Colour reverse learning and animal personalities: the advantage of behavioural diversity assessed with agent-based simulations.

A. G. Dyer[1*], A. Dorin[2*], V. Reinhardt[3] and M.G.P. Rosa[4]

[1] adrian.dyer@rmit.edu.au, Media and Communication, RMIT University, Australia 3001

[2] alan.dorin@monash.edu, Faculty of Information Technology, Monash University, Australia 3800

[3] gitarrengoettin@gmx.de, Institut für Zoologie III (Neurobiologie), Johannes, Gutenberg-Universität, Mainz, Germany 55099

[4] marcello.rosa@monash.edu, Faculty of Medicine, Nursing & Health Sciences, Monash University, Australia 3800

[*] These authors contributed equally to this work.

Abstract

Foraging bees use colour cues to help identify rewarding from unrewarding flowers, but as conditions change, bees may require behavioural flexibility to reverse their learnt preferences. Perceptually similar colours are learnt slowly by honeybees and thus potentially pose a difficult task to reverse-learn. Free-flying honeybees (N = 32) were trained to learn a fine colour discrimination task that could be resolved at ca. 70% accuracy following extended differential conditioning, and were then tested for their ability to reverse-learn this visual problem multiple times. Subsequent analyses identified three different strategies: 'Deliberative-decisive' bees that could, after several flower visits, decisively make a large change to learnt preferences; 'Fickle-circumspect' bees that changed their preferences by a small amount every time they encountered evidence in their environment; and 'Stay' bees that did not change from their initially learnt preference. The next aim was to determine if there was any advantage to a colony in maintaining bees with a variety of decision-making strategies. To understand the potential benefits of the observed behavioural diversity agent-based computer simulations were conducted by systematically varying parameters for flower reward switch oscillation frequency, flower handling time, and fraction of defective 'target' stimuli. These simulations revealed that when there is a relatively high frequency of reward reversals, fickle-circumspect bees are more efficient at nectar collection. However, as the reward reversal frequency decreases the performance of deliberative-decisive bees becomes most efficient. These findings show there to be an evolutionary benefit for honeybee colonies with individuals exhibiting these different strategies for managing resource change. The strategies have similarities to some complex decision making processes observed in humans, and algorithms implemented in artificial intelligence systems.

1.0 Introduction

Foraging for nutrition in the form of nectar in natural environments presents a variety of potential dilemmas for free-flying bees. For example, there may be a number of flowers possessing similar identifying cues that may offer nectar as a nutritional reward [1], there could also be mimics like orchids that offer no reward [2], and there may be different temporal phases in which flowers vary their rewards [3-6]. It is also possible that plant species that do have rewarding flowers will sometimes present empty flowers simply because other foragers have recently depleted the flowers [6-9].

Many social bees, like honeybees and bumblebees, exhibit flower constancy and will remain constant to one type of learnt rewarding flower, as long as it continues to present rewards [10]. This type of pollinator behaviour may be evaluated in reverse-learning experiments [11, 12] and has been investigated in honeybees for both olfactory [12, 13] and saliently different colour discrimination [14, 15] tasks; and for bumblebees on sensorimotor learning tasks [16]. When considering saliently different colours like 'orange' and 'blue', free-flying honeybees can quickly learn within five trials to choose a rewarding colour with accuracy greater than 80%, and then quickly switch these learnt preferences after a further 1-2 trials if the reward paradigm is changed. However, with this short training, honeybees can only reverse decisions up to three times before discrimination falls to chance levels [15]. This finding was confirmed in a separate study that trained honeybees to discriminate between saliently different 'blue' and 'yellow' colour stimuli which were learnt in three trials to an accuracy greater than 80%, and the bees could then quickly switch preferences after 1-2 further trials when the reward paradigm was switched [14]. In this case of a short learning opportunity, honeybees also chose between the colour stimuli at random levels if the reverse training continued for more than three reversals; however, if the training was extended to 10 rewards on a particular colour stimulus then reverse-learning was very robust for at least nine reversals [14]. This indicates that length of training is important to the capacity of honeybees to robustly reverse-learn a salient colour task.

Recent work examining how harnessed honeybees reverse-learn olfactory stimuli has revealed that different individual bees may possess different strategies for reverse-learning [12]. If harnessed honeybees are presented with two different odorants to discriminate between using a standard proboscis extension reaction (PER) experimental setup [17], some bees can reverse-learn the discrimination up to three times [12], which is consistent with the work on salient colour discrimination with free-flying honeybees [14, 15]. However, an important difference in the recent work on olfactory reverse-learning was the observation of individual differences between how honeybees were able to perform the reverse switching task [12]. It was observed that there were three categories of honeybees: 'efficient' reversers that could quickly change preferences when experimental conditions changed; bees that did learn the initial discrimination task but then appeared unable to reverse-learn the task; and a third category that failed to learn the initial olfactory discrimination task [12]. The existence of the last category of bees implies that there was a reasonable degree of perceptual difficulty involved in the initial learning of this olfactory discrimination task. This finding of individual differences for perceptually difficult olfactory learning in harnessed bees agrees with other recent work reporting differences in performance levels when individual free-flying bees solve perceptually difficult colour discrimination tasks [18, 19].

When considering colour stimuli, recent work on honeybees [20, 21] and bumblebees [22] has revealed that the difficulty of a task can be controlled by varying the perceptual similarity of colour stimuli, especially since there is a soft sigmoidal function that describes the probability with which colour differences can be judged by bees [23-25]. Colour difference can be conveniently specified in a colour space like a Hexagon colour model which allows for a Euclidean distance between stimuli to be quantified [26]. Recent research has shown that when either honeybees [20, 21, 23, 27] or bumblebees [22, 28] learn colour information in isolation (termed *absolute*

conditioning), individuals only demonstrate a coarse level of colour discrimination (about 1.5 hexagon units [22, 29]); whilst individuals that learn a target colour in the presence of perceptually similar distractor stimuli (termed *differential conditioning*) can learn to make relatively fine colour discriminations (<0.10 hexagon units [22, 29]). However, colour learning with differential conditioning takes considerably longer for bees than learning coarse level colour discrimination with absolute conditioning. For example, a colour distance of about 0.04-0.08 hexagon units typically takes either honeybees or bumblebees about 50-60 visits to learn the discrimination to a level of about 75% accuracy [1, 18, 22, 29, 30]. There is evidence that this type of perceptually difficult discrimination places increased cognitive load on the information processing of both honeybee and bumblebees since individual bees will slow down to maintain accuracy for fine colour discrimination problems [18, 30, 31], and/or will not perform at a high level of accuracy unless incorrect choices are punished with a bitter tasting substance [20, 31, 32]. Since perceptually similar colour stimuli potentially place increased cognitive load on bee colour judgements, it is important to understand the extent to which bees can reverse-learn such fine discrimination tasks, and to determine if there are differences in decision-making behaviour between individuals processing similar colours.

To understand decision-making in bees for difficult visual tasks, it is useful to combine approaches of behavioural testing with computer modelling to interpret the biologically relevant factors that may influence how and why individuals make certain decisions, and subsequent benefits at a colony level [18, 33]. In particular, where bee behaviour varies between individuals or where local environmental conditions influence individual decision-making, *agent-based models* (ABMs), also called *individual-based models*, offer a powerful approach for understanding the intricate interactions and emergent outcomes of these complex systems in the context of behavioural ecology [34-40]. ABMs have been used to understand bee behaviour since the 1980s [41]. For example, ABMs have been used to understand bee foraging strategies (where these are assumed to be homogeneous within a population) in keeping with empirical data whilst considering recruitment, homing and memory of food source location [42]; and ABMs have been used to show that the benefits of recruitment by honeybees is heavily dependent on the density of flowers within an environment [43]. In this current study we combine the powerful approaches of behavioural testing with free-flying bees whilst solving perceptually difficult colour reversal tasks, and the use of ABMs to understand the potential colony level benefits of the diversity of observed behaviours.

