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 When an environment changes rapidly any information that is gained by tracking it becomes outdated very soon.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 encounters any one of the two options at discrete time intervals. It 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*. This is the discrete probability that the variable option stays the same from one encounter to the next. 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.  
Our study consisted of two experiments inspired by the Stephens model which aimed 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 gradually. We termed the latter a ‘fluctuating’ option instead of a ‘variable’ option, to differentiate it from an option that could only be in one of two states, good or bad. We varied the rate of change of the fluctuating option, and the contrast between the fixed and fluctuating options 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, consistent with the assumption that the bats’ perception of reward volumes was according to Weber’s Law. that is, we fixed it at the variable option’s hypothetical ‘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.

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 was inspired by Stephens’ model but not designed to test it. It was 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, somewhat along the lines of Dunlap, Papaj, and Dornhaus (2017): allocating choice behaviour by matching to the relative reward volumes of the two options. 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. The ideally tracking bat with a perfect knowledge of the state of the environment would always choose the fluctuating output when it was larger than the fixed, and choose the fixed when it was larger than the fluctuating. Real foragers however, often obey a ‘matching law’ in their behavioural allocation, defined by Kacelnik and Krebs (1991) as follows: “the animal allocates its behaviour between options in proportion to the rewards it has obtained from them.” ((Seth 1999), (Houston, Trimmer, and McNamara 2021)). Essentially, a ratio of the animal’s choices matches a ratio of the possible rewards. Thus we calculated the goodness of fit between the animals’ choice behaviour and a ratio of the reward volumes of the two options calculated as follows:

$$\displaystyle \frac {\rm Fluctuating \ volume}{\ (\rm Fixed \ volume + Fluctuating \ volume)}$$

This is slightly different from the formula specified by the original matching law (Herrnstein 1970), which is the ratio of the reward magnitudes without summing them. This was in order to account for the fact that rewards are simultaneously available at the two options (see figure 1).

(ref:predictions) Plot of the different ways an animal’s behaviour might correspond to the state of the environment. The ordinate indicates the proportion of visits to the fluctuating output, the abscissa the time progression of the experimental night. The black line indicates what the choice allocation would be if the bats’ behaviour exactly corresponded to the fluctuating volume; the orange line if the bats’ behaviour corresponded to the ratio of the fluctuating volume to the sum of the fixed and fluctuating volumes; and the dark grey line if the bats chose exactly the option that yielded a higher reward volume at any point in time.

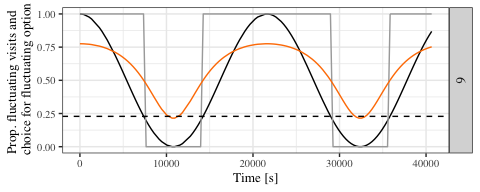


Figure 1: (ref:predictions)

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. Any volume of the fluctuating option could be preceded by volumes that were higher, meaning the downward trend in the sine function was occurring; or it could be preceded by volumes that were lower, meaning the upward trend was occurring. For example, a bat might visit the fluctuating option and receive 10 L of nectar reward, but the reward volume it received at that option on its last few visits might have either been 15, 14, 13, 12 and then 11 L, or 5, 6, 7, 8 and then 9 L. If bats made a higher proportion of visits to the fluctuating option in the former case, the bats might be said to be ‘trailing’ in their estimate of the rewarding properties of this option. If the bats made a higher proportion of visits to the fluctuating option in the latter case, the bats might be said to be ‘predicting’ greater rewarding properties of this option. We referred to these hypotheses as the ‘trailing’ and ‘predicting’ hypotheses respectively (see figure 6 for a schematic summary). Thus we explored the effect of the trend in the reward volume of the fluctuating option on the proportion of choices the bats made to the fluctuating option.

# 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 2 and figure 3). 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 4). Each Teflon tube was connected to a short piece of soft peroxide-silicone tube that ran through a pinch-valve.

(ref:operant-wall) Photograph of an operant wall with its two flowers. Each bat was allowed exclusive access to one such wall.

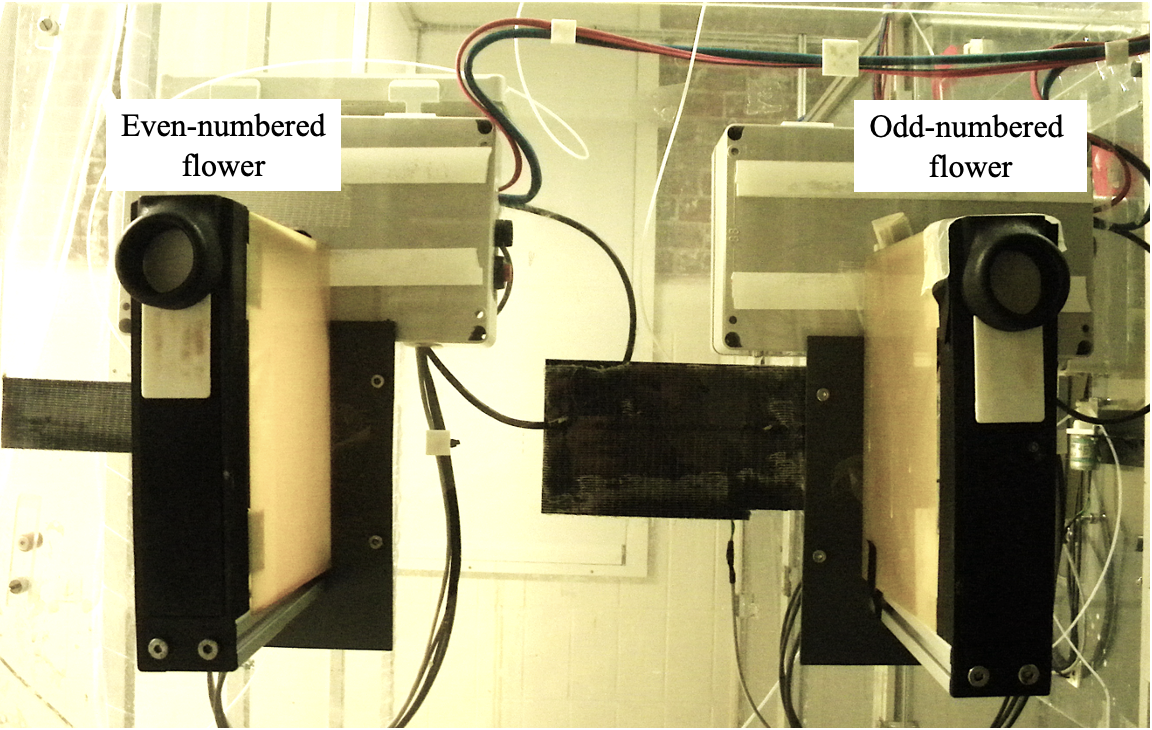


Figure 2: (ref:operant-wall)

