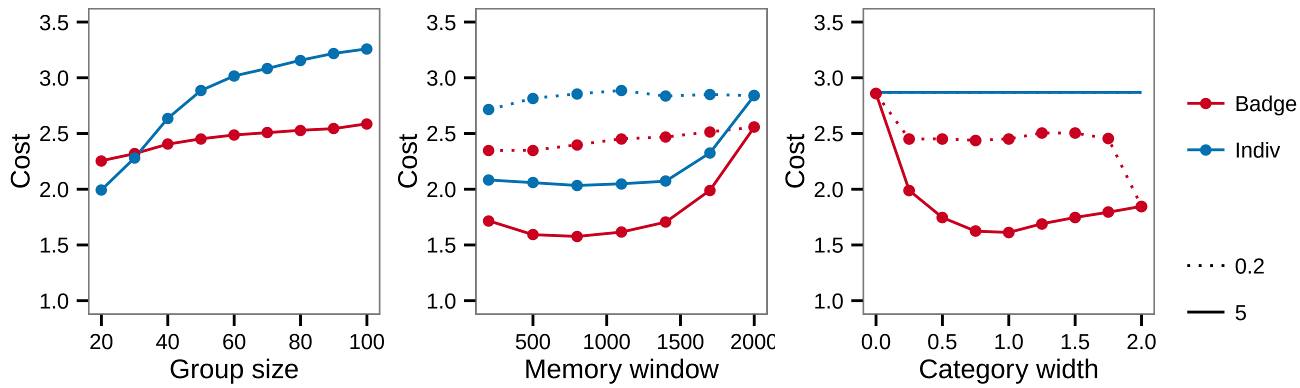

Figure 4: **Overall costs are lowest in small groups and with intermediate memory window and category width, when costs increases at high cognitive abilities ($\alpha > 1$).** When costs increase at low cognitive abilities ($\alpha < 1$), memory window w does not affect overall costs and large category width δ minimizes overall costs. In A,B, and C, we show overall costs C as a function of group size N, memory window w, and category width δ respectively. In each panel, blue lines show animals using individual recognition and red lines show animals using the badge system. The solid lines correspond to $\alpha = 5$ and the dotted lines correspond to $\alpha = 0.2$. Parameters: unless the parameter is being varied $\delta = 0.5$, $\ell_{\mathsf{i}} = 0.2$, $N = 50$, $p_{\mathsf{o}} = 0$, $\rho = 0.9$, $\sigma_{\mathsf{b}} = 0.2$, $\sigma_{\mathsf{i}} = 0.01$, $\sigma_{\mathsf{q}} = 0.5$, $T = 10000$, $w = 1100$.

