Simulations of Learning and Behaviour in the Hawkmoth *Deilephila elpenor*

Anna Balkenius
Vision Group
Cell and Organism Biology
Lund University
Helgonavägen 3
S-223 62 LUND
anna.balkenius@zool.lu.se

Almut Kelber
Vision Group
Cell and Organism Biology
Lund University
Helgonavägen 3
S-223 62 LUND
almut.kelber@zool.lu.se

Christian Balkenius
Lund University Cognitive Science
Kungshuset, Lundagård
S-222 22 LUND
christian.balkenius@lucs.lu.se

Abstract

We describe a behavioural experiment with the hawkmoth *Deilephila elpenor* and show how its behaviour in the experimental situation can be reproduced by a computational model. The aim of the model is to investigate what learning strategies are necessary to produce the behaviour observed in the experiment. Since very little is known about the nervous system of the animal, the model is mainly based on behavioural data and the sensitivities of its photoreceptors. The model consists of a number of interacting behaviour systems that are triggered by specific stimuli and control specific behaviours. The ability of the moth to learn the colours of different flowers and the adaptive processes involved in the choice between stimulus-approach and place-approach strategies is also modelled. The behavioural choices of the simulated model closely parallel those of the real animal. The model has implications both for the ecology of the animal and for robotic systems.

1. Introduction

Deilephila elpenor is a hawkmoth that feeds nectar from flowers. It is most active at night. While many other insects land on flowers when foraging, *D. elpenor* hovers in front of them while extending its long proboscis to retrieve the nectar. It is extremely good at compensating for drift while hovering under windy conditions. It has small wings and flies fast with a high wing beat frequency. Since this flight behaviour is very energy consuming, it is essential that it can feed effectively. To do this, it must continuously adapt to changes in its environment and learn where to find nectar.

D. elpenor has superposition compound eyes with three different photoreceptor types, an ultraviolet-, a blue- and a green sensitive receptor instead of blue, green and red as in humans.

Like most insects, *D. elpenor* can adapt their behaviour to accommodate changes in the environment. Although the total size of an insect brain is about one cubic millimetre or less, they are still able to show many of the types of learning that have been studied in mammals. However, the situations where each type of learning occurs is much more restricted than in mammals. The stimuli and responses must be selected carefully if any learning is to be shown.

For example, the moth *Spodoptera littoralis* can be classically conditioned to associate an odour with the proboscis extension reflex (PER) when rewarded with sucrose solution (Fan, 2000). *S. littoralis* can also learn discrimination and discrimination reversals as well as feature positive and negative discriminations (Fan & Hansson, 2000). However, classical conditioning was not possible with *S. littoralis* when colour stimuli were used instead of odour (Fan, Kelber & Balkenius, unpublished study).

During classical conditioning, the moths were constrained in a plastic tube. This prevents *D. elpenor* from being used in classical conditioning experiments of this type since they must be hovering to extend the very long proboscis. For free-flying animals, instrumental conditioning is more tractable. A suitable response is the approach of an artificial flower that is rewarded with sucrose solution (Kelber, Warrant & Balkenius, in preparation). Instrumental conditioning has been shown both in the moth *Macroglossum stellatarum* and in *D. elpenor* (Kelber, 1996, Kelber & Pfaff, 1997, Kelber, Warrant & Balkenius, in preparation, Balkenius, 2001).

We have performed experiments with *D. elpenor* where they were trained to search for food at differently coloured artificial flowers. The positions and colours of the flowers were manipulated to investigate how the moths would adapt. By constructing a computational model of *D. elpenor* we hope to generate hypotheses about its behaviour and learning ability that can later be tested in experiments. An additional goal is to find principles that can be used in constructing artificial animals and robots.

