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:** [york.winter@hu-berlin.de](mailto:york.winter@hu-berlin.de)

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

# Abstract

Animals that show flexibility in their behavioural responses to environmental change have a strong advantage in foraging for food. We aimed to explore this ability in Commissaris’s long-tongued bat through a spatial serial reversal learning task. Wild 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 that the bats responded swiftly to the reversals of reward properties: when a food location suddenly dried up the bats quickly abandoned visiting it and switched to the alternate option, showing a near-exclusive preference for the rewarding option. After experiencing multiple such serial reversals, the bats chose the rewarding flower more frequently overall, but less frequently over the course of a single night. The bats also switched to the rewarding flower more swiftly as they experienced multiple reversals, but only up to a point. Our results are consistent with the occurrence of second-order learning or rule-learning.

# Introduction

Many animals live and forage in environments that change frequently and often unpredictably. The foraging environment of nectar-feeding animals mainly consists of flowering plants and the food resource they provide: nectar. Though flowers – and their nectar contents – are stationary, they bloom seasonally and single flowers on plants themselves wither and die every day or every few days, altering their efficacy with time as food resources. Thus, nectar-feeding animals face the challenge of exploiting resources that continually change with time but are relatively predictable in space. These animals need to detect the changing reward contingencies in their environment and adjust their behaviour accordingly.

Behavioural flexibility is an ability to cope with such changes. The word ‘flexibility’ has been used to mean many different things in the animal behaviour literature (often inconsistently – Audet and Lefebvre 2017), and one interpretation of the word is similar to the concept of elasticity: behavioural patterns that can be repeatedly and readily reversed (Bond et al., 2007). One experimental protocol that has been widely used to test for and demonstrate this sort of behavioural flexibility is reversal learning [is there a review on the reversal learning task?].

[The following paragraph is very wordy; I tried to shorten it]

[Also: The primary rule is in the environment not in the animal. It is in the environment in which according to some rule the reward contingencies change. The question then is can the animal detect this rule for environmental change, and learn which CUE indicates the change in the environment and learn to respond to this “cue of change”. Responding to a CHANGE stimulus by changing one’s own behaviour is why it is termed “second-order” learning]

Reversal learning essentially involves two separate stages of learning. An animal must first learn about multiple stimulus response associations and to discriminate in its behaviour amongst them. These can be two spatial locations that potentially provide a food reward. When only one of the two options is rewarding the animal should only respond to this. When reinforcement at this location stops the animal should abolish its current behaviour and instead perform another stimulus-response action that in the past had proven rewarding. It is this ability to abruptly terminate a repeated behaviour and switch to another behaviour that the animal needs to have for efficient exploitation of the available resources. In a serial reversal learning procedure the reward contingencies reverse repeatedly. Successful performance critically depends on the animal‘s ability to switch abruptly after a reversal.

Improvement in the reversal learning task is clearly demonstrable and has been used to compare the performance of different species (Hermer et al. 2018, Bond et al., 2007). In this respect, two levels of learning should be distinguished. First-order learning is – the stimulus-response association – with changes in behaviour according to the strength of this association . Second-order learning is the learning of rules or regularities of the environment that underly patterns of change in reward contingencies. This requires that an animal learns that a specific stimulus or stimulus configuration announces this change in the environment. Here, it is the information contained in the absence of an expected reward that becomes a stimulus in itself. A successful animal associates this non-reward stimulus with the appropriate response which is the inhibition of its previous behaviour. This provides the basis for then quickly reorienting towards the alternative.

Animals that learn to respond to stimuli that predict change by changing their own behavior, i. e. second-order learning, can quickly abandon a suddenly unsuccessful response that previously resulted in reward. The optimal behaviour pattern has been termed ‘win-stay; lose-shift’ behaviour. a behaviour is repeated as long as it produces a reward, but on absence of reward another behavioural response is switched to immediately. Ideally, under a 1:1 fixed ratio schedule an animal makes exactly one ‘error’ per reversal. Underlying this is standard reinforcement learning based on the acquisition of a stimulus-response association and leading to reward as a reinforcer. It is termed second-order because this is learning about predictable changes in the environment and responding to a change stimulus with a change in one’s own behavior. A non-reward stimulus is responded to by inhibiting a behavioural response. The first instance of a behaviour leading to a lack of reward (the error) is sensed as the stimulus that signals the change of environmental contingency. On this experience , the animal inhibits its previously successful action and changes to exclusively respond to the other stimulus that is now paired with reward.

[I wonder if we should somewhere place this short paragraph. Maybe it should come somewhere after reversal learning has been introduced.]

Especially flower visitors may often experience reversal situations in their natural environment. A flower full of nectar may remain rewarding for multiple visits before it is empty. The flower visitor should then seek another flower. However, at a later time point, when ongoing secretion has replenished nectar, a previously visited flower is rewarding again.

Animals are rarely, if ever, so finely-tuned in their behaviours in real life. Large and deterministic changes are uncommon in nature and more gradual shifts in behaviour may be better suited to natural environmental changes.

For this reason, win-stay lose-shift in practice means an animal will repeat the behaviour that produces a reward after one or a few experiences of reward, and switch to another behavioural response after one or a few errors leading to experiences of no reward. Errors in reversal learning tasks are common and can occur in different ways. An animal can persevere, continuing to choose a non-rewarding cue after the reversal has occurred; or the animal can make anticipatory errors, changing its behaviour and choosing the other cue before a reversal has occurred. Progressive ‘improvement’ in this task, where an animal makes fewer and fewer errors of either type [No, I do not think that both types of errors can be explained by the same mechanism.] per reversal is indication that higher-order learning, or ‘learning to learn’, is occurring (Shettleworth, 2010b.