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

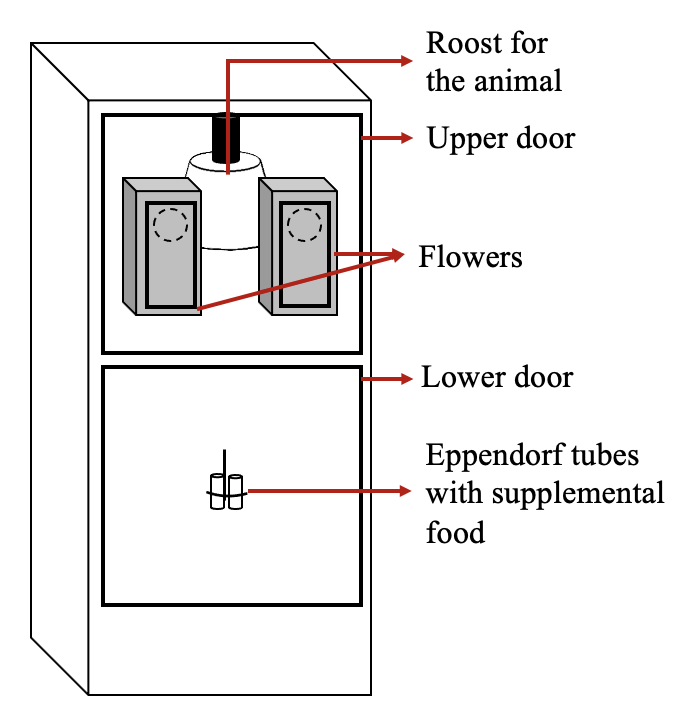


Figure 3: (ref:flower-cage-schematic)

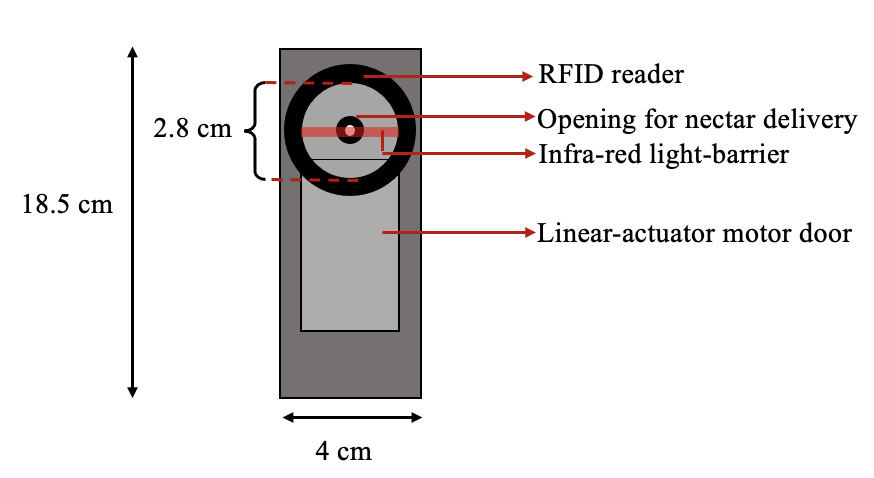


Figure 4: Diagram of a flower with its parts

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 (Nachev, Stich, and Winter 2013):

$$\ \displaystyle \frac{\rm\ volume\_1 - volume\_2 }{\rm (volume\_1 + volume\_2)/2} $$

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 (in the high contrast experiment) or 13.5 L (in the low contrast experiment). 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, calculated 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 option. Then the bat experienced the peak of the wave, i.e., the highest possible fluctuating volume (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 volume (see **Supplementary Information**).

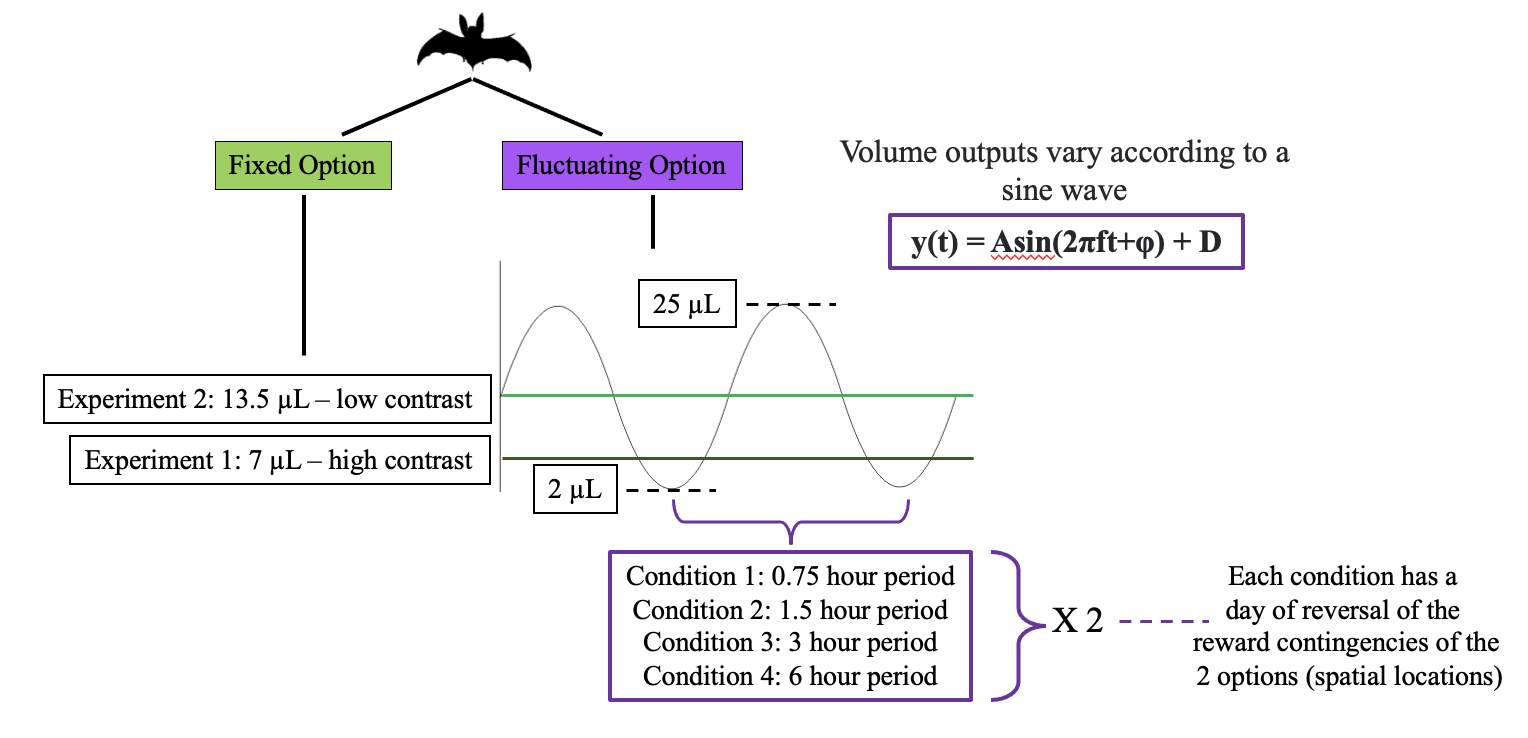


Figure 5: Schematic of the design of the two experiments in the study

## Data analysis

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.

