Tracking a continually changing environment

Shambhavi Chidambaram1,2, Alex Kacelnik3, Vladislav Nachev1, York Winter1,2**\***

1 Institute of Biology, Humboldt University, Berlin, Germany

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

3 Department of Zoology, University of Oxford

**\*For correspondence:** york.winter@hu-berlin.de

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

# Introduction

In many foraging environments the properties of the available food resources change. In order to adjust to and exploit changing resources, an animal can potentially learn and remember those properties. An animal can never learn everything about its environment; any learned information will always be incomplete, no matter how much effort is spent in obtaining it. One might then ask - when does the benefit of the learned information outweigh the cost of gaining it? The value of information lies in whether it can tell a forager something that changes its behaviour (Stephens, 2007). When a forager’s behaviour allows it to experience environmental change, and so gain information about the current state of the world, this might be termed ‘tracking.’ The information gained can then be translated into appropriate actions (Dunlap and Stephens 2012).

It can easily be seen that the value of seeking information depends on temporal parameters. When an environment changes rapidly any information that is gained by tracking it becomes outdated very soon. When an environment changes so slowly that there is no consequence in the animal’s lifetime, any effort spent in tracking would not yield usable information. Furthermore, the benefit of tracking lies in allowing a forager to choose the best of the options available in an environment - for example, the option that results in the highest caloric gain. For certain combinations of environmental rates of change and differences in the quality of the available options, environmental tracking is both possible and beneficial, in the sense of resulting in a higher energetic net yield. Under some other circumstances it may be preferable to adopt a ‘one size fits all’ or averaging approach, where a forager applies one behavioural response that does best on average over all the possible environmental states (Stephens and Dunlap 2008). One might then ask: what sort of environmental change is tracking, rather than applying an averaging response, beneficial?

An early attempt to model such a situation was done by Stephens (1987), attempting to answer the question of whether, and to what extent, a forager should modify its behaviour in response to a change in its environment. In this simple model the environment has a ‘variable’ option and a stable ‘alternative’ option. The latter has a single value, , and the former can vary between a good state, , and a bad state, , such that > > . The forager can recognise the type of resource (variable vs alternative) upon encounter, but must consume a resource to know its sub-type (good vs bad). The mechanism through which tracking happens is sampling, i.e., visiting the variable option when the last experience of it was the bad state, with the intention of learning what state it is in at the present time. The probability that the variable option stays the same from one encounter to the next is *q*. The probability that the variable option stays the same from one encounter to the next is q. A forager can make two kinds of errors, (i.e., choices for the less rewarding option) in this environment: an overrun error if the forager visits the alternative option when the variable is in its good state, and a sampling error if the forager visits the variable when it is in its bad state. The relative cost of these two errors is the ratio .

$$\mathrm{\epsilon} = \displaystyle \frac{\rm Cost \;of\; a\; sampling\; error\; }{\rm Cost \;of\; an\; overrun\; error\;} = \frac{\ v\_a - v\_b }{\ v\_g - v\_a} $$

Thus the optimal sampling period, i.e., tracking, could be solved for in terms of these two variables. This simple model had several predictions. First, tracking behaviour should decrease with a decrease in . This is because sampling errors become more costly and overrun errors become less costly: increases. When , tracking behaviour should stop completely. Second, and conversely, tracking behaviour should increase with , as overrun errors become more costly, and decreases. Third, tracking behaviour should decrease as *q* decreases, as the states of the variable option become more stable.

These predictions were partially held up by some experimental studies. Hummingbirds were found to decrease their sampling rates as the probability of change of the varying option decreased, as predicted, but did not avoid the variable option when the value increased (Tamm 1987). Similarly, the behaviour of pigeons qualitatively conformed to the predictions of the model, but quantitatively best explained by a model of choice where reward rate is maximized on a moment-to-moment basis based on scalar expectancy (Shettleworth et al. 1988). These experiments manipulated but not *q*. When *q* was manipulated in an experiment with blue jays, presented with either a high and a low rate of change, both sampling and learning rates - i.e., tracking - were found to increase at faster rates of environmental change (Dunlap and Stephens 2012). Similarly, bumblebees sampled the variable resource more frequently when the probability of change was high, as predicted, but did not consistently choose the more rewarding option except when the probability of change was low and the potential reward was very high (Dunlap, Papaj, and Dornhaus 2017).

The merit of the Stephens model is that it outlines the minimum theoretical basis of the issue of environment tracking in order to generate quantitative predictions in a real environment. In a real-world context, however, it is instructive to consider the limitations of the model. A very important assumption of the model is that the forager not only knows the values of the parameters *q* and , but also knows the structure of the environment: that the variable option switches between a good and a bad state. A real foraging animal can only have a distribution of values as an estimate for the parameters, and can never know the whole structure of its environment. Indeed, since *q* is the probability of change at every encounter with the variable option, knowing the current state of the variable option does not say anything about what its state will be at the next encounter.

Another caveat that affects the predictions of the model is that sampling, by definition, should never occur when the forager is exploiting the variable option. When the state of variable option is known, the state of the environment is known, so a subsequent visit to the stable alternative option will not yield any further information. Thus, the model’s predictions only apply to what a forager does when it is at the alternative option. Thus, basic assumption of the model, namely that (a) the forager never visits the stable option when exploiting the variable one in its good state, and (b) that the forager immediately switches to the stable option when the variable option switches to its bad state are not met in any of the systems that have been studied. This is a serious issue because it means that while the conceptual contributions and rationale of the model are still valuable, its quantitative predictions are not valid because its assumptions are not fulfilled. These different kinds of foraging errors are discussed in a study by Commons, Kacelnik, and Shettleworth (2013), which offers a series of models for a similar situation but in which the strategies are based on the observation that assumptions of the Stephens model are not met in real datasets.  
In this study, consisting of two experiments, we attempted a empirical implementation of the model to study the tracking behaviour of the nectar-feeding bat *Glossophaga mutica* (Calahorra-Oliart, Ospina-Garcés, and León-Paniagua 2021). The natural foraging environment of these animals consists of mainly of flowers that contain varying levels of nectar. From the point of view of an individual bat that encounters a flower, the nectar levels change constantly: increasing gradually according to the flower’s nectar secretion rate and decreasing according to how many competitors are present in the environment. Bats must constantly compare flowers in different states: full, partially full, or empty.

In our experiments we placed the bats in an environment containing exactly two ‘flowers’: a flower that always yielded the same volume of reward - a fixed option - and a flower that yielded a reward whose volume changed as a sine function of time, increasing and decreasing. We termed the latter a ‘fluctuating’ option instead of a ‘variable’ option, to differentiate it from an option that could only be in two states, good and bad. While most previous empirical tests of tracking models manipulated either the rate of environmental change (*q*) or the relative cost of the two kinds of errors (), we varied the equivalents of both parameters systematically.

The average relative cost of sampling the two options was determined by the volume of the fixed option. An additional factor is that behaviour may not be driven directly by the absolute real values, but by how they are perceived, and it may be useful to take into account how perception works. In many foraging situations, animals discriminate between relevant variables such as reward magnitudes and time costs according to Weber’s Law, that states that the just-noticeable difference to a stimulus is proportional to the magnitude of the stimulus (Fechner (1860); see Kacelnik and Brito e Abreu (1998) for its application to foraging). In our first experiment the fixed option yielded a reward at the arithmetic mean of the maximum and minimum volumes of the fluctuating option. In the second, the fixed output was smaller than the arithmetic mean. By fixing it at the geometric mean of the variable extremes, we aimed at making the fixed volume equally discriminable from the minimum and maximum values of the fluctuating option, that is, we fixed it at the variable option’s ‘subjective’ mean.

