Bees in two-armed bandit situations: foraging choices and possible decision mechanisms

Tamar Keasar,[a] **Ella Rashkovich,**[b] **Dan Cohen,**[c] **and Avi Shmida**[c]
[a]Department of Life Sciences, Ben Gurion University, Beer Sheva, and Achva College, Mobile Post Shikmim 79800, Israel, [b]QBI Ltd., Ness Ziona, Israel, and [c]Department of Evolution, Systematics and Ecology and Center for Rationality and Interactive Decision Theory, The Hebrew University, Jerusalem, Israel

In multi-armed bandit situations, gamblers must choose repeatedly between options that differ in reward probability, without prior information on the options' relative profitability. Foraging bumblebees encounter similar situations when choosing repeatedly among flower species that differ in food rewards. Unlike proficient gamblers, bumblebees do not choose the highest-rewarding option exclusively. This incomplete exclusiveness may reflect an adaptive sampling strategy. A cost–benefit analysis predicts decreased sampling levels with increasing differences in mean profitability between the available food sources. We simulated two-armed bandit situations in laboratory experiments to test this prediction. Bumblebees (*Bombus terrestris* L.) made 300 foraging visits to blue and yellow artificial flowers that dispensed sucrose solution according to seven probabilistic reward schedules. Reward schedules varied in profitability differences between the two feeding options. As predicted, the bees specialized more on the higher-rewarding food type (and thus sampled the alternative less) when the mean reward difference between the feeding options was larger. Choice ratios of individual bees were linearly related to the reward ratios they had experienced. It has been suggested that the behavioral mechanism underlying incomplete exclusiveness may involve simple rules of thumb that do not require long-term memory. However, the bees' response to recent foraging experience (rewarded and non-rewarded visits) differed between the beginning and the end of observation sessions and between treatments. Simulations of the Rescorla-Wagner difference learning rule reproduced the main trends of the results. These findings suggest that the observed incomplete exclusiveness results from associative learning involving long-term memory. *Key words:* bees, foraging, learning, probability matching, two-armed bandit. *[Behav Ecol 13:757–765 (2002)]*

I n multi-armed bandit situations, gamblers must decide which arm of a slot machine to play in a sequence of trials in order to maximize their rewards. The reward probabilities and magnitudes of the several available arms are not known to the gamblers in advance. The gamblers' optimal strategy is to find the arm with the highest expected return as soon as possible, and then to keep gambling using that arm. Multi-armed bandit situations have been modeled and studied experimentally as examples of decision-making problems with incomplete information in humans (Auer et al., 1995). Bees frequently face similar situations while foraging for nectar and pollen in flowers. Specifically, they have to choose food plants in a sequence of foraging visits when the relative profitability of the several available plant types in the field is not known in advance (Real, 1992). Bumblebees do not follow the strategy of maximizing expected rewards: After a learning period, they typically direct most of their foraging visits to one flower species that is most abundant, morphologically accessible, and rich in food reward (the "major"). Yet they continue to visit a few other species occasionally (the "minors"). Heinrich (1979) suggested that this "majoring-minoring" behavior reflects a trade-off between resource exploration and exploitation, which should be favored by natural selection under changing foraging conditions. According to this hypothesis, exploration of several food sources can help bees detect changes in resource profitability over time, although it de-

creases the bees' food intake if foraging conditions never change (Cohen, 1993; Shettleworth et al., 1988; Stephens, 1987).

How should an animal optimally allocate its efforts between exploitation and exploration in different foraging situations? Several theoretical models deal with this question for patch choice (Cohen, 1993; McNamara and Houston, 1985a,b; Stephens and Krebs, 1986) and prey choice (Estabrook and Jespersen, 1974) situations. These models predict that optimal exploration levels would be affected by the variance in quality within and among food resources, their predictability in time, and the total time available for foraging. According to a cost–benefit analysis (Cohen, 1993), the profit of sampling for foragers is the possibility to detect a food source that is better than their current choice. The costs of sampling involve travel and missed opportunity costs. The model predicts that the benefit of sampling to foragers should increase if all food sources available to them are similar in mean profitability— that is, if there is much overlap between the frequency distributions of food rewards in the different sources. In this case, repeated sampling is needed to detect which food source offers the highest mean. Such sampling will lead to reduced exclusiveness on the best food source.

