Serial reversal learning in nectar-feeding bats

# Introduction

*‘Take some more tea,’ the March Hare said to Alice, very earnestly.*

*‘I’ve had nothing yet,’ Alice replied in an offended tone, ‘so I can’t take more.’*

*‘You mean you can’t take LESS,’ said the Hatter: ‘it’s very easy to take MORE than nothing.’*

***- Alice’s Adventures in Wonderland*, Lewis Carroll**

The ability to recognize and respond appropriately to change is crucial to animals’ survival. This ability is often referred to as cognitive or behavioural flexibility (Tello-Ramos et al. 2019), and can be seen in a range of behaviours in many different species: locating food in ants (Czaczkes and Heinze 2015); spatial navigation in male guppies (Lucon-Xiccato and Bisazza 2017); parental care in poison frogs (Ringler et al. 2015); and migratory behaviour in many vertebrates (Winkler et al. 2014). One experimental protocol that has been widely used to demonstrate cognitive/behavioural flexibility is reversal learning.

Reversal learning is essentially a specific type of discrimination learning, which is when animals must learn a specific, arbitrary response to each of multiple stimuli. In a reversal learning experiment an animal is first faced with a simultaneous choice between two stimuli, only one of which is paired with a reward. After a certain number of trials has occurred and the animal has likely learned the association between the correct stimulus and reward, the reward contingencies of the two stimuli are reversed. In a serial reversal learning procedure the reward contingencies reverse repeatedly. An animal that chooses the rewarded stimulus more frequently than the non-rewarded stimulus is judged to perform better on the task; a choice to the non-rewarded stimulus is, in the context of the task, an ‘error.’

The serial reversal learning protocol can be adapted to the behaviour and sensory physiology of many different species, thus allowing comparative research. It has been done using visual stimuli in bumblebees (Strang and Sherry 2014) and guppies (Boussard et al. 2020); visual and spatial stimuli in corvids (Bond, Kamil, and Balda 2007); spatial stimuli in rats (Boulougouris, Dalley, and Robbins 2007), great tits (Hermer et al. 2018a) and gray squirrels (Chow et al. 2015); and olfactory stimuli in rats (Kinoshita et al. 2008). Reversal learning, specifically serial reversal, has been used as an explicit comparative measure of animal ‘intelligence’ (Bitterman 1964): ‘higher’ animals like pigeons, rats and monkeys showed a progressive improvement on the task and ‘lower’ animals like turtles and fish did not. Though the idea of such a hierarchy is outdated, comparative research using reversal learning can reveal important differences in behaviour and learning that have evolved under the selection pressures faced by different species.

Learning in the reversal learning task is clearly demonstrable, and is therefore a meaningful criterion when comparing the performance of different animals. First-order learning happens when an animal learns the stimulus-reward association and changes its behaviour according to the strength of this reinforcement. Higher-order or second-order learning is the learning of rules or strategies; in serial reversal learning the same stimuli are successively paired with a reward and then not paired with a reward, so a strategy can be very profitable. The optimal rule in reversal learning is ‘win-stay; lose-shift,’ which means in practice one ‘error’ per reversal. After learning the task, the perfectly optimal animal will first exclusively choose the stimulus that is paired with reward. At the first choice of this stimulus that does not give a reward (the error), the animal will change its preference and exclusively choose the other stimulus which is now paired with a reward. Progressive ‘improvement’ in this task, where an animal makes fewer and fewer errors per reversal is indication that the animal is learning the rule of reversal, or ‘learning to learn’ (**Shettleworth, 2010** - how to cite a book?).

Comparative research from the primate literature reveals a very interesting difference in these two types of learning (first- and second-order). Thirteen different species of primates were compared on a visual reversal learning task (Rumbaugh, Savage-Rumbaugh, and Washburn 1996), where they were trained to discriminate a pair of stimuli to either 67% or 84% choice for the rewarded option, and then given a single reversal of reward contingencies. If the primates’ behaviour was driven mainly by first-order learning, they should make fewer errors after the reversal when the criterion was 67% rather than 84%, and *vice versa* if second-order learning or rule-learning was occurring. The results showed that Prosimian species tended to perform better when trained to 67%; apes when trained to 84%; and monkeys were intermediate.

Progressive improvement on the serial reversal task as more reversals are experienced, evidence of rule-learning, has been shown in many different species: bumblebees (Strang and Sherry 2014), two different species of great tits (Hermer et al. 2018a) and three different species of corvids (Bond, Kamil, and Balda 2007). Corvids in fact show significant transfer across stimulus modality, which is strong evidence for rule-learning.

What performance on the serial reversal task says about the deeper cognitive mechanisms at work, and whether the task is a measure of cognitive or behavioural flexibility, are not completely settled questions. Cognitive flexibility cannot be directly observed; it is inferred to have occurred through changes in behaviour, or behavioural flexibility (Tait et al. 2018). However, just because behavioural flexibility has been observed does not necessarily indicate cognitive flexibility (Dhawan, Tait, and Brown 2019). The term ‘behavioural flexibility’ itself has been used widely but inconsistently, applied to many traits that have different underlying neural mechanisms or do not co-vary (Audet and Lefebvre 2017). Behavioural flexibility in animals has evolved in response to selection pressures from different foraging environments: the flexibility required to deal with seasonal changes in fruit availability is not the same kind of flexibility required to deal with capturing a prey animal intent on escaping.

There is a sense in which the foraging ecology of some nectar-feeding animals is a natural analogue to the serial reversal learning task. The Neotropical bat species *Glossophaga commissarisi* relies primarily on flower nectar for energy. These bats have remarkably high metabolic rates for their body mass (C. C. Voigt and Winter 1999; Christian C. Voigt, Kelm, and Visser 2006), due to the energetic demands of hovering flight (Winter and Helversen 1998; O. v. Helversen and Reyer 1984). As flowers yield only small droplets of nectar each time they are visited (Christian C. Voigt, Kelm, and Visser 2006), the bats make several hundred flower visits per night. Many plants visited by bats put out only a few flowers every night that bloom for a long time (Kunz and Fenton 2005). A certain time after a flower is emptied the nectar-levels are replenished, so bats can visit the same flower multiple times, relocating it primarily through their excellent spatial memory (Winter 2005; Toelch et al. 2008). The longer the bat waits, the more the flower refills but the higher the likelihood that a competitor could find and exploit the flower first. To make repeated, profitable visits to a flower, a bat must remember both the location of the flower and estimate the flower’s expected reward value. The serial reversal learning task requires an animal to respond to a change in the profitability of available options, and remember all potentially rewarding options: the kind of behavioural flexibility, in short, required in the typical foraging bout of a nectar-feeding bat.

We carried out a serial reversal learning task with wild *G. commissarisi* individuals. Over three nights the bats were given two potentially rewarding options to choose between. At the start of the night, only one of the options was rewarding and the other was not. After a certain number of visits had been made by the bats, the reward contingencies reversed without any cue: the previously rewarding option was now unrewarding and the previously unrewarding option was rewarding. This reversal occurred five times a night on every night.

Our aims with this experiment were as follows. Firstly, we wanted to test whether that the bats were capable of reversal learning. We believed this to be extremely likely as the behavioural requirements of the task are typical features of the animals’ foraging ecology. Secondly, if the bats demonstrated the ability to respond to the reversals, we wanted to explore how this was reflected in their decision-making. What changes occurred in the relative number of visits made to the rewarding and the non-rewarding options? Thirdly, we wanted to see if the bats were capable of second-order learning, or ‘learning to learn.’ Could the bats learn the rule behind the change in their environment and use the optimal strategy of one error per reversal?

After the analyses described above were done and the data and results examined, we performed further analyses to explore the conclusions of our confirmatory analyses. The difference between these results must be clearly noted. Firstly, we reasoned that there is a difference between the first visits of a night, before any experience of a reversal, and all the subsequent visits after at least one reversal had occurred. We statistically tested for this difference in the bats’ choice behaviour. Secondly, we examined the effect of the asymptotic level of performance (the highest stable proportion of visits to the rewarding option after a reversal) on the performance immediately after a reversal.