The environmental rate of change in our experiment was determined by the period of the sine function governing the fluctuating output: the smaller the period, the faster the change. In both experiments the bats experienced the same four periods. It is important to note that in this study the rate of environmental change does not correspond exactly to *q*, as the fluctuating option changes, not probabilistically, but systematically. From the point of view of the bats the reward on an encounter with the fluctuating option changes from the last encounter when it is discriminably different. The lower the period, the more likely it is that the the fluctuating output is different for a given encounter rate, and so is an equivalent of the parameter *q*.

Stephens’ model applies to a situation where a foraging agent that is perfectly informed about its environment would follow the model’s predictions. This is because system described by the model is intrinsically stochastic: it behaves according to some probability of changing state. Therefore, even an ideal forager would show errors in its behaviour in such a system. In our experiment however the system is deterministic, so an ideal agent would in fact behave optimally without any error at all, allocating its behaviour entirely to whichever option was yielding a higher reward at any point in time. A realistic agent on the other hand, does not know everything, even in a deterministic scenario. From the point of view of a real agent, the system does behaves as if it were stochastic. For these reasons, our experiment is inspired by Stephens’ model but not designed to test it. It is an empirical study aimed at understanding how and if bats exploit fluctuations in their environment.

We redefined tracking behaviour in our experiment as an outcome, along the lines of Dunlap, Papaj, and Dornhaus (2017): allocating choice behaviour by matching the option yielding the larger reward at time of each choice **(see Figure whatever - I will insert an explanatory figure later in the Methods where it is appropriate)**. This is in contrast to the original mathematical model and some previous studies which put tracking in terms of sampling as its mechanism. A closer match between an animal’s choice behaviour and the state of the environment meant that the animal was tracking better: a perfectly tracking bat would always choose the fluctuating output when it was larger than the fixed, and choose the fixed when it was larger than the fluctuating.

We predicted that tracking would be better when a) the period of the sine function was larger, i.e., the environment was changing more slowly and b) when the contrast between the fixed and fluctuating options was higher. The latter condition was satisfied, not whent the fixed output was the arithmetic mean, but when it was the subjective mean. By definition the subjective mean was equally discriminable from the best and worst fluctuating outputs, and so the arithmetic mean was less discriminable from the best fluctuating output than from the worst. We referred to the experiment where the fixed option was the subjective mean as the ‘high contrast’ experiment and where the fixed option was the objective mean as the ‘low contrast’ experiment.

We also investigated how much the bats had learned the structure of their environment. We did not expect the bats to learn the complex rule of the environment, i.e., that fluctuating output varied sinusoidally. Instead, we thought it was possible for the bats to detect an increasing or decreasing trend in the fluctuating output and for this to influence their choice behaviour. Thus we compared the choice for fluctuating volumes when these volumes were part of a downward trend, to the same volumes when they were part of an upward trend.

# Materials and Methods

## Subjects and housing

Both experiments were done at the Cognitive Neurobiology Lab at the Humboldt Universität zu Berlin: the high contrast experiment in December, 2019; the low contrast experiment in June and July, 2020. The experiments were performed with two different sets of individual bats, and were identical in their design and procedure except for the one critical difference of the volume of reward delivered by the fixed option (see **Experiment Schedule** below).

Bats of the species *Glossophaga mutica* from a captive colony at the Humboldt Universität were used for the experiment. The colony was a breeding population housed at 18-24C and 45-70% humidity on a 12-hour light-dark cycle (light phase: 0200 to 1400 CET; 0300 to 1500 CEST). In this colony every bat older than approximately a year (judged through the ossification of the finger joint - **Brunet-Rossinni and Wilkinson, 2009**) was assigned a permanent ID number, which shall be referred to from now on in order to distinguish the individuals. The bats that were selected for the experiment were a mix of animals that had previously been exposed to the experimental apparatus, and naive ones. None of the bats had participated in such an experiment, or a similar one, before. 16 animals completed the high contrast experiment: 11 females and 5 males. 18 animals completed the low contrast experiment: 10 females and 8 males.

## Experimental Setup

The experimental setup was common to both experiments.

### Reward

The reward received by the bats during the experiment was also their main source of food. The reward was a 17 ± 0.2% by weight solution of sugar dissolved in water (prepared fresh everyday or every other day), hereafter referred to as ‘nectar.’ The sugar consisted of a 1:1:1 mass-mixture of glucose (“Traubenzucker,” Müller’s Mühle GmbH, Germany), sucrose (“Zucker,” Belbake, Südzucker AG, Germany) and fructose (“Fruchtzucker,” Hamburger Zuckerhandelsgesellschaft mbH, Germany). The nectar was thus similar in composition and concentration to the nectar produced by wild chiropterophilous plants (Baker, Baker, and Hodges 1998).

### Experimental Apparatus

The animals were placed in individual, adjacent cages (0.7 x 2.2 x 1.5 m) for the duration of the experiment. As there were six cages in total the experiment was carried out in batches of six bats at a time, and each individual progressed through the experiment independent of all the others. Each cage had an operant wall with two electronic reward-dispensing devices spaced approximately 30 cm apart, hereafter referred to as ‘flowers’ (figure 1 and figure 2). Each flower had a circular head and a door controlled by a linear-actuator motor that could move up and down. Just inside the head of the flowers was an infra-red light barrier, and at the back of the flower was a Teflon tube that supplied the nectar to the flower(figure 3). Each Teflon tube was connected to a short piece of soft peroxide-silicone tube that ran through a pinch-valve.