In the present experiment we recorded the foraging behavior of bees on two types of artificial flowers that were characterized by display color and reward probability. Thus, we mimicked two-armed bandit situations, the simplest case of the multi-armed bandit problem. The first aim of the experiment was to characterize the bees' choices and to relate them to the prediction of the optimal sampling model above. To this end, we tested whether the degree of specialization on the high-rewarding flower type is affected by the ratio and/ or difference in mean reward probability between the two flower types. The second aim of the experiment was to learn

about the decision-making mechanisms underlying the bees' choices.

Conflicting views exist regarding the proximate molecular mechanism that leads to incomplete exclusiveness by foraging bees. According to one view, foraging preferences are formed through associative learning that reflects the bees' long-term foraging experience. Food sources are chosen in proportion to their relative profitability, as perceived by the foragers. In support of this view, modifications of Rescorla and Wagner's (1972) difference rule accounted for foraging choices in two previous laboratory studies that studied incomplete exclusiveness by bees (Fischer et al., 1993; Greggers and Menzel, 1993).

Alternatively, bees may base their choices on rules of thumb that assume no long-term learning, yet lead to incomplete specialization. It was suggested that such rules lead to efficient foraging and are thus adaptive for organisms whose nervous system may be too simple to use more complex choice mechanisms (Real, 1991). A number of such rules have been proposed in theoretical analyses (Milano, 1994; Thuijsman et al., 1995). For example, Thuijsman et al. (1995) described two rules of thumb, the ϵ-sampling rule and the failures rule, that can result in partial exclusiveness on the high-rewarding food source in a two-armed bandit situation. The ϵ-sampling rule assumes that bees switch food source type at a constant probability, regardless of reward. If the reward obtained after the switch exceeds a fixed threshold, the bee continues foraging at its new food source (with the same fixed probability of switching back). Otherwise it switches back immediately. The second rule of thumb, the failures rule, involves switching food source type after a bee experiences a certain number of consecutive rewards that are below threshold. For example, the bees should switch flower type after encountering a fixed number of empty flowers of the same type in succession.

The second aim of our experiment was to compare the adequacy of simple rules of thumb versus simple models of associative learning for describing the bees' choices. The ϵ-sampling rule and the failures rule were used as example rules of thumb, while the Rescorla-Wagner rule was used as a representative mechanism of associative learning. We tested whether the experimental results could be reproduced by following the rules of thumb and through simulations of the Rescorla-Wagner rule. We show that long-term associative learning produces a better fit with experimental data than the short-term rules of thumb.

METHODS

Experiments were carried out in a flight room measuring 4 $\times$ 2.5 m. Temperature ranged from 26° to 30°C, and relative humidity was 40–70%. The room was illuminated from 0630 to 1830 h with six D-65 neon light tubes. We conducted observations throughout 1995 and 1996, between 0800 and 1630 h. Each bee was typically observed for 2–3 h within this time frame.

Colonies of naive *Bombus terrestris* (L.) bumblebees were obtained from Kibbutz Yad Mordechai, Israel. The queens of the colonies that were used during the winter months were treated by the suppliers to forego hibernation. Colony development and the behavior of workers in these colonies resembled colonies reared in summer. All individuals in the colony were marked within 3 days of emergence. Pollen was supplied ad libitum, directly to the colony. Sucrose solution was dispensed by artificial flowers and by a petri-dish feeder, as detailed below. We used 72 workers from 4 colonies for experiments.