# Methods

## Study site and subjects

The experiment was done from the 28th of June to the 25th of July, 2017, at La Selva Biological Field Station, Province Heredia, Costa Rica. Male and female individuals of the species *Glossophaga commissarisi*, were captured from the wild for the experiment. The bats were attracted to a particular location in the forest using sugar-water (see **Reward** below) as bait and then caught in mist-nets. The bats were sexed and the selected individuals were were then taken to two flight-cages (4 x 6 m). The flight-cages had mesh walls and therefore the same climatic conditions as the surrounding environment. A group of four bats at a time were put into a flight cage. All the individuals in a group were the same sex. The bats were weighed and radio frequency identification (RFID) tags that were uniquely assigned to each bat were placed around their necks as collars. The bats were then released into the flight-cages so they could fly within them freely.

Before the start of the experiment the procedure was tested with four females and refinements were made to the procedure. The data from these individuals were not analyzed. 16 bats participated in the main experiment. At the end of the experiment, the RFID collars were removed and the bats were weighed to make sure they were still at a healthy weight. No blinding was done as all the data collection was completely automatized. Two of the bats did not drink a sufficient amount of sugar-water to meet minimum energy requirements and were released before the end of the experiment and not replaced. The data from these two individuals were not analyzed. Thus, 14 bats in total (seven males and seven females) completed the experiment and the data from these animals were analyzed.

Animal experimental procedures were reviewed and permission for animal experimentation and RFID-tagging was granted by Sistema Nacional de Areas de Conservación (SINAC) at the Ministerio de Ambiente y Energía (MINAE), Costa Rica.

## Experimental Setup

### Reward

The reward received by the bats during the experiment was also their main source of food. The reward was a 17% by weight solution of sugar dissolved in water (prepared fresh everyday), hereafter referred to as ‘nectar.’ The sugar consisted of a 1:1:1 mass-mixture of sucrose, fructose and dextrose. The nectar was thus similar in composition and concentration to the nectar produced by wild chiropterophilous plants (Baker, Baker, and Hodges 1998).

### Flower and pump setup

Each flight cage had a square plastic frame in the center (2x2x1.5m). Eight reward-dispensing devices - hereafter referred to as ‘flowers’ - were fixed in a radial pattern on this frame, two on each side of the square (see **Figure 1**) with a distance of 40 cm between adjacent flowers. This is a distance the bats can discriminate (Thiele and Winter 2005). Each flower had the following parts: an RFID reader mounted on a plastic cylinder around the head of the flower; an infra-red light-barrier beam; an electronic pinch valve through which a PVC tube was placed and fixed to the head of the flower

A stepper-motor pump was placed in the center of the plastic frame in each cage. The pumps contained a 25 mL Hamilton glass syringe (Sigma Aldrich). The precision of the two pumps differed slightly: the pump in Cage 1 delivered 2.11 L per step of the stepper-motor, and the pump in Cage 2, 3.33 L per step. The glass syringe was connected to the tubing system of the flowers through five pinch valves. The pinch valves controlled the flow of liquid from the pump to the system and from a reservoir of liquid to the pump. The reservoir (500 mL thread bottle, Roth, Germany) was filled with fresh nectar everyday and connected to the syringe through the valves.

When a tagged bat approached a flower, the individual RFID number was read by the reader. If the bat then poked its nose into the flower and broke the light barrier, it triggered the release of a reward. The pinch valve opened and the pump moved the correct number of pre-programmed steps to dispense nectar to the head of the flower. The bat could easily hover in front of the flower and lick up the nectar. Only when both events occurred, i.e., the RFID reader detected a bat and the light-barrier was broken, would a reward be triggered.

The flowers and the pump were connected to a Lenovo ThinkPad laptop computer, which ran the experimental programs and the programs used to clean and fill the systems: PhenoSoft Control 16, PhenoSoft GmBH, Berlin, Germany. The raw data were recorded to this computer as comma-separated values (CSV) files.

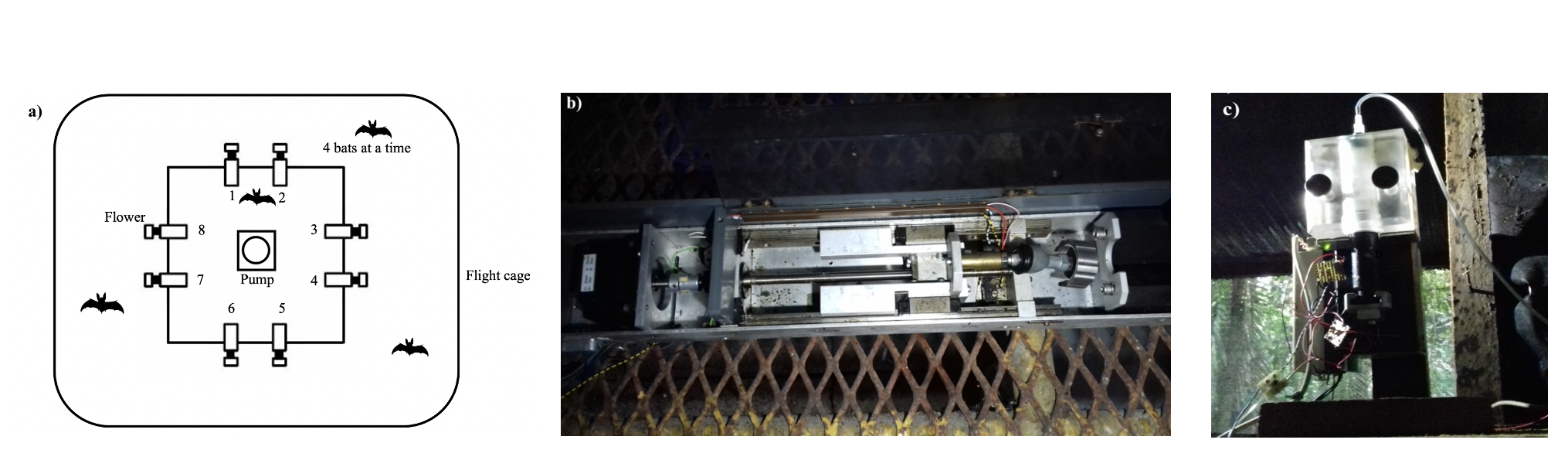


Figure 1: a) Schematic of the cage and flower set-up b) Horizontal pump c) Vertical pump

## Experimental procedure

Every day at around 10 h the old nectar was emptied from the system. The system was rinsed and filled with plain water until 15 h , when it was filled again with fresh nectar. Twice a week the system was filled with 70% ethanol for an hour to prevent microbial growth, then repeatedly rinsed with water.

Four bats were placed in a flight-cage in a group, and all the bats were the same sex. There were four such groups in total, and data were collected simultaneously from two groups, one in each flight-cage. Each bat was uniquely assigned two adjacent flowers on the same side of the square frame, out of the array of eight. These flowers were programmed to reward only one of the four bats in the cage. After the system was filled with fresh nectar at approximately 17 h, the program was left running for data-collection till the next morning. Thus, the bats could begin visiting the flowers to collect a reward whenever they chose, which was at approximately 18 h every night.

During the course of the night, when the syringe of the pump had been emptied, the pump re-filled automatically. This event happened only once every night. On the main experimental days this process took 4.5 minutes (SD = ±0.18) for the horizontal pump, and 2.43 minutes (SD = ±0.04) for the vertical pump. About 1 % (SD = ±0.74) of all visits made by the bats over all three experimental nights happened during the pump refill events, and the bats did not receive any reward on these visits, even if they were made to the rewarding flower.

Every night the bats were also given ad-libitum supplemental food: 3.5g of hummingbird food (NektarPlus, Nekton) in 100 mL of water and 3.5g of milk powder (Nido 1+, Nestle) in 100 mL of water. They were also given a small bowl of locally-sourced bee pollen.

## Experimental design

The experiment proceeded through the following stages.