2.0 Methods: Behavioural experiments

Behavioural experiments were conducted with free-flying honeybees (*Apis mellifera* Linnaeus) at the biological gardens of Johannes Gutenberg University (Mainz, Germany) between July and September 2009. A hive of honeybees was maintained 10 m from a gravity feeder that provided 5% (vol.) sucrose solution. Individual bees were collected from the feeder site on a small Plexiglas® spoon, and were transferred to a test site situated 15 m from the feeder, and 20 m from the hive. At the test site individual bees were marked with a colour code on their thorax. Each bee was tested individually, and testing typically took 5-6 h/bee.

The training apparatus consisted of a vertical rotating screen of 50 cm diameter. This apparatus presents stimuli on four 6×8 cm hangers, each with a small landing

platform, so that the spatial position of stimuli can be continuously changed during training, and hangers can be easily exchanged for fresh ones to exclude olfactory cues [44, 45]. Two of the hangers presented target stimuli, and two hangers presented distractor stimuli in pseudo-random positions. As the rotating screen apparatus could be freely rotated, this apparatus simulates a complex natural scenario where bees must forage from several flowers of potentially similar colour. The hangers enable well-defined data collection by counting choices (touches to the landing stage of stimuli). In an experiment, bees were rewarded with 25% sucrose for making correct choices on a designated target stimulus, whilst a distractor stimulus only presented plain water. When a bee landed on a target stimulus it was presented with an additional sucrose drop on a plexiglass spoon so that it could be moved 1 m away to behind a small screen so that stimuli could be exchanged [44, 45]. A photograph of the rotating screen is presented in a previous study [45].

Stimuli were 6×6 cm coloured cards (Tonpapier no. 32 & 37, Baehr, Germany) that were of turquoise and blue appearance to a human observer respectively. Stimuli spectral properties were measured with a spectrophotometer (Spectro 320; Instrument Systems, München, Germany) between 300 and 650 nm. The colour visual system of the honeybee is different to human vision, and is based on three spectrally different types of photoreceptors maximally sensitive in the ultraviolet (UV), blue (B) and green (G) regions of the electromagnetic spectrum [28]. To quantify the colour difference between the turquoise and blue stimuli we used a colour hexagon model designed for hymenopteran trichromatic vision [26], considering the previously measured electrophysiological recordings for honeybee photoreceptors [46], data for standard daylight conditions converted to photon flux units [47], and assuming the visual system was adapted to the grey plastic background of the training apparatus [26]. The colour distance was 0.06 hexagon units, which is a colour discrimination task that is predicted to be sufficiently close so as to require differential conditioning for bees to learn [20, 24].

A total of 32 bees were tested. The experiment was counterbalanced in a pseudo-randomised fashion so that for half of the bees the 'turquoise' stimuli were the initial targets, whilst the 'blue' stimuli were the initial targets for the other bees. Each individual bee was first provided with a form of absolute conditioning to the target stimulus for 30 decisions where the distractor hangers only presented the dissimilar grey background material. This ensured the bees were familiar with the experimental apparatus, and also enabled within-subject testing of whether absolute conditioning enabled any colour learning of the perceptually difficult colour discrimination task.

Following this initial absolute conditioning, each bee received differential conditioning to the similar turquoise and blue stimuli for 60 decisions. At the conclusion of the 60th decision, a test bee was satiated on the initial target stimulus and allowed to return to the hive. When the bee returned to the apparatus, all four hangers presented the initial distractor stimulus to allow for priming to the reverse contingency as these four hangers also presented sucrose solution. The bee was allowed to collect sucrose from four landings and was satiated on the 4th hanger so that it would return to the hive. When the test bee returned again it received differential conditioning with the initial reward situation switched. This reverse training lasted for 20 decisions. The rewards were reversed again for another 20 decisions as described above. Next, in the final phase of the experiment, the rewards

were reversed for 60 decisions. Thus each bee had to solve a complex set of reverse-learning tasks that involved perceptually difficult discriminations where individual discrimination decisions were potentially ambiguous. For statistical analysis data was checked for normal distribution and arcsine transformed where necessary.

3.0 Results: Behavioural experiments

To evaluate if the bees had learnt the target stimulus following the initial absolute conditioning for 30 decisions, we considered the frequency of correct choices for the target in the first 10 decisions of the differential conditioning phase. For the 'turquoise' target group the frequency of correct choices for the target was 52.4% (± 16.9 s.d.) which was not significant from chance expectation (1-sample t-test, N=16, t= 0.560, df15, p= 0.584). For the 'blue' target group the frequency of correct choices for the target was 47.9% (± 15.5 s.d.) which was not significant from chance expectation (1-sample t-test, N=16, t= 0.560, df15, p= 0.584). These pooled results were not significantly different from chance expectation (independent sample t-test, t=0.774, df30, p=0.455). Thus the two colour stimuli were initially perceived as very similar by the bees.

With differential conditioning the bees learnt the visual task gradually such that the mean frequency of correct choices was 73.5% ± 14.2 s.d. after 60 choices (Fig. 1). With the first reversal there was an indication that some bees could switch their decisions quickly (Fig. 1), but the choices for the initial target (45.0% ± 23.72 s.d.) was not what would be expected if the bees had completely switched their initially learnt colour preference (100-73.5 = 28.5%). Thus, considering previous evidence that there could potentially be individual differences between honeybees in how they reverse-learn a perceptually difficult task [12], we next tested for evidence of bees having different abilities to reverse-learn.
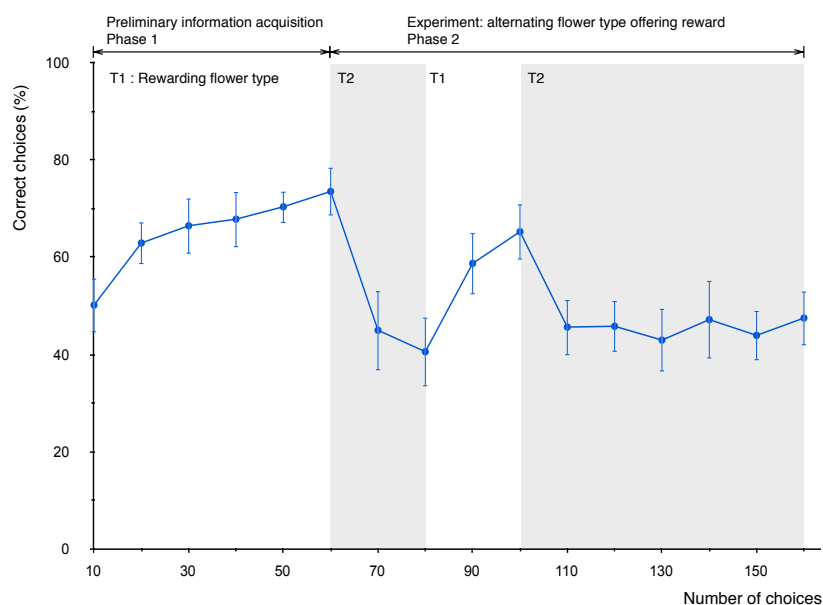

Fig. 1. Mean (±95% confidence interval) of all 32 bees during differential conditioning and the three reversals occurring after 60, 80 and 100 choices.

