Nectar-feeding bats learn the rule behind serial reversals

Shambhavi Chidambaram1,2**\***, Sabine Wintergerst3, Alex Kacelnik4, York Winter1,2, Vladislav Nachev1

1 Institute of Biology, Humboldt University, Berlin, Germany

2 Berlin School of Mind and Brain, Humboldt University, Berlin, Germany

3 Fairchild Tropical Botanic Garden, Miami, USA

4 Department of Zoology, University of Oxford

**\*For correspondence:** [shambhavi.chidambaram@hu-berlin.de](mailto:shambhavi.chidambaram@hu-berlin.de) [I am not sure yet, if I should not be the corresponding author; you are already on your way to new waters and the topic stays with me.]

**Present Address:** Institute of Biology, Humboldt University, Philippstr. 13, 10115 Berlin, Germany

# Abstract

Flowering plants stay in one place, but the amount of nectar a flower contains varies as it is emptied, replenishes nectar, or withers and dies [maybe the the three word “wither and die” distracts from the central topic]. Nectar-feeding animals such as the Neotropical bat *Glossophaga commissarisi [maybe not introduce species name at this place. It is a distracting detail when the reader still wants to understand what the scientific question of the study is ]* benefit from perceiving and responding flexibly to these changes of local food availability [is this a “strong enough” scientific goal for this study? Can this not be formulated more succinctly?]. We aimed to demonstrate this through a spatial serial reversal learning task with wild *G. commissarisi* individuals. The bats were trained to obtain rewards from two artificial flowers. At any given time only one of the flowers provided a reward. After the bats had experienced the rewarding properties of the flowers for some time, a reversal happened: the rewarding flower became non-rewarding and vice versa. These reversals of rewarding properties occurred repeatedly.

We found strong evidence that the bats learn the rule underlying reversing: when a food location suddenly dried up they quickly/abruptly abandoned visiting it and switched to the alternate option. This shows that they can ‘learn to learn.’ After experiencing multiple such serial reversals, bats switched faster and faster to the rewarding option after a reversal, and made more choices for the rewarding flower overall. In fact, this increase in the number of choices for the rewarding flower became high enough that it could not increase much further: a ceiling effect was seen in the animals’ preferences. In the trade-off between exploiting options known to be rewarding and exploring [this experimental setting does not really have “new” options] options known to be potentially rewarding, the animals’ behaviour approached a ‘best of both worlds’ strategy. The bats showed near-exclusive preference for the rewarding option, to which they changed rapidly after a reversal as they had learned and relied on the rule behind the change in their environment. Whether nectar-feeding bats can understand other predictable but more complex patterns caused by the determinisitic process of floral nectar secretion , or even potentially anticipate and predict them, is exciting grounds for further research. [last sentence may need additional work; again, it should be much more succinct; “environmental changes” is too vague, especially since we know a lot about the driving processes]

# Introduction

Nectar-feeding bats face the challenge of exploiting resources that continually change with time, but are relatively predictable in space. Floral nectar content varies within any given night, as well from night to night; flowering plants are stationary, however, they bloom seasonally and single flowers on plants themselves wither and die every day or every few days. Bats need to detect the changing reward contingencies in their environment and adjust their behaviour accordingly. This ability of animals to recognize and respond to changes is crucial for behavioural efficiency, and includes cognitive or behavioural flexibility (Tello-Ramos et al. 2019). One experimental protocol that has been widely used to demonstrate cognitive/behavioural flexibility is reversal learning, a task that is a simple experimental analogue to some of the decision-making bats do in their natural environment.

Reversal learning essentially involves two separate stages of learning. An animal must first learn about multiple stimulus response associations. These can be two spatial locations that potentially provide a food reward. The animal is then confronted with the next step which is a situation where in fact only one of the two options is profitable. In this phase of the experiment, it is optimal to only pay attention to this one option and ignore other possibilities. However, after a longer series of such repetitive stimulus response actions, reinforcement stops. The optimal strategy for the animal is to then immediately abolish its recent behavior and instead perform another stimulus response action that in the past had proven profitable. It is this ability to abruptly terminate an ongoing behavior in favor of switching to another behavior that the animal should learn for efficient exploitation. In a serial reversal learning procedure the reward contingencies reverse repeatedly. An animal that responds to the rewarding stimulus more frequently than to the non-rewarded stimulus receives food more often, and performs more efficiently on the task; a choice to the non-rewarded stimulus is, in the context of the task, an ‘error.’ This performance critically depends on the animal ‘s ability to switch abruptly after a reversal.

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 both corvids (Bond, Kamil, and Balda 2007) and cowbirds (Lois-Milevicich et al. 2021); 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): animals considered then to be ‘higher,’ like pigeons, rats and monkeys showed a progressive improvement on the task while species considered to be ‘lower,’ 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.

Improvement in the reversal learning task is clearly demonstrable, and is therefore a meaningful criterion when comparing the performance of different species. First-order learning happens when an animal perceives a stimulus and learns which behavioural response leads to a reward - 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. This is the ability to actually choose efficiently among potential behavioural responses in any given situation. In serial reversal learning the same stimuli are successively paired with a reward and then not paired with a reward. Under such conditions animals must learn the second-order rule to quickly abandon a suddenly unsuccessful response that previously resulted in reward. A rule that can maximize reward in a deterministic, reversal learning schedule is ‘win-stay; lose-shift’ (WSLS): repeat the behaviour which produces a reward until it stops being rewarding, and then immediately switch to another behavioural response. Such a rule means in practice that the animal makes exactly one ‘error’ per reversal. After learning the task, the perfectly optimal animal will first exclusively respond to the rewarded stimulus. After the first experience of a lack of reward (the error), the animal will change its actions and exclusively respond to the other stimulus which is now paired with a reward.

Animals are rarely, if ever, so optimal in real life. Large and deterministic changes are uncommon in nature and more gradual shifts in behaviour maybe better suited to natural environmental changes. Errors in reversal learning tasks are common, and can occur in different ways. An animal can perseverate: continue to choose a non-rewarding cue after the reversal has occurred or they can make anticipatory errors: changing their behaviour and choosing the other cue before a reversal has occurred. 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).

Primates show a very interesting difference in between first- and second-order learning. In one study, 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 up to either a 67% or 84% level of correct choices for the rewarded option. At these two different levels of training they were then given a single reversal of reward contingencies. If the primates’ behaviour was driven mainly by first-order learning, they should switch to the newly-rewarding stimulus more quickly when their training had not yet reached a high level of performance, i.e., 67%. The opposite would be true if second-order learning or rule-learning was occurring: knowing the reversal rule, the primates would more quickly adapt to the reversed option when they had reached a high level of performance in their training, i.e., 84%. The results showed that Prosimian species tended to perform better when trained to 67% and apes when trained to 84% level of correct choices before the reversal.