### Training to use a flower

On the night the naive bats were captured and placed into the flight cages they could receive a reward from any of the flowers whenever they visited them throughout the night. To enable the bats to find the flowers a small cotton pad was placed on the flowers, soaked in di-methyl di-sulphide. This is a chemical attractant produced by many bat-pollinated flowers (O. von Helversen, Winkler, and Bestmann 2000). A small drop of honey was applied to the inside of the flowers to encourage the bats to place their heads inside, break the light-barrier and trigger a nectar reward. By the end of the night all the bats had found the flowers and learned to trigger rewards quickly.

### Training to use two specific flowers

After the bats had learned to trigger rewards, the next stage of training involved assigning the bats uniquely to two out of the eight flowers in the array. For an individual animal only the two flowers assigned it would be rewarding from this stage of training until the end of the experiment. Because the bats had already learned to trigger a reward at the flowers, the flowers were not provided with a cotton piece with the chemical attractant and honey was not applied to them. This stage was similar to the previous one, except the bats could only trigger a reward at their assigned flowers.

### Alternation

To ensure that the bats were familiar with both flowers assigned to them they went through one final stage of training: forced alternation between the two assigned flowers all night long.

### Main Experiment

In this serial reversal learning task the bats had to choose between a flower that gave 40 L of nectar and one that gave no reward at all. The location of the rewarding flower was not cued, but through the Alternation phase of training each bat knew the locations of both flowers that were potentially rewarding to it. After a bat had made 50 visits in total to the two flowers a reversal occurred: the previously rewarding flower became the non-rewarding flower and *vice versa*. Importantly, only visits to the two flowers assigned to a bat counted towards the visit tally, not visits to any of the other flowers which were unrewarding to that particular bat. The batch of 50 visits that occurred between two consecutive reversals (when the locations of the rewarding and unrewarding flowers remained stable) was termed a ‘reversal block,’ including the first 50 visits of a night when the bats had not experienced any reversal at all that night. This occurred at regular intervals of 50 visits until the bat either stopped making visits or reached a maximum of 300 visits in a night. After the bat had made 300 rewarded visits it could no longer receive a reward on that experimental night. There were five reversals per night. This stage of the experiment was repeated for three nights in a row. The same flower was the first to be rewarding at the start of every night. Thus, because there were five reversals every night (six blocks of 50 visits), if a bat completed the maximum of 300 visits on a night, the last flower to be rewarding that night was non-rewarding at the start of the next night.

## Statistical analysis

All the models were fit in a Bayesian framework using Hamiltonian Monte Carlo in the R package brms (Bürkner 2017), which is a front-end for rstan (Stan Development Team, 2020).

All the visits made by the bats during a night, up to a maximum of 300, were included in the analyses. There were three experimental nights, divided into six blocks of 50 visits each. At the end of the first five blocks a reversal occurred and the end of the last block was the end of data-collection for the night. Each block was further divided into five bins, each consisting of ten visits, in order to examine the bats’ behaviour within each block.

We defined a perseverative visit as a visit to the previously-rewarding option just after the occurrence of a reversal, until the first visit to the newly-rewarding option. By definition this could not happen in the first block of a night. A generalized linear mixed-model was used to investigate the effect of experimental night and reversal block on the number of perseverative visits. A negative-binomial likelihood function was used for this model. Experimental night, reversal block and their interaction were fixed effects and random slopes and intercepts were used to fit regression lines for each individual animal.

We also examined the proportion of visits made to the rewarding flower. This was defined as the ratio of the number of visits to the rewarding flower divided by the total number of visits in a bin. The total number of visits in a bin only included visits made to the two flowers assigned to a bat and any visit to a flower that was not assigned was not considered in the analysis. The model was fit using a binomial likelihood function, with experimental night, block, bin and their interactions as fixed effects; random slopes and intercepts were used to fit regression lines for the individuals.

After examining the above results, further analyses were done. It is important to note that these were *exploratory*, and the ideas were suggested to us by the results of the intended analyses described above.

A second model was fit to the proportion of visits to the rewarding flower to take into account the fact that the first night and the first block of each night were qualitatively different from the others. On the first night the animals had had no prior experience of any reversals, and during the first block of every night they had not experienced any reversals on that night, and this was reflected in the fit of the posterior predictions made from the first model. The second model of these data was identical to the first except for the addition of experimental night and block as factor variables, with the first night and the first block of every night as one level and the other nights and other blocks of each night as the other level. The two models were compared using leave-one-out cross-validation, implemented in brms using the package loo (Vehtari, Gelman, and Gabry 2017).

We reasoned that a comparison of the bats’ behaviour just before and after a reversal might reveal something of the learning mechanisms at work. If a higher proportion of visits to the rewarding flower just before a reversal is predictive of a higher proportion of visits to the rewarding flower just after, that might potentially indicate that the bats are learning the ‘rule’ behind the reversals. On the other hand, if there is no rule-learning, and the animals’ choice is driven by how much reinforcement was received at the two options, we would expect the opposite: the proportion of visits to the rewarding flower before the reversal is predictive of a *lower* proportion of visits to it just after as the animals take longer to ‘reverse’ their choices from a highly reinforced option. We took the proportion of visits to the rewarding option, averaged over the last three bins of a reversal block for each individual as the ‘asymptote’ of the bats’ choice behaviour. We fit a generalized linear-mixed model with the proportion of visits to the rewarding option the first bin just after a reversal as the response variable with the fixed effects asymptote, a continuous variable, and night, a factor variable. Random slopes and intercepts were used to fit regression lines for each individual animal.

Weakly informative priors were used. The random intercepts and slopes were given a Normal distribution with a mean of 0, and a standard deviation drawn from a Cauchy distribution with a mean of 0 and a standard deviation of 1. All the models were estimated using 4 chains with a thinning interval of 3, with 1200 warm-up samples and 1300 post-warm-up samples for the model with the first experimental night and block additionally treated differently; 2000 warm-up samples and 2000 post-warm-up samples for the model of the first bin of 10 visits after a reversal; and 1000 warm-up samples and 1000 post-warm-up samples for the others.

Visual inspection of the trace plots, the number of effective samples, the Gelman-Rubin convergence diagnostic () and the calculation of posterior predictions for the same clusters were all used to assess the fit of the models. In all of the models the was equal to 1 for all the chains.

The data from all 14 bats that participated in the three experimental nights were included in these models, even though some individuals did not complete all 300 visits on every single night.

All statistical analyses and creation of plots were done in R.

## Data availability

All data and analysis code are available online at …..

# Results

## Confirmatory Analyses

### Explanatory paragraph

**This paragraph is just for our guidance and will be removed later: it is an overall summary of the conclusions we draw from all of the analyses, i.e., the story we are building.**

**Confirmatory analyses**: As bats experience serial reversals they show a more rapid switch to the rewarding option after a reversal, and an overall increase in the proportion of visits made to the rewarding option. This overall increase was smaller with each successive reversal.

**Exploratory analyses**: The largest proportion of visits to the rewarding option was made every night before the bats had experienced any reversals. The higher the asymptote the longer the bats took to reverse, but this negative relationship varied by night and almost disappeared on the second night.

**In conclusion**: Experiencing serial reversal of reward, nectar-feeding bats learn that reversal can occur in their environment: we know this because they switch more rapidly; reach a higher proportion of visits to the rewarding option with a ‘ceiling effect’; and sample a little more once they have experienced a reversal on a day

### a) Bats made a very high proportion of their visits to the rewarding option after reversals

At the start of the first night, in the very first bin of ten visits when the bats did not yet have any information about the available options and had never experienced a reversal, the proportion of visits to the rewarding option averaged over the bats was close to chance: 54.5 %, 95% CIs [46.8, 63]. Within the next ten visits, however, this proportion increased to 92.1 %, 95% CIs [87.9, 96.4] and by the last bin of this first block was 99.3 %, 95 % CIs [97.9, 100]. Immediately after the first experience of a reversal, this proportion dropped down to 13.6%, 95% CIs [8.4, 19.5] in the first ten visits, but came back up to 96.4%, 95% CIs [92.9, 99.3] by the last bin of this block. This pattern of a decrease in the average proportion of visits to the rewarding option of all the bats immediately following a reversal, and then a rapid increase, was consistently true for all the reversals on the first night (**Figure 5**). Indeed, the proportion of visits to the rewarding option in the last 3 bins of a reversal block was 93 %, 95% CIs [91.7, 94.1] on the first night, even excluding the first block before any reversal experience. At the very start of the second and third nights, in the first bin of visits before any experience of a reversal *on that night*, the average proportion of visits to the rewarding option of all the bats was 69.8 %, 95% CIs [64.6, 74.7]. However the average proportion of visits to the rewarding option still showed the pattern of a decrease immediately after a reversal and then increase to a high proportion. On nights 2 and 3, after the bats had experienced the first reversal of the night, the average proportion of visits to the rewarding option in the last three bins of a reversal block was 94.8 %, 95% CIs [94.1, 95.5], comparable to the 93 %, 95% CIs [91.7, 94.1] on the first night.