To analyse the decision-making of the bees we considered the variation in choices at a particular reversal (R_{1-3}) for a given colour when it changed from being a rewarded target to an unrewarding distractor. For each bee a value representing the change in its individual frequency of correct choices was calculated for the three respective reversals (R_1 for 60-70, R_2 for 80-90 and R_3 for 100-110 choices). To statistically examine the magnitude of the reversal values (R_{1-3}) we considered the standard deviation (s.d.) of all bees during the initial learning phase of the visual task (s.d. was 14.1 and 14.2 for the 'turquoise' and 'blue' groups respectively) as the measure of behavioural variability. We then considered three categories of decision-making. Category 1 was no significant variation in choices (R_{1-3}) following a particular reversal defined as within 1.645 s.d. (90% confidence interval); these bees were termed '*Stay*' bees. Category 2 was a change greater than 1.645 s.d. and less than 1.960 s.d. (95% confidence interval); these bees were termed '*Fickle-circumspect*' bees. Category 3 was a change greater than 1.960 s.d. in choices during a reversal; these bees were termed '*Deliberative-decisive*' bees.

The subsequent analysis of the experiment revealed that 12% of bees (4/32) were consistent *Deliberative-decisive* bees (Fig. 2), and 12% of bees (4/32) were consistent *Stay* bees (Fig. 3). There were 24 bees that did not always remain faithful to just one strategy. Of these, seven bees initially exhibited a *Stay* bee strategy but moved to a *Deliberative-decisive* bee strategy, and nine bees were initially *Deliberative-decisive* and moved to a *Stay* bee strategy. There were eight bees that showed no clear pattern of loyalty to a strategy, and four of these eight sometimes fell into the *Fickle-circumspect* strategy.

To understand whether there might be a biologically plausible reason why colonies have bees with different strategies for dealing with switching between colours that are perceptually similar, we next used agent-based computer simulations.
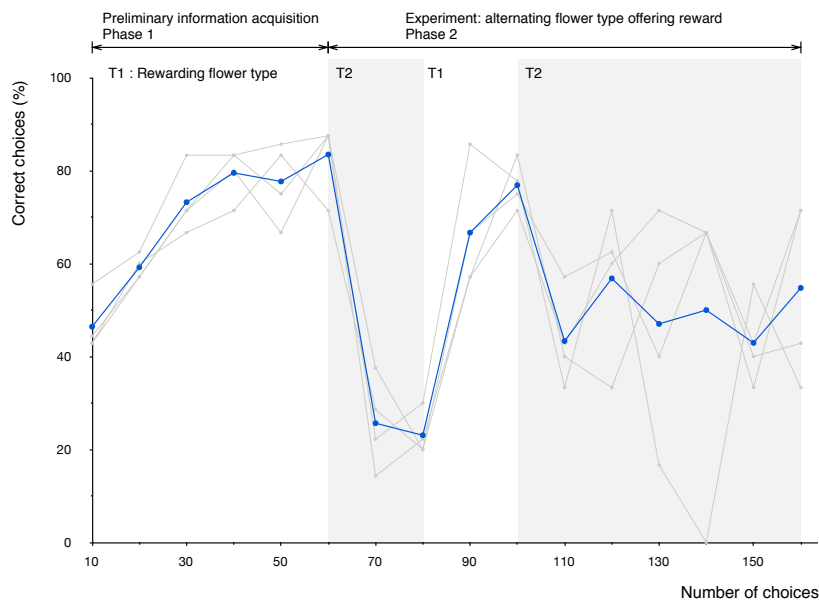

Fig. 2. Mean (dark, blue line) of the four bees (grey lines) that exhibited a Deliberative-decisive strategy during differential conditioning and the three reversals following 60, 80 and 100 choices.
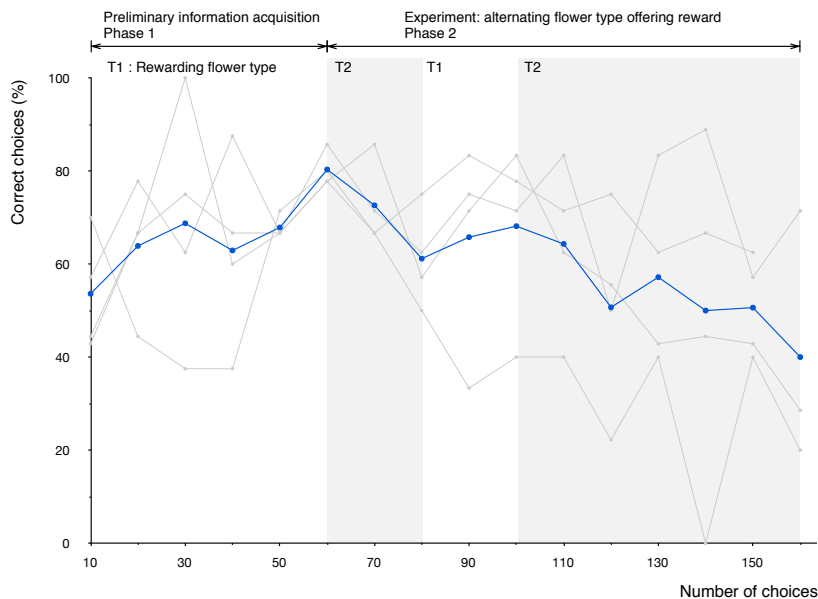
Fig. 3. Mean (dark, blue line) of the four bees (grey lines) that exhibited a Stay strategy during differential conditioning and the three reversals following 60, 80 and 100 choices.

4.0 Methods: Simulation experiments

The agent-based model used for our simulation represents a foraging patch with dispersed flower-agents, and a colony of bee-agents that search the patch to collect nutritional rewards from the flowers. We conducted a number of different experiments using the simulation by varying parameters determining the distribution of the foraging strategies of bees in a colony, the availability of flower nutritional rewards, and the reliability of flower nutritional rewards.

The simulation was updated in discrete time steps. In each time step every bee-agent was updated synchronously, computing an action to be executed based on its current internal state and its perception of its local environment. Each flower-agent's reward availability was also updated at every time step. Details are provided below.

4.1 Flower distribution and reward availability

The foraging patch was modelled as a torroidal grid. Each cell represented 0.35 m square in a foraging site of 200 m square (571 × 571 cells). Bee-agents searched this patch for flower-agents. Flower-agents of two equally abundant types, T1 and T2, were positioned according to a uniform random distribution, at most one per grid cell at density 13,071 / 326,041 cells (~1/25) for all simulations. This simulates a temperate environment in which resources are typically randomly spaced [48, 49]. Flower-agents may offer a reward (1 nutrition unit) to visiting bee-agents, or not, depending on the experiment. When a bee-agent visited a rewarding flower-agent it acquired the entire 1 unit. The reward was replenished after a bee-agent collected it, however, an individual bee-agent cannot return to a flower-agent within a single bout so that a particular reward is only available once for each bee-agent in each bout. This assumption allowed us to isolate bee-agents from one another to avoid forager population density effects that were not central to the particular research questions under consideration in some experiments.

In control experiments we introduced 'defective' flower-agents that did not offer a reward even when their type was expected to do so. By varying the percentage of defective flowers we isolated population density effects.

4.2 The bee-agent and its colony

Each bee-agent contained data: (x,y) grid cell position, variable preferences used as flower-agent landing probabilities for each flower-agent type (0, 100%) that totalled 100%, current load (0, 100 nutrition units) as well as visual acuity and memory parameters (discussed below).