Progressive improvement on the serial reversal task as more reversals are experienced, which is evidence of rule-learning, has been shown in many different species: bumblebees (Strang and Sherry 2014); great tits from different environments (Hermer et al. 2018a); three different species of corvids (Bond, Kamil, and Balda 2007); rats (Castañé, Theobald, and Robbins 2010); and marmosets (Clarke, Robbins, and Roberts 2008). A decrease in the number of errors to reach a criterion (a pre-decided number of correct choices) with multiple reversals is seen in mice (Caglayan, Stumpenhorst, and Winter 2021), even when each reversal is done with a different pair of stimuli, indicating the formation of a learning set.

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, behavioural flexibility 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 (Voigt and Winter 1999; Voigt, Kelm, and Visser 2006), due to the energetic demands of flight (Winter and von Helversen 1998; von Helversen and Reyer 1984). Many plants visited by bats put out only a few flowers every night, but their flowering seasons that can last for several months. Bat-pollinated flowers can secrete up to 1-2 mL of nectar a night (von Helversen 1993) but as only small droplets are available each time they are visited (Voigt, Kelm, and Visser 2006), the bats make several hundred flower visits per night. Indeed, a certain time after a flower is emptied the nectar-levels are replenished, so bats can visit the same flower multiple times. The bats relocate the flowering plant primarily through their excellent spatial memory in addition to local echo-acoustic cues to find the individual flowers (Thiele and Winter 2005; von Helversen and von Helversen 1999). The longer the bat waits, the more the flower refills but the higher the likelihood that a competitor could exploit the flower first. To make repeated, profitable visits to a flower, a bat must remember both the location of the flower and, ideally, 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. Though this is a simplified situation with no inter-individual competition, the behavioural flexibility required by the experimental protocol is the kind of flexibility required in the natural foraging circumstances of a nectar-feeding bat.

To investigate the flexibility of responding to depleting food locations we carried out a serial reversal learning task with wild *G. commissarisi* individuals. The bats were given two potentially rewarding options to choose between. At the start of the night, only one of the options was rewarding, the ‘S+’ option and the other was not rewarding, the ‘S-’ option. After a certain number of visits had been made by the bats, the reward contingencies reversed without any signal or cue to the bats: the previously rewarding option was now unrewarding and the previously unrewarding option was rewarding, and this reversal happened five times in a night.

Our aims with this experiment were as follows. First, we wanted to test how the bats would alter their preference between the two reward sources according to their transient rewarding properties. Second, if the bats demonstrated flexibility in responding to the reversals, we wanted to explore how this was reflected in their decision-making. What would be the relative number of visits made to the rewarding and the non-rewarding options, and how swiftly would behavioural allocation between the two options change after a reversal? Third, 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 converge to the maximal payoff of one error per reversal achievable through win-stay, loose-shift [avoid abbreviations wherever possible; this is easier on the reader and “printing space” is not a problem anymore in today’s digital publications]?

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. We reasoned that there is a difference between the first visits of a night, before any experience of a change in contingencies, and all the subsequent visits after at least one reversal had occurred. We statistically tested for this difference in the bats’ choice behaviour and compared the results of the confirmatory and exploratory analyses.

# Methods

## Study site and subjects

The experiment took place at La Selva Biological Field Station, Province Heredia, Costa Rica. Bats of the species *Glossophaga commissarisi* were captured from the wild and retained in a flight cage through the experiment. The bats were attracted to a particular location in the forest using sugar-water (see **Reward** below) in chicken feeders with dimethyl disulfide odor on cotton swabs as bait (von Helversen, Winkler, Bestmann, 2000 *J Comp Physiol A 186: 143-53*) and then caught in mist-nets. The bats were sexed on capture and housed in two outdoor, meshed flight-cages (4 x 6 m) under ambient light conditions. All individuals were weighed and marked with radio frequency identification (RFID) tags placed as collars around their necks.

See comment below: Maybe this is the place to tell the reader that experiments were conducted with groups of bats together in the same experimental cage.

16 bats participated in the main experiment. At the end of the experiment, the RFID collars were removed and the bats were released back into the wild . 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 (seven males and seven females) completed the experiment. Permission for this research 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, hereafter referred to as ‘nectar.’ The sugar consisted of a 1:1:1 mass-mixture of sucrose, fructose and glucose. The nectar was thus similar in composition and concentration to the nectar produced by wild chiropterophilous plants (Baker, Baker, and Hodges 1998). Every night groups of six bats were also given ad-libitum access to 10 ml of supplemental food: 3.5g of hummingbird food (NektarPlus, Nekton) and 3.5g of milk powder (Nido 1+, Nestle) in 10 mL of water. They were also given a small bowl of locally-sourced bee pollen. .

### 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. At this distance bats can easily discriminate neighbouring flowers (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 photo gate; an electronic pinch valve through which a silicon 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 step volume 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 a set of five pinch valves [can you cite a publication with a schematic of the pump setup? A paper with Vladi?]. 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 every day and connected to the syringe through the valves. A syringe refilled automatically from this reservoir during an experimental night.

*Move text on “Every day at around 1000h …” from below, to up here. It is better to keep the whole topic of “experimental nectar supply” in one place.*

When a tagged bat approached a flower, the individual RFID number was read. If the bat then poked its nose into the flower and interrupted the light beam, 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 identified a bat and the photo gate was triggered, would a reward be given. The flowers and the pump were connected to a Windows PC, which ran the experimental programs and the program used to automatically flush, clean and fill the pump and tubing system ( PhenoSoft Control, PhenoSys). The raw data were recorded as comma-separated value (CSV) files.

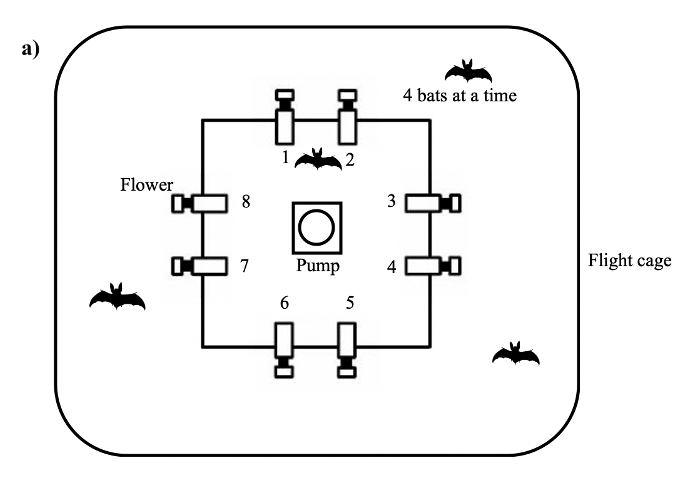


Figure 1: Schematic of the cage and flower set-up

## Experimental procedure

*Every day at around 1000 h the old nectar was emptied from the system. The system was rinsed and filled with plain water until 1500 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. -> move this remaining information about operating the nectar pump system to above*

Four bats of the same sex participated in a flight cage in an experiment together . Two groups were run in parallel, 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 to only one of the four bats in the cage. After the system was filled with fresh nectar at approximately 1700 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 nightfall at approximately 1800 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), and 2.4 minutes (SD = ±0.04) for the two pumps.