To quantify tracking behaviour the data were first binned into intervals of five minutes. This bin-size was selected because each bat made on average at least one visit in each five minute bin in the vast majority of cases. A goodness of fit measure was calculated between the proportion of visits to the fluctuating option and ratio of the fluctuating volume and the sum of the fixed and fluctuating volumes in each five-minute bin. The goodness of fit measure was the normalized root mean squared error:

$$ \displaystyle \rm NRMSE = \frac {\rm \sqrt[2][{(\sum\_{i = 1}^{n}({Observed\_{i} - Predicted\_{i}})^2})/n\_{}]} {\rm \sigma\_{Observed}} $$

A single NRMSE value was calculated for each bat for each experimental condition: the four different wave periods in each of the two experiments with two different fixed volumes. Then a ‘Gini coefficient’ was calculated for the area under the cumulative distribution curve of the NRMSE values in each condition. This was the ratio between the area under the bumulative distribution curve of the calculated NRMSE values and the maximum possible area under the curve. We termed this quantity the Tracking Index.

While the majority of the bats responded to the reversal of reward contingencies between the two flowers, about 25% did not. The latter made almost exclusive visits to one of the flowers, i.e., to only one of the two spatial locations where reward was available. This meant that on any experimental night these bats experienced almost only the fixed or the fluctuating volume. This was a behaviour so qualitatively different from the rest of the bats that we set a criterion to exclude these animals from the statistical analysis. NRMSE values were calculated for the fit of all the bats’ choice behaviour to the choice predictions of a ‘location preference strategy.’ A bat that exclusively made visits to only one location would have an NRMSE value of 0 on that experimental night. Our criterion for including a bat in the statistical analysis was that it must not have an NRMSE value of 0 on either night of two or more experimental conditions.

To analyse whether the bats’ choice behaviour was trailing behind or predicting the fluctuating volumes (see figure 6), we fit a generalized linear mixed model (GLMM) to the data (see **Supplementary Information** for the technical details of the model fitting). The model was fitted in a Bayesian framework using Hamiltonian Monte Carlo in the R package brms (Bürkner 2017), a front-end for rstan (Carpenter et al. 2017).

The fluctuating volumes were grouped into bins, and the proportion of visits to the fluctuating option was calculated in each bin. The two extreme volumes, 2 and 25 L were excluded from the analysis. The lowest volume bin thus started at 3 L. The size of the bin was the theoretically just discriminable difference from 3 according to Weber’s Law: 3.16 L. Thus the smallest volume bin was [3 - 6.16]. This binsize was applied consistently for the range of volumes to create the volume bins. We did not expect that the bats would be able to discriminate between the volume bins at the higher volumes as well as they did at the lower ones, but reasoned that this would be evident in the data.

When a fluctuating volume was preceded by higher volumes it was classified as part of a downward trend in the sine function. When it was preceded by lower volumes it was classified as part of an upward trend. We did not directly look at the bats’ choice history, but classified the volumes into either one of these two trends based on the programmed volume of the fluctuating option.

The response variable in our model was the proportion of visits to the fluctuating option in the volume bins. The predictor variables were the fluctuating volumes, trend, the period of the sine function and the volume of the fixed option. We investigated the interaction of these four predictors on the response variable as a fixed effect, and allowed the intercepts and slopes due to the effect of volume alone to vary randomly with individual bat. We presented plots of the conditional effects of the predictor variables with the mean as a measure of central tendency and the 89% quantile-based credible intervals for the intercept and slope coefficients (89% boundaries are the default for reporting credible intervals – McElreath (2020)).

(ref:trailing-predicting-schematic) Schematic representation of the hypotheses underlying the analysis of the effect of fixed volume, fluctuating volume, trend of the fluctuating volume and period of the sine function on the visits to the fluctuating option. If the fluctuating volume at a time T0 is higher than the volume at T-1, the fluctuating volume is decreasing, i.e., the sine function is trending downwards. If the fluctuating volume at a time T0 is greater than the volume at T-1, the sine function is trending upwards. Any fluctuating volume except the maximum and minimum could be part of either a downward or an upward trend. We explore two hypotheses with different predictions of the bats’ choice behaviour for at a time T1. If the bats’ estimate of the fluctuating volume is driven by their recent reinforcement history, i.e., they were ‘trailing’ in their estimate, we expected to see a higher proportion of visits to a fluctuating volume at the time T1 when it was part of a downward trend compared to an upward trend. We expected the opposite if the bats were ‘predicting’ the increase in the fluctuating volume.

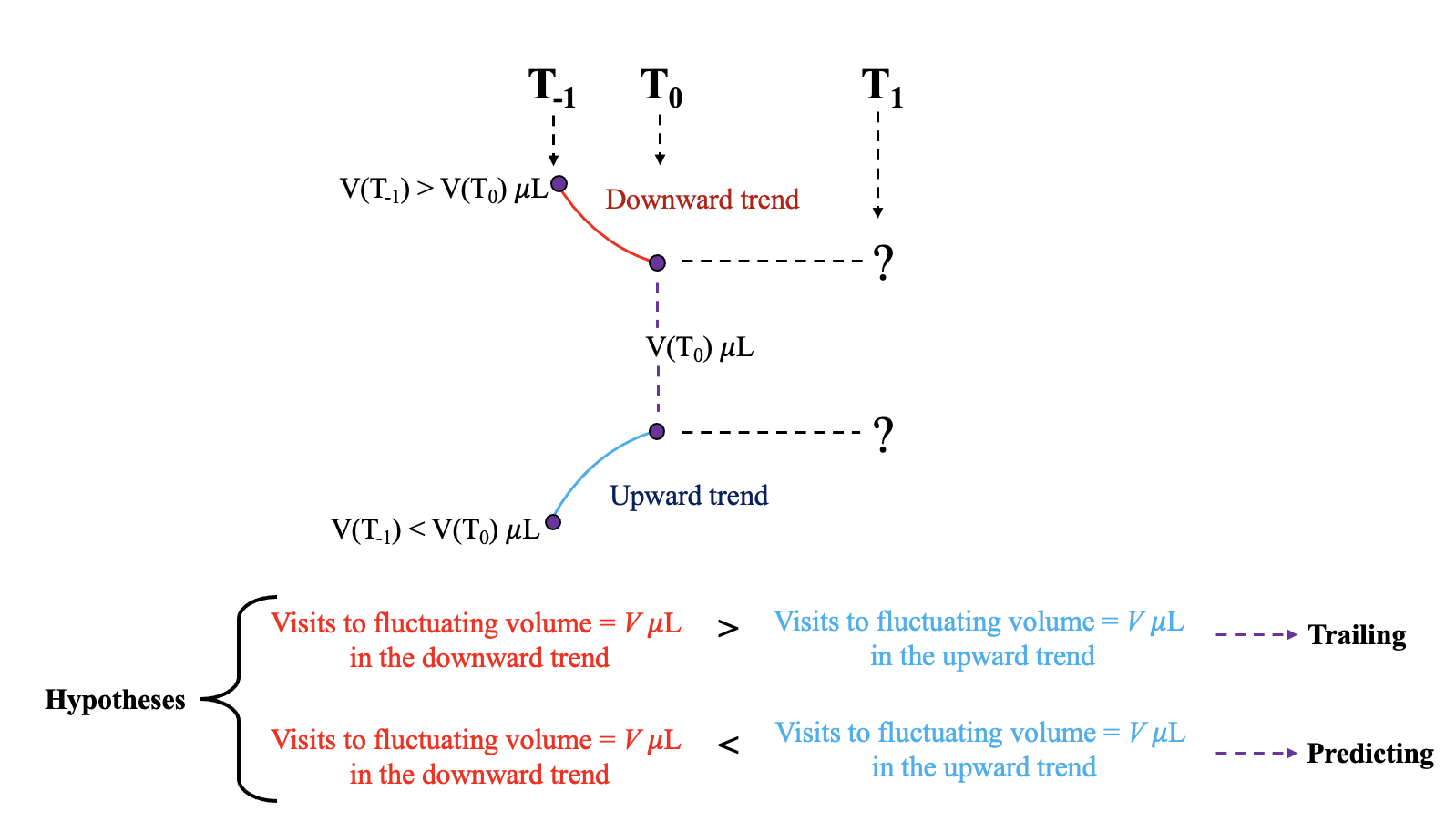


Figure 6: (ref:trailing-predicting-schematic)