To eliminate locality artefacts the beehive was not allocated a position. Instead bee-agents were randomly placed on the grid to begin and end foraging bouts from any location. Since we are interested in *relative* foraging strategy performance, bee-agent flight speeds were irrelevant and were therefore uniformly set to one grid cell (in the direction of the cardinal points of the compass) per simulation time step.

If a bee-agent occupies a grid cell containing a single flower-agent it will determine (using a strategy discussed below) whether or not to land and attempt to collect a reward. Each bee-agent memorises the flower-agents it just visited, and will not revisit a flower-agent in a particular bout. At the conclusion of a foraging bout, bee-agents 'forget' the locations of previous visits. The justification for bee-agent forgetting is that animals have different memory phases, and are likely to just use short-term memory for recalling the recent choices made in a complex foraging environment [50]. Apart from for the four most recently visited flowers, a bee-agent must enter a flower-agent's grid cell to determine if it has already visited the flower-agent. To ensure that flower-agents are not repeatedly re-approached, bee-agents were modelled with accurate memory of the whereabouts of the last four flower-agents approached [51, 52].

At each simulation time step, each bee-agent examines its neighbourhood for flower-agents. The bee-agent randomly chooses an unapproached flower-agent as its destination, or, if no unapproached flower-agents are detected, the bee-agent chooses a random cell in its neighbourhood. We recognise that bee foraging path selection is probably not entirely random [52], however, we are interested only in *relative* flower selection strategy performance, not absolute success of particular navigational approaches. Hence, as long as all bee-agents choose a path in the same way – randomly in our case – we obtain the information required to understand the relative success of different foraging strategies.

4.3 Flower detection

Bees can use multiple cues (vision and olfaction) to find flowers [53, 54]. Our model only considered that bee-agents distinguish between the two similar flower-agent types, T1 and T2, using visual cues.

Bee spatial acuity is relatively poor compared with a vertebrate lens eye [55]. For example, in real life bumblebees can only detect a plant's cluster of 3-5 flowers (each flower of 2.5 cm diameter) at a distance of less than 0.7 m [25, 56]. Detection appears to be a step function [25, 57] so we modelled bee-agent acuity as distances > 0.7 m — not detected, distances < 0.7 m — 95% chance of detecting flowers. Thus a bee-agent

can detect the presence of a flower-agent in a neighbouring grid cell (up to 0.35 m + 0.35 m = 0.7 m away) with 95% accuracy, but bees cannot detect flowers at a greater distance [25, 57].

If a bee-agent is sharing a grid cell with a detected flower-agent the bee-agent has a probability of accurately discriminating a rewarding from an unrewarding flower-agent. This probability is modelled as the learnt percentage preference for a particular flower-agent. This value represents the probability with which the bee brain correctly determines if a given flower is a target or distractor. With these values it is possible that a bee-agent detects a target flower-agent, but chooses not to land if its preference for that type causes it to mistakenly perceive it as a distractor. Conversely a bee may choose to land on a distractor due to a perceptual error, causing it to mistakenly perceive it as a target. This models the probabilistic way that bees discriminate between similar flower colours depending upon experience [23].

In this simulation, bee/flower-agent visits to rewarding or unrewarding flower-agents require the same time. The visit includes landing on and handling the rewarding or unrewarding flower-agent, as well as extracting or attempting to extract a reward. Flower-agent handling occurs within a single simulation time step along with flight between neighbouring locations and decision-making time unless otherwise noted for specific experiments.

All bee-agents in our experiments undergo two phases of flower-agent preference adjustment behaviour based on experimental results (Section 3.0). These phases allow a bee-agent to adjust its internal preferences for flower-agents based on previous successful and failed reward collection attempts. During *phase 1*, bee-agents conduct standard learning behavior (as described above for real bees, Section 2.0). During *phase 2*, bee-agents adopt either a Deliberative-decisive, Fickle-circumspect or Stay behavior based on the strategies we outlined above for the real bees. The simulation of these phases is discussed below.

4.4 Simulation phase 1 – bee-agent preliminary information acquisition

At the start of an experiment, all bee-agents commence phase 1 with equal 50% preferences for each flower type T1 and T2. Thus, initially a bee-agent has a 50% probability of choosing to land on the first flower-agent it encounters. This corresponds directly to how real bees initially generalise similar colour flowers [23, 28].

During phase 1, each time a bee-agent visits a rewarding flower-agent its preference for that type (T1 or T2) increases by 1% and its preference for the other type drops by 1%. Each time a bee-agent visits an unrewarding flower-agent preference for its type *decreases* by 1% and its preference for the other type of flower-agent *increases* by 1%. This learning model fits how real bees have been shown to learn in experiments [23, 28]. In this simulation, a bee-agent's flower-agent preferences are clamped between 80 and 20% in keeping with realistic learning behaviour of bees for perceptually difficult tasks [23, 28] and our own results (Fig. 1).

All bees undergo phase 1 described entirely by the strategy just outlined. At some point in time, entry into phase 2 occurs, triggering a change in bee-agent foraging behaviour.

4.5 Transition from simulation phase 1 to phase 2

A simulation always begins in a scenario where T1 flower-agents are rewarding and T2 are not. This holds true for the whole of phase 1. Phase 1 continues until an average of 70% preference for T1 (and 30% preference for T2) is reached for the entire bee-agent colony. Then phase 2 begins.

4.6 Simulation phase 2 – Deliberative-decisive, Fickle-circumspect or Stay behaviour

Throughout phase 2, each bee-agent adopts one of the following foraging strategies depending on the experiment.

Deliberative-decisive. A Deliberative-decisive bee-agent decisively reverses its preferences if, over a period of deliberation time its preferred flower type is consistently found to have become unrewarding and another type is detected as being consistently rewarding. When a switch occurs in the rewards offered by a flower-agent type (e.g. T1 flower-agents were rewarding and suddenly become unrewarding), a Deliberative-decisive bee-agent that encounters four unrewarding T1 and four rewarding T2 flower-agents, can invert its preference accuracy. For instance, if a Deliberative-decisive bee-agent had learnt through experience to have 70% preference for T1 and 30% for T2 flower-agents, after encountering four unrewarding T1 and four rewarding T2 this can become a 70% preference for T2 and 30% for T1. This bee-agent has discriminatory ability and sufficient neural flexibility to quickly adjust to a new reward situation. This bee-agent's foraging closely fits the profile of the free-flying bees shown in figure 2.

Fickle-circumspect. A bee-agent may learn to make accurate decisions about flower-agent types based on experience as described for experimental phase 1. The Fickle-circumspect bee type continues to learn in phase 2 with a 1% change per flower-agent visit as it did initially in phase 1. Whenever the reward situation changes, this bee type incrementally adjusts its preferences. It is fickle in the sense that it is always willing to adjust preferences based on new evidence. It is circumspect in the sense that the changes it makes are slight. This bee can adapt, but only slowly. Fickle-circumspect bee preferences are clamped between 80 and 20% in keeping with the findings of our behavioural experiments (Fig. 1).

Stay. A Stay bee's preference 'stays' loyal to the flower-agent type for which it originally acquired a strong preference, even if that flower-agent type changes from rewarding to unrewarding. For example, once a Stay bee-agent reaches ~70% preference for a flower-agent type, this preference becomes hard-wired, even if the availability of rewards offered by the flower-agents changes. This bee-agent's foraging closely fits the profile of the free-flying bees shown in figure 3.

4.7 Simulation experiments

We conducted experiments to determine the impact of varying different conditions on the foraging success of colonies, each containing 60 bee-agents, utilising the following four foraging strategy compositions: all Fickle-circumspect bees, all Stay bees, all Deliberative-decisive bees, 1/3 mix of each bee type. Experimental method details are provided next.