Figure 1: *Deilephila elpenor* hovering while foraging in the natural habitat (courtesy of M. Pfaff).
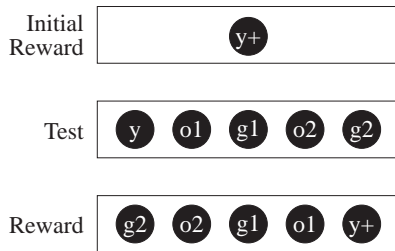

Figure 2: Positions of colours in the stimulus array during initial training, test and reward trials.

2. Experimental Study

This section describes the experimental study with *D. elpenor* and the results that were used to derive the computational model presented in section 3.

2.1 Materials and Methods

Moths were collected in July 1999 and kept at 4°C for hibernation. Three weeks before the experiment they were placed on a 12:12 hour light-dark cycle in a flight cage with a temperature around 20°C. Experiments were performed when the eyes were in the dark adapted state. The room was shielded from daylight and the cage was illuminated from above with a white lamp. The stimuli used were five artificial flowers with different colours. The rewarded colour was yellow (y). The other colours were yellow-orange (o1), orange (o2), light-green (g1), and green (g2). For a moth, these colours look relatively similar (see Fig. 6). The stimuli were presented at different positions in a vertical array on the wall of the flight cage.

Six moths were used in the study. Experiments started on day two after eclosure. The moths were initially trained to associate a single colour with a reward of 20% sucrose solution and later tested when five test stimuli were present. During initial training, each moth was fed the sucrose solution administrated through a 3 mm wide hole in the centre of a yellow artificial flower using a tube connected to a syringe. After one or two days of training, the moths had learned to forage at the artificial flower and the test phase began.

During tests, no food was present. Each animal was tested once every day. Between experiments, animals were released into the flight cage with a day-night cycle. Each test trial consisted of the presentation of five differently coloured flowers. For this paper we chose 6 experimental days when the colours where in the positions shown in the Fig. 2. The animals were allowed to approach the wall of the cage with the artificial flowers four times in a row. Each time the animal touched an artificial flower with its proboscis was counted as one visit. Each trial thus started with the approach of the flowers followed by one or several visits to the different flowers.

After four trials, the positions of the colours were temporarily changed as shown in Fig. 2 and the moths where fed at location 5 at the yellow flower. Without these rewards, the animals would loose interest in the experiment and stop flying.

2.2 Results of Behavioural Tests

Fig. 3 shows a typical example for the behaviour of the moth during the experiment. The first day, the moth is fed at a single yellow flower in the middle of the stimulus array until it has learned to approach and forage at the artificial flower.

When a moth was released in the cage at the test day, it first warmed up before it started to fly and approach the stimuli. It would stop at approximately 3-5 cm distance from the stimulus array and move sideways before choosing one of the stimuli. After a visit, it would either leave or choose a neighbouring stimulus.

In Fig. 3, the visits of flowers within a single trial are connected with a solid line. Dotted lines indicate that the moth left the stimulus array and approached again. This counted as the start of a new trial.

After the first reward, a moth would possibly visit the rewarded colour first but it would more often visit the artificial flower in the position where it received the reward (Fig. 3). This would be even more obvious after the second reward.