### b) Bats switch to the rewarding option faster as they experience more reversals

A visit to a non-rewarding flower after a reversal, before the first experience of a reward at the newly-rewarding option was called a perseverative visit. Our analysis showed that both experimental night and block had a negative effect on the number of perseverative visits made by the bats. As the bats experience more reversals on more nights the number of perseverative visits decreased. The interaction of night and block, however, was positive: the decrease in the number of perseverative visits was smaller as the bats experienced more experimental nights (**Figure 3**).

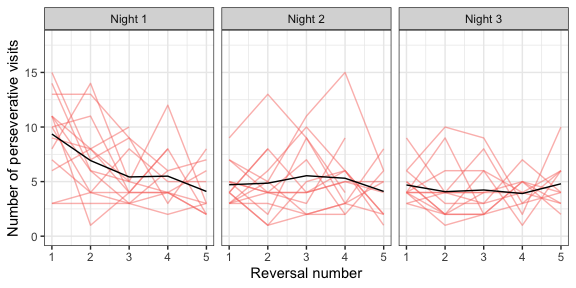


Figure 2: Number of perseverative visits made by the bats after each reversal across all three nights. A perseverative visit was defined as visit to the previously-rewarding option before the first visit to the newly-rewarding option after a reversal. By definition there were no perseverative visits in the first block of a night. The red lines are the perseverative visits made by the individual bats after each reversal of a night; the black lines are the number of perseverative visits averaged over the bats

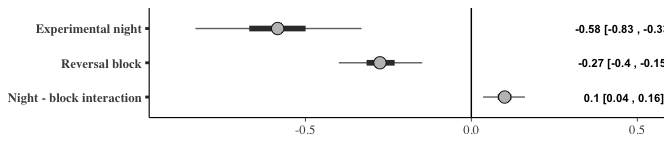


Figure 3: Forest plot of the estimates of the effect of night and block on perseverative visits; a negative effect indicates the perseverative visits decreased with successive nights and blocks. The circles represent the slope estimates, the thick horizontal lines represent the 50% credible intervals and the thin horizontal lines 95% credible intervals. The numbers are the slope estimates and the 95% credible intervals of each effect

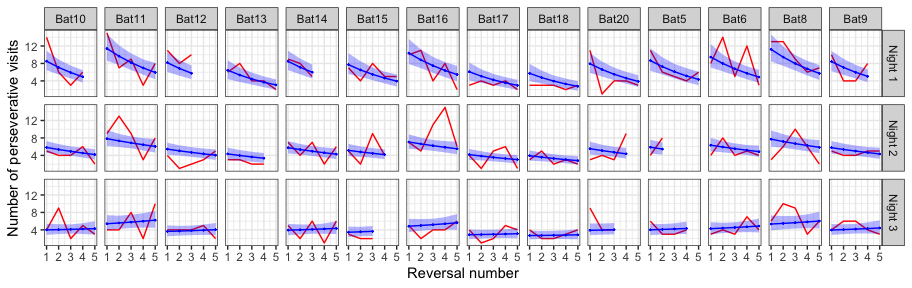


Figure 4: A comparison of the posterior predictions of the generalized linear mixed-effects model of perseverative visits and the empirical data from the bats. The red line indicates the number of perseverative visits per reversal made by the bats; the blue line indicates the corresponding posterior prediction of the model with the blue shading indicating 95% confidence intervals

### c) The bats make more visits to the rewarding option as they experience more reversals

The proportion of visits to the rewarding option was modeled as an effect of experimental night, block and bin. This analysis showed that night, block and bin all had a positive effect on the proportion of visits to the rewarding option: as the bats experience more reversals on more nights, and the more visits made since the occurrence of a reversal, the proportion of visits made to the rewarding option increased. Interestingly the slopes of the interactions of experimental night and block; block and bin within the block; and experimental night and bin within the block were all negative. That is, the increase in proportion of visits to the rewarding flower gets smaller as the bats experience more reversals on more nights. There is a smaller ‘improvement’ in performance, likely through a ‘ceiling effect,’ given how high the proportion of visits to the rewarding option was (**Figure 6**).

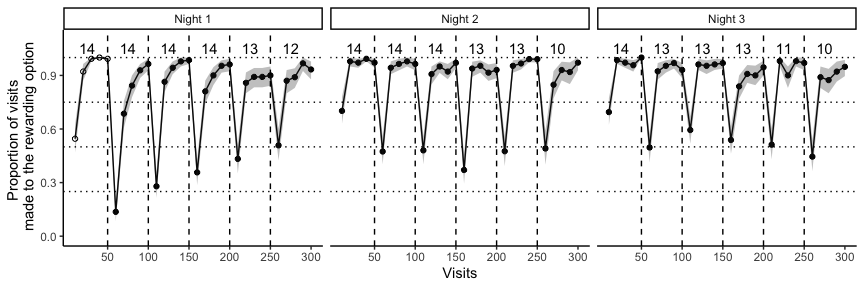


Figure 5: Average proportion of visits to the rewarding option in each bin of ten visits, averaged over all the individuals, across the three experimental nights. This bin average is indicated by white points in the first block on the first night before the bats experienced any reversals at all; the bin averages of all the other blocks are indicated by black points. The small numbers indicate the number of bats that participated in a particular reversal block, i.e., the sample size of a block. The solid black line connecting the points indicates the change in the average proportion of visits to the rewarding option with the shaded areas marking the 95% confidence intervals. The dashed vertical lines indicate the points when the reversals occurred

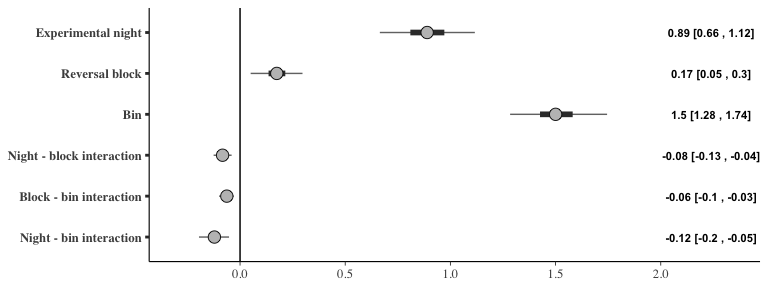


Figure 6: Forest plot of the estimates of the effect of night, block and bin on the proportion of visits to the rewarding option; a positive effect indicates the visits to the rewarding option increased with successive nights, blocks and bins (within each block). The circles represent the slope estimates, the thick horizontal lines represent the 50% credible intervals and the thin horizontal lines 95% credible intervals. The numbers are the slope estimates and the 95% credible intervals of each effect