Alternating reward experiments

Real flowers typically offer rewards for only a period of time, and a flower that is rewarding in the morning may not do this later in the day (and vice versa) [3-6]. Thus we considered scenarios where the variation in the availability of rewards offered by flower-agents of type T1 and T2 alternates cyclically in periods ranging from 100 time steps per cycle (on average, 19 changes of the flower reward situation during phase 2)[1] to 1900 time steps per cycle (on average, only 1 change of the flower reward situation during phase 2) of the experiments.

Throughout phase 1 of these experiments, T1 is rewarding and T2 is not. From the start of phase 2 for *n* time steps, T2 is rewarding and T1 is not. For the next *n* time steps T1 is rewarding and T2 is not. This switching process continues, where *n* is a number between 100 and 1900. See Figure 4a for a visual explanation of this process for n=1000, and Figure 4b for n=1900.

Defective reward control experiments

When bees forage together with coworkers or bees from other colonies in the same patch, there is some likelihood that a bee lands on a flower that should be offering a reward at the time of the visit, but the reward has recently been acquired by another bee [3-6].

By varying the fraction of a rewarding flower-agent type that is defective (in the sense that they do not offer rewards when they should), it was possible to test the relative impact of these conditions on the bee-agents' foraging strategies. Phase 1 of these defective reward experiments were run as described above. During phase 2 of these experiments the proportion of the T1 flower-agents that were defective and contained no reward was systematically varied. T2 flower-agents remained unrewarding throughout the experiment.

Handling time control experiments

Flower handling time may impact on the relative success of different bee foraging strategies [3, 18]. Even if flower handling times are identical for rewarded and unrewarded landings, bees making poor landing decisions when compared with bees skipping over poor landing opportunities to search further could be expected to waste more time on flower handling for no reward [18, 19]. We tested for the impact of flower handling times by running experiments with a single reward swap at the transition from phase 1 to phase 2, using a series of flower handling times from 0 to 20 time steps. In this experiment bee-agents sat idle for a number of simulation time steps corresponding to flower handling time when they visited a flower-agent, regardless of whether or not a reward was collected.

Dependent variable

The dependent variables of interest are the relative amounts of nutrition collected during each simulation run and returned to the hives by the four bee colony compositions under the experimental conditions. Nectar collection rate is a variable

---

[1] We found phase 1 of the experiments to have an average length of 1131 time steps (46 s.d.) when flower handling time was 0. Since the experiments all cease at time step 3000, this leaves an average length of 1869 time steps for phase 2.

likely to be of high importance to colony survival in natural conditions where resources may be constrained at certain times of the year [18, 58].

Environment	
Patch size	571 × 571 cells, torroidal boundary
Patch grid cell size (modelled)	0.35 × 0.35 m
Flower-agents	
T1 : T2 flower relative abundance (mean)	1 : 1
Total number of flowers	13071
Bee-agents	
Colony sizes and compositions	60 agents/colony, composed as: all Fickle-circumspect, all Stay, all Deliberative-decisive, 1/3 mix of each
Flower detection accuracy	95% from a neighbouring cell or one shared with a flower
Storage capacity	100 reward units
Recently approached flower memory length	4 most recent approaches
Visited flower memory length	Every flower visited on a single bout
Learning increment (for all bees during phase 1, and for Fickle-circumspect bees during phase 2)	±1% point per flower visit
Preference switch threshold (for Deliberative-decisive bees during phase 2)	Visit 4 surprising rewarding flowers *and* 4 surprising unrewarding flowers without encountering any unsurprising flowers
Simulation length	
Duration of a simulation run (phase 1 + phase 2)	3000 time steps
Duration of phase 1	Until colony mean preference for T1 flowers ≥70%

Table 1. Summary table of simulation parameters.

4.8 Simulation model verification and validation

Simulation verification involved checking the correctness of our model by ensuring that the simulation behaved in accordance with the behaviours detailed above. Checks included that:

- Learn bee-agents changed preferences in ±1% increments appropriately.
- Bee-agent approach and visit memories functioned correctly.
- Stay bees did not change preferences during phase 2.
- Flower distribution mean was 50% T1, 50% T2.

This verification process also allowed us to determine some derived values for the behaviour of the system that were important for interpreting our results. These are given below.

Check that Deliberative-decisive bees make a flower-agent preference switch after receiving four rewards from sources that were not expected to offer a reward, and four failed attempts to extract a reward from flowers where a reward was expected.

There is always some delay after a flower-agent reward availability swap until a Deliberative-decisive bee-agent makes a preference switch due to its need to visit a minimum of eight flower-agents. Mean time until bee-agent preference switch was 765 time steps (273 s.d.) after the flower-agents changed their rewards (median switch time was 756 time steps). Values computed from a sample of 60 Deliberative-decisive bee-agents.

Check transition condition from phase 1 to phase 2 occurs when the mean bee-agent flower-agent preference for T1 reaches 70%.

We found that the mean time to reach a preference of T1 = 70% (T2 = 30%) was 1131 time steps (45 s.d.). At this time, the bee-agent colonies had foraged on average a total of 667 units of reward (18 s.d.). Values computed from 40 randomised simulation runs with 60 bee-agents per colony.