About 1 % (SD = ±0.74) of all visits made by the bats over all three experimental nights were wrongly unrewarded: the bats did not receive a reward during these visits even when they were made to a flower assigned to them that was rewarding at the time. This happened either during the pump refill times or when the pump was moving to reward a visit made by another bat that happened almost at the same time. Such events did not count towards the total of 50 visits in a block.

Reviewers have asked about this before. Somewhere should be the explicit statement. “The data collected during this study were the computer logged events of feeder visits with their time stamp, animal ID, photo gate interruption duration, and amount of nectar dispensed.”

## Experimental design

The experiment proceeded through the following stages.

### Training

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 dimethyl disulphide. This is a chemical attractant produced by many bat-pollinated flowers (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 *[just a comment: I think that this is actually a relatively “useless” procedure. It only works well in our lab where our captive bats are familiar to drinking honey water and know the smell; “wild bats” have not made that experience]*, interrupt the photo gate and trigger a nectar reward. By the end of the night all the bats had found the flowers and learned to trigger rewards.

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 these two flowers assigned to it would be rewarding from this stage of training until the end of the experiment. This stage was similar to the previous one, except the bats could only trigger a reward at their assigned flowers, and the chemical attractant and honey were not used.

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

### Serial Reversal Learning Task

In the serial reversal learning task the bats had continuous access to two flowers, one that gave a 40 L nectar reward and one that remained empty. 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 (regardless of relative allocation) 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 in the flight cage which were always unrewarding to that particular bat. Each block of 50 visits to the assigned two flowers, either at the start of each night [why is the first block of 50 visits in a night termed “reversal block” -> this sentence does not make sense to me] or between reversals, was termed a ‘reversal block’. There were six blocks and five reversals per night, unless the bat ceased visiting earlier. This was repeated for three consecutive nights. The same flower started the sequence every night. Consequently the last flower to be rewarding one night was non-rewarding at the start of the next.

## Data analysis

All the statistical models were fitted 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). The only exception was one model in the Supplementary section, calculating the correlation between behaviour just before and after the reversal, which was fitted using the R package MCMCglmm (Hadfield 2010); the details of this model are provided in the Supplementary section. Model comparison was done using leave-one-out cross-validation, implemented in brms using the package loo (Vehtari, Gelman, and Gabry 2017).

Except when stated otherwise, all the visits made by the bats to their two assigned flowers - up to 300 – during a night were included in the analyses (some of the bats did not complete all 300 visits on every night). The bats made some visits and approaches to the flowers that were not assigned to them but they were the minority, and not considered for the analysis (see **Supplementary Material** for details). At the end of each 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 of ten visits, in order to examine a bats’ behaviour within each block. Generalized linear mixed-models were used for the analyses (see **Supplementary Material** for details of the model fitting). We report here the mean as a measure of central tendency and the 89% quantile-based credible intervals for the parameters. 89% boundaries are the default for reporting credible intervals and are no more arbitrary than the convention of using 89% [if these two values are supposed to be identical then I do not understand the message of this sentence] boundaries (McElreath 2020).

We investigated the effect of experimental night and reversal block on the number of perseverative visits, i.e., consecutive visits to the previously-rewarding option just after the occurrence of a reversal; and the effect of experimental night, block and bin on the proportion of visits to the rewarding flower. The proportion of visits to the rewarding flower was calculated as the number of visits to the S+ divided by total number of visits to both the S+ and S-, and we denote this as the Proprew:

A picture containing table

Description automatically generated

After examining these results, we conducted further post-hoc exploratory analyses. We investigated the effect of night, block, bin, and the first night and first block of the night as separate predictor variables for Proprew: 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.

A few of the results report 95% confidence intervals, as opposed to credible intervals, and these are noted specifically. The confidence intervals were calculated by non-parametric bootstrapping without assuming a normal distribution of the data, using the Hmisc package, which is part of the tidyverse system of R packages (Wickham et al. 2019).

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

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

The bats made a very high number of visits to the rewarding option, quickly abandoning an option when it stopped being rewarding after a reversal and switching to the other (Figure 2). A consistent pattern emerged over all three nights: a sharp decrease in the proportion of visits to the formerly rewarding option immediately following a reversal, concomittant with a rapid increase in visits to the newly rewarding option. There are two particularly interesting points about this overall pattern: the behaviour during the first bins of each block; and the changes between the first night and the following two nights.

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 Proprew (the proportion of visits to the rewarding option) averaged across individuals was close to chance: 54.5% [95% CI 46.8, 63]. Within the next ten visits however, Proprew increased to 92.1% [95% CI 87.1, 96.4] and by the last bin of this first block was 99.3% [95% CI 97.8, 100]. Immediately after the first experience of a reversal, the Proprew dropped down to 13.6% [95% CI 8.4, 18.8] in the first ten visits, but came back up to 96.4% [95% CI 92.9, 99.3] by the last bin of this block.

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 Proprew of all the bats was 69.8% [95% CI 64.9, 75.3]. This was significantly higher than random choice and higher than the Proprew in the corresponding bin of the first night. However, the Proprew on the second and third nights also showed the pattern of a decrease immediately after a reversal and then increase to a high proportion overall - 94.8% [95% CI 94.1, 95.5], comparable to the 93% [95% CI 91.8, 94.1] on the first night.

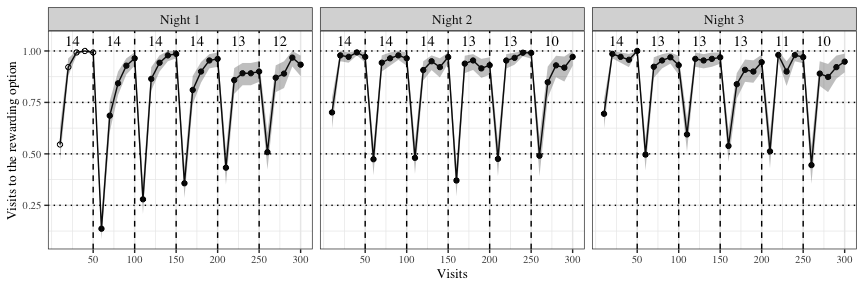


Figure 2: Visits to the rewarding of two options across the three experimental nights. Data are average proportions for bins of ten visits averaged over all 14 individuals. Data are indicated by white points in the first block on the first night before the bats had experienced any reversals; the bin averages of the other blocks are indicated by black points. Numbers indicate the bats that participated in a block. Shading shows 95% confidence intervals. Dashed lines show reversals -> dashed lines should be positioned between points; currently the last bin of a block is incorrectly marked by a dashed line as “reversal”. This because your points are not centered above the mean count of a block. i.e. first block centered above 5, second above 15 etc. This should be changed.

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

As the bats experienced more reversals, they made their first visit to the rewarding flower faster and faster: immediately after a reversal, the number of visits to the previously-rewarding flower (perseverative visits) decreased as the animals experienced more reversals (Figure 3). The results of the statistical analysis (Figure 4) show that both experimental night (1 to 3) and block within a night (1 to 5) had a negative effect on the number of perseverative visits made by the bats, and the effect of the interaction of night and block was positive. Examining the data however, indicates that this change does not seem to occur on the second and third nights. It would appear therefore, that the bats get faster at responding to reversals only the first night until they reach a plateau, which persists over the second and third nights.