# Results

## 75% of the bats responded to the reversal of the location of the two options

Two qualitatively different kinds of choice behaviours 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 high contrast experiment, and 3 out of the 18 bats in the low contrast experiment, made near-exclusive visits to the same flower (i.e., the same spatial location) on both nights of a condition or more than one condition, regardless of whether that flower was the fixed or the fluctuating option. We designated these bats the ‘reversal non-responsive’ bats, and these were the bats that did not pass the criterion for inclusion in the main statistical analyses set by the NRMSE value.

Figure 7 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 reward volumes of the fluctuating option and was thus ‘uninformed’ of all the available options.

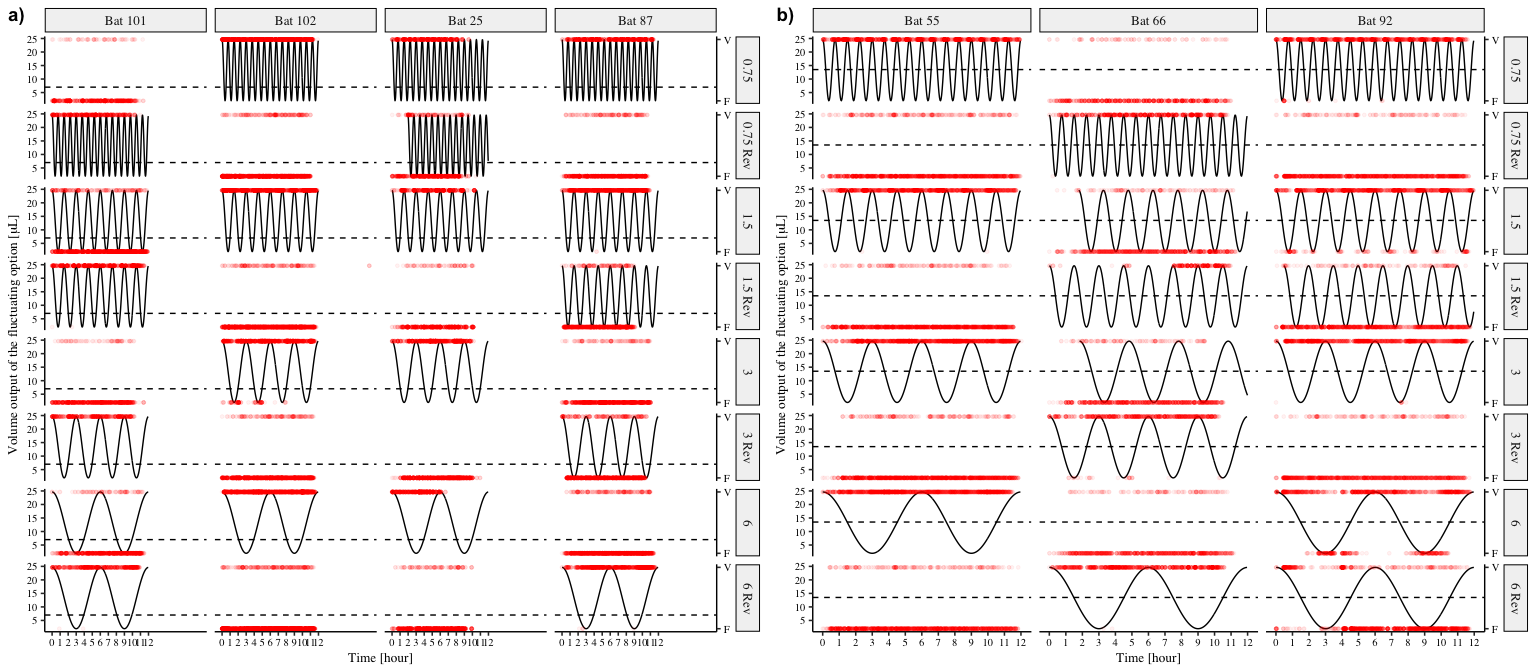


Figure 7: Choice behaviour of all the reversal non-responsive bats in the two experiments. Each row is one night of experimental condition, i.e., the two nights for each of the four experimental conditions, and each column an individual bat. The solid black line represents the volume of the fluctuating 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 horizontal line represents the volume output of the fixed option. a) Reversal non-responsive bats in the high contrast experiment b) Reversal non-responsive bats in the low contrast 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 8.

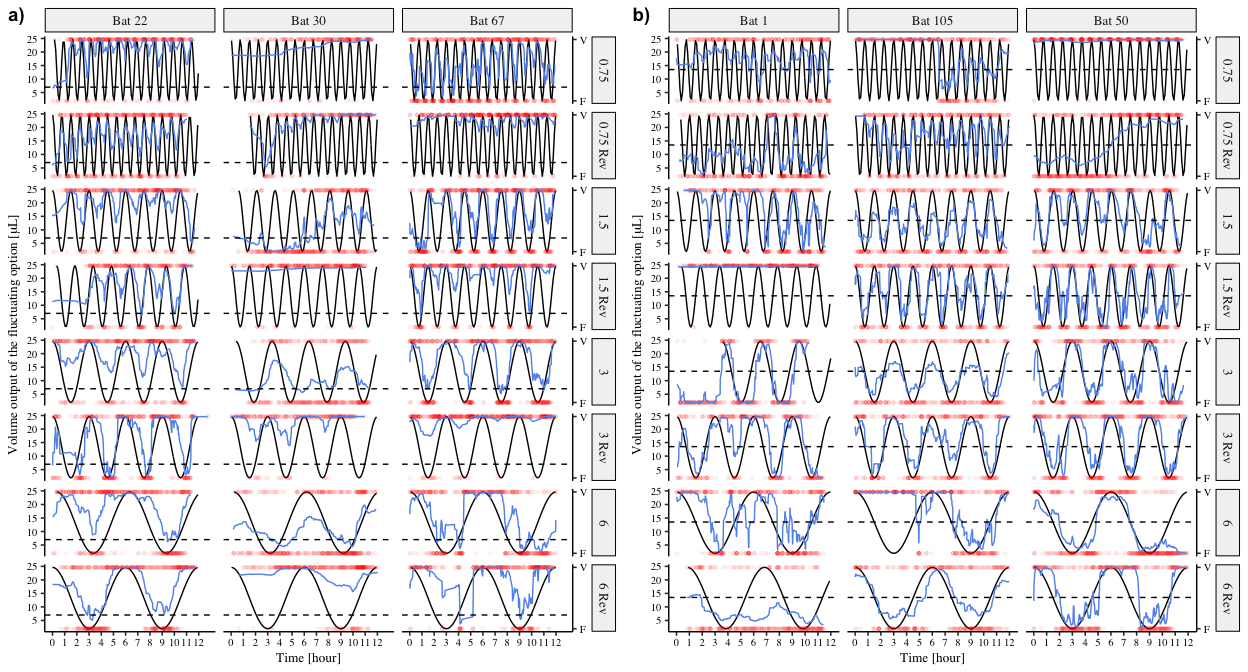


Figure 8: Choice behaviour of three representative reversal responsive bats from each of the two experiments. Each row is one night of experimental condition, i.e., the two nights for each of the four experimental conditions, and each column an individual bat. The solid black line represents the output 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 high contrast experiment b) Three of the reversal responsive bats in the low contrast experiment