The gradual increase in a bee-agent's preference for T1 flower-agents during the initial learning period (phase 1) is illustrated in figures 4 a and b. Consistent with fine colour learning in foraging honeybees and bumblebees, after both unsuccessful and successful visits to flowers, the bee-agents' preferences for T1 increase. Reward foraged is plotted on the lower (green) lines. T1 preferences are plotted on the upper (blue) lines.
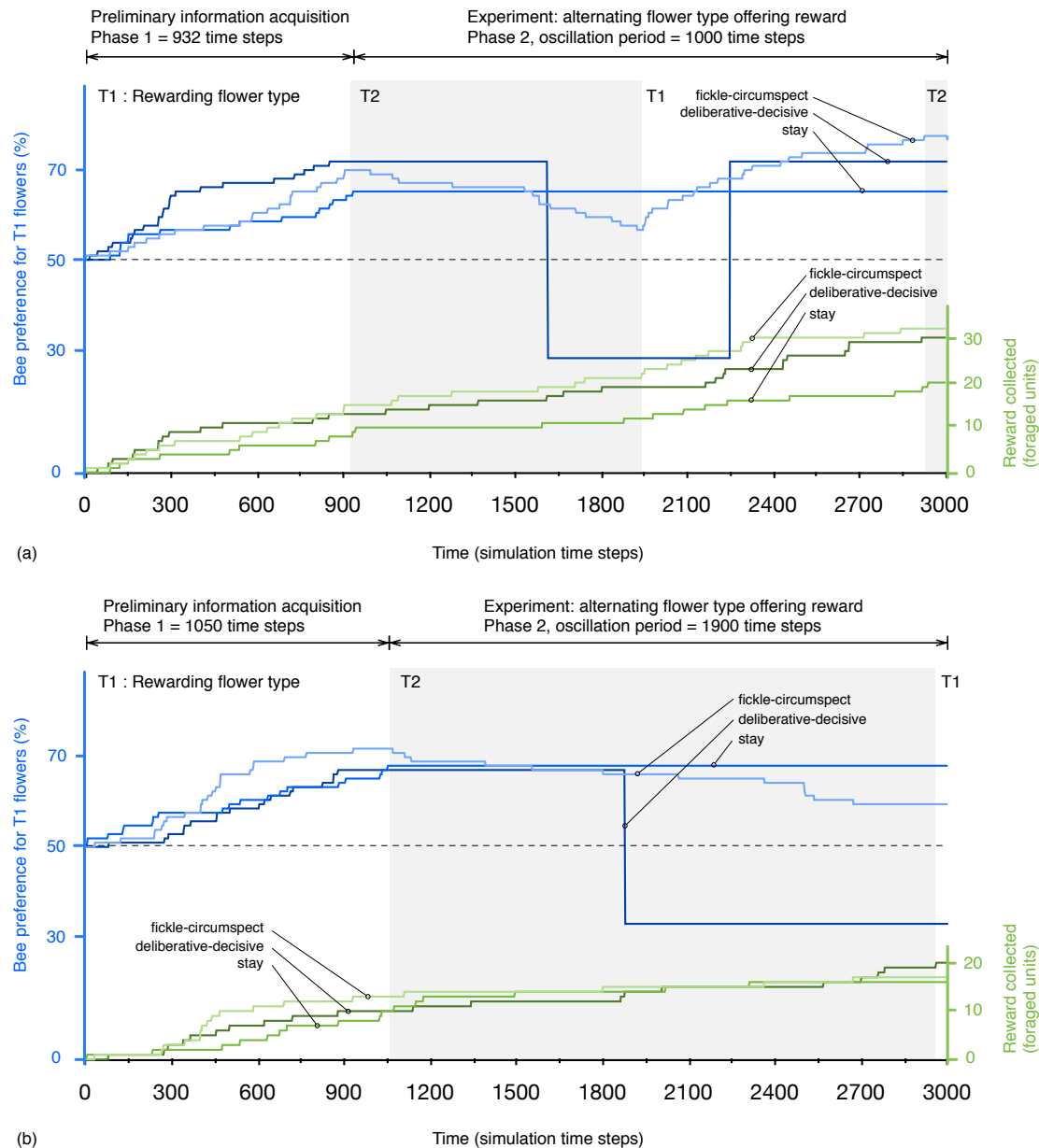
Figure 4a & b. Sample *alternating reward* experiments for single bee-agents of each strategy. The upper plots (blue) are the preferences of each bee-agent for T1 rising from an initial 50% value (vertical, left hand axis). The lower plots show units foraged by the individual bee-agents during the depicted simulation runs (vertical, right hand axis). In these examples the preliminary acquisition period during which the bee-agent colony reaches a mean preference for T1 of 70% occurs from time step 0 to (a) 932, (b) 1050. From then on, in these particular experiments, the period of reward availability swapping between T2 and T1 is set to (a) 1000 time steps (b) 1900 time steps. The availability of rewards from T2 and T1 is illustrated by the alternating grey bands that occur at regular intervals from the beginning of phase 2 until the end of the experiments at time 3000.

5.0 Results: Simulation experiments

5.1 Alternating reward experiments

The impact on mean reward foraged of varying reward swapping periods during phase 2 of the simulations is illustrated in figure 5. Swapping periods from 100 time steps (approximately 19 swaps during the 1900 time step long phase 2) to 1900 time steps (1 single reward swap at the commencement of phase 2 and no subsequent swaps) were simulated.

At no single value for swapping period in this range could we find any indication that the mixed beehive composition (illustrated in solid (red), Fig. 5) was most effective. The all Fickle-circumspect strategy hives were consistently the best performers for reward swapping periods until the all Deliberative-decisive strategy hives took over as the most effective foragers when reward swap period reached 1600 time steps.
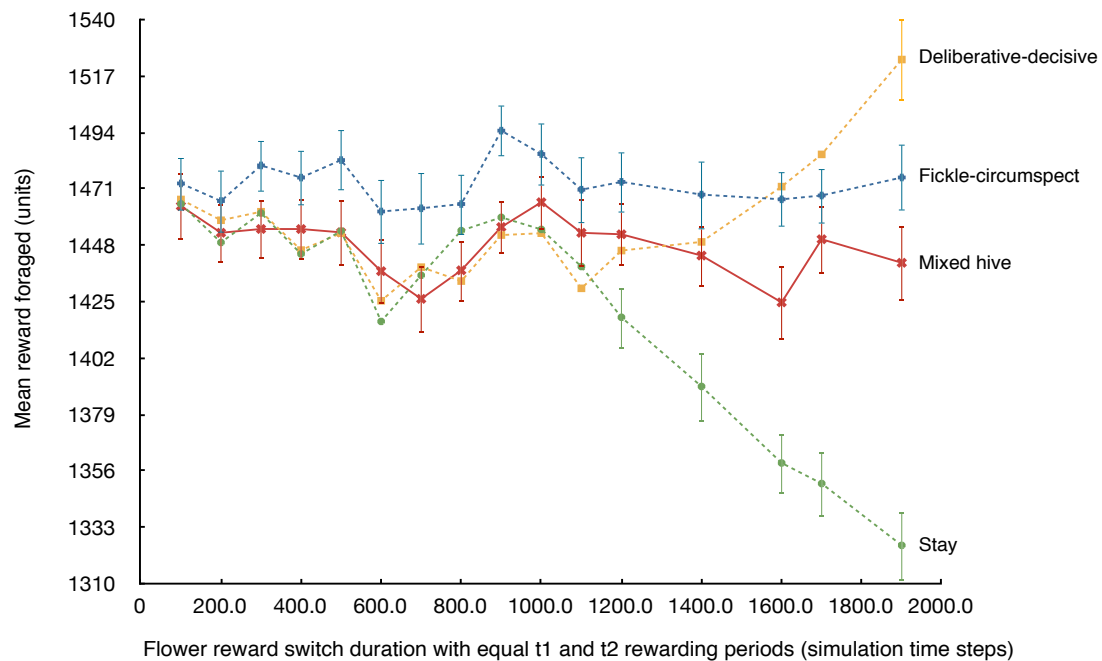
Figure 5. Mean reward foraged by hives of different composition after 3000 time steps, as compared to the length of time (measured in simulation time steps) for which rewards are alternately offered by flower-agents of T1 and T2 during simulation phase 2. Mean values based on 50 simulation runs. Error bars indicate 95% confidence intervals. (Note, figure 4(a) shows the experiment above conducted at oscillation period=1000 for 3 individual bees of each strategy. Figure 4(b) shows the experiment above conducted at oscillation period=1900 for 3 individual bees of each strategy.)

5.2 Control experiments

Defective reward experiments: These simulation runs (N=20) determine the impact of defective flower-agents on foraging strategies by providing a source of noise that could potentially confuse bee-agents as to which flower type is rewarding. We separately considered the 0, 30, 50, 70 and 90% defective flower distributions by testing how the four different hive compositions collected nutrition for each distribution.

There was no significant difference in the dependent variable of mean reward collected for any of the defective reward experiments [One-way ANOVA, d.f. (3,76); 0% condition $F = 1.992$, $p=0.122$; 30% $F = 0.886$, $p=0.452$; 50% $F = 1.181$, $p=0.323$; 70% $F = 0.087$, $p=0.967$; 90% $F = 1.277$, $p=0.268$], showing that the relative foraging of the bee-agent hives with different strategies was independent of defective flower density effects.

Handling time experiments: In these simulation runs (N = 20) a single reward swap occurred at the conclusion of simulation phase 1. From then on, all bee-agents enacted their phase 2 change-of-preference strategies. In these runs, every landing on a flower

caused a bee-agent to sit idle during the designated flower handling time.

We did not identify any condition where the mixed beehive composition was most effective for the flower handling times tested. Flower handling time did not impact on the ranking of the different hive compositions for any test condition.