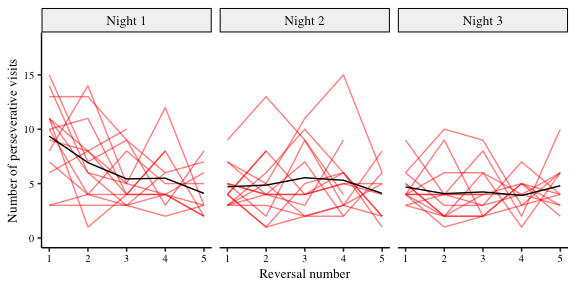


Figure 3: Number of perseverative visits made by the bats after each reversal across all three nights. A perseverative visit is a visit to the previously-rewarding option after a reversal and before any visit to the newly-rewarding option. By definition there were no perseverative visits in the first block of a night. The red lines show individual data , the black lines the group average (N = 14). Reversals were between the two flowers assigned to an individual bat

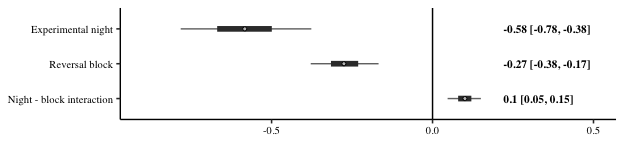


Figure 4: Forest plot of the estimates of the effect of night and block on perseverative visits; a negative value indicates a decrease of perseverative visits with successive nights and blocks. Circles represent the means of the posterior distributions, thick horizontal lines represent 50% credible intervals, and thin horizontal lines 89% credible intervals. The numbers are the means and 89% credible intervals of each effect

### Bats make more visits to the rewarding option overall as they experience more reversals

As the experiment proceeded and the bats experienced more and more reversals, the overall proportion of the animals’ choices for the rewarding flower increased. This was true within each reversal block, within each experimental night, and across nights: as the bats had experienced more reversals on more nights, and had made more visits after a reversal had occurred, the proportion of those visits made to the rewarding option increased. Though this increase consistently occurred, the size of the increase became smaller. That is, there was a progressively smaller ‘improvement’ in performance. This was likely due to a ceiling effect, as the proportion of visits to the rewarding option [nearly] reached 1 (Figure 2).

The results of the statistical analysis are shown in Figure 5. All three variables, night, block and bin, had a positive effect on the Proprew. The coefficients of the interactions of experimental night and block, block and bin within the block, and experimental night and bin within the block, were small and all negative.

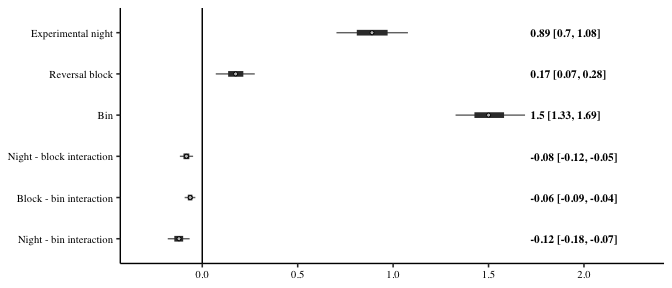


Figure 5: Forest plot of the estimates of the effect of night, block and bin on the proportion of visits to the rewarding option; a positive effect indicates the visits to the rewarding option increased with successive nights, blocks and bins (within each block). Circles show the means of the posterior distributions, thick horizontal lines the 50% credible intervals and thin horizontal lines the 89% credible intervals. Numbers are slope estimates with 89% credible intervals

## Exploratory analysis

### 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 very first experimental night: before this night the bats had never experienced a reversal at all . Therefore, we performed an exploratory analysis to specifically explore the effects of the first block of a night and the first night, expanding on our previously-described confirmatory analysis.

In addition to the effect of experimental night, block and bin captured by the model described earlier, the exploratory analysis showed that the bats made more visits to the rewarding flower in the first block of every night (before they had experienced any reversals that night), but *not* in the first night overall [I think it is better to put this second statement into a separate sentence; I wonder of this second sentence in its current form is not too condensed and thus not easy to understand]. However, the highest number of visits to the rewarding flower (the highest Proprew) was in the first block of 50 visits of the very first night, before any reversals had ever been experienced. Interestingly, this was despite the Proprew in the first ten visits of this block being close to random chance (Figure 2).

You describe the statistical results. But they have only been computed to corroborate the biological results. The way you do this here is something that is often done. But I find this difficult to read. Could you paraphrase all of this in terms of effects that increased or decreased correct choice after reversals. In the end, this is what the reader is interested in. The reader is not interested in the statistics per se.

I suggest that we go together through the following paragraph.

The analysis showed that the main effects of night, block and bin, and their interactions on the Proprew were similar to the previous model (Figure 6). The variable ‘block type’? [“block type” is not mentioned in Fig 6] (i.e., whether a block was the first reversal block of the night or not), had a large positive coefficient, but ‘night type’? [also the parameter “night type” is not mentioned in Fig. 6] (whether a night was the first experimental night or not) did not. Block type and night type had an interaction effect with a positive coefficient, showing the effect of the very first block of the whole experiment.

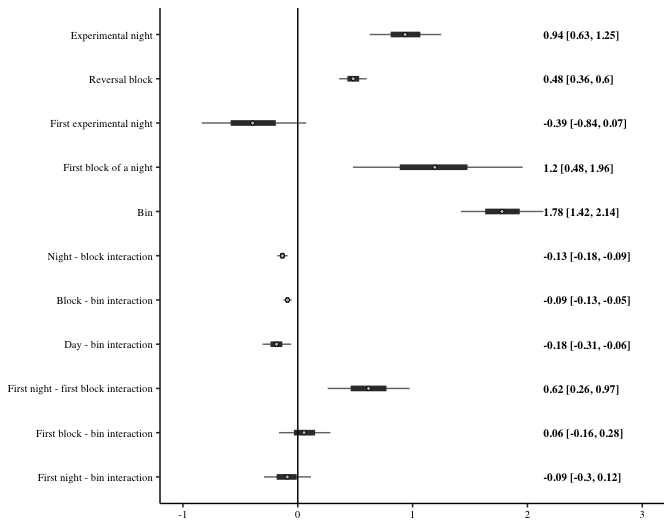


Figure 6: Forest plot of the estimates of the effect of night, block, and bin; and the differential effect of the first night and the first block (‘night-type’ and ‘block-type’ [both terms are not in the graph which I would have expected if you give them here in parentheses]) on the proportion of visits 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. Circles represent the means of the posterior distributions, the thick horizontal lines represent the 50% credible intervals and the thin horizontal lines 89% credible intervals

We compared the predictive accuracy of the model that included night type and block type (Fig. 6) to the model that did not include these (Fig. 5), using leave-one-out cross-validation (LOO). The LOO criterion was lower for the former than the latter, indicating better predictive accuracy of the model including night-type and block-type (Table 1).