168 ## Discussion

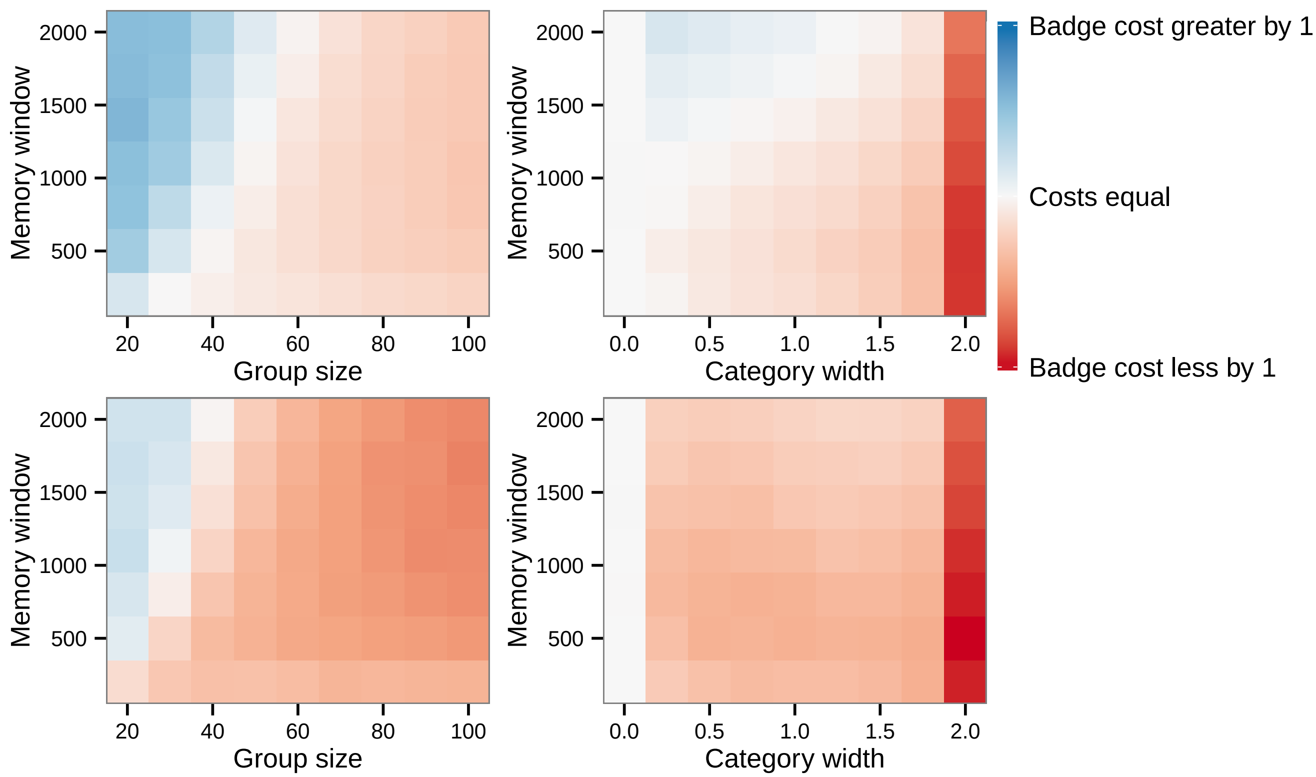

Figure 5: **The overall costs of using individual recognition are lower than the costs of using the badge system at low group size, high memory window, and low category width.** When the signal-quality correlation ρ increases, animals using the badge system incur lower costs for more combinations or parameters. Here we show the difference in the overall costs (C) associated with individual recognition and with the badge system as a function of A,C group size N and memory window w and B,D category width w and memory window w. Red indicates the badge system is less costly and blue indicates individual recognition is less costly. Parameters: in A and B $\rho = 0.5$, in C and D $\rho = 0.9$; unless the parameter is being varied $\delta = 0.5$, $\ell_{\text{i}} = 0.2$, $N = 50$, $p_{\text{o}} = 0$, $\sigma_{\text{b}} = 0.2$, $\sigma_{\text{i}} = 0.01$, $\sigma_{\text{q}} = 0.5$, $T = 10000$, $w = 1100$.

References

[1] Sheehan, M. J. and Bergman, T. J. (2015). A quality signaling?recognition trade-off at the level of the type of interaction not species: a response to comments on sheehan and bergman. *Behavioral Ecology*.

[2] Sheehan, M. J. and Bergman, T. J. (2016). Is there an evolutionary trade-off between quality signaling and social recognition? *Behavioral Ecology*, 27(1):2–13.