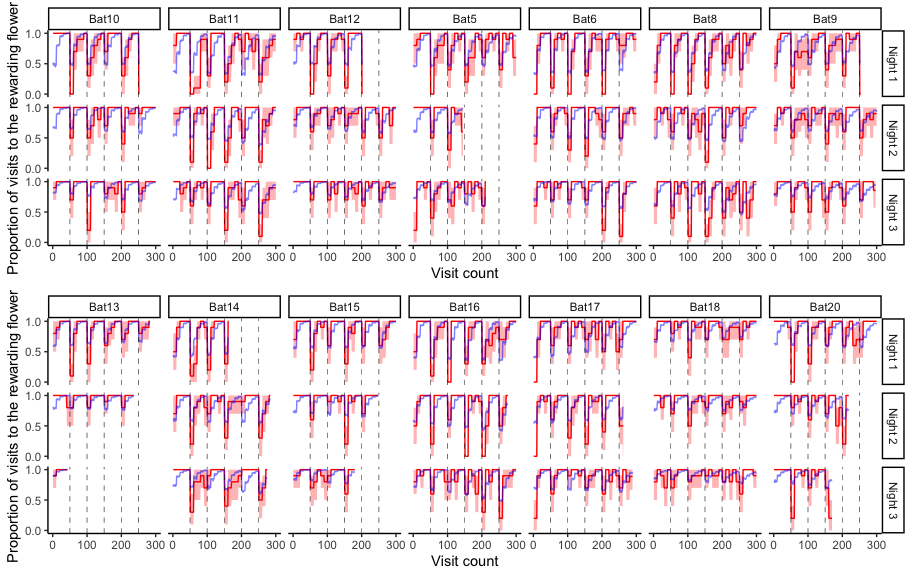


Figure 7: A comparison of the posterior predictions of the generalized linear mixed-effects model of the visits to the rewarding option and the empirical data from the bats. The red line indicates the average proportion of visits to the rewarding option per bin made by the individual bats, with the red shading indicating 95% confidence intervals; the blue line indicates the corresponding posterior prediction of the model

## Further exploratory analyses

### a) The effect of the first experimental night and the first block of each night

The first block of an experimental night was qualitatively different from the other blocks, as this was the only part of the night when the bats had not yet experienced a reversal. A similar argument can be made about the first experimental night: on this night the bats experience reversals for the first time. As it seemed the posterior predictions of our model with the effects of experimental night, block and bin did not completely capture the behaviour of the bats we did an exploratory analysis to specifically explore the effect of the first block and the first night.

The results of the model that accounts for the first night and block are presented in **Figure 8**. The main effects of night, block and bin, and their interactions were similar to the previous model. Additionally, the variable ‘block type,’ (i.e., whether a block was the first reversal block of the night or not), had a positive slope. The bats made more visits to the rewarding flower before they had experienced any reversal at all. While the ‘night type’ did not seem to affect the proportion of visits to the rewarding flower, block type and night type had an interaction effect with a positive coefficient: the first block of the first night, before any reversals had ever been experienced even once, had the highest proportion of visits to the rewarding flower compared to any other block on any other night.

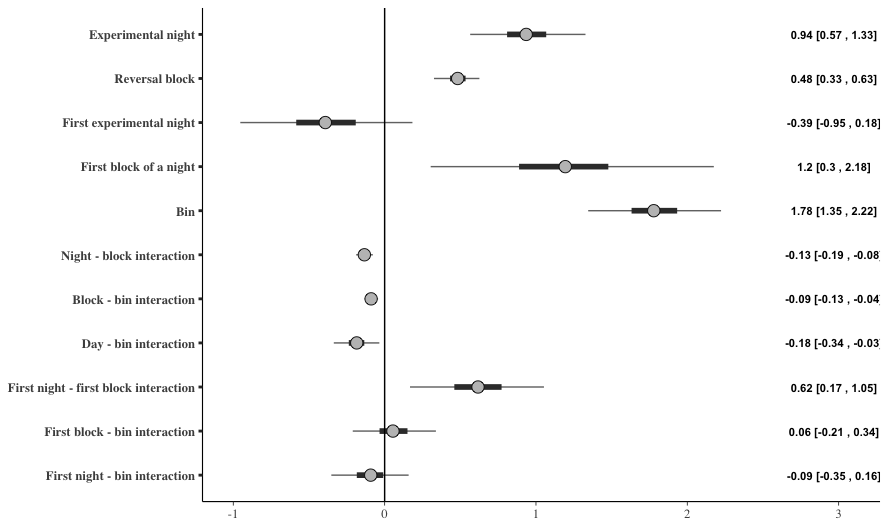


Figure 8: Forest plot of the estimates of the effect of night, block, and bin; and the differential effect of the first night and the first block (‘night-type’ and ‘block-type’) on the proportion of visits to the rewarding option; a positive effect indicates the visits to the rewarding option increased with successive nights, blocks and bins (within each block), or there were more visits to the rewarding option on the ‘block-type’ or ‘night-type’ indicated. The circles represent the slope estimates, the thick horizontal lines represent the 50% credible intervals and the thin horizontal lines 95% credible intervals

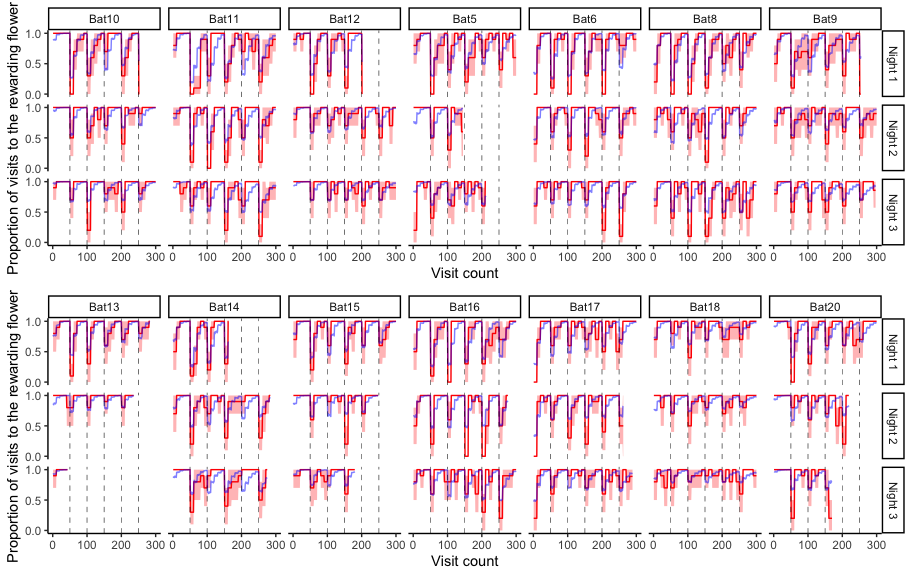


Figure 9: A comparison of the posterior predictions of the exploratory generalized linear mixed-effects model of the visits to the rewarding option and the empirical data from the bats, including the effect of ‘night-type’ and ‘block-type.’ The red line indicates the average proportion of visits to the rewarding option per bin made by the individual bats to show the change, with the red shading indicating 95% confidence intervals; the blue line indicates the corresponding posterior prediction of the model

We compared the predictive accuracy of the model that included night type and block type to the model that did not include these, using leave-one-out cross-validation. The LOO criterion was lower for the former than the latter, indicating better predictive accuracy (**Table 1**). Indeed, the posterior predictions of the model including night-type and block-type were a better fit for the first block of a night for many of the bats than those of the model that did not include them (**Figure 9**).

Table 1: Comparison of the two models of the proportion of visits to the rewarding option using leave-one-out cross-validation

| **Model** | **Difference in LOO estimates** | **Standard error** |
| --- | --- | --- |
| **Model accounting for the first night and first block separately** | 0.00 | 0.00 |
| **Night and block as continuous only** | -108.21 | 14.14 |

### b) The effect of asymptotic performance on the visits to the rewarding option just after a reversal

The proportion of visits to the rewarding flower just after a reversal made by the bats was compared to the asymptote just before. The asymptote, or asymptotic performance, was calculated as the average proportion of visits to the rewarding flower in the three bins before a reversal. The asymptotic performance before a reversal and the corresponding visits to the rewarding option just after that reversal are shown in the scatter-plot in **Figure 10**, with regression lines fit for each individual animal.