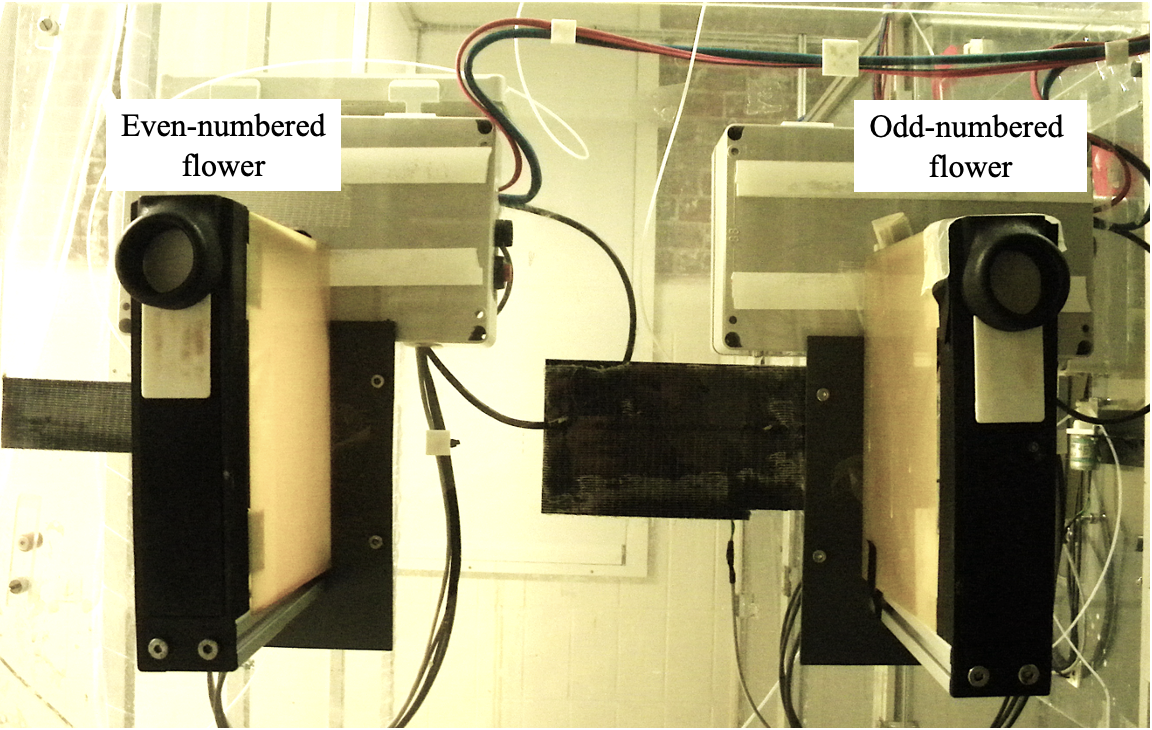


Figure 1: Photograph of operant wall

(ref:flower-cage-schematic) Schematic of cage and operant wall with flowers

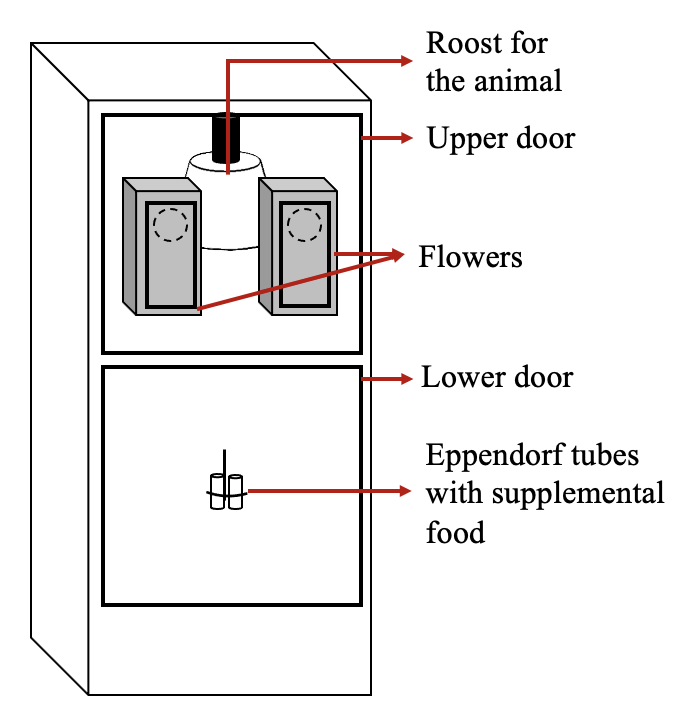


Figure 2: (ref:flower-cage-schematic)

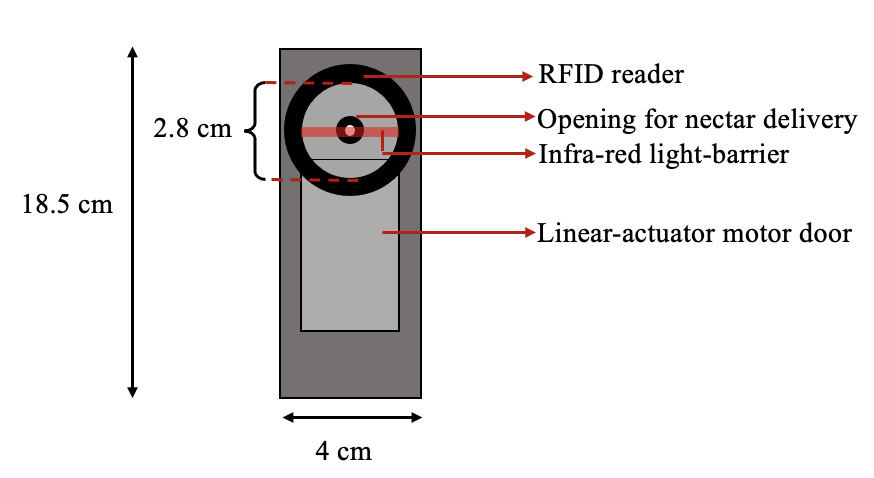


Figure 3: (ref: flower-parts)

The Teflon tubes were connected to a syringe pump in a branching design that ensured the length of tube between every flower and the pump was exactly equal to 470 cm. The pump was placed outside the cages on a shelf, inaccessible to the bats. The syringe of the pump was a Hamilton 25 mL glass syringe (Sigma Aldrich, Germany) and connected to the tubing system of the flowers through five pinch valves on the pump. These 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.

The flowers and the pump were connected by ethernet cables to a laptop computer (ThinkPad, IBM) that stood outside the cages. This computer ran the experimental schedule and the program used to clean and fill the systems using the PhenoSoft Control program (Phenosys, GmBH, Germany). To trigger a reward a bat had to place its nose inside the flower and break the infra-red light barrier. This sent a signal to the computer, which triggered the pinch-valve to open and the pump to move the correct number of steps.

## General Experimental Procedure

Data-collection was completely automated and happened for 12 hours every day. The experimental animals were kept on the same light-dark cycle as the bats in the colony and were active during the dark phase, which is when the data were collected. The experiment was prepared everyday in the morning during the light phase. The animals were inspected everyday to make sure they were healthy and flying well. Then a preliminary analysis of the data from the previous night was done everyday on the laptop running the experimental program using a Shiny App written in R, to make sure the program had been executed correctly and the bats had drunk sufficient nectar. The minimum quantity of nectar was an amount that yielded 25 kiloJoules of energy. Any bat that drank less than this amount was given honey water for an hour before the start of the experiment.

The old nectar was flushed from the system using the automated PhenoSoft program and fresh nectar refilled. Twice a week, the pump and tubing system was thoroughly rinsed with 70% ethanol and de-calcified water to remove pathogens.

At approximately 1800 h the data were checked to see if all the bats had made at least two visits to the flowers, and thus learned to trigger rewards. If bats had not made visits, they received *ad-libitum* honey water for the rest of the experimental night and they were replaced with another animal on the next night.

The bats were given supplemental food in addition to the nectar from the flowers. 0.2 g of a powdered nectar mixture (NEKTAR-Plus, NEKTON, Germany) and 0.3 g of milk powder (Milasan “Folgemilch 2,” Sunval Baby Food, Germany) mixed in approximately 1 mL of water, and 2 mL of plain water were given to each bat. These supplements were put into Eppendorf tubes attached to the operant wall of the cage, about 87 cm below the flowers. The additional food was such that the bats would prefer to visit the flowers instead, both because the flowers were at a more comfortable height for the animals and because the nectar had a higher sugar content and was preferred to the milk powder-nectar supplement mix. The additional food was given firstly to supply micronutrients to the bats while they were in the experiment, and secondly to ensure the animals received a sufficient number of calories in case there was a technical system failure or the bats did not make a sufficient number of visits to the flowers. No technical failures occurred during either experiment.

Once an animal had completed the experiment, it was removed from the cage, weighed to see if it had lost weight since the start of the experiment, released back into the colony and replaced with another bat.