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 |

# Discussion

In our experiment wild nectar-feeding bats participated in a spatial 2-AFC serial reversal learning task where the locations of rewards changed after blocks of 50 trials, five times a night.

[Do you really want to bring up the 8 flower array again down here? I wonder if this is not confusing. It might be easier to ignore this and only discuss a bat’s choices between two options]

Bats detected and responded to the changing reward contingencies, evidencing first-order learning. As the animals experienced more reversals [no change between experimental nights; only on the first night], their performance improved in two significant ways.

## A faster switch

With each successive reversal they were quicker to switch from the previously-rewarding option to the newly-rewarding option. The number of perseverative visits decreased, meaning that the first visit to the newly-rewarding option happened earlier as more reversals were experienced (Fig. 3 and Fig. 4). Bats improved until the fith reversal. [see comment above] The difference between a large reward and no reward, (‘something’ versus ‘nothing’) is the easiest possible discrimination. 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 this particular type of environmental change. -> would this be a possible place to continue with the paragraph currently at the bottom of the discussion?

## Improved ‘performance’: the effect of experience at different time-scales on the exploration-exploitation tradeoff

When foraging efficiently in a natural context, such as the one we simulated in our experiment, animals face the classic trade-off between exploration and exploitation. The greater the behavioural allocation to the currently better alternative, the lower the ability to immediately detect an increase in reward elsewhere. The balance between these alternatives may be affected by experience. To explore this possibility, we examined whether experiencing repeated alternations had an impact on this trade-off at two different time scales: within a single night and repeatedly across nights.

A visual examination of our results summarized in Figure 2 shows that, to some extent, both effects were [results always in the past tense; only after this study has been published, do its findings become established knowledge and you can refer to them in the present tense as something that has been scientifically established for these animals] present. In the first night, exploitation as measured by the proportion of visits directed at the rewarding feeder [not sure I ever read that term before; let us stay simple rather than coining new terminology]declined at the end of each block of 50 visits

[Is this a significant finding? Can you guide me to the corresponding statistics? In Fig. 2 for the first night I only see a “decline” in one single block, in all the other blocks the proportion of visits to the rewarding feeder still increase at the end of the block.]

. As a consequence the avoidance of the previously rewarding feeder [what type of behaviour is “correction”?] after each reversal became faster. By the second and third nights the bats appeared to show further behavioural changes to have the ‘best of both worlds.’ [What does this “best of both worlds” mean? It needs to be explained, since this is not obvious.] They could reach near-total visitation of the rewarding feeder rapidly within a block, eventually reaching a ceiling, and rapidly avoiding the previously rewarding feeder after a reversal. This approach to a ‘best of both worlds’ [well, one might argure that this is an artifact of your position as an outside observer; from the point of view of the animal there is only one behavior. Since there are two options at which this behaviour can be directed , reducing one option increases the other option; one might argue if this is one or two “worlds”. ]strategy is strong evidence of second-order learning, or learning the rule behind the environmental change. -> I wonder if you should drop the “best of two worlds” argument. In any case it is not needed to explain second order learning. Instead, I become confused when the explanation of a learning mechanism is mixed with considerations of evolutionary value (proximate / ulitimate explanations of behaviour). First, you need to establish that some (proximate) phenomenon is present. After that is done you can speculate how this fits in terms of evolutionary (ultimate) arguments. But do not mix the two different levels. Maybe, at a later stage you could have a paragraph on behavioural ecological implications where you could bring this up. But I am not sure.

One could also argure more along the lines of the underlying mechanism. Throughout a block of trials the animal repeatedly experiences reinforcements at one spatial location. Occasional visits to the alternate location confirm this is barren. These rewards are strong reinforcers of the current behaviour. With second order learning the bats have learnt that even just a single instance or short sequence of absent rewards is the cue to abruptly abandon the recent location of reinforcement in favor of the newly rewarding alternate location.

It might also be interesting to point out that the bats needed about 5 blocks of reversals to reach that performance. This corresponded roughly to one night. This is an important finding that should be explicitly mentioned. This is a parameter where different animal species will differ. After the first night, performance did not further improve.

The results of the statistical analysis support this interpretation. Figure 5 shows the strength of improvement in foraging performance as a function of each of these two time scales. Unsurprisingly, we see that the frequency of choosing the rewarded option increased within each block (as a function of bin). Much more interesting is the fact that this increase occurred also over a night (as a function of reversal block) and -> I am not sure you can claim that since there is a negative interaction with number of night, indicating adaptive foraging on both time scales. -> To avoid that problem with the interpretation of the data in the light of the significant interaction term, I would point out the performance change during the first night only. This is sufficient for your argument.

[Mabe you can write it like this]

Much more interesting is the fact that this increase occurred also over a night (as a function of reversal block) as can be seen in Fig. 2 for the successive blocks of the first night. Thus, foraging efficiency adapted on both time scales within a block of trials and throughout the whole night. There was no evidence for a further increase in nights 2 and 3 as the interaction term xxxxx was negative.

The fact that experience has an effect on the trade-off between exploration and exploitation is underlined by the results of the exploratory analysis. It is evident from Figure 2 that the highest rate of visiting the rewarding location was in the first block of the first night before any reversal had been experienced [maybe not for the manuscript, but for my own interest: could you provide a table with the total numbers of visits to the rewarding feeders for each block? It is not obvious to me when looking at Fig. 2 that the area under the curve is larger for the first block of night 1 as compared to the first block on night 2]. At this point the trade-off is strongly in favour of exploitation and not exploration, as only one flower has repeatedly been experienced as rewarding: the proportion of visits went rapidly from around 50% to nearly 100%. The bats had no information yet about how the environment would change; the possibility that exploration could be profitable arose when the first reversal happened and the hitherto rewarding flower was suddenly empty.

This was true, but to a lesser extent at the start of the second and third nights. Although reversals had happened on the previous night, none had yet happened during the current night. The results of the exploratory analysis represented in Figure 6 demonstrate this. Choice of the rewarding feeder was higher/highest? during the first block of the night, particularly on the first night when reversals had not been experienced?. Second-order learning was only possible /could only happen after reversals had occurred, and before that first-order learning led to the highest possible amount of exploitation.

## 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 show improvement primarily through a reduction in the perseverative errors (Strang and Sherry 2014; Chittka 1998) on a colour reversal task. Notably, this ability to improve at the task seems to be achieved through the large number of trials, just as we had in our experiment. When the task was done with a small number of trials between the reversals, 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) show both a decrease in perseverative errors as well as an increase in preference for the rewarding stimulus as they experienced successive reversals. The improvement in performance however is seen only in a colour reversal task and not a spatial reversal task. Great tits, sampled from two different locations, do 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). A similar performance is seen in pigeons on a colour reversal task (Diekamp, Prior, and Güntürkün 1999). Among mammals, a decrease in perseverative errors is 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). The bats’ improvement on the serial reversal learning task thus seems to follow a similar pattern to the improvement of several other animal species, potentially indicating similar learning mechanisms.