To investigate the flexibility of responding to depleting food locations we carried out a serial reversal learning task with Commissaris’s long-tongued bat, (*Glossophaga commissarisi*), which primarily feeds on flower nectar. The bats were given two potentially rewarding options to choose between. The two options were separated in space and their spatial location was the cue to indicate their reward properties at any moment. *G. commissarisi* is known to have excellent spatial memory (Thiele and Winter, 2005) so no other cue was necessary. 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 aim with this experiment was as follows. It is well known that bats alter their preference between reward sources according to their transient rewarding properties and thus show first-order learning for spatial locations. [This is basic and already established so that we should not point it out as a goal of this study.]

Here, we tested if the bats can also learn to respond to stimuli that indicate that their environment has changed by then changing their own behaviour. This is second-order learning. For this, we presented bats with multiple experiences of abrupt changes in the spatial location of reward. We examined how the occurrence of successive reversals affect their behaviour. For example, the behavioural allocation between the two options might change more swiftly as the bats experienced more reversals. We were especially interested to see to what extent bats could approach the optimal strategy in such an environment: one error per reversal, achievable through a win-stay, lose-shift pattern of behaviour. Exploratory analyses were done to take the confirmatory analyses described above further. We reasoned that there is a difference between the first visits of a night before any experience of a change in reward 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 in June-July 2017. 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 chicken-feeders filled with sugar-water (see **Reward** below) as bait. The feeders had cotton swabs soaked in dimethyl disulphide on them, a chemical attractant produced by many bat-pollinated flowers (von Helversen et al., 2000). 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.

A total of 16 bats participated in the main experiment and the first stage of the experiment began on the same night the bats entered the cages. A group of four experimental bats of the same sex were placed in a flight cage together. Two such groups were run in parallel, one in each flight-cage so the data were collected simultaneously. 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. These two bats were released before the end of the experiment and not replaced, and the data from these two individuals were not analyzed. Thus, 14 bats (7 males and 7 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 et al., 1998). Every night, the bats were also given ad-libitum access to 10 mL of supplemental food: 3.5 g of hummingbird food (NektarPlus, Nekton, Germany) and 3.5 g of milk powder (Nido 1+, Nestle, Switzerland) in 10 mL of water. The bats in each cage 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 (2 × 2 × 1.5 m). 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 (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; and 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, Germany). The step volume of the two pumps differed slightly: the pump in Cage 1 delivered 2.11 L per step of the stepper-motor, whereas the pump in Cage 2 delivered 3.33 L per step. The glass syringe was connected to the tubing system of the flowers through five pinch valves (Nachev and Winter, 2012). 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 was connected to the syringe through the valves.

Every day at around 1000 h, the old nectar was emptied from the system, which 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.

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, this could trigger 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. A reward was given only when both events occurred, i.e., the RFID reader identified a bat and the photo gate was triggered. 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, Germany).

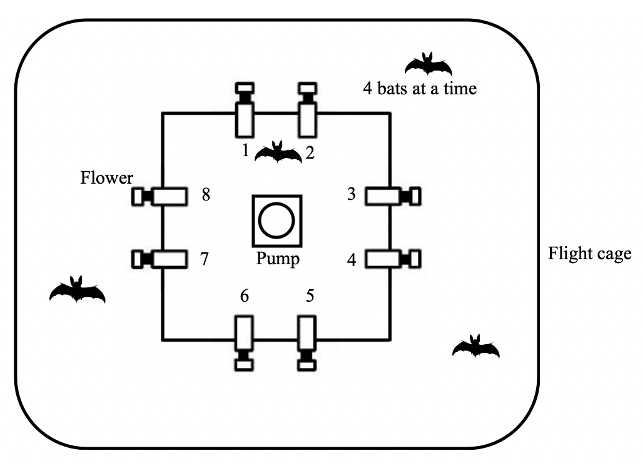


Figure 1: Schematic of the cage and flower set-up (not drawn to scale)

## Experimental procedure

Out of the array of eight flowers, each bat was uniquely assigned two adjacent flowers on the same side of the square frame, programmed to reward 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 main experiment, the bats could make a maximum of 300 visits a 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 2.4 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 were wrongly unrewarded: the bats did not receive a reward during these visits even when the visits 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 300.

## Experimental design

The experiment proceeded through the following stages.

### Training

On the night the naïve 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 soaked in dimethyl disulphide was placed on each flower. A small drop of honey was applied to the inside of the flowers to encourage the bats to place their heads inside, 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.

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 to it would elicit rewards 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, the bats went through one final stage of training: forced alternation. The bats received a reward at one of the two flowers for one trial, and then could only receive reward at the other flower for the next trial. In this way the bats had to alternate between the two flowers every single trial.

### Serial Reversal Learning Task

In the serial reversal learning task, the bats had continuous access to two flowers: one that gave a 40L nectar reward, and one that remained empty. The location of the rewarding flower was not cued; however, after completing the alternation training phase, 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. Each set of 50 visits to the assigned two flowers, either at the start of each night or between reversals, was termed a ‘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

The raw data collected during this study were the computer logged events of feeder visits. Each event included the time stamp, animal ID, photo gate interruption duration and the volume of nectar dispensed. R was used for all statistical analyses and creation of plots.

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 (Carpenter et al., 2017). The details of these models are provided in the **Supplementary Material**.

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; however, these were the minority, and were not considered for the analysis (see **Supplementary Material** for details). At the end of each of the first five blocks, a reversal occurred; 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 the 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 – McElreath, 2020). To aid in the interpretation of the parameter values of the models we also present plots of the conditional effects of the predictor variables.