Artificial flowers

Twenty morphologically identical computer-controlled artificial flowers were used for experiments. The flowers were tube-shaped and 8 cm tall. A removable, round plastic landing surface of 5.8 cm diameter was placed on top of each flower. Ten flowers were marked with a human-blue landing surface, and the other 10 flowers were marked human-yellow. Blue and yellow flowers were placed in alternation in a 4 $\times$ 5 grid. The city-block distance between neighboring flowers in the grid was 19 cm, except for two rows that were separated by 16 cm only.

The design of the electronic flowers is described in detail in Keasar (2000). Briefly, each flower consisted of a cylindrical container that held a 30% sucrose solution, and a 1-μl miniature cup that was refilled when programmed to dip into the sucrose-solution container. Only foragers that landed on the top part of the flowers and probed them correctly were able to access this cup and feed. Each artificial flower was equipped with a photodetector that was activated when the foraging bee inserted its head into the flower. The photodetector signals were recorded, allowing us to track flower visitation sequences for each forager.

All flowers were full with 1 μl 30% sucrose solution at the beginning of each observation session. This was done in order to motivate the bees to forage by providing positive reinforcements during their first flower visits. After a flower was depleted by a bee, it was either refilled with the 1-μl nectar serving or left empty, according to a predetermined refilling probability. We used a stochastic refilling program that was run independently for each flower. Therefore, flowers in the experimental array refilled asynchronously and independently of their neighbors. Refilling required about 2 s; time intervals between foraging visits were generally longer than 3 s (Keasar et al., 1996). Thus, refilling did not interfere with the bees' foraging activity.

Experimental design

Each bee, foraging singly, made 300 visits to the artificial flowers. Refilling probabilities and number of bees are shown in Table 1. Refilling probabilities differed for blue and yellow flowers in test conditions p8p2, p6p2, p8p4, and p4p2. About half of the individuals in each test condition were rewarded at a higher probability by yellow flowers, while the remaining bees were rewarded at a higher probabilities by blue flowers. This procedure aimed to control for possible color preferences. The higher reward flowers (whether blue or yellow) were numbered 1–10 in conditions p1p0, p8p2, p6p2, p8p4, and p4p2, and the lower reward flowers were assigned numbers 11–20. In conditions p1p1 and p5p5, where both flower types were equally rewarding, we arbitrarily numbered yellow flowers 1–10 and blue flowers 11–20. Conditions p8p2, p6p2, p8p4, and p4p2 tested the bees' level of exclusiveness under probabilistic reward schedules. In control conditions, we eliminated either the difference in profitability between the two feeding options (condition p5p5), reward variability (condition p1p0), or both (condition p1p1). Each bee was used in one condition (one observation session) only. We changed the colored plastic disks that served as landing surfaces to prevent effects of odor marking (Giurfa, 1993; Goulson et al., 1998) before another individual was allowed to approach the artificial flowers. There was no control over the possible use of scent marks by the same bee that had produced them. All artificial flowers were covered at the end of each experimental session. We then allowed all bees in the colony to feed without restriction from a petri-dish feeder that was placed on the table for 2–3 h. The feeder was then removed and the bees

Table 1
Outline of experiments

Treatment	Refilling probability		No. of bees observed	Main aim of treatment
	Flowers 1–10	Flowers 11–20		
p1p0	1	0	10	Control for visit frequency to non-rewarding feeders
p1p1	1	1	10	Effect of reward variability on choice
p5p5	0.5	0.5	9	
p8p4	0.8	0.4	9	Effect of ratio of reward probabilities on choice
p6p2	0.6	0.2	10	
p8p2	0.8	0.2	14	
p4p2	0.4	0.2	10	Effect of magnitude of probability differences on choice (compare to treatment p8p4)

were starved until the next observation session, on the following morning. Thus, experimental bees had previously fed from the petri-dish feeder, but not from any artificial flower.