The role of the sensory modality of the experimental stimuli must not be overlooked: we suggest that animals are likely to perform best on a version of the task that uses a sensory modality relevant to their natural foraging ecology and this is consistent with the results of serial reversal experiments with multiple species. Indeed, the transfer of improved performance across stimuli (as seen in the Corvids), is extremely strong evidence of rule-learning, and suggests an important follow-up experiment to the one we have carried out.

Flexibility towards predictable resource fluctuations This study using the serial reversal task evaluated behavioural flexibility of nectar-feeding bats to fluctuations in food resource availability. The results are consistent with previous work of the same species under natural conditions in the same environment (La Selva Biological Field Station, Costa Rica). Free-flying, ID-tagged wild bats interacted with 50 ID-sensor equipped artificial flowers placed over 100 x 100 m in the open forest that varied in their rate of nectar production. The spatial locations of flower types changed with the same pattern every night (Thiele 2006). Within each night bats adapted to the changes in resource availability. However, they needed four nights before they had adapted to the underlying recurring, predictable pattern of resource variability.

Under natural conditions bats exploit the flowers of many different species of plants that vary in their 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 the flowers were artificially replenished so they never got depleted, 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.* I am not sure if these findings are relevant to the experiments. It also seems too trivial: bats go where the food is.

Bats can distinguish flowers that differ in nectar secretion rates and return to more profitable flowers sooner (Tölch 2006), utilizing their excellent spatial memory to find the flower again and adjusting the time interval between successive visits based on the secretion rate of the experimental flowers. Such more frequent visitation of more profitable flowers can be interpreted as matching visitation rate to different fixed ratio reward schedules. It may not necessitate the ability to actually estimate the temporal component of secretion.

In the wild flowers alternate between being rewarding and non-rewarding.

Well, in most cases a flower is emptied with a single visit. Thus a “win-stay” phase during forating is not the norm.

This below paragraph needs rewriting. We can also discuss: which message do you want to bring across?

You could certainly make a case that there are plant species with flowers that hold copious nectar such as *Ochroma* or *Agave*. Also, flowers such as *Vriesea* [now a new genus name] that have remained undetected throughout a night may hold accumulated nectar that requires 20 or more single hovering visit to deplete (the small glossophagine bats normall avoid landing on flowers but rather make repetitive brief hovering visits. Thus win-stay situations, while not the norm, do occur in the natural environment.

Therefore, the ability to sometimes adopt a ‘win-stay, lose-shift’ strategy is likely part of the bats’ natural foraging ecology, just as in a serial reversal task. But the experimental task is not a perfect analogue of the natural environment.

Flower nectar levels in nature are more likely to decline at a perceptible rate, rather than suddenly drop to zero (Lemke 1984) -> No, in the majority of cases flower nectar levels drop abruptly., and [the following is a really different topic. I would put it in a separate sentence] the perception of flower nectar volume is subject to Weber’s Law (Tölch and Winter 2007). Dueto 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. A more complex behavioural strategy might be more profitable in this more complex environment.

## Learning mechanisms in a reversal learning task

This key interpretation of the results must be presented at an earlier stage of the discussion.

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 an 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, and the estimates are updated according to the animal’s learning rate. Our results, supported by both the confirmatory and the exploratory analyses, with bats improving performance within blocks of trials are in line with the expectations of choice-history dependent reinforcement learning (Worthy and Maddox 2014).

As bats were continually confronted with reversals, however, they learned a second reinforced response that was conditioned to a novel cue. This novel cue was the single or repeated absence of expected reinforcement. This absence of reinforcement became a conditioned stimulus that was coupled to the novel conditioned response of switching to the other reward location. This conditioned response of switching location was reinforced. The two conditioned responses acquired through reinforcement learning were therefore: the win-stay response and the lose-shift-response. It was the ability of abruptly switching from one type of response, the win-stay response, to another type of response, the lose-shift response, that required behavioural flexibility. This flexibility allowed the action selection of both conditioned responses in alternation. The result of this could be observed as the win-stay lose-shift behavior of the bats.

A pure win-stay, lose-shift behaviour would result in optimum performance. The bats’ behaviour became a close approximation of win-stay, lose-shift already after the fith reversal by the end of the first night of the experiment.

## Conclusions

I have not gone through the conclusions yet, since this may change.

We demonstrate that nectar-feeding bats bats learn to take the occurrence of spatial reversals into account as they discriminate between potentially rewarding options, showing second-order learning. The bats’ strategy approached the optimum balance between exploration and exploitation: they show a high preference for the rewarding option but also make the occasional exploratory visit to the non-rewarding option.

Nectar-feeding bats are thus not only capable of higher-order learning, but of flexibly applying different behavioural strategies in response to a predictably changing environment.

## Acknowledgements

We thank Alexej Schatz for the programming of the PhenoSoft Control software. We thank the members of the Winter lab, \_\_\_, \_\_ and \_\_\_ for many useful discussions and our colleagues at La Selva Biological Field Station for all their support. We also thank \_\_\_ for their comments and suggestions for the improvement of the manuscript.

## Author Contributions

**SC**: [you participated in the role of helper but not in the role of responsibility; the person responsible for these two points was only Sabine; we should not forget that this was solely her project] formal analysis, data curation, writing - original draft, writing - review and editing. **SW**: conceptualization, experimental methodology, data collection. **AK**: formal analysis, writing - review and editing, supervision. **YW**: conceptualization, experimental methodology, resources, formal analysis, writing - review and editing, supervision. **VN**: formal analysis, data curation, writing - review and editing, supervision.

## Funding

Open Access funding enabled by …

## Availability of data and code

All data and code are available in the Zenodo repository …

# Declarations

## Funding

This work was funded partly by a scholarship from the Deutscher Akademischer Austauschdienst (DAAD) to SC. Support was provided by EXC ….. and EXC ….. [need to look up the two project numbers]

## Conflict of interest

YW owns PhenoSys equity

## Code availability

All data and code are available in the Zenodo repository …

## Open Access

Licenses

# Electronic Supplementary Material

## Visits and approaches to the unassigned flowers

Only two out of the array of eight flowers were assigned uniquely to each bat but all the flowers were accessible to all the animals. The number of approaches to and attempts to get a reward from all the flowers, both assigned and not assigned, is shown in Figure 7.