The distribution of visits to the different colours is shown together with the simulation results in Fig. 4. Fig. 4a shows the choices made by the moths during the first trial, before they were rewarded. The yellow flower is at position 1, and the generalisation to the other locations depends on the similarity between the colour at each position and the yellow colour. To the moth, the yellow-orange (o1) at position 2 and the orange (o2) at position 4 are more similar to the learned yellow than the light-green (g1) at position 3 or the green (g2) at position 5 (Compare positions of colours in the colour triangle, Fig. 6). The moth clearly uses colour to select which flowers to visit.
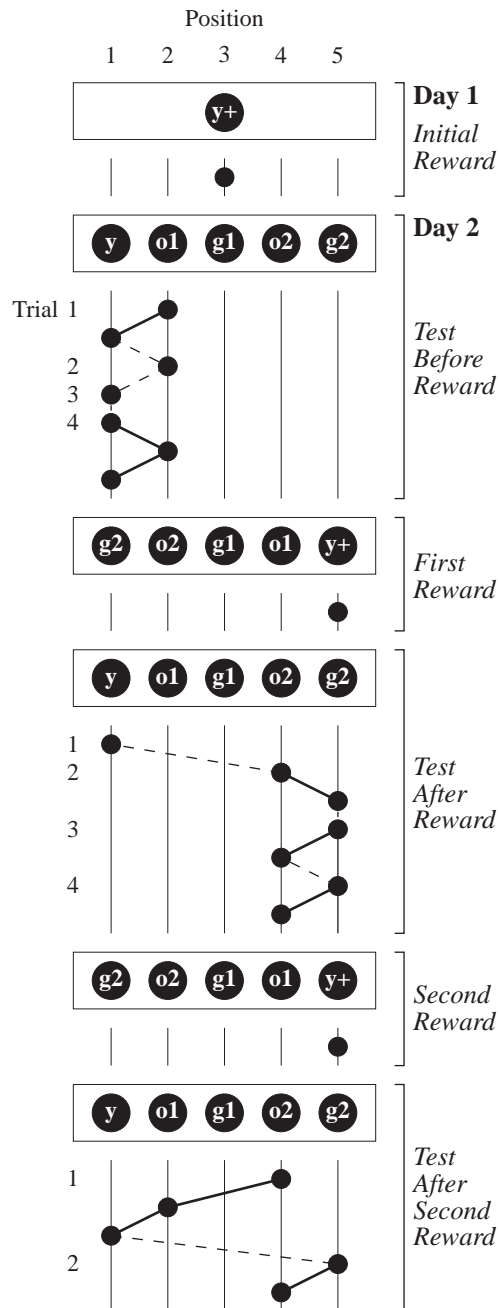
Figure 3: Typical behaviour of a moth during the experiment. See text for explanation.
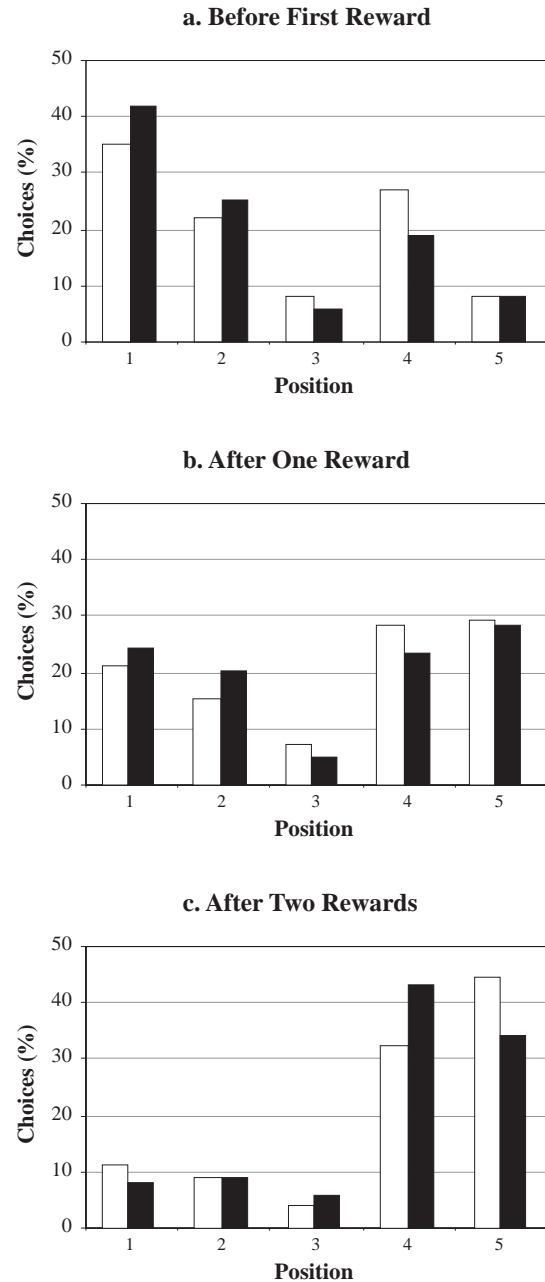

Figure 4: The behavioural choices of the model moth (white) and the real moths (black). The bars correspond to the sums of all visits for all moths. (a). Choices before first reward. 48 choices by 6 animals. (b). Choices after one reward. 79 choices by 6 animals. (c). Choices after two rewards. 35 choices by 6 animals.