We investigated the effect of experimental night and 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:

[I would think that this is too trivial to merit an equation. This is standard procedure and the text before the equation already says it all. ]

After examining these results, we conducted further post-hoc exploratory analyses. We investigated the effect of the first night and first block of a night as predictor variables separate from the later two nights and the later five blocks on Proprew. This was because the bats had had no prior experience of any reversals on the first night, and during the first block of every night, the bats had not yet experienced any reversals that night. Thus, night and block were taken as categorical variables with levels ‘first night’ and ‘later nights’; and ‘first block’ and ‘later blocks’ respectively. Then the analysis of the effect of experimental night, block and bin on the Proprew described above was repeated, but the data from the first block of every night was removed to see if the decrease in the Proprew due to the effect of block still remained.

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 (Harrell and Dupont, 2007).

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

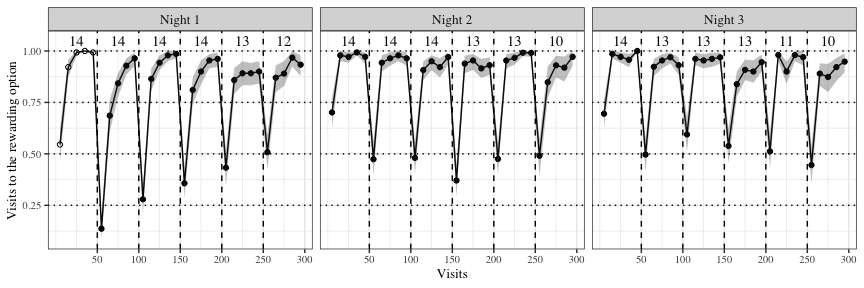
### Bats made the majority of their visits to the rewarding option

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, and a simultaneous 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 at chance level: 54.5% [95% CI 46.8, 62.3], about 5 out of the 10 visits. 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 97.9, 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 64.6, 74.7], about 7 out of the 10 visits. This was significantly higher than random choice and higher than the Proprew in the corresponding bin of the first night. For the remainder of these nights the Proprew showed a similar pattern to the first night: a decrease immediately after the reversal and then an increase to near 100%: 94.8% [95% CI 94.1, 95.5], comparable to the 93% [95% CI 91.8, 94.1] on the first night.

No figure 1?



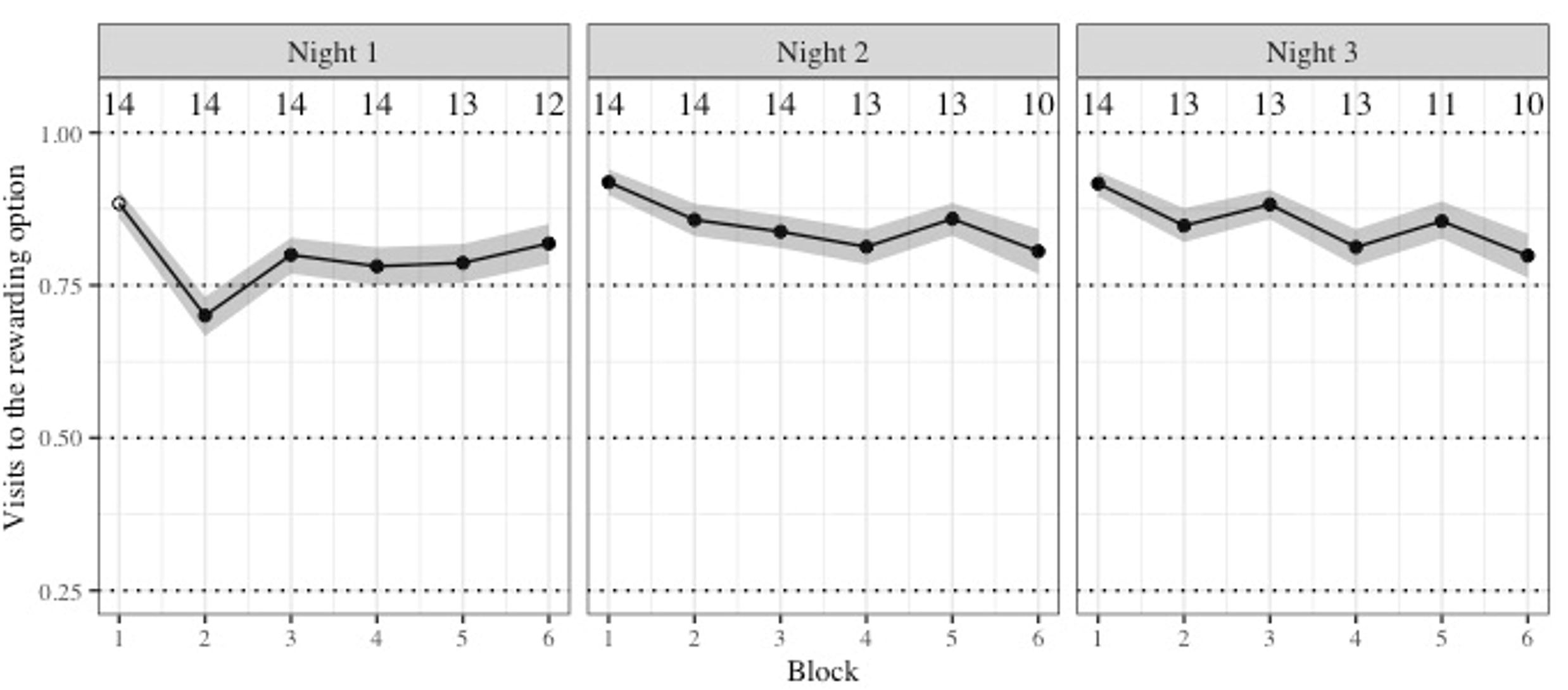
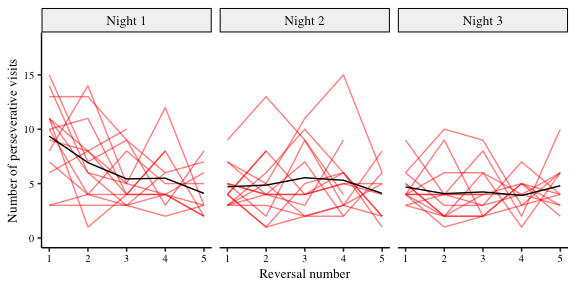