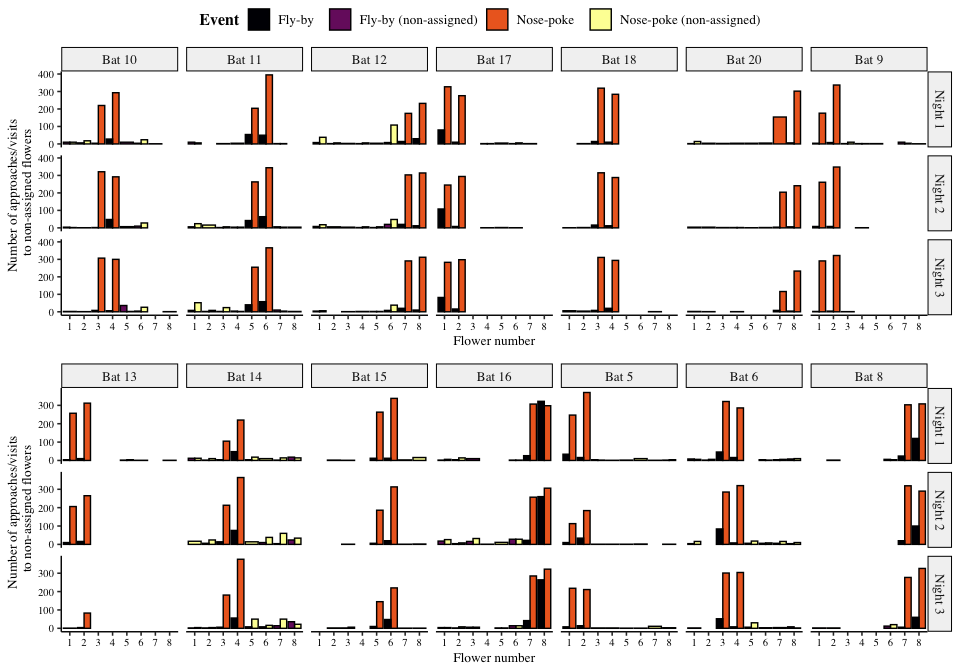


Figure 7: Visits made by the bats to all the flowers, including the ones that were not assigned to them. Yellow bars are nose-pokes at the assigned flowers, where the bats attempted to get a reward by breaking the light-barrier. Purple bars are ‘fly by’ events near the assigned flowers where the bat flew near the flower but did not attempt to get a reward. Orange bars are nose-pokes at the non-assigned flowers and black bars are fly-bys at the non-assigned flowers.

The number of approaches or attempts to get a reward at the non-assigned flowers was a small proportion of the overall number of approaches and reward-attempts at the flowers, less than 10% every night on average as Figure 8 shows.

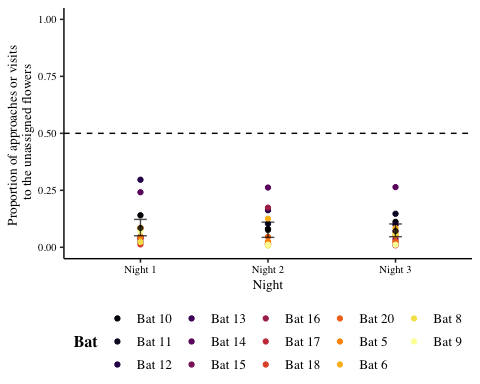


Figure 8: Proportion of visits or approaches to the un-assigned flowers out of the total number of visits or approaches to flowers. Coloured points are data from individual bats. Black points are the mean proportion per night and the error bars are 89% CIs

## Details of the statistical analyses

Weakly informative priors were used for the generalized linear mixed-models in brms. 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.

For the model investigating the effect of experimental night and reversal block on the number of perseverative visits, a negative-binomial likelihood function was used. 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.

The model investigating the effect of experimental night, block and bin on the Proprew (calculated only over the two flowers assigned to a bat) used 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. The model of the change in Proprew in the exploratory analysis was identical to that in the confirmatory analysis, except for the addition of experimental night and block as factor variables. The first night and the first block of every night was treated as one level of the factor variables and the other nights and other blocks of each night as the other level.

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.

## Posterior predictive checks of the models of the Proprew

We carried out a posterior predictive check of the two models of the Proprew in order to check the models’ descriptive adequacy - the agreement of the models with the empirical data, which is a basic goal of modeling (Shiffrin et al. 2008). Posterior samples from the posterior predictive distribution were computed for the data used to fit the models and plotted in Figure 9 and Figure 10. A visual comparison of the two reveals a closer correspondence between the data and the posterior samples of the second model compared to the first, especially in the first block of the experimental nights.

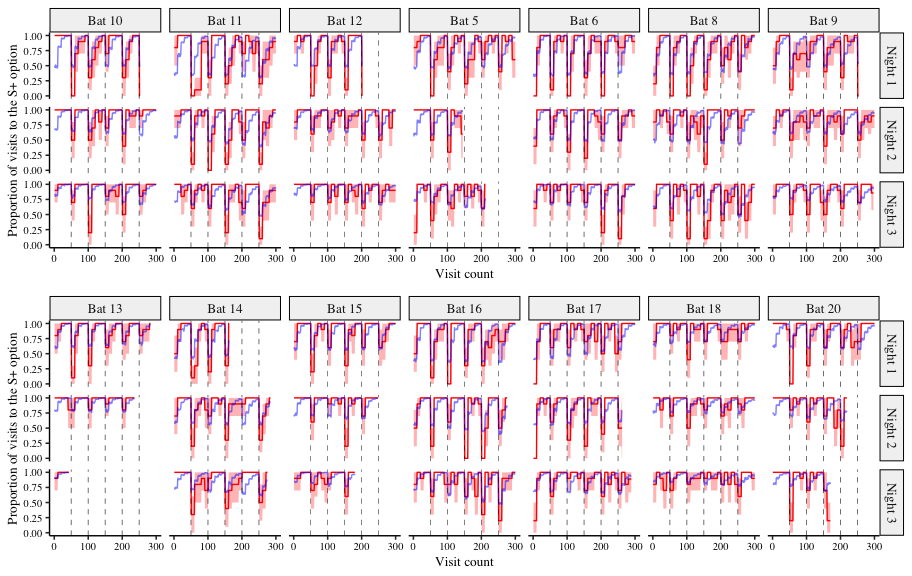


Figure 9: A comparison of the posterior predictions of the generalized linear mixed-effects model of the visits to the rewarding option and the empirical data from the bats. The red line indicates the average proportion of visits to the rewarding option per bin made by the individual bats, with the red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

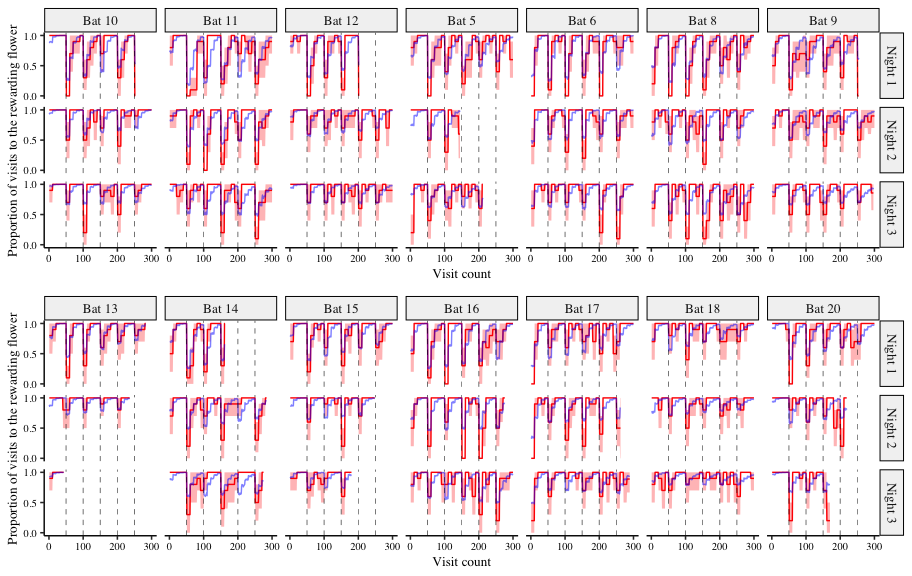


Figure 10: A comparison of the posterior predictions of the exploratory generalized linear mixed-effects model of the visits to the rewarding option and the empirical data from the bats, including the effect of ‘night-type’ and ‘block-type.’ The red line indicates the average proportion of visits to 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. Note the differences to Figure 9: there is a better fit to the data from the first block, especially on Nights 1 and 2 for many of the bats