The behaviour of the moths after they had been rewarded once is shown in Fig. 4b. The visits shift from the rewarded yellow colour in favour of the position where the animals were rewarded, in this case, position 5. Finally, in Fig. 4c, the distribution of visits after two rewards are shown. The animals now select stimuli to visit according to the rewarded position most of the time. The distribution of choices before reward (Fig. 4a) and after two rewards (Fig. 4c) are significantly different (G-test, P<0.001).

The position of the yellow flower during reward trials was the same throughout the experiment and after two rewarded trials the moth had learned that it always received the reward at a specific position and started to ignore colour. This shows that *D. elpenor* can use a place strategy to select flowers.

The difference between the stimulus-approach strategy and the place strategy can easily be seen during the experiment since they are qualitatively very different. The
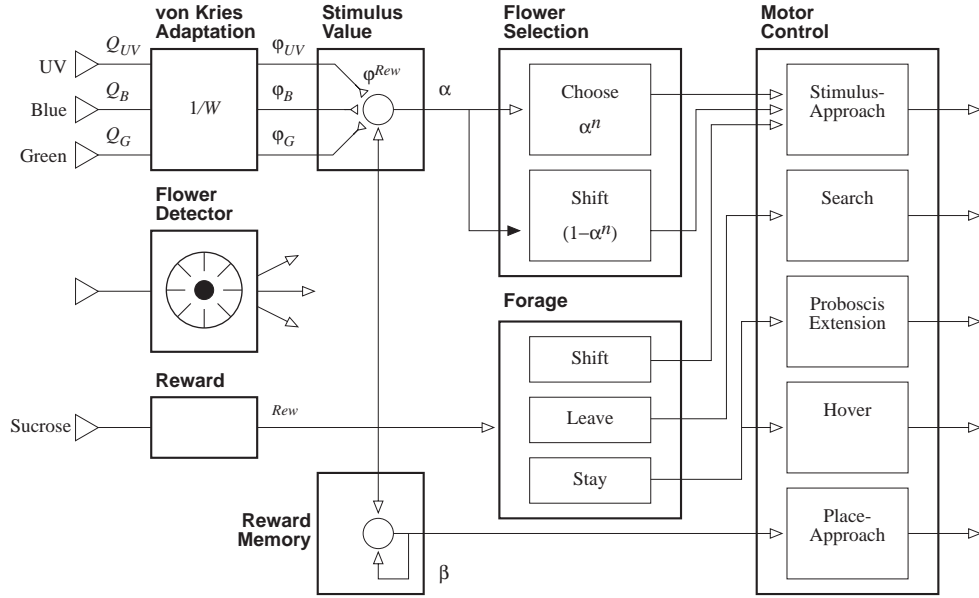
Figure 5: The model of *Deilephila elpenor*. The boxes to the left correspond to different types of stimulus processing. The next column are the learning processes. The boxes in the middle control different behaviours and the boxes to the right correspond to motor control systems.

stimulus-approach behaviour stops well before any flower is reached and is followed by what looks like an evaluation phase where the moth moves sideways in front of the stimulus array. The place-approach behaviour, on the other hand, is much faster and does not stop until the moth is directly ahead of a flower.

Interestingly, when the moths were tested again the next day, they no longer used the place strategy. Instead they returned to using colour to select flowers. The choice to use position rather than colour is temporary. In summary, the experiment shows that moths can use two different strategies when it chooses artificial flowers. They can be instrumentally conditioned to choose a flower according to its colour or position.