During the experimental night, when the syringe of the pump had been fully emptied, the pump had to refill with nectar from the reservoir. This event happened on average 3.85 times per night (SD = ± 0.26), taking 6.6 minutes each time (SD = ± 1.63). During this time, if the bats made visits to the flowers, they did not receive any reward.

## Experiment Schedule

In both experiments, one option was the ‘fixed’ option and the other was the ‘fluctuating’ option. The fluctuating option delivered a reward that varied as a sine function of time, starting at its maximum volume when a bat made its first visit to the fluctuating option, and proceeding through the sine-function regardless of where the bat made its subsequent visits. In the high contrast experiment the reward delivered by the fixed option was selected so that the volume pairs of the fixed option and the minimum output of the sine-wave, and the fixed option and the maximum output of the sine-wave were, in principle, equally discriminable. This was based on the relative intensity of the volume pairs, calculated as follows:

In the low contrast experiment, the output of the fixed option was the arithmetic mean of the peak and trough volumes of the fluctuating option, and so was less discriminable from the peak than the trough. The maximum volume of the fluctuating option, i.e., the peak of the sine-wave, was 25 L, and the minimum was 2 L, so the output of the fixed option was 7 L in the high contrast experiment and 13.5 L in the low contrast experiment.

The experiment proceeded through the following stages:

### Pre-training

On the first day of the experiment the bats were placed inside the cages and allowed to acclimatize to the new environment. The flowers were covered with a towel to prevent the animals accessing them, and containers of honey water were placed on top of the covered flowers, which the bats found easily. On this day alone no other food was given, not even the supplementary mixture. Food was only available at the location of the flowers. No data were recorded by the computer on this day, and the amount of honey-water consumed was not monitored.

### Training

Shortly before 1400 h, the towels were removed from the flowers so the bats could access them. To teach the bats to put their noses into the flower head and trigger the reward, a drop of honey was applied to the back of the flower and a drop to the top of the flower.

The training proceeded in five phases that repeated throughout the night. Whenever the bats completed 50 visits to both flowers in total, the phase ended and the next began.

1. **Initial:** The doors in front of the flowers remained open, and the bats could pay a visit to whichever flower they wanted. The bats received a reward volume of 25 L at both flowers.
2. **Forced 1:** This was a phase of forced alternation. At the start of this phase, the door in front of one of the flowers moved up to prevent access to it, forcing the bat to visit the other one. After a visit was made and the reward collected, the door of the visited flower would move up to block access to it, and door of the other flower would open. In this way the bat was forced to alternate its visits to the two flowers and so ensure that the locations of both flowers were learned. In this phase there was a difference in reward volume between the two flowers. Two pairs of volumes were possible: the fixed output and 2 L; or the fixed output and 25 L. Depending on which experiment it was, the fixed output was either 7 L (the subjective mean) or 13.5 L (the objective mean). Half the bats were given one volume pair, and the other half the other volume pair. The flower on which the higher volume was given was counter-balanced across animals.
3. **Free 1:** This was a phase of *ad-libitum* reward similar to the Initial phase: both flower doors were open so both flowers were freely accessible to the bats. The volume differences of the Forced 1 phase were maintained. As the bats were free to visit both flowers, the preference of the bats for the flower that gave the higher volume was taken as indication of the discriminability of the volumes.
4. **Forced 2:** This phase was the same as the Forced 1 phase except the volume pairs were different. Those bats that received the fixed output vs. 2 L volume pair in the Forced 1 phase now received 25 L vs. the fixed output and *vice versa*. Half the bats received the higher volume at the same flower as Forced 1 and the other half at the other flower.
5. **Free 2:** This was similar to the Free 1 phase, in that both flowers were accessible and reward was *ad-libitum*, but the reward volumes at the flowers were the same as those in the phase Forced 2. In this way the bats’ preferences for the higher volume of both volume pairs was determined.

After the bats had completed all five phases, the schedule repeated itself except for the Initial phase. This continued for the rest of the night. If a bat learned to trigger rewards and made visits, but not a sufficient number to experience all five phases at least once it had to repeat the Training stage on the next night. If the bat did not complete all five phases even on the second day of Training it was removed from the experiment and replaced.

### Main Experiment

The bats experienced four experimental conditions, corresponding to four periods of the sine wave:

* 0.75 hours
* 1.5 hours
* 3 hours
* 6 hours

The period of the wave was the time interval between two consecutive peaks or troughs. During each experimental night the bats were given free choice between the fixed option and the fluctuating option whose output varied by a sine function of time, calculaed as follows:

$$ \mathrm{y(t)} = {\rm Asin(2\pi ft + \varphi) + D} $$

where:

* A is the Amplitude of the wave, or the distance between the peak and the mid-value of the wave
* f is the frequency of the wave, or the reciprocal of the wave period in seconds
* t is the time point in seconds since the start of the wave
* is the Phase, specifying in units of radians where the wave is when t = 0
* D is the Displacement, or a center Amplitude that is not 0

The bats first experienced a condition for a night, during which the fixed and fluctuating options were assigned to a flower location each, and this location did not change. On the following night there was a reversal of options, i.e., a reversal of the reward contingencies of the flowers: the flower that had previously been the fixed option was now the fluctuating, and *vice versa*. This was done to control for a location preference by the bats. After the bats had experienced a condition on two successive nights in this way, the next condition was given, so there were 4x2 or 8 experimental nights in total (in addition to the training). The order of the conditions was pseudo-randomized across animals.

On the first night of the main experiment the fluctuating option was assigned to the flower that each bat had made more visits to overall on the previous training night, as it was assumed that the animals now had a slight preference for this flower. From then on the reversal of reward contingencies between the two flowers happened every night. At the start of each experimental night, the sine-function that determined the fluctuating output did not begin until the bat made a visit to the fluctuating output. Then the bat experienced the peak of the wave, i.e., the highest possible fluctuating output (25 L). This was a large reward, and designed to motivate the bats to make repeated visits to the fluctuating option so they could experience the change in the output (see **Supplementary Information**).