Figure 2: Visits to the rewarding option out of the two potentially rewarding options across the three experimental nights. Data are average proportions for bins of ten visits averaged over all 14 individuals. Data are indicated by points that are white in the first block on the first night before the bats had experienced any reversals. Numbers indicate the bats that participated in a block. Shading shows 95% confidence intervals. Dashed lines show reversals

[keep axis legends brief: Y-axis “Visits rewarding option”]

### Bats switched to the rewarding option faster as they had experienced more reversals

As the bats experienced more reversals they made their first visit to the rewarding flower sooner and sooner (data not shown). 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 analysis showed however that this change occurred mostly on the first night (Figure 4 and Figure 5). The bats responded to the reversals faster with each block, i.e., made fewer perseverative visits only on the first night until they reached a constant level in their performance, and this level persisted over the second and third nights.

Chart

Description automatically generated with medium confidenceFigure 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 and 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. Circles represent the means of the posterior distributions of the intercept and slope coefficients, thick horizontal lines represent 50% credible intervals, and thin horizontal lines 89% credible intervals. The numbers in bold are the means of the posterior distributions and 89% credible intervals*

Chart

Description automatically generated

*Figure 5: a) Conditional effects of experimental night on the number of perseverative visits b) Conditional effects of block on the number of perseverative visits c) Conditional effects of the interaction of experimental night and block*

[- can you increase font size for a), b), c)

### Bats made more rewarded visits over all three nights, but fewer over the course of each night

[I think we need to discuss this following sentence. ]

As the experiment proceeded and the bats experienced more reversals on more nights and had also experienced for a longer time that only a single feeder was active at a time the overall proportion of the animals’ choices for the rewarding flower changed (Figure 2).

Within each block the change in the Proprew was very large, showing a rapid increase from less than 0.7 at the start of a block, when the animals had either experienced the start of the experimental night or a reversal, to more than 0.9 by the second bin of 10 visits.

This change in performance within a block was strongest in the very first night as indicated by the significant night block interaction. Rausnehmen, oder woanders hin.

We saw no evidence of anticipatory errors. [different position?] Across the three nights the animals’ choices for the rewarding flower also increased. However, there seemed to be a decrease within each night in the Proprew: the bats made fewer visits to the rewarding flower at the end of the night compared to the start of the night (Figure 6 and Figure 7). However, as the significance of the night-block interaction indicates this main effect was not consistent over nights. In fact, for the very first night the effect was even the opposite with an increase of Proprew over the course of the night. This is shown in more detail further below.

[Pulle up from below] On the first night alone, the bats made more visits to the rewarding option with each successive block.

In other words, the bats made an increasing number of visits to the rewarding option as each block progressed from its start to the next reversal, and as they experienced more experimental nights; however, this increase was dampened over the course of each individual night. This effect, however, was not true for the very first night. Here, the visits to the rewarding option increased -> wait for new plot

Chart

Description automatically generated

*Figure 6: Forest plot of the estimates of the effect of night, block, bin and their two-way interactions on the Proprew. Circles represent the means of the posterior distributions of the intercept and slope coefficients, thick horizontal lines represent 50% credible intervals, and thin horizontal lines 89% credible intervals. The numbers in bold are the means of the posterior distributions and 89% credible intervals*

Diagram

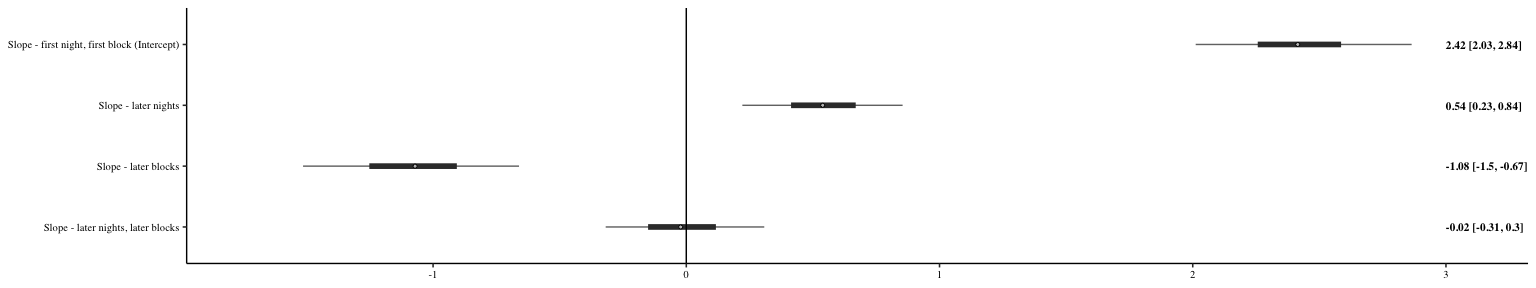
Description automatically generated

*Figure 7: a) Conditional effects of experimental night on the Proprew b) Conditional effects of block on the Proprew c) Conditional effects of bin on the Proprew d) Conditional effects of the interaction of night and block on the Proprew e) Conditional effects of the interaction of block and bin on the Proprew f) Conditional effects of the interaction of night and bin on the Proprew*

## Exploratory analyses

### 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 except for the strict alternation training prior to the experiments. Therefore, after examining the results of the analyses described above, we did further analyses to specifically explore the effects of the first block of a night and the first night.