3. Simulation Study

The result of the experiment described above has been used to design a computational model of the behaviour in *D. elpenor*. Section 3.1 describes the model, and its performance is described in section 3.2.

3.1 A Computational Model

The behaviour selection of *D. elpenor* depends on both external and internal factors. The external factors used in the model are the colour of the flower in front of it, the location of the moth relative to the stimulus array, and whether it is currently being rewarded. The internal factors are the learned colour and position of the rewarded flower and a memory for how many rewards it has recently received.

Colour

To model the colour vision system of the animal, we calculated the receptor responses corresponding to the different colours used in the experiment. Let $Q = \langle Q_{UV}, Q_B, Q_G \rangle$ be the vector formed by the number of light quanta absorbed by the three photoreceptor types of the animal. The light reflected from the flower in front of it is assumed to excite each receptor type Q_i according to the formula,

$$Q_i = \int_{\lambda=300nm}^{700nm} I(\lambda)S(\lambda)R_i(\lambda)d\lambda,$$

where I is the spectrum of the illumination, S is the reflectance spectrum for a surface, and R_i is the spectral sensitivity of the photoreceptor of type i and λ ranges over the wavelengths visible to the moth.

Fig. 6 shows the location of the colours used in the experiment in the Maxwell colour triangle for the moth. Each corner corresponds to one of the three photoreceptor types of the moth eye. The location of a colour in the triangle represents the relative excitation of the three receptor types for that colour. As can be seen, the five colours are very similar.
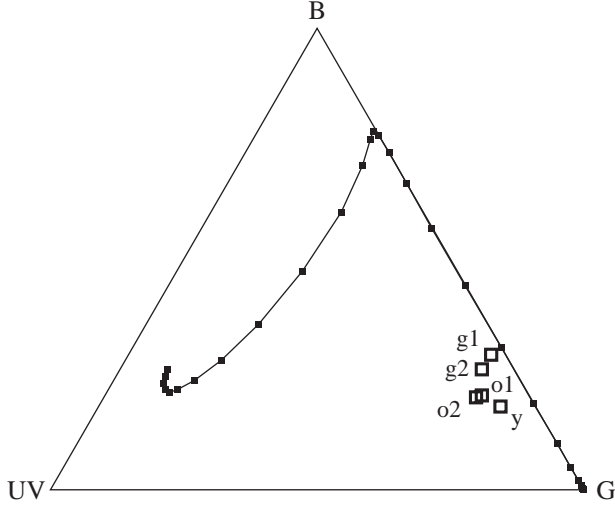
Figure 6. The Maxwell colour triangle for *Deilephila elpenor* with the five different colours used in the experiment. The curve illustrates the location of the monochromatic lights. The corners represent the three different photoreceptor types.
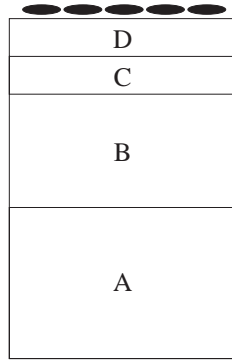

Figure 7. The cage seen from above with the stimulus array and the four states corresponding to the different locations.

The loci in the colour triangle were calculated using the spectral sensitivity curves for the photoreceptors of *D. elpenor* (Höglund, 1973) and the spectral reflectance of each test colour measured using a spectrophotometer (S2000, Ocean Optics).