Supplemental Information

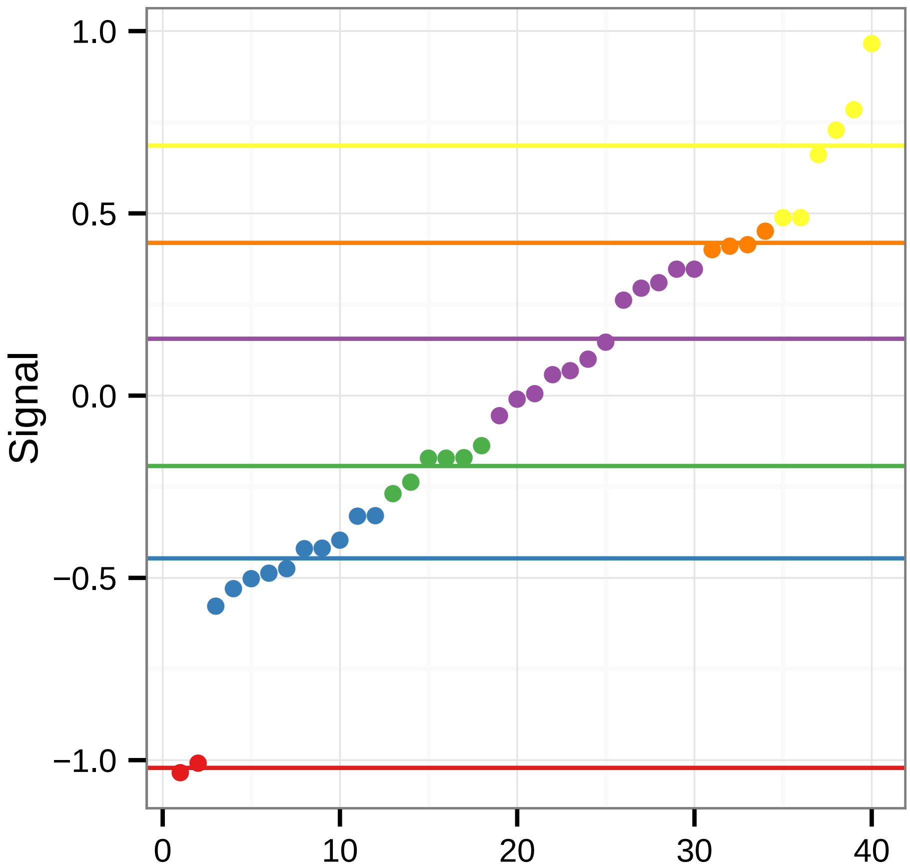

Figure 6: **Example of how categories are defined.** Here we show how a group of size $N = 40$ might be divided into categories with maximum width 0.5. Each point corresponds to an animal in the group. We show the signal of each animal on the y-axis and the order of the animals from lowest to highest signal on the x-axis. The color of the point indicates the category to which it belongs and the horizontal lines show the median signal value for each category. Categories were formed by choosing an animal i at random, putting all other animals whose signals were within $\delta/2$ of i's signal s_i into the same category, then choosing another uncategorized animal at random, and continuing until all animals were categorized.

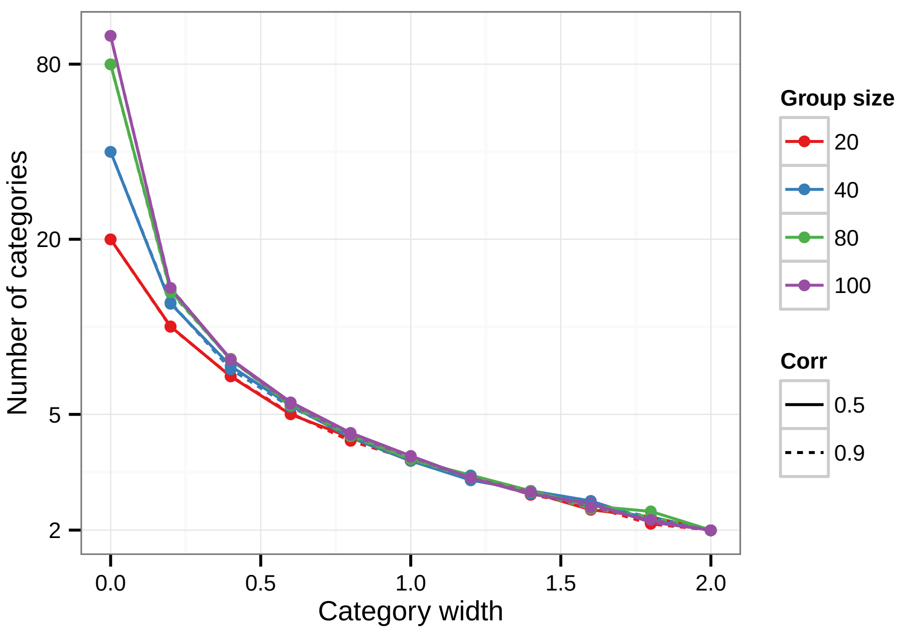

Figure 7: **As category width δ decreases the number of categories into which the group is divided increases.** For a given group size N and category width δ we generated a group with N signal values $\{s_i\}$ and categorized the group according to the procedure in the text and Figure 6 100 times and took the average number of categories across these 100 groups. Here we show how the average number of categories into which a group is divided depends on group size N and category width δ. Each colored line corresponds to a group size. There are overlapping solid and dotted lines in each color since the signal-quality correlation ρ does not affect the number of categories formed. When $\delta = 0$ there are as many categories as there are animals in the group. When $\delta = 2$ there are on average 2 categories, although it can happen that all animals are put into the same category.