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

The proportion of visits to the fluctuating option made by the bats changed as the fluctuating volume changed. In every experimental condition, apart from the 0.75 hour period in the high contrast experiment, the bats made a higher proportion of visits to the fluctuating option when the fluctuating volume was higher than the fixed volume, and lower proportion when the fluctuating volume was lower than the fixed volume (figure 9). For the shortest period in the high contrast experiment alone, the bats made a higher proportion of visits to the fluctuating than the fixed option regardless of the fluctuating volume.

In the low contrast experiment, as the fluctuating volume increased, the proportion of visits to the fluctuating option matched the ratio of the fluctuating volume to the sum of the fluctuating and fixed volumes. This was true in all four experimental conditions, i.e., all four periods of the sine function. In the high contrast experiment, the proportion of visits to the fluctuating volume for the fastest rate of change - the 0.75 hour period - was higher than the ratio of the fluctuating volume and the sum of the two volumes. In the other three experimental conditions there was a closer matching between the choice behaviour of the bats and the volume ratio, though the match decreased at the higher fluctuating volumes.

(ref:matching-proportions) a) Average proportion of visits to the fluctuating option with the change in the fluctuating volume for each experimental condition. The purple, blue, green and yellow lines indicate the experimental condition, i.e., period of the sine function. The orange lines indicate the ratio of the fluctuating volume to the sum of the fixed and fluctuating volumes, capturing state of the environment. Vertical dashed lines indicate the fixed volume and horizontal dotted lines indicate a proportion of 0.5, or an equal number of visits to both options. b) An example of the choice behaviour of a single bat - ID#22 in the 6-hour period condition in the high contrast experiment. The black line indicates the fluctuating volume. The blue line indicates the proportion of visits made by the bat to the fluctuating option. The orange line indicates the ratio of the fluctuating volume to the sum of the fluctuating and fixed volumes. The horizontal dashed line indicates the fixed volume. The top panel is the first night of the experimental condition and the bottom panel is the second night of the condition with the locations of the two options reversed.

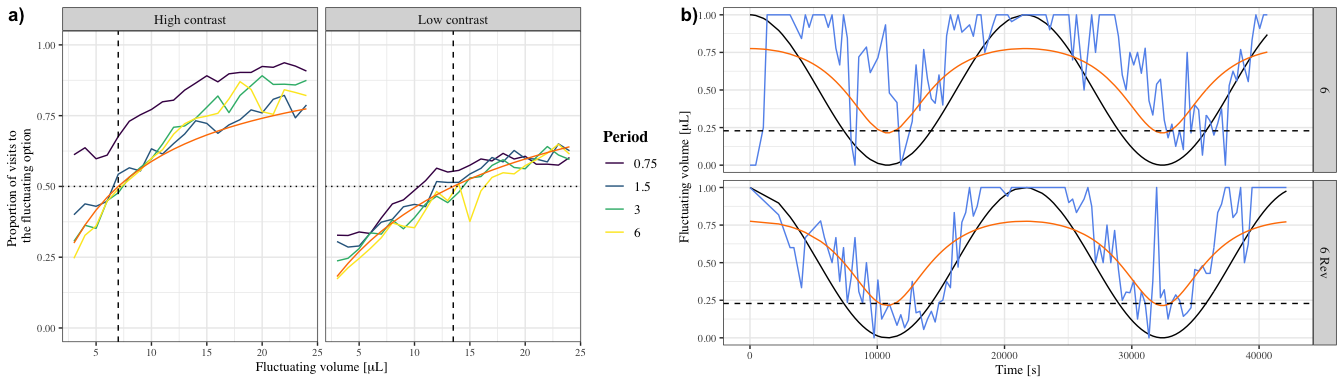


Figure 9: (ref:matching-proportions)

The goodness of fit measure, i.e., the NRMSE value, between the proportion of visits to the fluctuating volume and the ratio of the fluctuating volume to the sum of the volumes was calculated. The distribution of these values for each experimental condition is shown in figure 10. The peak in the distribution of the values is lower for the higher periods, indicating a closer match between the choice data and the ratio of the volumes.

(ref:cdf-values-match) a) The distribution of the goodness of fit values, i.e., NRMSE values, for each experimental condition, indicated by the four coloured lines. These values were calculated between the bats’ proportion of visits to the fluctuating option and the ratio of the fluctuating volume to the sum of the fixed and fluctuating volumes. b) The cumulative distribution of the NRMSE values for each experimental condition, indicated by the four coloured lines. The horizontal black line is the line under which is the maximum possible area under the curve for the range of the NRMSE values.

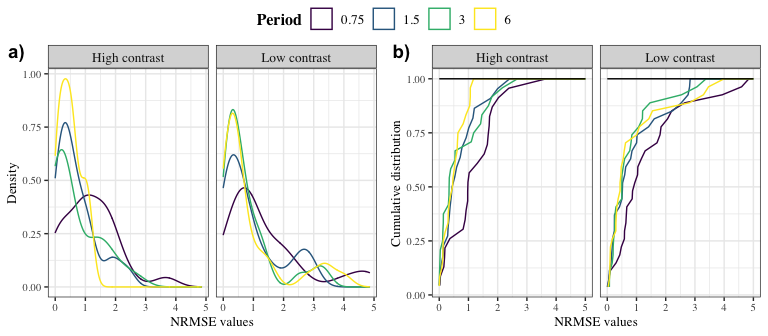


Figure 10: (ref:cdf-values-match)

The area under the curve of the cumulative distribution of NRMSE values was calculated and the Tracking Index, or the ratio between this calculated area and the maximum possible area under the curve (indicated in figure 10b by the black horizontal line) is plotted in figure 11. The value of the Tracking Index increased as the period of the sine function increased in both experiments, though it reached a plateau in the two slowest periods in the low contrast experiment.

(ref:gini-value-calculation) The Tracking Index values for each experimental condition, calculated as the ratio of the area under the curve of the cumulative distribution of NRMSE values and the maximum possible area under the curve.

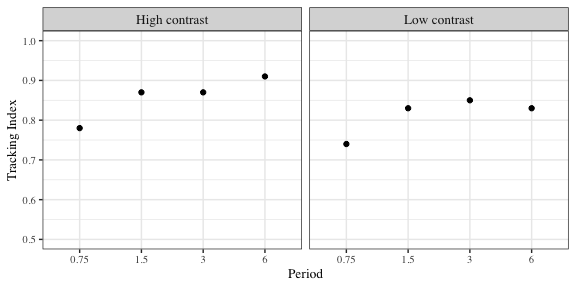


Figure 11: (ref:gini-value-calculation)

## Bats make a higher proportion of visits to a fluctuating volume when it is part of a downward trend in most experimental conditions, consistent with the trailing hypothesis

In the high contrast experiment the bats made a higher proportion of visits to a fluctuating volume when it was part of a downward rather than an upward trend all in all the experimental conditions except the fastest period, 0.75 hours. In this condition there was a small difference between the posterior estimates of the proportion of fluctuating visits in the downward and upward trends, but an overlap in the 89% credible intervals (figure 13).