To compensate for fluctuations in the illumination, the animal is assumed to use a von Kries adaptation mechanism working at the receptor level (von Kries, 1902). This mechanism scales the sensitivity of each receptor with the average activation of the receptor in the environment. The von Kries adaptation mechanism leads to an incomplete form of colour constancy (von Kries, 1902). If we assume that the average reflectance in the environment is white (as is the case in the experimental cage) and that the receptor responses for white are given by the constants W_i, the von Kries coefficients ρ_i are given by,

$$\rho_i = \frac{1}{W_i}.$$

We can now calculate the colour coordinates φ_i for a flower as,

$$\varphi_i = \rho_i Q_i$$

Locations and States

We distinguish between four different states of the moth that are not meant to be internal states but roughly correspond to locations in the flight cage: In state A (and location A) the moth is far away from the test stimulus array and does not take notice of it. In location (and state) B the moth is close enough to the array to take notice and react to the stimuli. Location C corresponds to the distance of 3 to 5 cm where the moth usually stops a stimulus approach (see section 2.2) and location D is directly in front of the flowers where the moth can reach them with the proboscis. For a description of the corresponding behaviour see the section on behaviour selection.

Reward

The reward signal that reaches the model is either 0 or 1. A value of 1 indicates that the moth receives sucrose solution. This is only possible when the model has extended its proboscis in front of a simulated flower.

Colour Memory and Matching

When the moth is rewarded at a flower, the colour coordinates for that flower φ are stored in the variable φ^{Rew}. This is the simplest possible learning mechanism that can account for the behavioural data. This variable acts as a colour memory for the rewarded colour and is subsequently used to calculate the similarity between other colours and the rewarded colour. This similarity is used to derive the probability of visiting the flower in front of the moth.

The normalised scalar product is used as a similarity measure for the two vectors corresponding to the rewarded colour φ^{Rew} and the colour of the stimulus in front of the moth,

$$\alpha = \frac{\varphi^{Rew}\varphi}{\left|\varphi^{Rew}\right|\left|\varphi\right|}$$

A perfect match will thus give the value $\alpha=1$, while two orthogonal colour vectors would give the value $\alpha=0$. In practice, however, two colour vectors are never completely orthogonal since the spectral sensitivities of the different receptor types overlap. Since the match is always between 0 and 1, it can easily be used as a probability.

Place Memory

When *D. elpenor* is rewarded, it becomes more likely to fly directly to the position where it was rewarded rather than to use the colour of the flowers. This implies that the moth has a memory for the position where it was last rewarded. To model this, we use a position variable p^{Rew} to hold the position where the moth was rewarded. We do not attempt to model how the moth knows where it is or the sensory processing involved in navigation.

Reward Memory

Since the moth becomes more likely to select a flower based on position than colour each time it is rewarded we let a value β indicate the probability that a place strategy will be used and increase this value each time the moth is rewarded. This probability starts out at 0 and is increased by 0.3 each time the moth is rewarded with the restriction that $0 < \beta < 1$. An increase of 0.3 gives a good fit to the experimental data. To model that the moth returns to a colour based stimulus-approach strategy with time, the value β is assumed to decay slowly with time at a rate that makes sure that it has reached 0 the next day. The variable β is thus essentially a memory for recent rewards.

Behaviour Selection

The behaviour selection of the model depends on the values of α and β together with the current state of the moth. As described above, α represents the similarity between the colour of the stimulus ahead and the rewarded colour and β described the probability of flying directly to the rewarded position instead of evaluating the colours of the flowers. The different behaviours of the moth are summarised in Table 1.

In state A, the model moth is in its search phase where it can either decide to fly directly to the place where it has previously been rewarded or continue to fly around until it finds flowers. If the moth does not start a place-approach behaviour, the model moth will either find flowers or continue searching with equal probability as shown in Table 1.

In state B, the moth has found flowers during its search phase, and approaches them, which will lead it to location C.

In state C, the model moth is flying in front of the flowers and needs to determine whether to try to forage from the flower in front of it or not. This choice depends on the similarity between the colour of that flower and the learned rewarded colour as explained above. If its chooses the flower, it will approach it and enter state D. If it chooses not to approach the flower, it will either shift to the flower to the left or right in the stimulus array.

To derive the probability of choosing the stimulus in front, the similarity α was raised to a power n to sharpen the choice between the different colours. A value of $n=4$ gave a good fit to the experimental data.