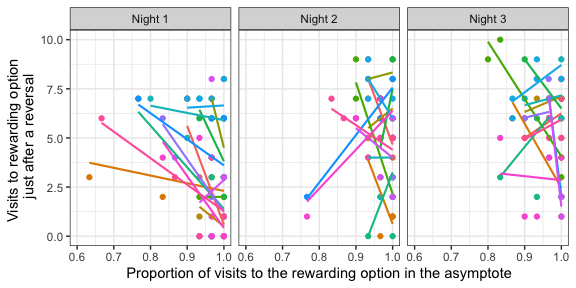


Figure 10: A comparison of the number of visits to the rewarding flower in the first bin of ten visits just after a reversal to the asymptotic proportion of visits to the rewarding flower just before that reversal. The asymptote was calculated as the proportion of visits to the rewarding flower averaged over the three bins of ten visits just before a reversal. Individual points are data from each individual bat for each one of the five reversals that occurred on a day. Colours indicate the individual bat. Regression lines are fit for each individual on each night

The analysis of the effect of the asymptotic performance on the visits to the rewarding option just after the reversal showed that the effect of the asymptote depended on the night. That is, there was a large interaction effect between the night and asymptote value. On the first night the slope of the effect of the asymptote was highly negative, indicating that the higher the asymptote value the lower the number of visits to the rewarding option just after a reversal. On the second night the slope was close to 0, though still negative, indicating a weaker relationship than on the first night. On the third night the slope highly negative again, similar to that on the first night (**Figure 11**).

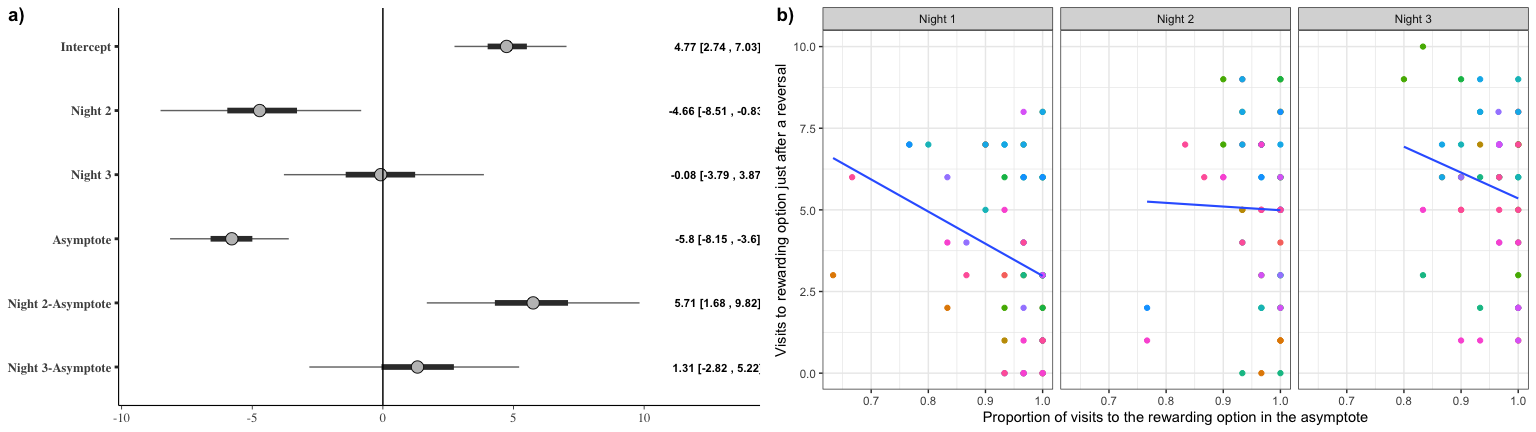


Figure 11: a) Forest plot of the coefficients of the effect of night and the value of the asymptote before a reversal on the number of visits to the rewarding option just after a reversal. The circles represent the intercept and slope estimates, the thick horizontal lines represent the 50% credible intervals and the thin horizontal lines 95% credible intervals. b) A comparison of the number of visits to the rewarding flower in the first bin of ten visits just after a reversal to the asymptotic proportion of visits to the rewarding flower just before that reversal. Individual points are data from each individual bat for each one of the five reversals that occurred on a day. Colours indicate the individual bat. Regression lines are fit for each night

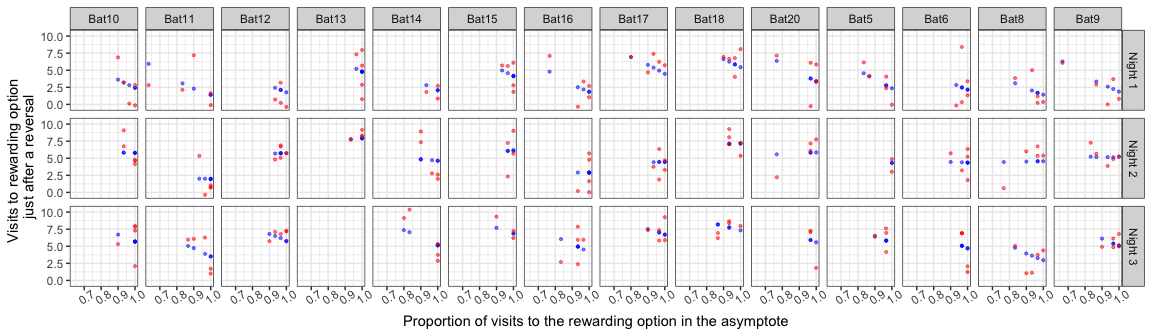


Figure 12: A comparison of the posterior predictions of the generalized linear mixed-effects model of the visits to the rewarding option after a reversal as an effect of night and asymptotic performance. Red points are the empirical values from the individual bats and the blue points indicate the corresponding posterior predictions of the model

# Discussion

In our experiment wild nectar-feeding bats participated in a serial reversal learning task, the first time, to our knowledge, such an experiment has been done with bats. It was clear that the bats detected and responded to the changing reward contingencies, evident first-order learning. As the animals experienced more reversals on more experimental nights, their performance improved in two significant ways.

## A faster switch

After each reversal they are quicker to switch from the previously-rewarding option to the newly-rewarding option. This was demonstrated in two ways. Firstly, after a reversal, the number of perseverative visits, i.e., the visits the bats make to the previously-rewarding option before their first visit to the newly-rewarding option decreases (as shown by the confirmatory analysis). Secondly, the proportion of visits to the newly-rewarding option in the first bin of ten visits after a reversal increases (as shown in the exploratory analysis - **Figure 11**).

It is true that the bats faced a relatively simple discrimination task: there were only two available options, so if one of them was perceived to be non-rewarding, there was only one possible alternative. The change from one option to another by the animals is less remarkable than the fact that this change happened more and more rapidly: more experience of the reversal led to better exploitation of the particular type of change the environment was undergoing.

## Improved ‘performance’

The bats’ preference for the rewarding option, i.e., their task performance, increased overall until it reached a ‘ceiling.’ This improvement in performance is strong evidence of second-order learning, or learning the rule behind the environmental change. The conclusion that the bats learn that reversals can occur is further supported by the results of the exploratory analysis: the highest proportion of visits to the rewarding option occurs before any reversals have been experienced at all. In a way this is unsurprising, as at this point in the experiment the bats are only discriminating between an option that delivers a substantial reward and an option that delivers nothing - the easiest possible discrimination - and have no information yet about the way the environment changes.

## The asymptote and immediate post-reversal performance

The exploratory analysis shed a little more light on the effect of the reversal on the animals’ performance just before and after the reversal. While the proportion of visits to the rewarding option at the end of a reversal block progressively increased, and the the number of perseverative errors decreased, the correlation between the performance just after a reversal and just before a reversal was negative on the first and third nights, and the correlation was close to zero on the second night. The correlation was most negative on the first night compared to the other two nights, suggesting that by the second night at least some learning had occurred.

Thus, the bats learn to take the regular occurrence of reversals into account as they discriminate between potentially rewarding options: they show a high preference for the rewarding option with the occasional exploratory visit to the non-rewarding option.