Bees within each experimental treatment originated from at least two colonies. This procedure diminished possible confounding effects of the bees' source colony (Ney-Nifle et al., 2001). Observation dates of individuals within each treatment spanned 3–9 months, reducing possible behavioral effects of season.

Data analysis

We divided each bee's visitation sequence into groups of 50 consecutive visits and calculated the choice proportion of flowers 1–10 for each group of visits. We averaged these proportions over all bees in a condition to produce average learning curves. Learning curves produced in this way do not represent individual learning curves because very different patterns of individual learning could produce the same average learning curve. We used F tests to compare variability in choice performance between conditions and between the beginning (visits 1–100) and the end (visits 201–300) of each observation session. We tested whether bees within each test condition were homogeneous in their choice performance by using replicated goodness-of-fit tests (Sokal and Rohlf, 1981). The null hypothesis was that all bees within a treatment chose flowers 1–10 at the same frequencies during visits 201–300 of the observations.

Because of the stochastic refilling of the artificial flowers, the bees often experienced reward ratios that were somewhat different from those planned. This was particularly evident for treatment p5p5, where the realized mean reward ratio was about 0.6 (see Figure 2). We calculated the reward ratios that were actually experienced by each bee (realized rewards) during visits 201–300 by dividing the number of rewards obtained in flowers 1–10 by the total number of rewards collected. The reward ratio experienced by bees that visited only flowers 1–10 was defined as 1, while the realized reward ratio for exclusive visits to flowers 11–20 was defined as 0. We then compared the bees' choice ratios with the realized reward ratios within each treatment using paired t tests. We also computed the t statistic to test whether the mean choice ratio for each treatment differed from 0.5 (indiscriminate choice) and 1 (complete exclusiveness on the higher rewarding food source).

Simulations

We used computer simulations to test the applicability of the Rescorla-Wagner learning rule for reproducing the experimental results. This rule is a variation of Bush and Mosteller's (1951) linear operator mechanism. It describes the changes in associative strength between a conditioned and an unconditioned stimulus (such as a feeder's color and its food reward) as a function of a bee's experience. The rule is described by the function

$$\Delta V_A = \alpha_A \beta (\lambda - V_A). \qquad (1)$$

V_A is the associative strength of the conditioned stimulus A at the start of a learning trial (a foraging visit); ΔV_A is the change in associative strength produced by reward or non-reward; λ is the asymptotic associative strength; β is a learning rate parameter of the unconditioned stimulus (the reward); and α_A is the learning rate parameter of the conditioned stimulus. Both parameters range 0–1. The equation describes an asymptotic acquisition function for stimulus A with a maximal rate (ΔV_A) at $\alpha_A \beta \lambda$. Until the maximum associative strength is reached, there is a difference between the expected reward and the actually experienced reward, which causes updating of the function on each foraging visit.

We used the following procedure in our simulations:
1. The initial strength of association between sucrose reward and the blue and yellow displays was set to 0.5.
2. We randomly determined the first color choice of a bee (blue or yellow).
3. We determined whether the visited flower was rewarding or empty according to reward probability in the chosen color for the simulated condition.
4. We updated the strength of the association between the chosen color and reward using the difference-rule function (Equation 1 above).
5. We set the probability for choosing yellow in the next visit to $V_y/(V_y + V_b)$ and the probability of choosing blue to $V_b/(V_y + V_b)$. V_y and V_b are the updated associative strengths for yellow and blue, respectively.
6. We repeated steps 3–5 for 300 flower visits.
7. We calculated the choice proportions of the higher rewarding flower types for bins of 50 simulated visits.