State D represents the situation when the moth is hovering in front of a flower and has extended its proboscis. The behaviour in this situation depends on whether the moth is rewarded or not. If it is not rewarded it will leave the flowers half of the time and start a new search phase. In the other cases, it will either shift to the flower to the left or right. When it has been rewarded, it will leave the flower and start a new search phase.

Table 1: The probability of each action in the simulation.

State	Rew	Probability	Action
A	0	β	*fly to p^{Rew}*
		$0.50\,(1-\beta)$	*approach*
		$0.50\,(1-\beta)$	*stay*
B	0	1.00	*approach*
C	0	α^n	*choose*
		$0.50(1-\alpha^n)$	*left*
		$0.50(1-\alpha^n)$	*right*
D	0	0.50	*leave*
		0.25	*left*
		0.25	*right*
	1	1.00	*leave*

3.2 Simulation Results

Simulation A

We run a simulation of the model presented above in the experiment described in section 2. Data was collected from 500 simulated test sessions. Like the real moth, the model first learned to select the yellow flower before the simulated experiment started. The number of visits before the first reward on the test day is shown in Fig. 4a together with the data from the real moth. In the 4b, the behaviour of the model is shown after a single reward, and finally in 4c, the behaviour after two rewards is presented. As can be seen, the behaviour of the model closely matches that of the real moth. When the behavioural data from the moths were compared to the simulation results, there was no significant difference between the behaviour of the model and the real moths before reward (G-test, $p>0.4$), after one reward (G-test, $p>0.5$) or after two rewards (G-test, $p>0.4$).

Simulation B

In the second simulation, we changed the stimuli in the experiment to much more different colours. Five spectral colours with wavelengths of 350, 400, 450, 500 and 550 nm were used. These were arranged in the stimulus array as shown in Fig. 8.

Initial Reward	450[+]

Test	450 300 350 500 550

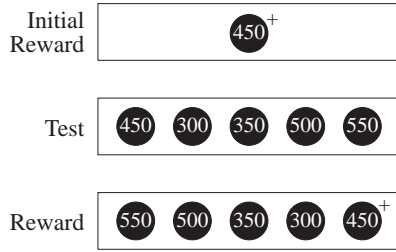|	
Reward	550 500 350 300 450[+]

Figure 8: Placement of the five spectral colour of different wavelengths in simulation B.

The simulation result is illustrated in Fig. 9. The graph shows the distribution of visits before the reward and after the first and second reward. As can be seen, the model predicts that the moth will not use a place strategy when the colours can easily be distinguished.
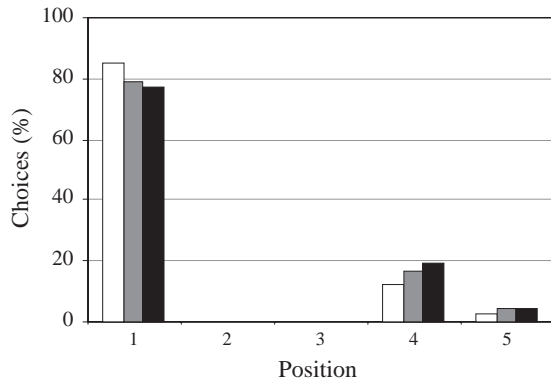

Figure 9: Simulation of the experiment with very different colours. White. Before the first reward. Grey. After one reward. Black. After two rewards.

4. Discussion

The simple model describes the behaviour of the moth very accurately although it makes minimal assumptions about colour matching and learning processes. The moth can not only learn to associate a food reward with a visual stimulus (here colour) but it can also change between two strategies as described in the model.

Other moths, like the diurnal hawkmoth *Macroglossum stellatarum*, visit a large number of artificial stimuli even when they are not rewarded. *D. elpenor* makes only few approaches without reward. This indicates that it needs to keep energy expenses as small as possible. For the same reason, the animal needs to fastly adapt its choice strategy.