In the low contrast experiment there was a clear difference in the effect of the downward and upward trends only in one experimental condition - the slowest period of 6 hours. In all the other conditions there was an overlap in the 89% credible intervals, and the estimate for the upward trend was slightly higher than that for the downward trend in the condition with the fastest period (figure 13).

The effect of the fluctuating volume was clear: the proportion of visits to the fluctuating option increased as the fluctuating volume increased. This was true for both the upward and downward trends (figure 14).

The proportion of visits to the fluctuating option was higher overall in the high contrast experiment, mainly due to a higher proportion of fluctuating visits at the higher fluctuating volumes (figure 14).

(ref:plotting-trends) Plot of the raw choice data from the bats, i.e., proportion of visits to the fluctuating option calculated for each of the volume bins of 3.16 L, upward and downward trends, and the different experimental conditions. Each line represents data from a single bat, red indicating choice data during the downward trend, blue during the upward trend. The vertical dashed line indicates the fixed volume, the horizontal dotted line indicates a proportion of 0.5.

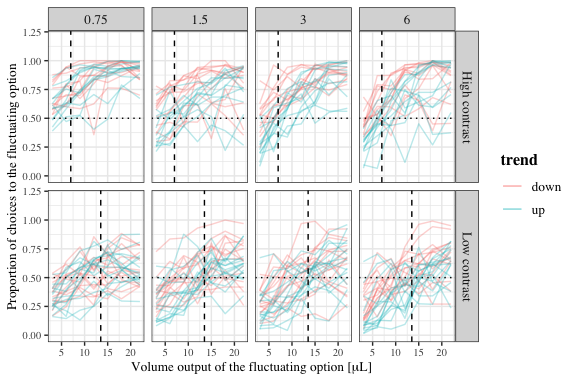


Figure 12: (ref:plotting-trends)

(ref:trends-model-conditional-effects-1) Conditional effects plot from the model of the effect of period, fixed volume, fluctuating volume and trend on the proportion of visits to the fluctuating option, showing the three-way interaction between period, trend and fixed volume. Circles indicate the mean of the posterior distribution and error bars the 89% credible intervals. a) High contrast experiment with fixed volume = 7 L b) Low contrast experiment with fixed volume = 13.5 L)

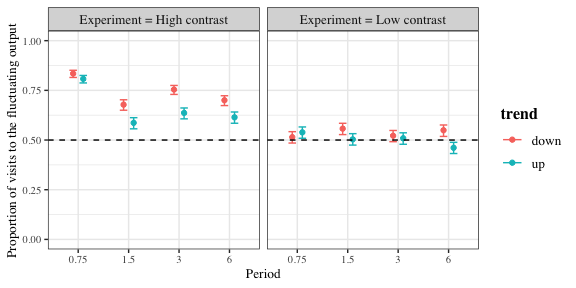


Figure 13: (ref:trends-model-conditional-effects-1)

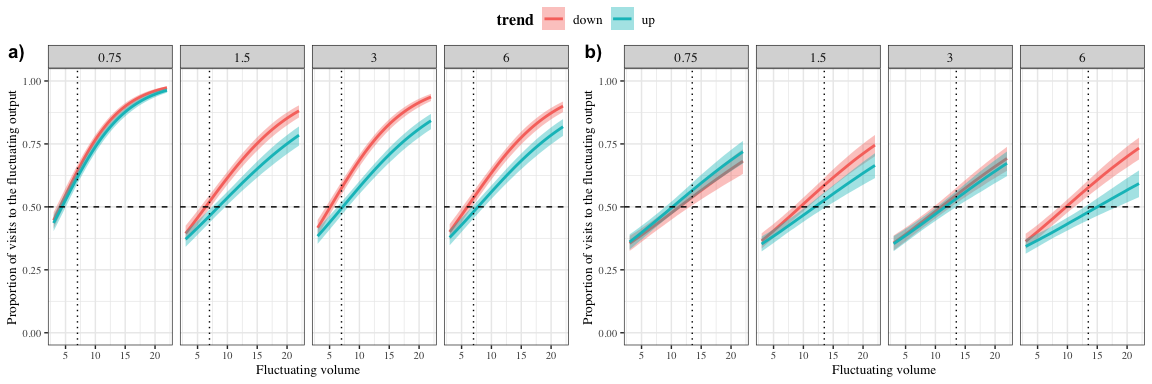


Figure 14: Conditional effects plot from the model of the effect of period, fixed volume, fluctuating volume and trend on the proportion of visits to the fluctuating option, showing the four-way interaction between period, trend, fluctuating volume and fixed volume. Thick lines indicate the mean of the posterior distribution and shaded areas the 89% credible intervals. a) High contrast experiment with fixed volume = 7 L b) Low contrast experiment with fixed volume = 13.5 L)

# Discussion

“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**

In this study captive bats of the species *Glossophaga mutica* were given the simultaneous choice between an option that yielded a nectar reward of fixed volume and an option that yielded a nectar reward whose volume fluctuated as a sine function of time. The rate of change of the environment and the contrast between the two options was systematically varied: each experiment had four conditions corresponding to four periods of the sine function (0.75, 1.5, 3 and 6 hours) and the fixed volumes were either the ‘subjective’ or the ‘objective’ mean of the maximum and minimum fluctuating volume.

The bats showed two qualitatively different kinds of choice behaviours. A majority of the bats, about 75% of them, made visits to both options and the reward volumes of the two options influenced the animals’ choice allocation (figure 8). The rest of the bats, four individuals in the high contrast experiment and three in the low contrast experiment, had a strong bias towards one of the two spatial locations where the reward sources were (figure 7). They made near-exclusive visits to their preferred location regardless of whether the fixed or fluctuating option was at this location. This side-biased choice allocation meant that on approximately half the experimental nights the bats never experienced the fluctuating volume, and so could not be treated the same way as the other bats that did experience it. Therefore, only the bats that made visits to both options were included in the statistical analyses.

Tracking, quantified as the goodness of fit between the choice allocation of the animals and the ratio of the fluctuating volume to the sum of the fixed and fluctuating volumes, was higher when the wave period was longer and the contrast between the options was higher (figure 9 and 11). The bats’ behaviour was thus sensitive to both rate of change and average payoff. The bats also made a higher proportion of visits to any given fluctuating volume when it was preceded by higher volumes, i.e., when it was part of a downward trend in the sine function, compared to when it was preceded by lower volumes as part of an upward trend (figure 13 and 14). This result was consistent with the hypothesis that the bats were ‘trailing’ in their estimate of the reward volumes in their environment, informed by recent reinforcement history. There was no evidence that the bats predicted a higher reward in the future.