We defined λ, the asymptotic associative strength, as 1 for rewarded visits and 0 for non-rewarded visits. This reflects the fact that the bees obtained a constant volume of sucrose solution on rewarded visits and no sucrose at all on unrewarded visits. Similarly, Greggers and Menzel (1993) set λ equal to the amount of sucrose obtained by the bees in their experiment. The learning parameter for the conditioned stimuli, α, was taken as 1 because bees learn colors easily and discriminate well between blue and yellow (Fischer et al., 1993; Greggers and Menzel, 1993). The learning parameter for the unconditioned stimuli, β, was taken as 0.02 because this value provided the best fit with empirical results.

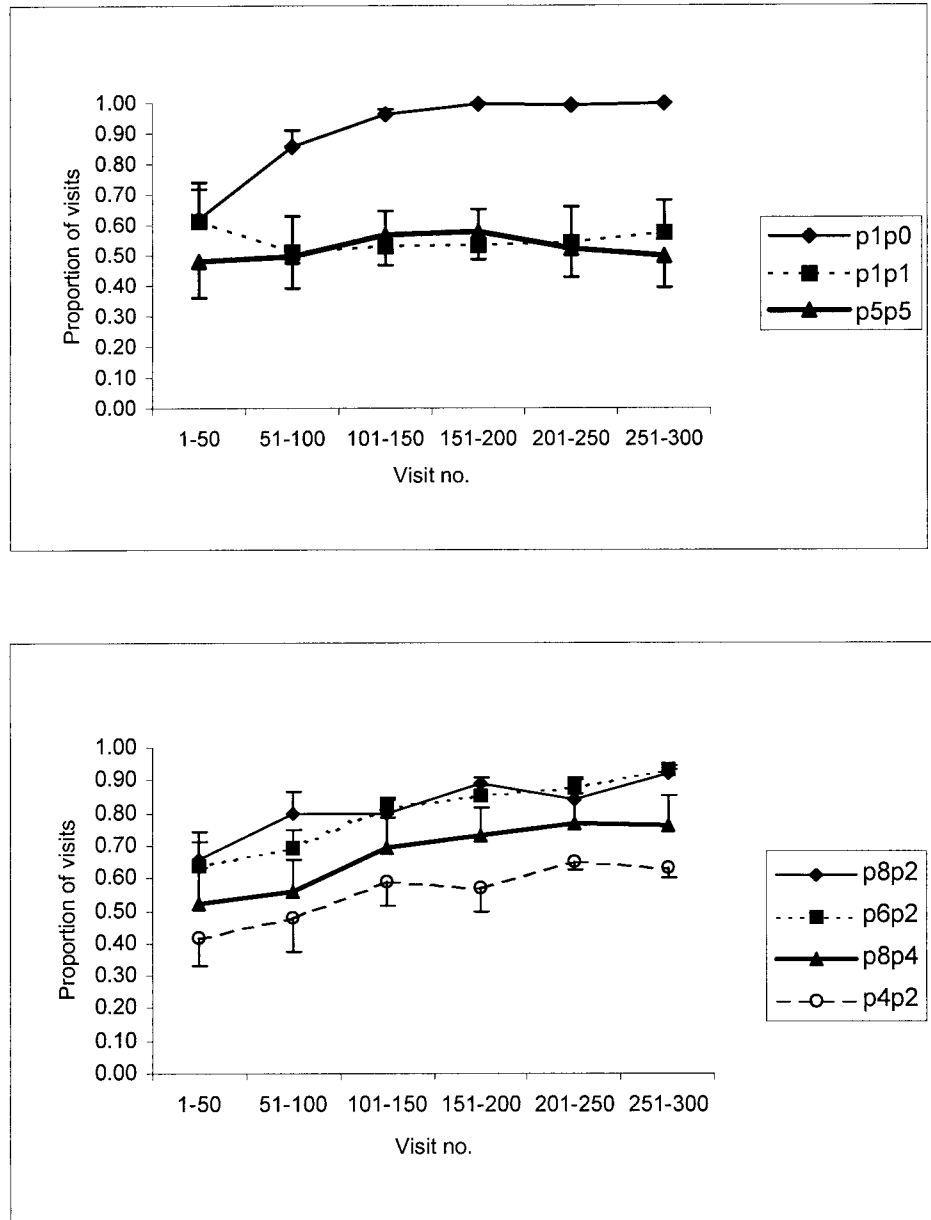We simulated the seven combinations of reward probabili-

Figure 1
Mean choice proportions of flowers 1–10 for the control (top) and test (bottom) experimental conditions. Error bars are 1 SE.

ties used in the experimental treatments. Simulations for each treatment were run 1000 times. We report on the means and standard deviations of these runs.