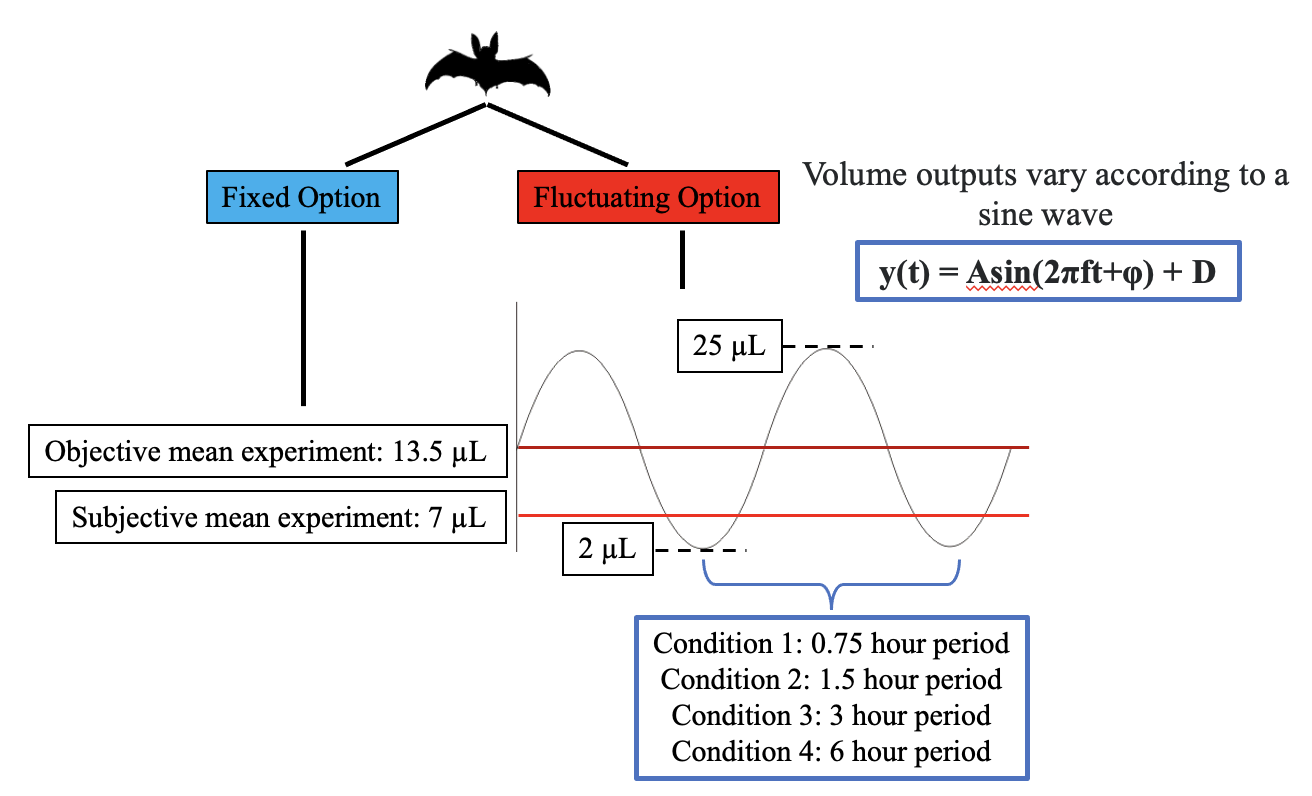


Figure 4: Schematic of the design of the subjective mean and objective mean experiments

## Data analysis - WRITE THIS SECTION ALONE LATER

The raw data from these experiments were logged as events by a computer and recorded in comma-separated value (CSV) files. Each event included the date and time of the event, the animal that made the event, the duration of interruption of the photo-gate and the volume of nectar dispensed. The CSV files were then read into R, which was used for all statistical analyses and creation of plots.

A bat had to experience the reward contingencies of both options on every night to be included for the statistical analysis. In practice this meant that the bat had to make at least one rewarded visit to both options every night.

The bats experienced the reward volumes in the fluctuating options as part of either a downward trend, when the fluctuating output was decreasing, or part of an upward trend, when the fluctuating output was increasing. In both cases the volume difference between the fixed and the fluctuating options was exactly the same but the difference was in the volume differences experienced just before. The bats could use their past experience in one of two ways: they could either estimate an option as being more rewarding based on their reinforcement at that option in the recent past; or they could estimate an option as being more rewarding based on their experience of an increasing reward output at that option, despite the recent past reinforcement being comparatively low. In the first case, we would expect the proportion of visits to any volume of the fluctuating option to be higher when that volume was part of a downward trend; in the second case, we would expect the proportion of visits to be higher when the volume was part of an upward trend. We also considered the specific case of the volume pair 7 vs. 13.5 L. In the subjective mean experiment this situation arose when the fluctuating output was 13.5 L because the fixed output was always 7 L; in the objective mean experiment it arose when the fluctuating output was 7 L as the fixed output was always 13.5 L. The volume pair was discriminable by the bats, and if there were no effect of trend, the preference for the higher volume should be higher than 50% in all the experimental conditions.

In both the experiments, we investigated the effect of trend, volume of the fluctuating output and rate of change of the fluctuating option on the proportion of visits to the fluctuating option. In both the experiments we also created separate models of one specific pair of volumes: 7 and 13.5 L. In the subjective mean experiment this was when the fluctuating output was 13.5 L, and in the objective mean experiment this was when the fluctuating output was 7 L. We investigated the effect of trend on the proportion of visits made to the higher volume in this pair (13.5 L) in both experiments. The proportion of visits to the fluctuating output was calculated as the number of visits to the fluctuating output divided by the sum of the number of visits to the fluctuating output and the number of visits to the fixed option in that category. The proportion of visits to the higher volume of a volume pair was calculated in a similar manner.

Generalized linear mixed-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 technical details of these models are provided in the **Supplementary section**. We present plots of the conditional effects of the predictor variables, with the parameter values of the models provided in the **Supplementary section**. We report 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)).

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

# Results

## A majority of bats responded to the reversal of location of the two options

Two behavioural strategies were observed in the main experimental phase. The locations of the fixed and fluctuating options were always reversed between the two flowers on the second night of a condition to control for the bats’ location preferences. While most of the bats made visits to both options on both nights, a minority did not. 4 out of the 16 bats in the subjective mean experiment, and 3 out of the 18 bats in the objective mean experiment, made near-exclusive visits to the same flower on both nights of a condition, regardless of whether that flower was the fixed or the fluctuating option. We designated these bats the ‘reversal non-responsive’ bats, as reversing the location of the fixed and fluctuating options induced no observable behavioural response.

**explain the NRMSE criterion for exclusion**

Figure 5 shows the overall activity of the reversal non-responsive bats. The first time point of the sine function was the first visit made by a bat to the fluctuating option. This meant that on those nights the fixed option was assigned to the preferred flower of a reversal non-responsive bat, the bat never experienced the changing output of the fluctuating option and was thus ‘uninformed’ of all the available options - these animals were excluded from statistical analyses.