A minor caveat in our study is that the experimental design was not completely within-subject. Each bat experienced only one out of two possible fixed volumes, though they all experienced all four periods of the sine function. For this reason the GLMM that we fit to the data was a multi-level model, which pooled information across the clusters in the data, improved the estimates about each cluster and retained the variation across the sub-samples (McElreath 2020).

In interpreting the results of our study it is important to remember that the environment was undoubtedly an artificial one, extremely simple compared to a natural foraging environment. In the latter, bats would have the choice of multiple patches of flowers, spread out over a large area. They would also face competition from conspecific individuals, the threat of predation and potential disruptions to foraging activity through unfavourable weather. The bats in our experiment only chose between two sources of reward, to which they had exclusive access without any competition. Nevertheless, the animals showed aspects of their behaviour that they might use even in a complex natural foraging environment: they can perceive the rate of change of their environment; they match their choice behaviour better to a slower rate of change; and they are sensitive to the average payoff of their environment. The behaviour of the animals observed in an artificially simple lab paradigm can inform predictions about their behaviour in a natural environment.

The reward volumes available in this experiment were large enough that even a bat that did not behaviourally respond to the change in reward volume received enough energy to sustain itself - as 25% of the animals did. Given the variability in the environment, one might think of the behaviour of the bats in terms of risk-sensitivity ((Shapiro, Schuck-Paim, and Kacelnik 2012), (Kacelnik and El Mouden 2013)). If a bat were risk-neutral, meaning that variation in reward availability did not affect its behaviour, its choice allocation would be exactly 50% to each option in the low contrast situation. In the high contrast situation the fluctuating volume was better than the fixed on average, so the choice allocation should be greater than 50% to the fluctuating volume. If instead the bats were risk-prone, the choice allocation to the fluctuating option would be higher than 50% in the low contrast experiment, and even higher than that in the high contrast experiment. These predictions based on risk-sensitivity hold even if the animals do not respond to the rate of environmental change. The fact that in the low contrast experiment the overall proportion of visits to the fluctuating option was 50% (figure 13) is consistent with risk-neutral behaviour in the bats. The fact that the proportion was higher in the high contrast than the low contrast experiment indicates that the bats are sensitive to the relative values of the two options: visits to the fluctuating option were higher when the fixed volume is smaller. A further indication that the bats are sensitive to reward volume (consistent with previous experiments - (Tölch and Winter 2007)) but probably not to risk or reward variance is that in each experimental condition the proportion of fluctuating visits is 50% when the fluctuating volume is equal to fixed volume (figure 9). The one exception was the shortest period in the high contrast experiment: the bats made more than 50% of their visits to the fluctuating option even when its volume was equal to or lower than the fixed. We predict that if fluctuating volume had changed even faster, say with a period of 0.25 h, the bats would show an even higher preference for the fluctuating option and insensitivity to its changing volume.

We mathematically represented environmental change as the ratio between the fluctuating volume and the sum of the fixed and fluctuating volumes. When the choice behaviour of the bats matches this ratio, what is observed is a molar distribution of behaviour between the two options. It is an *outcome*, rather than a mechanism in any sense, because it is not per se a behavioural algorithm to generate action from stimuli.

Animals learn about stimuli that indicate the state of their environment as part of a feedback mechanism. When animals perform actions that change the state of the world or change their own status, for example by receiving food and alleviating hunger, they learn how their own actions lead them towards a desired internal state. These mechanisms are embedded in a hierarchy of processing systems. At one end of this hierarchy is a “cognitive system” that performs information processing and plays an important role in finding novel behavioural responses that are adequate for a specific stimulus set–reward outcome pairing. At the other end of this hierarchy are “routine systems” that are associated with minimal information processing but perform some routine action that has been learnt in the past to be sufficiently successful. This continual switching from “cognition” to “routine” greatly alleviates the cognitive processing machinery from processing overload.

A changing environment is characterized by a situation where the stimulus – reward correlation changes continuously. This presents to the animal the challenge of how to and how quickly to respond to these changes, which would first involve higher-level cognitive systems, followed by “routines.” In our experiment one might ask: did the nectar-feeding bats learn about the cyclicity of environmental change so that they could develop a routine that “automatically” changed their behaviour even before a reward-prediction error signalled that changed? The difference between the visits for fluctuating volumes with different recent histories offers a possible answer to this question. When a fluctuating volume is immediately preceded by a richer recent history, the bats’ choice for this option is higher than the choice when the same volume is preceded by a leaner recent history (figure 13). This effect of recent history was pronounced when there was a high contrast between the two options, but less so when there was a low contrast. In the low contrast environment, the effect was less pronounced, and clearest when the environment had the slowest rate of change. These environments that changed slowly and had a high contrast between the options were where the bats tracked the change best (figure 11).

Putting all these results together, we might conclude that tracking by the bats, i.e., matching to their environment, relies on forming an estimate of the relative rewarding properties of the available options and this estimate includes the recent history of reinforcement at the options. When the ‘signal’ indicating where higher reward was available was strong, the bats made more choices for the higher reward. The discriminability of the options contributed to the signal strength but so did the rate of change: when the fluctuating volume changed more slowly the volume experienced at any visit was closer to those experienced at that option in the recent past compared to a faster rate of change.

While the volume of reward changed constantly in our study, and so consequently signal reliability, the probability of reward did not. Indeed there was a 100% certainty of reward at both the available options. These two properties - environmental certainty and signal reliability - can have dissociable effects on the value of information. McLinn and Stephens (2006) modelled the effect of these two parameters on on animal choice in the framework of statistical decision theory (Dall et al. 2005). The model predicted that animals should pay attention to the overall probability of reward, i.e., track the environment when environmental certainty exceeds signal reliability, and follow the signal when signal reliability exceeds environmental certainty. Blue jays qualitatively conformed to these predictions, but showed a bias towards environment tracking. In a nectar-feeding bat’s natural foraging environment flowers may be empty or contain nectar, meaning that environmental certainty is less than 100%. The flower’s smell and echo-acoustic signature indicate its identity as a bat-pollinated flower and therefore as a potential source of food, but a more reliable signal may be the quantity of nectar found during any single visit in combination with information about the nectar-secretion rate of the plant species. Bats avoid flowers where they have recently experienced a reward (Winter and Stich 2005), and faced with flowers with different nectar-secretion rates, i.e., different rates of replenishment, the bats adjust the time intervals between successive visits, taking longer to revisit a flower with a slower rate of replenishment (Tölch 2006). The theoretical model that best explained these data was one where bats integrated both information about the amount of nectar previously received at a flower, and past revisit time intervals into a reference memory based on several past experiences. This is consistent with the data from our study: bats that integrate several past experiences of reward volume would indeed show a higher preference for a fluctuating option with a richer rather than a leaner recent history.

A forager that perceives the structure of the environment, i.e, perceives that the fluctuating option alternately increases and decreases at a certain rate, could predict what will happen in the near future. Such a forager might withhold visits to the fluctuating volume until it near its maximum volume, and so save the energy required to make visits to the flower. This is exactly what the bats did not seem to do. While bats do adjust the time-intervals between one visit and the next based on rates of replenishment, the amount of reward experienced influences this behaviour. Furthermore, as each experimental treatment was only experienced for two nights, it is possible that the bats did not have sufficient experience of each rate of change to adjust the time-intervals between visits accordingly.