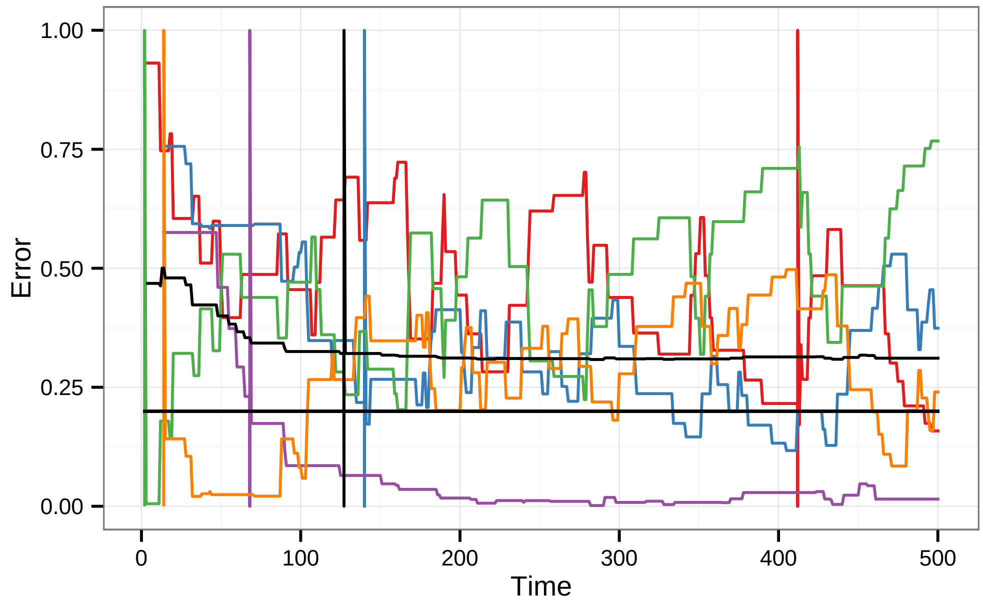

Figure 8: **Example of how to calculate learning time.** Learning time τ_i can be less than T, even if average error at time T is above the threshold 0.2. Each colored line shows how one the error in one animal's opinion of its five group mates changes over time. The black line is its average error about the other animals, $\epsilon_i(t)$. The horizontal black line is the error threshold, 0.2. Each vertical colored line is the first time when corresponding line drops below the threshold. The vertical black line is the average of these learning times, τ_i.

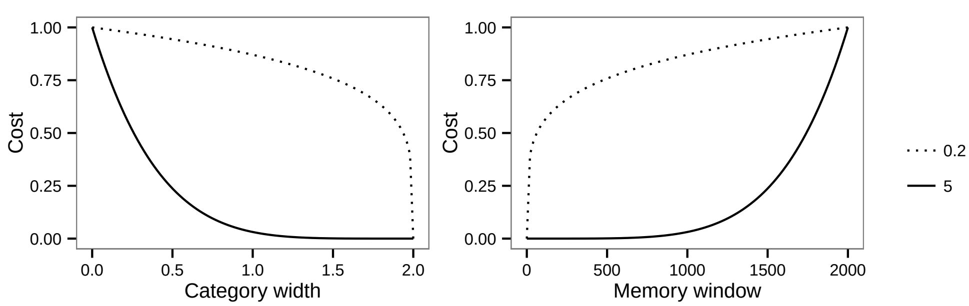

Figure 9: **Cost functions.** Here we show how the cost functions for the cognitive parameters category width and memory window. In each panel, the solid line corresponds to $\alpha = 5$ and the dotted line corresponds to $\alpha = 0.2$.