## How do other animal species compare with the bats?

There are several key points of similarity between the bats’ performance on the serial reversal task and that of other animals. Bumblebees showed improvement primarily through a reduction in the perseverative errors Chittka (n.d.) on a colour reversal task. Notably, this ability to improve at the task seems to be due to a large number of trials, just as we had in our experiment. When the task was done with very small trials numbers, both bumblebees (Couvillon and Bitterman 1986) and honeybees (Mota and Giurfa 2010) stopped discriminating and began responding to both the rewarding and non-rewarding stimuli at chance levels. Several different species of birds also showed performance on this task that was similar to the bats. Corvids (Bond, Kamil, and Balda 2007) showed both a decrease in perseverative errors as well as an increase in performance as they experienced successive reversals. The improvement in performance however was seen only in a colour reversal task and not a spatial reversal task. Two species of Great Tits did even better on a spatial reversal task than the Corvids: both trials within a reversal block and reversal number had a positive effect on the proportion of visits to the rewarding option (Hermer et al. 2018b), and similar performance was seen in pigeons on a colour reversal task (Diekamp, Prior, and Güntürkün 1999). Among mammals, a decrease in perseverative errors was seen both in marmosets on a visual reversal task (Clarke, Robbins, and Roberts 2008) and in rats on a spatial reversal task (Castañé, Theobald, and Robbins 2010).

## ‘Reveral learning’ in the wild

Under natural conditions bats exploit the flowers of many different species of plants that vary in the flowering season, flowering duration, the number of flowers that bloom per night, and the quantity of nectar they provide. Observations of foraging behaviour at *Agave desmettiana* flowers show that the visitation rates of bats to flowers depends heavily on nectar volume (Lemke 1984). Feeding rates are high in the first four hours of the night after sunset, and decline sharply when nectar volumes approach 50% of what they were at the start of the night. However when a nectar source is non-depleting, bats visit them at a significantly higher frequency, and for 3-5 hours longer than flowers that were depleted normally. The opposite pattern was seen when flowers were prematurely depleted: bats stopped feeding at these flowers earlier than at control flowers, and visited neighbouring flowers with a higher frequency. This is a sensible strategy, given that Agave flowers only replenish their nectar levels at a rate of 30% in 20 hours (equally over day and night). When rates of nectar replenishment are higher, bats are capable of detecting it and returning the same flowers sooner [**Ulf’s thesis** - how to cite a thesis?], utilizing their excellent spatial memory to find the flower again, adjusting the time interval between successive visits based on the secretion rate of the experimental flowers. It would seem that a ‘win-stay; lose-shift’ strategy is ideal for the bats’ natural foraging ecology. But flower nectar levels in nature are more likely to decline at a perceptible rate, rather than suddenly drop to zero; furthermore, the perception of flower nectar volume is subject to Weber’s Law (Toelch and Winter 2007). Thanks to these factors, a foraging bat would need to make more than one visit to a flower (as the optimal ‘win-stay; lose-shift’ strategy requires) to perceive that nectar levels have been depleted so much that future visits will not be profitable.

## Learning mechanisms in a reveral learning task

Previous work has shown that the bats’ behaviour can be described well by a choice-history dependent behavioural model (Nachev et al. 2017). In such a model, an animal has a estimate of the profitability of the available option, based on its memory of past reward at these options. The past experiences are weighted by recency: the most recent experience affects the estimate of the option more than those in the past. The animal updates these estimates faster or slower depending upon a parameter for learning rate. The animal then makes a choice based on these estimates using a Softmax decision rule. An additional parameter, the exploration rate or ‘temperature’ parameter determines how frequently an animal makes an exploratory visit, i.e., a visit to an option that is estimated at the time of choice to be the less rewarding or non-rewarding one.

Our results, both from the confirmatory and the exploratory analyses seem to indicate that that reinforcement learning, which is choice-history dependent (Worthy and Maddox 2014) is playing a role in the animals’ behaviour, especially at the beginning of the experiment before second-order learning has happened. As they learn the task the bats appear to approach the strategy of win-stay; lose-shift, wherein decision-making is dependent only on the last one experience and not on previous choice-history. The key word is *approach*. We suggest that through second-order learning, the bats shift from a reinforcement-learning strategy to a strategy that approximates win-stay; lose-shift. That is, the estimate of the profitability of an option is weighted more and more towards the most recent experience of that option as the animals learn the rule behind the reversals.

# References

Audet, Jean-Nicolas, and Louis Lefebvre. 2017. “What’s Flexible in Behavioral Flexibility?” *Behavioral Ecology* 28 (4): 943–47. <https://doi.org/10.1093/beheco/arx007>.

Baker, Herbert G., Irene Baker, and Scott A. Hodges. 1998. “Sugar Composition of Nectars and Fruits Consumed by Birds and Bats in the Tropics and Subtropics1.” *Biotropica* 30 (4): 559–86. https://doi.org/<https://doi.org/10.1111/j.1744-7429.1998.tb00097.x>.

Bitterman, M E. 1964. “THE EVOLUTION OF INTELLIGENCE.” *SCIENTIFIC AMERICAN*, 11.

Bond, Alan B., Alan C. Kamil, and Russell P. Balda. 2007. “Serial Reversal Learning and the Evolution of Behavioral Flexibility in Three Species of North American Corvids (Gymnorhinus Cyanocephalus, Nucifraga Columbiana, Aphelocoma Californica).” *Journal of Comparative Psychology* 121 (4): 372–79. <https://doi.org/10.1037/0735-7036.121.4.372>.

Boulougouris, Vasileios, Jeffrey W. Dalley, and Trevor W. Robbins. 2007. “Effects of Orbitofrontal, Infralimbic and Prelimbic Cortical Lesions on Serial Spatial Reversal Learning in the Rat.” *Behavioural Brain Research* 179 (2): 219–28. <https://doi.org/10.1016/j.bbr.2007.02.005>.

Boussard, Annika, Séverine D. Buechel, Mirjam Amcoff, Alexander Kotrschal, and Niclas Kolm. 2020. “Brain Size Does Not Predict Learning Strategies in a Serial Reversal Learning Test.” *The Journal of Experimental Biology* 223 (15): jeb224741. <https://doi.org/10.1242/jeb.224741>.

Bürkner, Paul-Christian. 2017. “**Brms** : An *r* Package for Bayesian Multilevel Models Using *Stan*.” *Journal of Statistical Software* 80 (1). <https://doi.org/10.18637/jss.v080.i01>.

Castañé, Anna, David E. H. Theobald, and Trevor W. Robbins. 2010. “Selective Lesions of the Dorsomedial Striatum Impair Serial Spatial Reversal Learning in Rats.” *Behavioural Brain Research* 210 (1): 74–83. <https://doi.org/10.1016/j.bbr.2010.02.017>.

Chittka, Lars. n.d. “SENSORIMOTOR LEARNING IN BUMBLEBEES: LONG-TERM RETENTION AND REVERSAL TRAINING,” 10.

Chow, Pizza K. Y., Lisa A. Leaver, Ming Wang, and Stephen E. G. Lea. 2015. “Serial Reversal Learning in Gray Squirrels: Learning Efficiency as a Function of Learning and Change of Tactics.” *Journal of Experimental Psychology: Animal Learning and Cognition* 41 (4): 343–53. <https://doi.org/10.1037/xan0000072>.

Clarke, H. F., T. W. Robbins, and A. C. Roberts. 2008. “Lesions of the Medial Striatum in Monkeys Produce Perseverative Impairments During Reversal Learning Similar to Those Produced by Lesions of the Orbitofrontal Cortex.” *Journal of Neuroscience* 28 (43): 10972–82. <https://doi.org/10.1523/JNEUROSCI.1521-08.2008>.

Couvillon, P. A., and M. E. Bitterman. 1986. “Performance of Honeybees in Reversal and Ambiguous-Cue Problems: Tests of a Choice Model.” *Animal Learning & Behavior* 14 (3): 225–31. <https://doi.org/10.3758/BF03200062>.

Czaczkes, Tomer J., and Jürgen Heinze. 2015. “Ants Adjust Their Pheromone Deposition to a Changing Environment and Their Probability of Making Errors.” *Proceedings of the Royal Society B: Biological Sciences* 282 (1810): 20150679. <https://doi.org/10.1098/rspb.2015.0679>.