6.0 Discussion

The behavioural experiments suggest a complex picture emerges for individual bees having to reverse-learn a difficult colour discrimination task. Whilst some bees were loyal to one of several strategies, other bees appeared to have flexibility to change between different strategies. A classic problem in understanding decision-making in insects is, why have colonies evolved different strategies for solving perceptually difficult problems [18, 33]? Agent-based simulations allowed us to analyse hypotheses in response to this question using our empirical results as a basis.

In the current study, the empirical analyses initially suggested that there were three types of bee strategy for collecting nutrition from similarly coloured flowers that may alternate in the availability of rewards offered. The potential strategies identified were Deliberative-decisive, Fickle-circumspect and Stay strategies. We found that Stay strategy bees *never* outperformed the other two strategies under the test conditions, often to the detriment of the hive's overall efficiency (Fig. 5). This leads us to question why evolution might maintain Stay bees. One possible reason is that by sticking with their hard-won knowledge in the case where there are occasional unrewarding target flowers, Stay bees may be effective foragers. However, the control experiment on defective flowers allows us to discount this hypothesis.

Another possibility is that the bees that we classified as Stay bees following our initial analysis of the empirical data, may have been Fickle-circumspect bees that just learnt at a very slow rate (Fig. 3). The individual-based simulation allows us to conclude that under the conditions we tested, bee colonies constrained to the requirement of collecting the maximum amount of nutrition per unit time should contain bees with only two types of decision-making strategy for complex colour tasks. Following the learning of a perceptually difficult task, our results suggest that under repeated and frequent changes in a reward situation, it is best to make frequent and small changes to preferences in order to keep abreast of the changing environment, without missing an opportunity to fine-tune the perceptual system. Even though this fine-tuning may not bring the perceptual system fully into line with the new reward situation (Fig. 4), the minor adjustments take little time to be swayed in the reverse direction, and so the bee may simply revert to earlier preferences when needed using the same mechanism. But under situations where changes occur less often, fine-tuning is too slow to capitalise. Instead, once sufficient evidence has been gathered that the reward situation has changed, a large and decisive change of preferences is a more effective way to bring a bee's perceptually based decisions into line with new foraging conditions.

Future work could consider whether there are unsimulated real-world conditions for which the presence of Stay bees is of benefit to a beehive. We note that the better the Deliberative-decisive bees learn during phase 1 to prefer T1 flowers, the longer it will take them to reverse their preferences after a reward swap in favour of T2, or any subsequent reversals. This occurs because their preferences for the currently

unfavourable flower will be so low that they seldom visit to gain evidence for any newly instigated favourability. Perhaps the presence of Stay bees in a hive together with Deliberative-decisive bees may be useful if the bees can communicate with one another in some circumstances [43]. In particular, a Stay bee loyal to one flower type might act as a 'watch-bee' consistently monitoring flowers that were known to be previously rewarding, ready to inform the hive once they become viable targets.

In complex real environments there may be no way for animals in general to know which conditions will be encountered, and thus it could be beneficial for a colony to possess individuals with multiple strategies, in agreement with bet hedging hypotheses [18, 19]. Figure 5 supports this hypothesis in the case of bees.

Classically it has been appreciated that decision-making in humans can be described as either *Rationalistic* in which a human subject "becomes aware of a problem, posits a goal, carefully weighs alternative means, and chooses among them according to his estimates of the respective merit with regard to the state of affairs he prefers" [59, 60]. This parallels our Deliberative-decisive bee strategy. Alternatively, human subjects may use an *Incrementalist* strategy. This strategy "seeks to adapt decision-making strategies to the limited cognitive capacities of decision-makers and to reduce the scope and cost of information collection and computation" [59, 60], by making a continual stream of micro-adjustments. In the field of Artificial Intelligence, which is concerned with decision-making agents, these different strategies are understood in terms of the concepts of *simple reflex agents* that select their behaviour based purely on their current situation; and *model-based reflex agents* that maintain an internal model of the part of the world that has been visible to them in the past in order to take this into account in their decision-making [61, pp. 48-52].

Our finding that bees demonstrate a variety of decision-making strategies is therefore consistent with theories of decision-making in both humans and artificial intelligence. This suggests that results from the current bee and simulation experiments have widespread implications across a number of fields. Indeed the beehive may, by maintaining a diversity of individual level behaviour, be acting as a super-organism [62, 63] with respect to its implementation of decision-making for complex problems that have no obvious, complete *a priori* solution. In addition, the behavioural experiments found some evidence that individual bees have a capacity to modulate their own decision-making approach, and modulation of decisions has previously been reported for bumblebees making speed/accuracy judgements for stimuli of similar colour [30, 31].

We conclude that for a perceptually difficult task, honeybees demonstrate a variety of methods for adjusting their flower preferences when environmental conditions change. Our agent-based simulations of this behaviour reveal that this diversity of strategies allows a hive to efficiently collect resources in a complex, dynamic environment.

Acknowledgements

Bibliography

[1]   Dyer AG, Chittka L. Biological significance of distinguishing between similar colours in spectrally variable illumination: bumblebees (*Bombus terrestris*) as a case study. J Comp Physiol A. 2004;190:105-14.

[2]   Dafni A. Mimicry and deception in pollination. Ann Rev Ecol Syst. 1984;15:259-78.

[3]   Chittka L, Gumbert A, Kunze J. Foraging dynamics of bumble bees: correlates of movements within and between plant species. Behav Ecol. 1997;8:239-49.

[4]   Waddington KD, Allen T, Heinrich B. Floral preferences of bumblebees (*Bombus edwardsii*) in relation to intermittent versus continuous rewards. Anim Behav. 1981;29(3):779-84.

[5]   Waddington KD, Heinrich B. Patterns of movement and floral choice by foraging bees. In: Kamil AC, Sargent TD, eds. *Foraging Behavior*. New York: Garland STPM Press 1981:215-30.

[6]   Townsend-Mehler JM, Dyer FC. An integrated look at decision making in bees as they abandon a depleted food source. Behav Ecol Sociobiol. 2012;66:275-86.

[7]   Giurfa M, Núñez J, Backhaus W. Odour and colour information in the foraging choice behaviour of the honeybee. J Comp Physiol A. 1994;175:773-9.

[8]   Chittka L, Schürkens S. Successful invasion of a floral market. Nature. 2001;411:653.

[9]   Heinrich B. Bumblebee economics. Cambridge: Harvard University Press 1979.

[10]   Chittka L, Thomson JD, Waser NM. Flower constancy, insect psychology, and plant evolution. Naturwiss. 1999;86:361–77.

[11]   Pavlov IP. Lectures on conditioned reflexes. New York: International Publishers 1927.

[12]   Mota T, Giurfa M. Multiple reverse olfactory learning in honeybees. Frontiers in Behavioral Neuroscience. 2010;4:1-9.

[13]   Komischke B, Giurfa M, Lachnit H, Malun D. Successive olfactory reversal learning in honeybees. Learning and memory. 2002;9:122-9.

[14]   von Helversen O. Dressurversuche mit Bienen. Biol Unserer Zeit. 1974;4:119-24.

[15]   Menzel R. Das Gedächtnis der Honigbiene für Spektralfarben. II. Umlernen und Mehrfachlernen. Z Vergl Physiol. 1969;63:290-309.

[16]   Chittka L. Sensorimotor learning in bumblebees: long term retention and reversal training. J Exp Biol. 1998;201:515-24.

[17]   Bitterman ME, Menzel R, Fietz A, Schäfer S. Classical conditioning of proboscis extension in honeybees (Apis mellifera). J Comp Psychol. 1983;97:107-19.

[18]   Burns JG, Dyer AG. Diversity of speed accuracy strategies benefits social insects. Curr Biol. 2008;18:R953–R4.