The first block of a night had an effect on the Proprew (Figure 8 and Figure 9): the bats made more visits to the rewarding option in the very first block of a night compared to all the later blocks. This effect was greater on the very first night compared to the later two nights, i.e., on the very first block of the very first night when the bats had experienced reward at only one option, and no reversals at all. This is consistent with the results of the confirmatory analysis, which showed that the number of visits to the rewarding flower decreased from the start to the end of a night and where the significant block-night interaction term indicated that this result was not universally the case for all nights.

Diagram

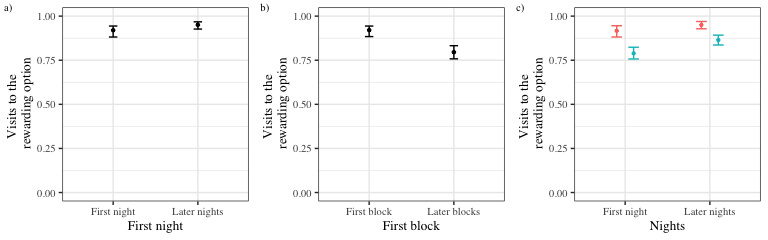
Description automatically generated*Figure 8: Forest plot of the estimates of the effect of night, block, bin and their two-way interactions on the Proprew. Circles represent the means of the posterior distributions of the intercept and slope coefficients, thick horizontal lines represent 50% credible intervals, and thin horizontal lines 89% credible intervals. The numbers in bold are the means of the posterior distributions and 89% credible intervals[I need an explanation of the data analysed to understand the parameters plotted]*

Figure 9: a) Conditional effects of first and later experimental nights on the Proprew b) Conditional effects of first and later blocks on the Proprew c) Conditional effects of the interaction of the first and later nights and first and later blocks on the Proprew

Additionally, when the data from the first block of every night were removed, the Proprew increased due to the effect of experimental night and bin(Figure 10 and Figure 11), consistent with the confirmatory analyses . The effect of block changed: the decrease in Proprew as an effect of block occurred on the second and third nights but not the first, primarily due to the bats’ behaviour in the second block of the first night, when the bats experienced a reversal for the first time in the experiment. On the first night alone, the bats made more visits to the rewarding option with each successive block. On the second and third nights, Chart, scatter chart

Description automatically generatedthe bats made fewer visits to the rewarding option with each successive block.

*Graphical user interface

Description automatically generated with medium confidenceFigure 10: Forest plot of the estimates of the effect of night, block (excluding the first block), bin and their two-way interactions on the Proprew. Circles represent the means of the posterior distributions of the intercept and slope coefficients, thick horizontal lines represent 50% credible intervals, and thin horizontal lines 89% credible intervals. The numbers in bold are the means of the posterior distributions and 89% credible interval*

*Figure 11: a) Conditional effects of experimental night on the Proprew b) Conditional effects of block on the Proprew c) Conditional effects of bin on the Proprew d) Conditional effects of the interaction of night and block on the Proprew e) Conditional effects of the interaction of block and bin on the Proprew f) Conditional effects of the interaction of night and bin on the Proprew*

# Discussion

With this study we examined the ability of nectar-feeding bats for second-order learning. This is ability would enhance their flexibility in a dynamically changing foraging environment. In our experiment wild nectar-feeding bats participated in a spatial serial reversal learning task with two potentially rewarding options that repeatedly alternated their rewarding properties.

We found that the bats quickly located the rewarding location and made most of their visits to it. The bats responded to their accumulating experience of reinforcement and lack of reinforcement in the serial reversal learning task by switching to the newly-rewarding option more swiftly and making a majority but decreasing number of their visits to the rewarding flower over the course of the night.

[I suggest to omit the following paragraph. At least it should not be introduced as a goal that we wanted to test for presence of reinforcement learning. That would be too basic.]

This seemed unambiguously to be the case (Figure 2). The difference between something and nothing (which was the choice presented by the experiment) is the easiest possible discrimination test and the bats showed a near-exclusive preference for the rewarding option on all three nights. As the rewarding properties reversed between the two flowers, the bats re-established a high preference for the rewarding flower by the first bin of ten visits after a reversal and maintained it until the next reversal.

The second aim of our experiment was to see if the bats’ behaviour indicated a change that was due to their repeated experience of reversing reward properties. We detected two such behavioural changes:

I think that at this point we need a new analysis. The prediction is that the ability to detect environmental change will reduce perseverative visits. This needs a detailed analysis of that phase in the experiment where perseverative visits occur. The more specific the analysis is to those data the better it should be able to detect changes.

But one should not ignore that there other processes of adaptation to the experimental environment that may be happening over the course of the experiment. It is “unnatural” that of two potentially rewarding flowers only one gives a reward. This is something that the bat has to learn. During the very first visits a bat may have learnt to simply ignore the non-rewarding flowers and this led to highest performance. But from the first reversal onwards a bat had experienced that in fact both flowers and locations can give nectar in the same night.

Also, the bats in this experiment with 40 ul received rather large portions of nectar. Thus, while at the beginning of the experiment where their behaviour was less optimal they might have sensed the environment as potentially scarce [or they may have carried that information over from prior experience in the natural environment], over experimental time they learnt that food is sufficiently available. This may have led to a state where their activity became less focussed on obtaining food and they spent more time simply “flying” about.