Dhawan, Sandeep S., David S. Tait, and Verity J. Brown. 2019. “More Rapid Reversal Learning Following Overtraining in the Rat Is Evidence That Behavioural and Cognitive Flexibility Are Dissociable.” *Behavioural Brain Research* 363 (May): 45–52. <https://doi.org/10.1016/j.bbr.2019.01.055>.

Diekamp, B., H. Prior, and O. Güntürkün. 1999. “Functional Lateralization, Interhemispheric Transfer and Position Bias in Serial Reversal Learning in Pigeons ( Columba Livia ).” *Animal Cognition* 2 (4): 187–96. <https://doi.org/10.1007/s100710050039>.

Helversen, O. v., and H. -U. Reyer. 1984. “Nectar Intake and Energy Expenditure in a Flower Visiting Bat.” *Oecologia* 63 (2): 178–84. <https://doi.org/10.1007/BF00379875>.

Helversen, O. von, L. Winkler, and H. J. Bestmann. 2000. “Sulphur-Containing ‘Perfumes’ Attract Flower-Visiting Bats.” *Journal of Comparative Physiology A* 186 (2): 143–53. <https://doi.org/10.1007/s003590050014>.

Hermer, Ethan, Maxime Cauchoix, Alexis S Chaine, and Julie Morand-Ferron. 2018a. “Elevation-Related Difference in Serial Reversal Learning Ability in a Nonscatter Hoarding Passerine.” *Behavioral Ecology* 29 (4): 840–47. <https://doi.org/10.1093/beheco/ary067>.

———. 2018b. “Elevation-Related Difference in Serial Reversal Learning Ability in a Nonscatter Hoarding Passerine.” *Behavioral Ecology* 29 (4): 840–47. <https://doi.org/10.1093/beheco/ary067>.

Kinoshita, Seijiro, Chihiro Yokoyama, Daiki Masaki, Tatsuhisa Yamashita, Hideto Tsuchida, Yasuhito Nakatomi, and Kenji Fukui. 2008. “Effects of Rat Medial Prefrontal Cortex Lesions on Olfactory Serial Reversal and Delayed Alternation Tasks.” *Neuroscience Research* 60 (2): 213–18. <https://doi.org/10.1016/j.neures.2007.10.012>.

Kunz, Thomas H., and M. Brock Fenton. 2005. *Bat Ecology*. University of Chicago Press.

Lemke, Thomas O. 1984. “Foraging Ecology of the Long-Nosed Bat, Glossophaga Soricina, With Respect to Resource Availability.” *Ecology* 65 (2): 538–48. https://doi.org/<https://doi.org/10.2307/1941416>.

Lucon-Xiccato, Tyrone, and Angelo Bisazza. 2017. “Sex Differences in Spatial Abilities and Cognitive Flexibility in the Guppy.” *Animal Behaviour* 123 (January): 53–60. <https://doi.org/10.1016/j.anbehav.2016.10.026>.

Mota, Theo, and Martin Giurfa. 2010. “Multiple Reversal Olfactory Learning in Honeybees.” *Frontiers in Behavioral Neuroscience* 4. <https://doi.org/10.3389/fnbeh.2010.00048>.

Nachev, Vladislav, Kai Petra Stich, Clemens Winter, Alan Bond, Alan Kamil, and York Winter. 2017. “Cognition-Mediated Evolution of Low-Quality Floral Nectars.” *Science* 355 (6320): 75–78. <https://doi.org/10.1126/science.aah4219>.

Ringler, Eva, Andrius Pašukonis, W. Tecumseh Fitch, Ludwig Huber, Walter Hödl, and Max Ringler. 2015. “Flexible Compensation of Uniparental Care: Female Poison Frogs Take over When Males Disappear.” *Behavioral Ecology* 26 (4): 1219–25. <https://doi.org/10.1093/beheco/arv069>.

Rumbaugh, Duane M., E. Sue Savage-Rumbaugh, and David A. Washburn. 1996. “Toward a New Outlook on Primate Learning and Behavior: Complex Learning and Emergent Processes in Comparative Perspective .” *Japanese Psychological Research* 38 (3): 113–25. <https://doi.org/10.1111/j.1468-5884.1996.tb00016.x>.

Strang, Caroline G., and David F. Sherry. 2014. “Serial Reversal Learning in Bumblebees (Bombus Impatiens).” *Animal Cognition* 17 (3): 723–34. <https://doi.org/10.1007/s10071-013-0704-1>.

Tait, David S., Eric M. Bowman, Lorenz S. Neuwirth, and Verity J. Brown. 2018. “Assessment of Intradimensional/Extradimensional Attentional Set-Shifting in Rats.” *Neuroscience & Biobehavioral Reviews* 89 (June): 72–84. <https://doi.org/10.1016/j.neubiorev.2018.02.013>.

Tello-Ramos, Maria C., Carrie L. Branch, Dovid Y. Kozlovsky, Angela M. Pitera, and Vladimir V. Pravosudov. 2019. “Spatial Memory and Cognitive Flexibility Trade-Offs: To Be or Not to Be Flexible, That Is the Question.” *Animal Behaviour* 147 (January): 129–36. <https://doi.org/10.1016/j.anbehav.2018.02.019>.

Thiele, Johannes, and York Winter. 2005. “Hierarchical Strategy for Relocating Food Targets in Flower Bats: Spatial Memory Versus Cue-Directed Search.” *Animal Behaviour* 69 (2): 315–27. <https://doi.org/10.1016/j.anbehav.2004.05.012>.

Toelch, Ulf, Kai Petra Stich, Clifton Lee Gass, and York Winter. 2008. “Effect of Local Spatial Cues in Small-Scale Orientation of Flower Bats.” *Animal Behaviour* 75 (3): 913–20. <https://doi.org/10.1016/j.anbehav.2007.07.011>.

Toelch, Ulf, and York Winter. 2007. “Psychometric Function for Nectar Volume Perception of a Flower-Visiting Bat.” *Journal of Comparative Physiology A* 193 (2): 265–69. <https://doi.org/10.1007/s00359-006-0189-3>.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. “Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC.” *Statistics and Computing* 27 (5): 1413–32. <https://doi.org/10.1007/s11222-016-9696-4>.

Voigt, C. C., and Y. Winter. 1999. “Energetic Cost of Hovering Flight in Nectar-Feeding Bats (Phyllostomidae: Glossophaginae) and Its Scaling in Moths, Birds and Bats.” *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology* 169 (1): 38–48. <https://doi.org/10.1007/s003600050191>.

Voigt, Christian C., Detlev H. Kelm, and G. Henk Visser. 2006. “Field Metabolic Rates of Phytophagous Bats: Do Pollination Strategies of Plants Make Life of Nectar-Feeders Spin Faster?” *Journal of Comparative Physiology B* 176 (3): 213–22. <https://doi.org/10.1007/s00360-005-0042-y>.

Winkler, David W, Christian Jørgensen, Christiaan Both, Alasdair I Houston, John M McNamara, Douglas J Levey, Jesko Partecke, et al. 2014. “Cues, Strategies, and Outcomes: How Migrating Vertebrates Track Environmental Change,” 15.

Winter, Y. 2005. “Foraging in a Complex Naturalistic Environment: Capacity of Spatial Working Memory in Flower Bats.” *Journal of Experimental Biology* 208 (3): 539–48. <https://doi.org/10.1242/jeb.01416>.

Winter, Y., and O. von Helversen. 1998. “The Energy Cost of Flight: Do Small Bats Fly More Cheaply Than Birds?” *Journal of Comparative Physiology B* 168 (2): 105–11. <https://doi.org/10.1007/s003600050126>.

Worthy, Darrell A., and W. Todd Maddox. 2014. “A Comparison Model of Reinforcement-Learning and Win-Stay-Lose-Shift Decision-Making Processes: A Tribute to W.K. Estes.” *Journal of Mathematical Psychology* 59 (April): 41–49. <https://doi.org/10.1016/j.jmp.2013.10.001>.