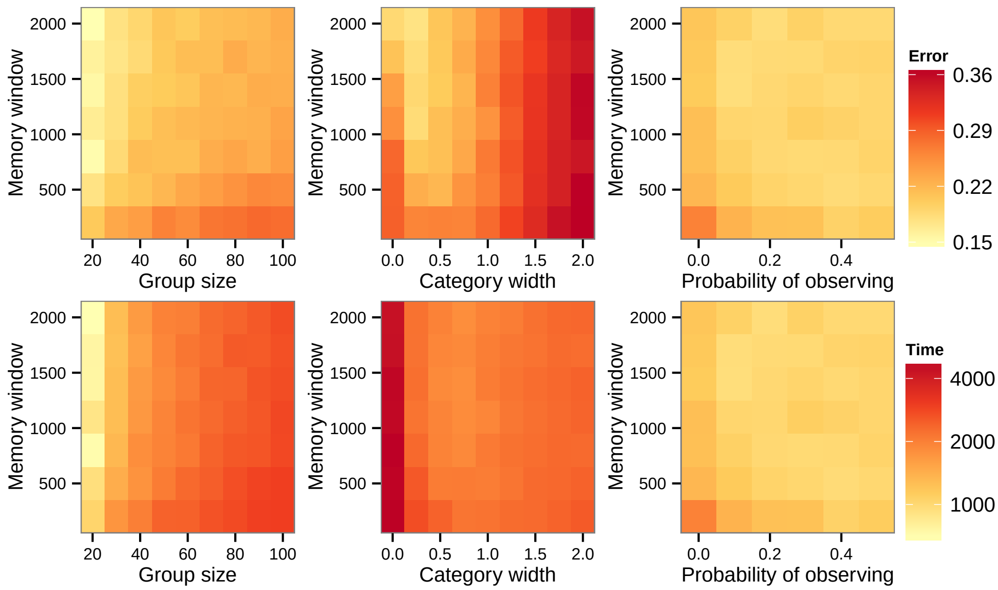

Figure 10: **Parameter interactions for animals using the badge system.** In A, B, and C we show average error $\bar{\epsilon}$ for animals using the badge system as a function of two parameters. In D, E, and F we show average learning time $\bar{\tau}$ as a function of two parameters. Parameters: unless the parameter is being varied $\delta = 0.5$, $\ell_{\mathsf{i}} = 0.2$, $N = 50$, $p_{\mathsf{o}} = 0$, $\rho = 0.9$, $\sigma_{\mathsf{b}} = 0.2$, $\sigma_{\mathsf{i}} = 0.01$, $\sigma_{\mathsf{q}} = 0.5$, $T = 10000$, $w = 1100$.

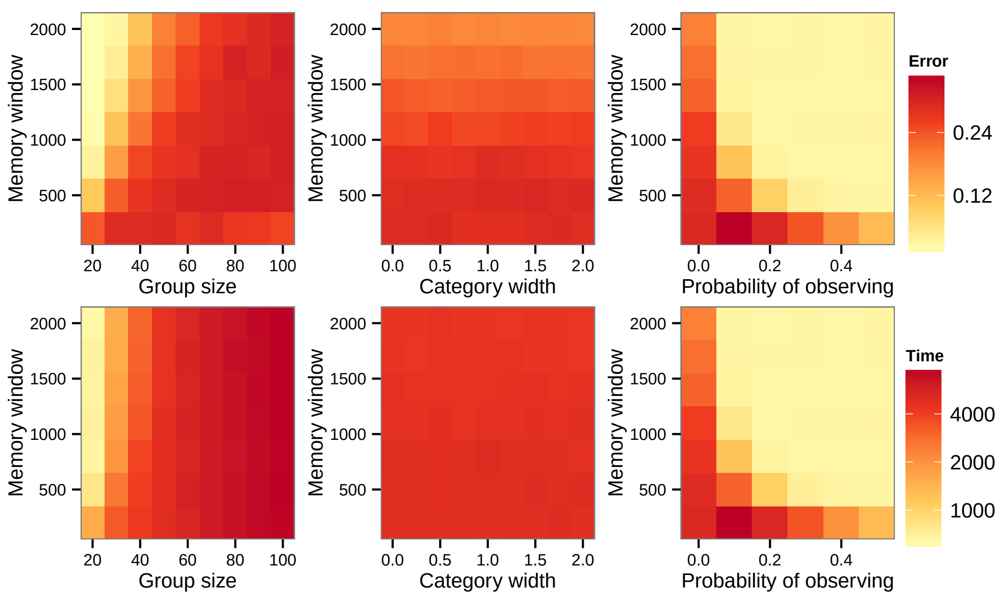

Figure 11: **Parameter interactions for animals using individual recognition.** In A, B, and C we show average error $\bar{\epsilon}$ for animals using individual recognition as a function of two parameters. In D, E, and F we show average learning time $\bar{\tau}$ as a function of two parameters. Parameters: unless the parameter is being varied $\delta = 0.5$, $\ell_{\mathsf{i}} = 0.2$, $N = 50$, $p_{\mathsf{o}} = 0$, $\rho = 0.9$, $\sigma_{\mathsf{b}} = 0.2$, $\sigma_{\mathsf{i}} = 0.01$, $\sigma_{\mathsf{q}} = 0.5$, $T = 10000$, $w = 1100$.