a decrease in the visits to the rewarding flower from the start to the end of each night; and an increase in the visits to the rewarding flower across the three experimental nights (Figure 6 and Figure 7). The latter was a small effect compared to the former, probably because of a ceiling effect – the bats were making a high enough number of visits to the rewarding flower that any further increase had to be relatively small, especially when there were only three nights in total. The critical parameter of this experiment is the rate of perseverative visits after a reversal and how this parameter changes with experience…… is the decrease in the number of visits to the rewarding flower over the course of an experimental night. Exploring this effect showed that on the first night, when the bats experienced reversals for the first time, at the start of the night the bats made almost all their visits to the rewarding flower [pointing out the behaviour before any reversal may be interesting but should occur at a different place than here](Figure 8 and Figure 9). [Somewhere you could emphasise that “The experiment began with the first reversal.”] When the bats experienced a reversal for the first time in the experiment, i.e., in the second block of the first night, initially the bats kept visiting the previously rewarding flower so that its rate of visiting the newly rewarding flower was very low. the number of visits to the rewarding flower decreased greatly, and then with subsequent experience of reversals in that night, increased again to a near-exclusive preference. On the second and third night, the largest number of visits to the rewarding flower occurred again in the first block before any reversals had occurred on that night. The decrease in visits to the rewarding flower when the first reversal was experienced was significant (Figure 9b and Figure 9c), but quite small [here, a different view on the data is needed. In all cases after a reversal were the initial visits directed at the formerly rewarding flowers. This did not change over the course of the whole experiment. But what changed is that these numbers of perseverative visits decreased with experience. -> It is this dynamic that you need to take a look at. This may need a visit-by-visit analysis and not a coarse binning of visits.] for the following reversals (Figure 7d and Figure 10d). The decreasing effect of block on the visits to the rewarding flower over the course of a night was likely artefactual -> you mean without biological explanation?, driven by the high number of rewarded visits made before the first reversal of each night -> but this would then be a biological reason and not “artefactual”. Alternatively, the decrease in visits to the rewarding flower with each successive reversal is consistent with two potential explanations: a response to the increasing experience of no reinforcement + experience of reinforcement at the formerly not-rewarding flower ….. as each reversal occurs; or an increase in the number of exploratory visits: visits to a location previously experienced as non-rewarding. -> what about “second-order learning”?

The [I would not call this an “aim of the experiment”. I would phrase it differently such as that it is interesting to evaluate to what extent the bats actually reached the theoretically optimal performance]

Not continued with the discussion beyond here as the discussion may still change as the analysis is changed.

of the experiment was to see whether the bats’ choice behaviour approached the optimal strategy of one error per reversal. The number of visits immediately after a reversal to the previously-rewarding flower before the first visit to the newly-rewarding flower, which we term perseverative visits, did in fact decrease on the first night but reached a plateau by the second night and did not decrease further (Figure 3 and Figure 4). The number of perseverative visits is distinct from the number of visits to the rewarding flower in the first bin of ten visits of a block, and we infer from the decrease in the former that a limited amount of higher-order learning occurred. This is interesting grounds for further exploration. A strong test of higher-order rule-learning is the transfer of improved performance at the serial reversal task from one sensory modality to another (Bond et al., 2007). A follow-up experiment would be to examine if bats that show a decrease in perseverative errors on a serial reversal task on one sensory modality (say spatial location) show a more rapid decrease in perseverative errors on the same task on another sensory modality (say olfactory cues), compared to bats that have only experienced the task on one of these modalities. The positive transfer between these sensory modalities would suggest the capacity to develop a generalized learning strategy.

When animals receive a reward, at a particular location or associated with a particular cue, they are more likely to return to that location or cue. As bats were continually confronted with reversals, they learned a second reinforced response that was conditioned to another cue: the single or repeated absence of expected reinforcement. Reinforcement learning is the likely mechanism underlying the win-stay, lose-shift strategy and the two conditioned responses acquired through reinforcement learning in this experiment were the win-stay response and the lose-shift response. In other words, an animal’s tendency to choose a location or cue becomes a near-exclusive preference when reinforcement is experienced once or a few times and drops close to zero when a *lack* of reinforcement is experienced once or a few times. We suggest that on the first night the bats learned the following: the association between the two flowers and the flowers’ potential to give reward (first-order learning); and the win-stay and lose-shift responses to reinforcement and no reinforcement respectively. The fact that fewer experiences of no reinforcement were required to provoke a lose-shift response with each successive reversal (Figure 3 and Figure 5) indicates that second-order learning occurred on the first night as well. On the second and third nights the bats were already acquainted with the potential for reward at two locations. The first- and second-order learning from the first night persisted over the later nights: the bats showed a high tendency to visit the rewarding flower and a lose-shift response to each reversal that was as swift as it was at the end of the first night.

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

In most cases under natural foraging conditions, flowers are emptied in a single visit, so win-stay is not often a profitable strategy. There are however certain plants such as species of *Agave* or *Vriesea*, that hold large amounts of nectar, which if undetected for a long time may require multiple hovering visits to deplete – “jackpot” rewards in other words (Ohashi and Thomson, 2005). Situations where lose-shift is the optimum response might be more common, but win-stay situations do occur in the bats’ foraging environments, and thus the ability to adopt a win-stay, lose-shift strategy is likely part of the bats’ natural foraging ecology as nectar-feeding animals.

The performance on the serial reversal task of animals that share similarities with bats in their foraging ecology is illuminating. Bumblebees in an experiment similar to ours with a large trial number (Strang and Sherry 2014) showed a reduction in their perseverative errors and an increase in the errors made in the last trials, though there was an overall decrease in the errors. The authors of this paper interpreted these findings as indicative of proactive interference, which occurs when previously-learned information interferes with the learning or remembering new information (Tello-Ramos et al., 2019).

Several animals that rely strongly on spatial memory have also been studied in reversal learning tasks, specifically birds that cache food at various locations that they must remember and return to. Birds that are known to have better or more long-lasting spatial memory such as black-capped chickadees (Hampton et al., 1998), Clark’s nutcrackers (Lewis and Kamil, 2006) and high elevation mountain chickadees (Croston et al., 2017) were worse at adapting to the new contingency after a reversal than the initial learning (reviewed in Tello-Ramos, et al., 2019). These data are consistent with the idea that there is a trade-off between acquiring new memories and retaining old ones, i.e., that proactive interference may be occurring in spatial reversal tasks, just like in bumblebees. It is known that Glossophagine bats not only have excellent spatial memory but that memory has been observed to last up to several weeks (Rose et al., 2016). It is certainly possible that proactive interference plays a role in the bats’ behaviour in the serial reversal task, but we believe that the simple reinforcement learning model outlined above explains the observed data adequately.