[19]   Muller H, Chittka L. Animal Personalities: The Advantage of Diversity. Curr Biol. 2008;18:R961-R3.

[20]   Avarguès-Weber A, de Brito Sanchez MG, Giurfa M, Dyer AG. Aversive reinforcement improves visual discrimination learning in free flying honeybees. PLOS One. 2010:e15370.

[21]   Giurfa M. Conditioning procedure and color discrimination in the honeybee *Apis mellifera*. Naturwiss. 2004 May;91(5):228-31.

[22] Dyer AG, Chittka L. Fine colour discrimination requires differential conditioning in bumblebees. Naturwiss. 2004 May;91(5):224-7.

[23] Dyer AG. Psychophysics of Honey Bee Color Processing in Complex Environments. *Honeybee Neurobiology and Behavior*. Netherlands: Springer 2012:303-14.

[24] Dyer AG, Neumeyer C. Simultaneous and successive colour discrimination in the honeybee (*Apis mellifera*). J Comp Physiol A. 2005;191(6):547-57.

[25] Dyer AG, Spaethe J, Prack S. Comparative psychophysics of bumblebee and honeybee colour discrimination and object detection. J Comp Physiol A. 2008;194:617-27.

[26] Chittka L. The color hexagon: a chromaticity diagram based on photoreceptor excitations as a generalized representation of colour opponency. J Comp Physiol A. 1992;170:533–43.

[27] Avarguès-Weber A, Deisig N, Giurfa M. Visual cognition in social insects. Annu Rev Entomol. 2011;56:423–43.

[28] Dyer AG, Paulk AC, Reser DH. Colour processing in complex environments: insights from the visual system of bees. Proc Royal Soc B. 2011;278:952-9.

[29] Dyer AG, Murphy AH. Honeybees choose "incorrect" colors that are similar to target flowers in preference to novel colors. Isr J Plant Sci. 2009;57(3):203-10.

[30] Dyer AG, Chittka L. Bumblebees (*Bombus terrestris*) sacrifice foraging speed to solve difficult colour discrimination tasks. J Comp Physiol A. 2004;190:759-63.

[31] Chittka L, Dyer AG, Bock F, Dornhaus A. Bees trade off foraging speed for accuracy. Nature. 2003;424:388.

[32] Dyer AG. The mysterious cognitive abilities of bees: why models of visual processing need to consider experience and individual differences in animal performance. J Exp Biol. 2012;215:387-95.

[33] Burns JG. Impulsive bees forage better: the advantage of quick, sometimes inaccurate foraging decisions. Animal Behav. 2005;70:e1-e5.

[34] DeAngelis DL, Mooij WM. Individual-based modelling of ecological and evolutionary processes. Annu Rev Ecol Evol Syst. 2005;36:147-68.

[35] Dorin A, Korb KB, Grimm V. Artificial-Life Ecosystems: What are they and what could they become? In: Bullock S, Noble J, Watson RA, Bedau MA, editors. Eleventh International Conference on Artificial Life; 2008: MIT Press; 2008. p. 173-80.

[36] Grimm V. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? Ecological Modelling. 1999;115:129-48.

[37] Grimm V, Railsback SF. Individual-based Modeling and Ecology. Princeton, N.J.: Princeton University Press 2005.

[38] Grimm V, Revilla E, Berger U, Jeltsch F, Mooij WM, Railsback SF, et al. Pattern-Oriented Modelling of Agent-Based Complex Systems: Lessons from Ecology. Science (New York, NY). 2005;310:987-91.

[39] Huston M, DeAngelis D, Post W. New Computer Models Unify Ecological Theory. BioSci. 1988;38(10):682-91.

[40] Judson OP. The rise of the individual-based model in ecology. Trends in Ecology and Evolution. 1994;9(1):9-14.

[41] Hogeweg P, Hesper B. The Ontogeny of the Interaction Structure in Bumble Bee Colonies: A MIRROR Model. . Behavioural Ecology and Sociobiology. 1983;12:271-83.

[42] de Vries H, Biesmeijer JC. Modelling collective foraging by means of individual behaviour rules in honey-bees. Behavioural Ecology and Sociobiology. 1998;44:109-24.

[43] Dornhaus A, Kluegl F, Oechslein C, Puppe F, Chittka L. Benefits of recruitment in honey bees: effects of ecology and colony size in an individual-based model. Behav Ecol. 2006;17:336-44.

[44] Dyer AG, Rosa MGP, Reser DH. Honeybees can recognise images of complex natural scenes for use as potential landmarks. J Exp Biol. 2008;211:1180-6.

[45] Dyer AG, Neumeyer C, Chittka L. Honeybee (*Apis mellifera*) vision can discriminate between and recognise images of human faces. J Exp Biol. 2005;208:4709-14.

[46] Peitsch D, Fietz A, Hertel H, de Souza J, Ventura DF, Menzel R. The spectral input systems of hymenopteran insects and their receptor-based colour vision. J Comp Physiol A. 1992;170:23-40.

[47] Wyszecki G, Stiles WS. Color science: Concepts and methods, quantitative data and formulae. New York: Wiley 1982.

[48] Dornhaus A, Chittka L. Why do honeybees dance. Behav Ecol Sociobiol. 2004;55:395-401.

[49] Visscher PK, Seeley TD. Foraging strategy of honeybee colonies in a temperate deciduous forest. Ecology. 1982;63:1790-801.

[50] Menzel R. Behavioral and neural mechanisms of learning and memory as determinants of flower constancy. In: Chittka L, Thomson JD, eds. *Cognitive Ecology of Pollination*. Cambridge: University Press 2001:21-40.

[51] Soltz RL. Foraging path selection in bumblebees: hindsight or foresight? Behav. 1986;99:1-21.

[52] Pyke GH, Cartar RV. The flight directionality of bumblebees: do they remember where they came from? Oikos. 1992;65:321-7.

[53] Streinzer M, Paulus HF, Spaethe J. Floral colour signal increases short-range detectability of a sexually deceptive orchid to its bee pollinator. J Exp Biol. 2009 May 1, 2009;212(9):1365-70.

[54] Leonard AS, Dornhaus A, Papaj DR. Flowers help bees cope with uncertainty: signal detection and the function of floral complexity. J Exp Biol. 2011;214:113-21.

[55] Land MF. Visual acuity in insects. Ann Rev Entomol. 1997;42:147-77.

[56] Wertlen AM, Niggebrugge C, Vorobyev M, Hempel de Ibarra N. Detection of patches of coloured discs by bees. J Exp Biol. 2008 July 1, 2008;211(13):2101-4.

[57] Giurfa M, Vorobyev M, Kevan P, Menzel R. Detection of coloured stimuli by honeybees: minimum visual angles and receptor specific contrasts. J Comp Physiol A. 1996;178:699-709.

[58] Mattila HR, Seeley TD. Genetic diversity in honey bee colonies enhances productivity and fitness. Science (New York, NY). 2007;317:362–4.

[59] Etzioni A. Mixed-Scanning: A "Third" Approach to Decision Making. Publ Admin Rev. 1967;27:385-92.

[60] Tarter cJ, Hoy WK. Toward a contingency theory of decision making. Journal of Educational Administration. 1998;36(3):212 - 28.

[61]   Russell SJ, Norvig P. Artificial Intelligence, a Modern Approach. 3rd ed. New Jersey: Pearson Education/Prentice Hall 2010.

[62]   Seeley TD. The honey bee colony as a superorganism. Am Sci. 1989;77:546-53.

[63]   Hölldobler B, Wilson EO. The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies. New York: WW Norton & Company 2009.