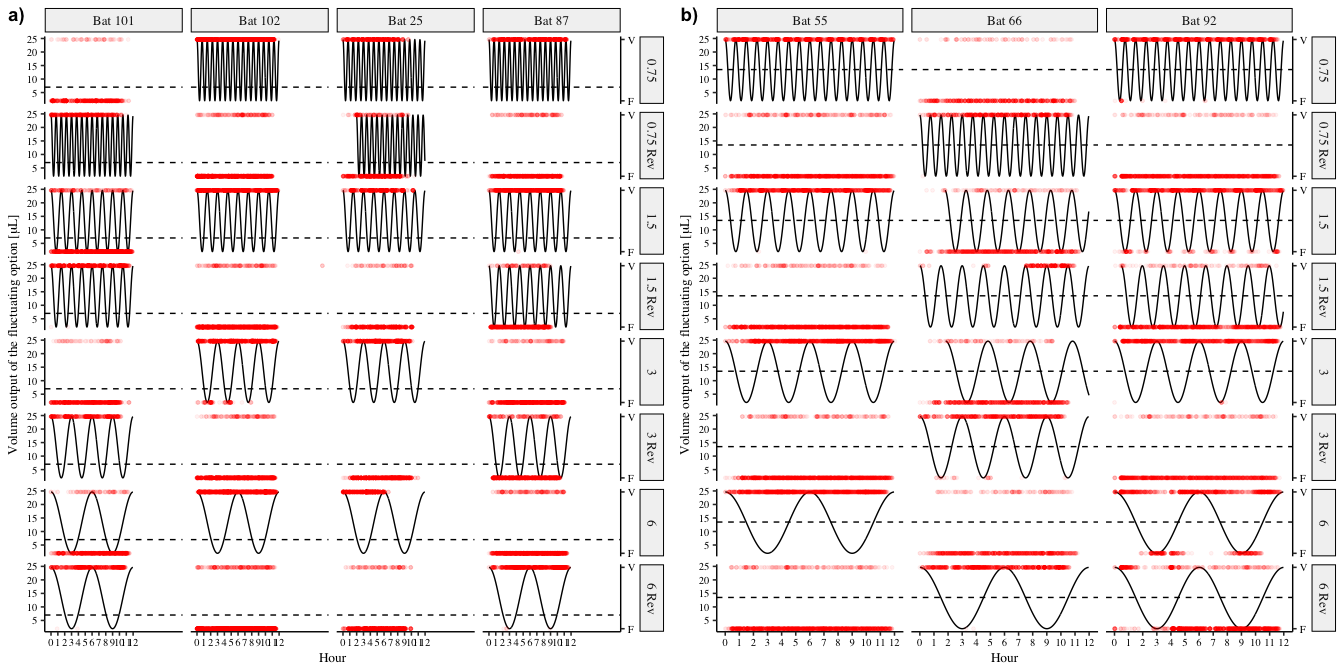


Figure 5: Choice behaviour of all the reversal non-responsive bats in the two experiments. Each row is one night of experimental condition, i.e., the two nights for each of the four wave periods, and each column an individual bat. The solid black line represents the output of the fluctuating option and the red points each individual visit made by a bat. The red points to the top of the plots are visits made to the fluctuating option and those at the bottom of the plots are visits made to the fixed option. The dashed horiontal line represents the volume output of the fixed option. a) Reversal non-responsive bats in the subjective mean experiment b) Reversal non-responsive bats in the objective mean experiment

The animals that did respond to the reversal showed a change in their choice behaviour corresponding to the output of the sine wave. This is represented in Figure 6.

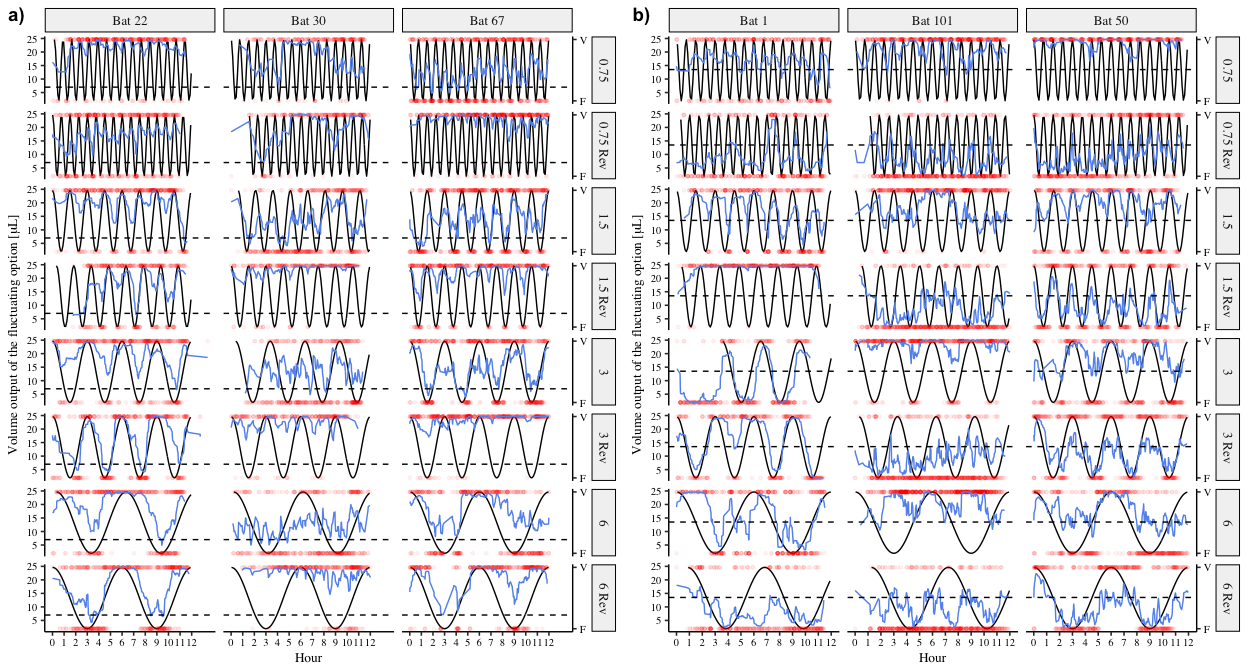
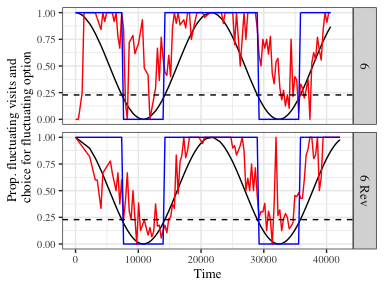
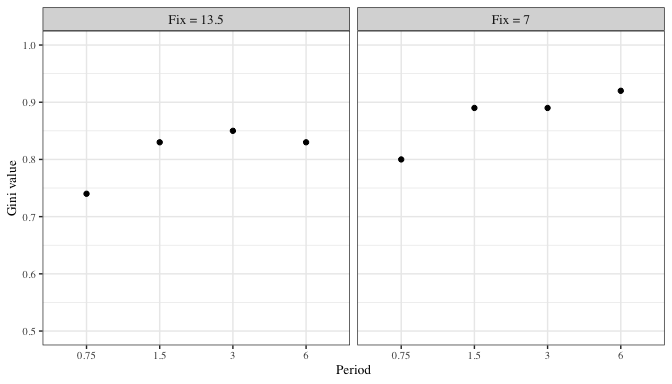
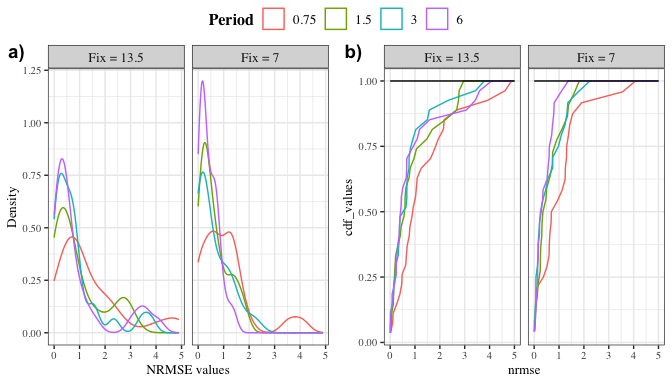