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

As an additional piece of exploratory analysis we calculated the correlation between the Proprew just before a reversal and the Proprew just after (Figure 11). This analysis was independent of the passage of time within a night, and so does not affect our interpretation of the occurrence of second-order learning: the correlation is another way of analysing where the bats’ behaviour fell in the balance between exploration and exploitation.

We ran a multivariate regression model with random slopes and random intercepts with experimental night and reversal number (1-5) within a night as fixed factors. The two dependent variables were the Proprew averaged over the last three bins of a reversal block for each individual as the ‘asymptote’ of the bats’ behaviour, and the Proprew over the first 10 visits just after a reversal. The very first visit after the reversal was excluded as the bats could not know at this point that a reversal had occurred at all. The multinomial family was used for both of these variables, and as it was not possible to fit such a model with this family in brms at the time of writing, this model alone was fit using the MCMCglmm package. We used 130,000 iterations with a thinning interval of 1000 and a burn-in phase of 30,000. Similar models were fit separately for each of the three experimental nights as well.

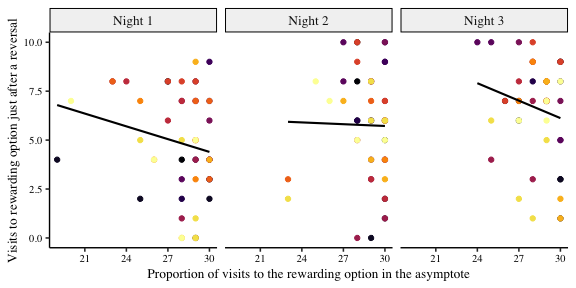


Figure 11: A comparison of the number of visits to the rewarding flower in the first bin of ten visits just after a reversal (excluding the very first visit after a reversal) to the asymptotic proportion of visits to the rewarding flower in the 30 visits just before that reversal. Individual points are 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.

Table 2: Correlation values between Proprew just before and after the reversals. The mode of the posterior distribution is given as the measure of central tendency with the 89% credible intervals

| **Model** | **Mode** | **Q94.5** | **Q5.5** |
| --- | --- | --- | --- |
| **Overall correlation** | -0.36 | -0.60 | -0.16 |
| **Night 1 correlation** | -0.61 | -0.80 | -0.19 |
| **Night 2 correlation** | -0.15 | -0.58 | 0.29 |
| **Night 3 correlation** | -0.61 | -0.85 | -0.14 |

The overall correlation as well as the correlation on the first and third nights were negative, but the correlation on the second night was not (Table 2). In general it appears that the higher the number of visits to the S+ before a reversal, the lower the number of visits just after: thorough exploitation of the rewarding option before a reversal seems to come at the cost of slower reversing after the reversal.

This result should be interpreted with caution. The range of asymptote values on the third night is much smaller than that on the first night as the bats’ overall preference for the S+ became higher, and some of the animals stopped making visits in the latter half of all three nights, reducing the sample size in the later reversals.

# References

The doctoral theses by Ulf Tölch and Johannes Thiele contain separated manuscripts on the different experiments. When citing the theses by, please, do not just cite the theses but instead cite the corresponding chapters. The citation form should be as you would cite the chapter from a book. If you cite multiple chapters use the 2006a, 2006b etc. format to distginguish.

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 Subtropics.” *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. Why capitalize journal name?

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 check upper case / lower case of latin species names; check all other citations too*).” *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>.

Caglayan, Alican, Katharina Stumpenhorst, and York Winter. 2021. “Learning Set Formation and Reversal Learning in Mice During High-Throughput Home-Cage-Based Olfactory Discrimination.” *Frontiers in Behavioral Neuroscience* 15 (June): 684936. <https://doi.org/10.3389/fnbeh.2021.684936>.

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. 1998. “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>.

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>.

Hadfield, Jarrod D. 2010. “MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package.” *Journal of Statistical Software* 33 (1): 1–22. <https://doi.org/10.18637/jss.v033.i02>.

Helversen, Dagmar von, and Otto von Helversen. 1999. “Acoustic Guide in Bat-Pollinated Flower.” *Nature* 398 (6730): 759–60. <https://doi.org/10.1038/19648>.

Helversen, O. von. 1993. “Adaptations of Flowers to the Pollination by Glossophagine Bats.” *Animal–Plant Interactions in Tropical Environments. Bonn: Zoologisches Forschungsinstitut Und Museum Alexander Koenig*, 41–59.

Helversen, O. von, 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>.

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

Lois-Milevicich, Jimena, Mariano Cerrutti, Alex Kacelnik, and Juan Carlos Reboreda. 2021. “Sex Differences in Learning Flexibility in an Avian Brood Parasite, the Shiny Cowbird.” *Behavioural Processes* 189 (August): 104438. <https://doi.org/10.1016/j.beproc.2021.104438>.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. 2nd ed. Boca Raton: Chapman; Hall/CRC. <https://doi.org/10.1201/9780429029608>.

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>.

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>.

Shettleworth, Sara J. 2010. *Cognition, Evolution, and Behavior*. Oxford; New York: Oxford University Press. <http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9780199717811>.

Shiffrin, Richard M., Michael D. Lee, Woojae Kim, and Eric-Jan Wagenmakers. 2008. “A Survey of Model Evaluation Approaches With a Tutorial on Hierarchical Bayesian Methods.” *Cognitive Science* 32 (8): 1248–84. <https://doi.org/10.1080/03640210802414826>.

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. 2006. Check German upper/lower case spelling; why quotation marks? “Nahrungssuchstrategien Der Nektarivoren Fledermaus *Glossophaga Commissarisi* (Phyllostomidae) Im Freiland - Eine Individuenbasierte Verhaltensstudie Unter Verwendung von Transpondertechnik.” Doctoral thesis. Universität München. Include location of publication, as with a book. I suggest you also include the edoc link.

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>.

Tölch, Ulf. 2006. *Bat Time Stories: Decision-Making in Spatio-Temporally Predictable Environments*. Doctoral thesis. Universität München. Include location of publication, as with a book. <http://edoc.ub.uni-muenchen.de/archive/00005130/>.

Tölch, 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, 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>.

Voigt, Christian C., and York 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>.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the Tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.

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>.