What performance on the serial reversal task says about the cognitive mechanisms at work is not completely settled. Cognitive flexibility describes the processes in the brain that underlie adaptive change in behaviour in response to changes in the internal or external environment, whereas behavioural flexibility is the modifiability of learned behaviour (Dhawan et al. 2019). Cognitive flexibility cannot be directly observed; it is inferred to have occurred through behavioural flexibility (Tait et al. 2018), and the reversal learning task is a test of behavioural flexibility, not cognitive flexibility (Dhawan et al., 2019). Our study with nectar-feeding bats revealed a high capacity for behavioural flexibility and a limited but intriguing amount of rule-learning, abilities that have evolved in the bats through a foraging ecology dominated by the search for nectar-rich flowers.

**The following concluding paragraph is sort of a place-holder, but it’s there for now. It can be rewritten after a bit more discussion**

Our study of serial reversal learning in nectar-feeding bats showed that the animals chose the rewarding option in the majority of trials, but these choices decreased in response to repeated experiences of the reversal, indicating that learning was occurring both in response to reinforcement and to non-reinforcement. Though the bats never reached the optimal behaviour of one error per reversal, the number of perseverative visits did decrease, indicating a limited amount of rule-learning.

## 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**: data collection, 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, and it was only these two flowers that could reward each bat. However, 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 12.

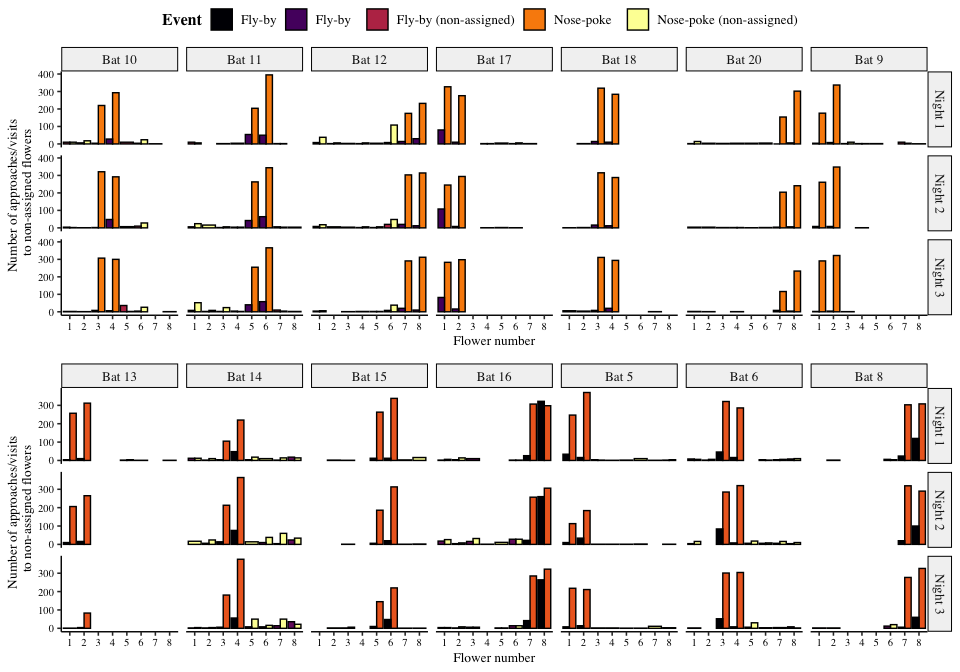


Figure 12: Visits made by the bats to all the flowers, including the ones that were not assigned to them. Yellow bars are nose-pokes at the assigned flowers, where the bats attempted to get a reward by breaking the light-barrier. Purple bars are ‘fly-by’ events near the assigned flowers where the bat 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 13 shows.

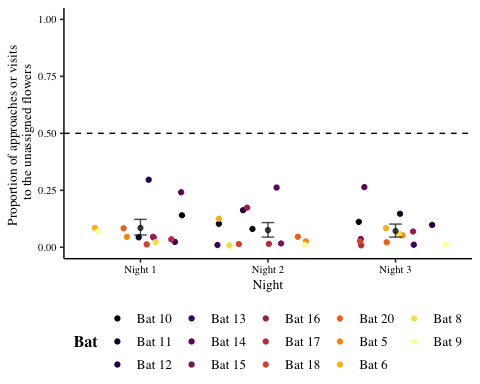


Figure 13: Proportion of visits or approaches to the un-assigned flowers out of the total number of visits or approaches to flowers. Coloured points are data from individual bats. Black points are the mean proportion per night and the error bars are 89% Cis

## 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 1200 post-warm-up samples for the model with the first experimental night and block 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 block on the number of perseverative visits, a negative-binomial likelihood function was used. Experimental night, 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 also used a binomial likelihood function with night and block as categorical predictors. The first night and the first block of every night was treated as one level of the categorical variables and the other nights and other blocks of each night as the other level. The first night and first block were the reference categories. Random slopes and intercepts were used to fit regression lines for the individuals.

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 the models the was equal to 1 for all the chains.

## Posterior predictive checks of the confirmatory model of the Proprew

We carried out a posterior predictive check of the confirmatory model 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 12. A visual comparison of the two reveals a close correspondence between the data and the posterior samples of the model.