When foraging, animals can either use a stimulus-approach or a place-approach strategy to move from one location to another (Balkenius, 1995). In the first case, the movement in space is guided by a single stimulus that the animal moves towards. In the second case, the goal location is given by the relative position of a number of spatial cues, for example, distal landmarks. These types of behaviours have been proposed as alternatives to traditional stimulus-response explanations of spatial behaviour. The adaptive advantage of these strategies is that they can make use of negative feedback from sensory organs to control the behaviour in a goal-directed way.

When both strategies are available, discrimination learning could be used to determine which strategy is most adaptive. The result of the experiment shows that *D. elpenor* can approach flowers using either strategy, but the learning involved is not discrimination learning. Since it is always rewarded at the yellow colour at position 5, but never at position 1, where it appears during test trials, the reward contingency of using colour or position does not differ. This implies that the choice between stimulus-approach and place-approach strategy depends on other factors. In the model, we hypothesised that it was the amount of reward received recently that was used to determine which strategy to use. This may reflect an innate win-stay strategy that is activated by reward (Gallistel, 1990).

In the simulations, the value β was increased by 0.3 each time the animal was rewarded. The value of 0.3 was used to reflect that the animals almost completely switched to using a place-approach strategy when they had been rewarded three times. This is a reasonable strategy when the colours of the flowers are very similar and cannot be distinguished by other means. However, when the different flowers can easily be identified based on their colour, it may be more adaptive to stay with a stimulus-approach strategy. This is the prediction made in simulation B, where the colours were very different and the effect of place learning became very weak. In agreement with this prediction, it has been observed that moths do not appear to use a place strategy in experiments where the colours are easily distinguishable (A. Kelber, unpublished observations). We are currently planning experiments to test the predictions of the model.

In the model, the normalised scalar product is used to measure colour similarity. This measure has the advantage that it stays between 0 and 1, but it does not have the properties of a metric. It would, of course, be possible to use also other forms of matching between the colours (Brandt & Vorobyev, 1996). When a metric is used to calculate colour similarity, a natural choice of probability function would be $G_\sigma(\|\varphi - \varphi^{Rew}\|_m)$, where G is the Gaussian of the distance between the two colour vectors with variance σ and m indicates the metric used. However, very accurate data would be needed to determine what kind of metric describes the colour space of the moth best.

The model reproduces the behavioural choices of the moth although there are essentially only two parameters: The increase of β and the exponent n used in flower selection. This indicates that a very simple learning mechanism can be used to explain the change in behaviour when the moths are rewarded which is reasonable for an animal with a very small brain.

Acknowledgements

References

Balkenius, A. (2001). Colour constancy in diurnal and nocturnal hawkmoths. In *International Conference on Invertebrate Vision*, p. 172, Bäckaskog Castle, Swden.

Balkenius, C. (1995). *Natural Intelligence in Artificial Creatures*, Lund University Cognitive Studies, 37.

Brandt, B. & Vorobyev, M. (1996). Metric Analysis of Threshold Spectral Sensitivity in the Honeybee. *Vision Research* 37, 425-439.

Fan, R.-J. & Hansson, B. S. (2000). Olfactory discrimination conditioning in the moth Spodoptera littoralis. *Physiology and Behaviour* (in press).

Fan, R.-J. (2000). *Learning and Memory in Moths*, PhD Thesis. Lund University, Sweden.

Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, MA: MIT Press.

Höglund, G., Hamdorf, K. & Rosner, G. (1973). Trichromatic visual system in an insect and its sensitivity control by blue light. *Journal of Comparative Physiology* 86, 265-279.

Kelber, A. (1996). Colour learning in the hawkmoth Macroglossum stellatarum. *Journal of Experimental Biology* 38, 1127-1131.

Kelber, A. & Pfaff, M. (1997). Spontaneous and learned preferences for visual flower features in a diurnal Hawkmoth. *Israel Journal of Plant Sciences* 45, 235-245.

von Kries, J. (1902). Chromatic Adaptation. In *Sources of Color Vision* (ed. D. L. MacAdam), pp. 109-119. Cambridge, MA: MIT Press.