In conclusion, our study demonstrated environment tracking by nectar-feeding bats, a novel finding, through matching of choice behaviour to the ratio of available reward magnitudes. The bats were sensitive to environmental rate of change and contrast between the options, tracking better at slower rates of change and higher contrast. Tracking relied on forming an estimate of the relative rewarding properties of the available options which included the recent history of reinforcement at the options.

# 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

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 15. 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.

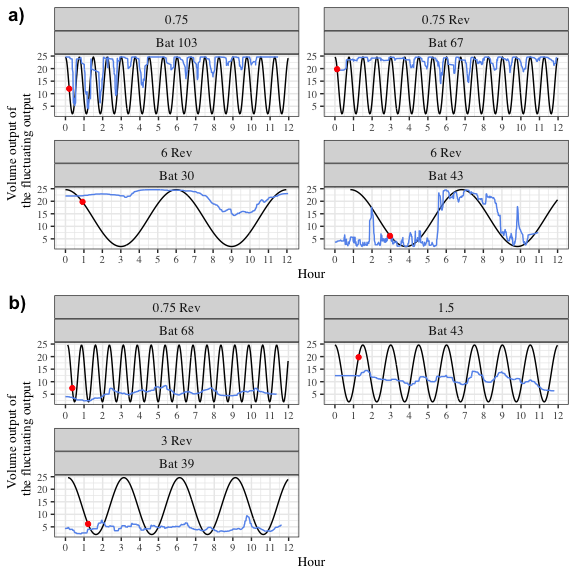


Figure 15: Volumes experienced by a small number of bats at their first rewarded visit to the fluctuating option in a) the high contrast experiment and b) the low contrast experiment. The black line represents the fluctuating output, the red dot represents when the first rewarded visit to the fluctuating option occurred.

## Details of the statistical analyses

The Bayesian generalized linear mixed-model 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 2000 warm-up samples and 2000 post-warm-up samples. A binomial likelihood function was used with the 4-way interaction of trend, wave period, fluctuating volume bin and fixed volume modelled as a fixed effect, with fluctuating volume bin as a continuous predictor and the other two as categorical predictors. Random intercepts and slopes as an effect of fluctuating volume bin 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 model. The was equal to 1 for all the chains.

## Coefficient values of predictor variables in the models

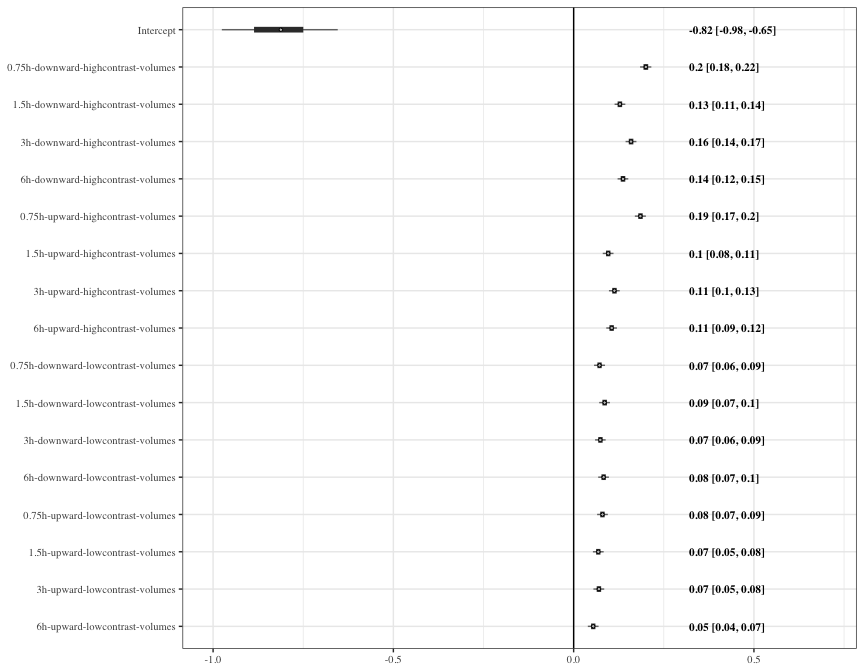


Figure 16: a) Forest plot of the estimates of the effect of Period, fluctuating volume and trend of the fluctuating volume on visits to the fluctuating option in the high contrast experiment. b) Forest plot of the estimates of the effect of rate of change, fluctuating volume and trend of the fluctuating volume on visits to the fluctuating option in the high contrast 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

# References

**Brunet-Rossini and Wilkinson**

**J.R. Krebs and A. Kacelnik. Decision making. In J.R. Krebs and N.B. Davies, ed- itors, Behavioural ecology, pages 105–137. Blackwell Scientific Publishers, Oxford, 1991. 3rd edition.**

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.

Dall, Sasha R. X., Luc-Alain Giraldeau, Ola Olsson, John M. McNamara, and David W. Stephens. 2005. “Information and Its Use by Animals in Evolutionary Ecology.” *Trends in Ecology & Evolution* 20 (4): 187–93. <https://doi.org/10.1016/j.tree.2005.01.010>.

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.

Herrnstein, R. J. 1970. “On the Law of Effect1.” *Journal of the Experimental Analysis of Behavior* 13 (2): 243–66. <https://doi.org/10.1901/jeab.1970.13-243>.

Houston, Alasdair I., Pete C. Trimmer, and John M. McNamara. 2021. “Matching Behaviours and Rewards.” *Trends in Cognitive Sciences* 25 (5): 403–15. <https://doi.org/10.1016/j.tics.2021.01.011>.

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>.

Kacelnik, Alex, and Claire El Mouden. 2013. “Triumphs and Trials of the Risk Paradigm.” *Animal Behaviour* 86 (6): 1117–29. <https://doi.org/10.1016/j.anbehav.2013.09.034>.

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>.

McLinn, Colleen M., and David W. Stephens. 2006. “What Makes Information Valuable: Signal Reliability and Environmental Uncertainty.” *Animal Behaviour* 71 (5): 1119–29. <https://doi.org/10.1016/j.anbehav.2005.09.006>.

Nachev, Vladislav, Kai Petra Stich, and York Winter. 2013. “Weber’s Law, the Magnitude Effect and Discrimination of Sugar Concentrations in Nectar-Feeding Animals.” Edited by Jacob Engelmann. *PLoS ONE* 8 (9): e74144. <https://doi.org/10.1371/journal.pone.0074144>.

Seth, Anil K. 1999. “Evolving Behavioural Choice: An Investigation into Herrnstein’s Matching Law.” In *Advances in Artificial Life*, edited by G. Goos, J. Hartmanis, J. van Leeuwen, Dario Floreano, Jean-Daniel Nicoud, and Francesco Mondada, 1674:225–35. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/3-540-48304-7_29>.

Shapiro, Martin S., Cynthia Schuck-Paim, and Alex Kacelnik. 2012. “Risk Sensitivity for Amounts of and Delay to Rewards: Adaptation for Uncertainty or by-Product of Reward Rate Maximising?” *Behavioural Processes*, Comparative cognition: Function and mechanism in lab and field., 89 (2): 104–14. <https://doi.org/10.1016/j.beproc.2011.08.016>.

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>.

Tölch, Ulf. 2006. *Bat Time Stories: Decision-Making in Spatio-Temporally Predictable Environments*. <http://edoc.ub.uni-muenchen.de/archive/00005130/>.

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

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