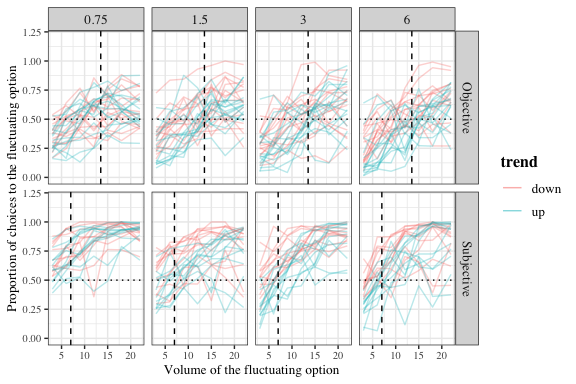
Figure 6: Choice behaviour of three representative reversal responsive bats from each of the two experiments. Each row is one night of experimental condition, i.e., the two nights for each of the four wave periods, and each column an individual bat. The solid black line represents the output of the fluctuating option and the red points each individual visit made by a bat. The red points to the top of the plots are visits made to the fluctuating option and those at the bottom of the plots are visits made to the fixed option. The blue lines are a smoothing function applied to the choices of the bats. The dashed horiontal line represents the volume output of the fixed option. a) Three of the reversal responsive bats in the subjective mean experiment b) Three of the reversal responsive bats in the objective mean experiment

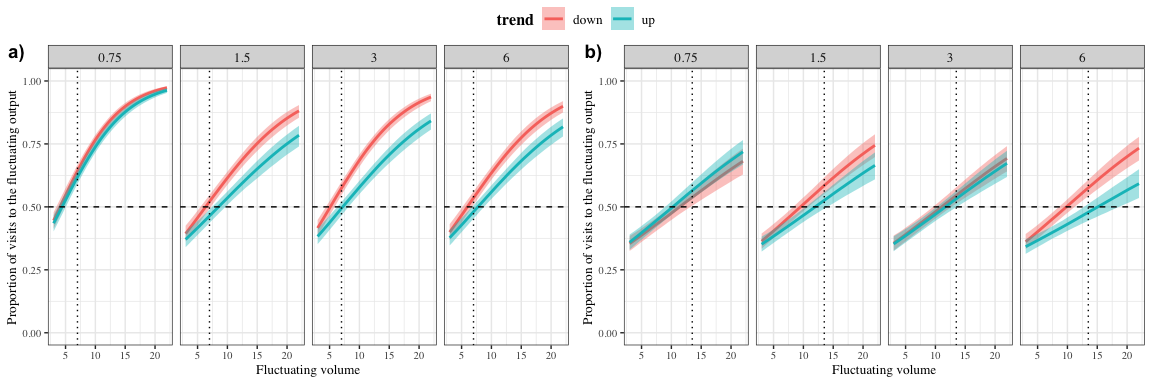
## Slower rates of change and higher contrast between options resulted in increased tracking

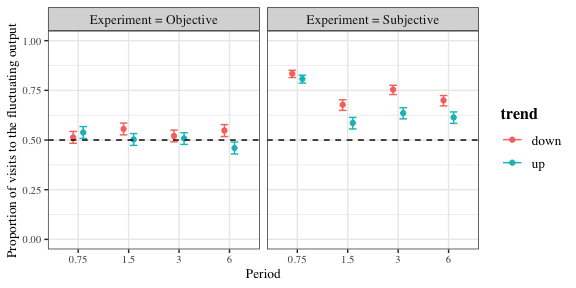




## Reinforcement history interacts with environmental parameters to influence choice







# Discussion and Conclusions

“Living backwards!” Alice repeated in great astonishment. “I never heard of such a thing!”

" — but there’s one great advantage in it, that one’s memory works both ways."

“I’m sure mine only works one way,” Alice remarked. “I can’t remember things before they happen.”

“It’s a poor sort of memory that only works backwards,” the Queen remarked.

***Alices Adventures in Wonderland*, Lewis Carroll**

## Verbal summary of the results

1. Bats respond to time-based change
2. How fine-grained this result is
3. More visits to the fluctuating option when the trend is downward rather than upward
4. A higher preference for a higher volume when that volume is part of a downward rather than an upward trend

## Interpreting the results and tying them up with the aims and rationale

1. Aim 1: to see if the a behavioural response corresponding to the state of the environment, i.e., is there tracking? -> Yes. Bats are capable of responding to an environment that changes in a time-based way on the order of hours, not just seconds
2. Aim 2: Refining aim 1, if there is tracking, how does the rate of change of the environment affect that? -> The faster the change, the worse the tracking. The contrast is reflected as one might expect in the choice behaviour, but there is worse tracking when the fixed option is the subjective mean.
3. Aim 3: What influences the bats’ expectations of the current state of the environment? -> There’s a higher pref for an option if its part of a downward trend rather than an upward one, meaning that the recent past experience of higher reward leads to higher expectation at an option.
4. The role of satiation and why this is cognitive and not purely physiological.

## Fitting it in with results of other papers

1. Model prediction for each of q and epsilon, results from bees, hummingbirds and pigeons, and then our bats
2. Potential explanations - eg. Weber’s Law
3. Overall picture that emerges

## What is known about the bats’ cognitive strategies

1. Connecting what the bats did to what was previously known about their cognitive strategies: volume discrimination; serial reversal learning; modelling in the Science paper – modelling was done (don’t oversell this one too hard); nectar secretion rates chapter from Ulf’s thesis.
2. Was timing necessary at all or just a simple tracking of environmental change? Timing *could* have been used but the bats didn’t use it.
3. A description of what we think is happening: outlining reinforcement learning in the bats

## What these results imply for the bats’ foraging ecology and nectar-feeding animals in general

1. How do nectar levels in flowers change?
2. Patch choice and the subjective/objective mean difference
3. Cannot project, so go by recent memory because of inter-individual competition; but we know they CAN project, like in the nectar-secretion rate experiment
4. Do bats keep a running tally of what was experienced recently in each patch?
5. Connect with what is known about patch choice strategies and the Stephens model

## Time-based environmental change

1. Intermediate time-scales can be tracked with constant environmental input
2. The state of the environment is in no doubt because it is always accessible in this case, which could be a discouragement to use an internal sense of time.
3. Animals can respond to changes on ecologically-relevant time-scales using cognitive strategies

# Acknowledgements

We thank Zlata Shishkina for all her help with the data-collection. We thank Alexej Schatz for the programming of the PhenoSoft Control software. We thank the members of the Winter lab, for many useful discussions. We also thank for their comments and suggestions for the improvement of the manuscript.

# Author Contributions

**SC**: conceptualization, experimental methodology, data-collection, formal analysis, data curation, writing - original draft, writing - review and editing. **AK**: conceptualization, formal analysis, writing - review and editing, supervision. **YW**: conceptualization, resources, formal analysis, writing - review and editing, supervision. **VN**: conceptualization, experimental methodology, 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 by a scholarship from the Deutscher Akademischer Austauschdienst (DAAD) to SC.

## 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

## First visits to the fluctuating option and initiating the sine function

**should we include this or does it imply the bats were endangered or something?** The first visit to the fluctuating option every night triggered the start of the sine function that determined the volumes of the fluctuating output. Thus, the bats’ first experience of the fluctuating option was the peak of the sine function, 25 L, from which point on the fluctuating output changed regardless of where the bats made visits. Most of the bats successfully triggered the start of the sine-wave and experienced the peak fluctuating volume as intended, but for a few individuals on a few experimental nights, their first visits to the fluctuating option were not properly recognised due to a technical error. This meant that the sine wave had begun and the fluctuating output was changing, but the bats had not experienced a reward at this option during their first visit to it. This raised the possibility that during their first rewarded visit to the fluctuating option the bats experience a low reward volume and so could have been less motivated to visit it again and experience the way the fluctuating output changed.

The bats that experienced this ‘false start’ to the fluctuating option are summarized in figure 7. These bats were all responsive, meaning that they made visits to both options on all the experimental nights; and each of them only experienced a non-rewarding first visit to the fluctuating option on one night each. It seemed therefore that this technical error had little to no consequence to the bats and they were included in the statistical analyses without differentiating them in any way.