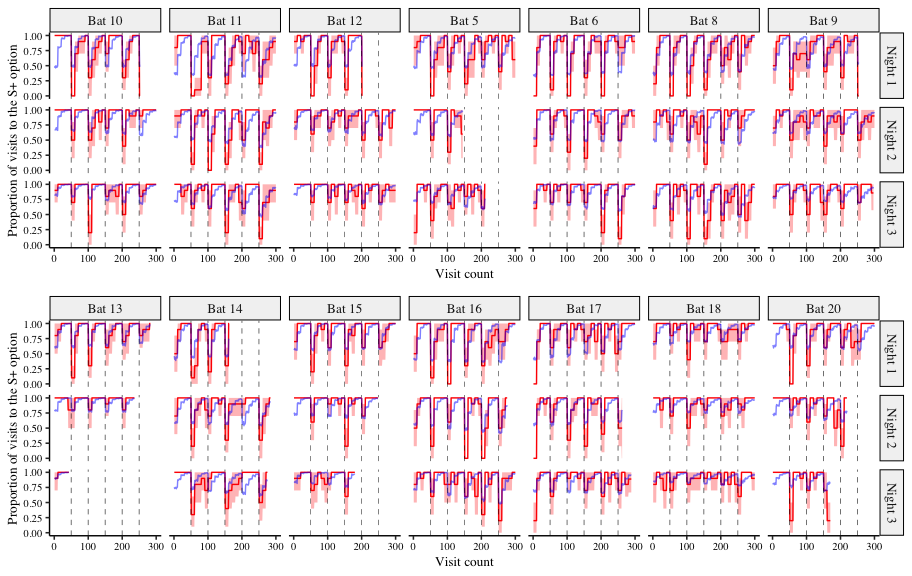


Figure 12: A comparison of the posterior predictions of the generalized linear mixed-effects model of the visits to the rewarding option and the empirical data from the bats. The red line indicates the average proportion of visits to the rewarding option per bin made by the individual bats, with the red shading indicating 95% confidence intervals; the blue line indicates the corresponding posterior prediction of the model

## 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 Subtropics.” *Biotropica* 30 (4): 559–86. <https://doi.org/10.1111/j.1744-7429.1998.tb00097.x>.

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>.

Bürkner, Paul-Christian. 2017. “**Brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*.” *Journal of Statistical Software* 80 (1): 1–28. <https://doi.org/10.18637/jss.v080.i01>.

Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "*Stan*: A probabilistic programming language." *Journal of Statistical Software* 76 (1): 1–32. https://www.osti.gov/servlets/purl/1430202

Croston, Rebecca, Carrie L. Branch, Angela M. Pitera, Dovid Y. Kozlovsky, Eli S. Bridge, Thomas L. Parchman, and Vladimir V. Pravosudov. 2017. "Predictably harsh environment is associated with reduced cognitive flexibility in wild food-caching mountain chickadees." *Animal Behaviour* 123: 139-149. [https://doi.org/10.1016/j.anbehav.2016.10.004](https://doi.org/10.1016/j.anbehav.2016.10.004" \t "_blank" \o "Persistent link using digital object identifier)

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: 45–52. <https://doi.org/10.1016/j.bbr.2019.01.055>.

Harrell Jr, Frank E., Maintainer Charles Dupont. 2007. "The design package." *R package version* 2 http://biostat.mc.vanderbilt.edu/s/Design

Hampton, Robert R., Sara J. Shettleworth, and Richard P. Westwood. 1998."Proactive interference, recency, and associative strength: Comparisons of black-capped chickadees and dark-eyed juncos." *Animal Learning & Behavior* 26 (4): 475-485. https://doi.org/10.3758/BF03199241

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

Lewis, Jody L., and Alan C. Kamil. 2006. "Interference effects in the memory for serially presented locations in Clark's nutcrackers, *Nucifraga columbiana*." *Journal of Experimental Psychology: Animal Behavior Processes* 32 (4): 407. [https://doi.org/10.1037/0097-7403.32.4.407](https://psycnet.apa.org/doi/10.1037/0097-7403.32.4.407" \t "_blank)

McElreath, Richard. 2020. “Statistical Rethinking: A Bayesian Course with Examples in R and Stan.” Chapter 4. 2nd ed. Boca Raton: Chapman; Hall/CRC. <https://doi.org/10.1201/9780429029608>.

Ohashi, Kazuharu, and James D. Thomson. 2005. “Efficient harvesting of renewing resources.” *Behavioral Ecology* 16 (3):592-605. <https://doi.org/10.1093/beheco/ari031>

Nachev, Vladislav, Kai Petra Stich, and York Winter. 2012. “Weber’s law, the magnitude effect and discrimination of sugar concentrations in nectar-feeding animals.” *PloS one* 8 (9): e74144. <https://doi.org/10.1371/journal.pone.0074144>

Rose, Andreas, Miriam Kolar, Marco Tschapka, and Mirjam Knörnschild. 2016. "Learning where to feed: the use of social information in flower-visiting Pallas’ long-tongued bats (*Glossophaga soricina*)." *Animal cognition* 19 (2): 251-262. <https://doi.org/10.1007/s10071-015-0930-9>

Shettleworth, Sara J. 2010a. “Cognition, Evolution, and Behavior.” Chapter 4. 2nd ed. Oxford; New York: Oxford University Press. <http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9780199717811>.

Shettleworth, Sara J. 2010b. “Cognition, Evolution, and Behavior.” Chapter 6. 2nd ed. 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: 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: 129–36. <https://doi.org/10.1016/j.anbehav.2018.02.019>.

Thiele, Johannes. 2006. “Nahrungssuchstrategien der nektarivoren Fledermaus *Glossophaga commissarisi* (Phyllostomidae) im Freiland - eine individuenbasierte Verhaltensstudie unter Verwendung von Transpondertechnik.” Dissertation. Chapter 4. Ludwig-Maximilians-Universität, Munich, Germany. https://edoc.ub.uni-muenchen.de/5566/1/Thiele\_Johannes.pdf

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>.

von Helversen, Otto, 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>. [must be sorted under “H”]