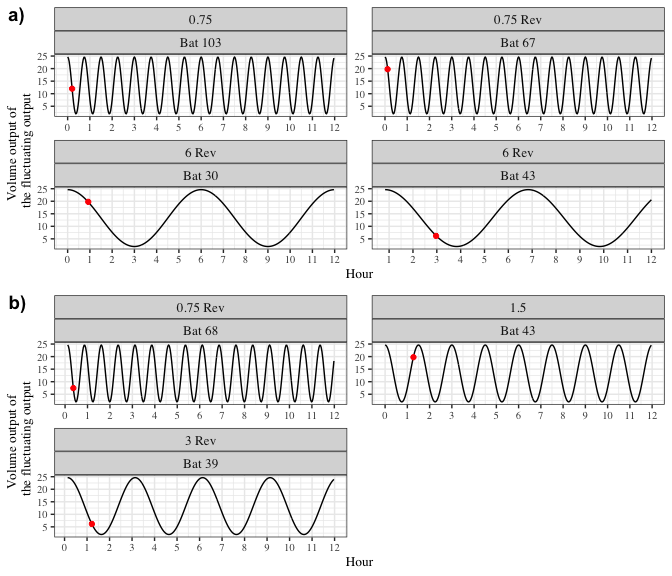


Figure 7: Volumes experienced by a small number of bats at their first rewarded visit to the fluctuating option in a) the subjective mean experiment and b) the objective mean experiment. The black line represents the fluctuating output, the red dot represents when the first rewarded visit to the fluctuating option occurred.

## Details of the statistical analyses

The Bayesian generalized linear mixed-models fitted in brms used weakly-informative priors. The slopes and intercepts 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.

The models investigating the effect of trend, fluctuating volume and rate of change on the proportion of visits to the fluctuating option used 1200 warm-up samples and 1800 post-warm-up samples. A Bernoulli likelihood function was used with trend, rate of change, fluctuating output and their 2-way interactions modelled as fixed effects, with fluctuating output as a continuous predictor and the other two as categorical predictors. Random slopes and intercepts were used to fit regression lines for the individual bats.

The models investigating the effect of trend and rate of change on the proportion of visits to the higher option of the volume pair 7 and 13.5 L used 1000 warm-up samples and 1000 post-warm-up samples. A Bernoulli likelihood function was used with the categorical predictors trend, rate of change and their interactions modelled as fixed effects. Random slopes and intercepts were used to fit regression lines for the individual bats.

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.

## Coefficient values of predictor variables in the models

**Do we need to display the coefficients of the fixed effects?**

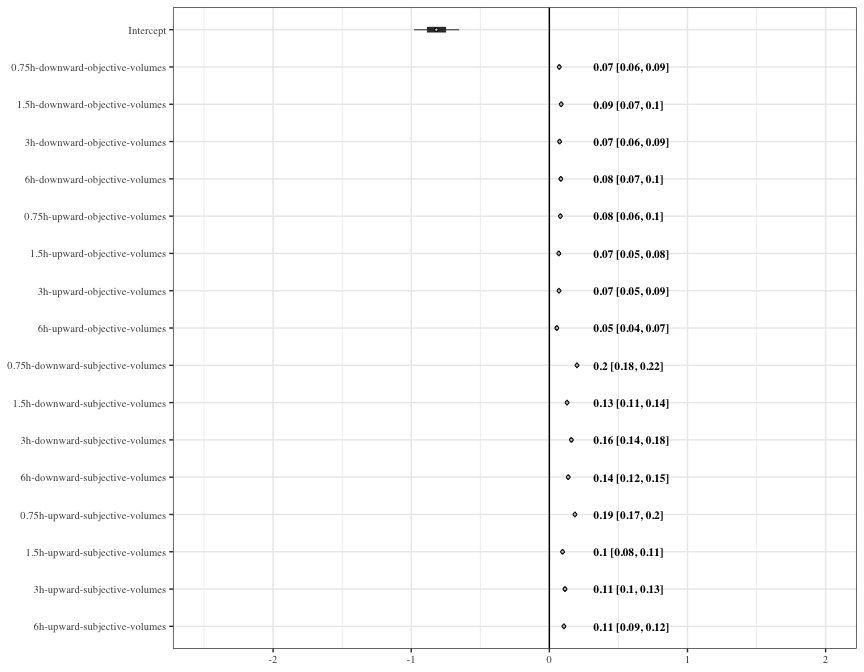


Figure 8: a) Forest plot of the estimates of the effect of Period, fluctuating volume and trend of the fluctuating volume on visits to the fluctuating option in the subjective mean experiment. b) Forest plot of the estimates of the effect of rate of change, fluctuating volume and trend of the fluctuating volume on visits to the fluctuating option in the subjective mean experiment. 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

## Reinforcement history of the fluctuating option - supplement to the trends analysis

# References

**Brunet-Rossini and Wilkinson**

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>.

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>.

Calahorra-Oliart, Adriana, Sandra M Ospina-Garcés, and Livia León-Paniagua. 2021. “Cryptic Species in *Glossophaga Soricina* (Chiroptera: Phyllostomidae): Do Morphological Data Support Molecular Evidence?” Edited by Amy Baird. *Journal of Mammalogy* 102 (1): 54–68. <https://doi.org/10.1093/jmammal/gyaa116>.

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). <https://doi.org/10.18637/jss.v076.i01>.

Commons, Michael L., Alejandro Kacelnik, and Sara J. Shettleworth. 2013. *Foraging: Quantitative Analyses of Behavior, Volume Vi*. Psychology Press.

Dunlap, Aimee S., Daniel R. Papaj, and Anna Dornhaus. 2017. “Sampling and Tracking a Changing Environment: Persistence and Reward in the Foraging Decisions of Bumblebees.” *Interface Focus* 7 (3): 20160149. <https://doi.org/10.1098/rsfs.2016.0149>.

Dunlap, Aimee S., and David W. Stephens. 2012. “Tracking a Changing Environment: Optimal Sampling, Adaptive Memory and Overnight Effects.” *Behavioural Processes* 89 (2): 86–94. <https://doi.org/10.1016/j.beproc.2011.10.005>.

Fechner, Gustav Theodor. 1860. *Elemente Der Psychophysik*. Breitkopf u. Härtel.

Kacelnik, Alex, and Fausto Brito e Abreu. 1998. “Risky Choice and Weber’s Law.” *Journal of Theoretical Biology* 194 (2): 289–98. <https://doi.org/10.1006/jtbi.1998.0763>.

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>.

Shettleworth, Sara J., John R. Krebs, David W. Stephens, and John Gibbon. 1988. “Tracking a Fluctuating Environment: A Study of Sampling.” *Animal Behaviour* 36 (1): 87–105. <https://doi.org/10.1016/S0003-3472(88)80252-5>.

Stephens, D. W. 1987. “On Economically Tracking a Variable Environment.” *Theoretical Population Biology* 32 (1): 15–25. <https://doi.org/10.1016/0040-5809(87)90036-0>.

Tamm, Staffan. 1987. “Tracking Varying Environments: Sampling by Hummingbirds.” *Animal Behaviour* 35 (6): 1725–34. <https://doi.org/10.1016/S0003-3472(87)80065-9>.