RESULTS

Initial choices and learning curves

Bees were not able to detect the higher rewarding flowers from a distance, as only 27 out of 53 bees that participated in test conditions chose the higher rewarding flower type on their first foraging visit. Preferences for the higher rewarding flower type developed gradually over the course of the experiment and usually approached stable values by visit 200 (Figure 1). Therefore, we used the bees' visits 201–300 for the analysis of steady-state performance.

Steady-state choice performance

Bees in the p1p0 condition visited the rewarding flower type almost exclusively. Bees that were rewarded equally by both

flower types visited both types, on average, with similar frequencies (treatments p1p1 and p5p5). Mean visit ratios were higher than the realized reward ratios in conditions p8p4 and p8p2, lower than the realized reward ratio in condition p5p5, and did not deviate significantly from the reward ratios in conditions p1p1, p4p2, and p6p2 (Figure 2). Choices in all conditions except p1p0 deviated significantly from exclusiveness on the higher rewarding flower type. Bees within the same condition experienced different reward ratios between the flower types because of the stochastic procedure of flower refilling. Individuals within each condition, except p1p0, also differed significantly from each other in the proportion of visits to flowers 1–10 during visits 201–300 ($G_H = 6.98$, 587.25, 358.57, 83.07, 83.85, 861.76 and 24.57 for treatments p1p0, p1p1, p5p5, p8p2, p6p2, p8p4, p4p2, respectively). For example, two bees in condition p5p5 foraged exclusively on flowers 1–10, one chose flowers 1–10 more often than flowers 11–20, and six showed the reverse choice pattern. We therefore also looked at the choices of individual bees, without pooling bees in the same condition. We plotted visit ratio ver-
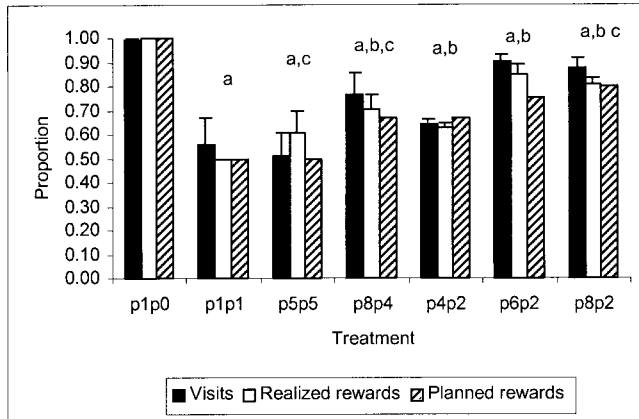
Figure 2
Mean proportions of visits (solid bars) and realized rewards (open bars) in flowers 1–10 during visits 201–300 in the seven experimental conditions. The hatched bars denote the planned reward ratios for the different conditions. Error bars are 1 SE. Mean visit proportions that differ significantly from 1, 0.5, and the realized reward ratio are denoted by a, b, and c, respectively.

Figure 3
The proportion of visits to flowers 1–10 versus the proportion of rewards obtained in these flowers during visits 201–300, plotted separately for each bee ($N = 72$, some data points are overlapping). The best-fit linear regression is described by the function $y = 0.974x + 0.051$, $r^2 = .54$.

sus realized reward ratio for each bee separately, for all experimental conditions combined (Figure 3). The plot is best described by a linear function that is not significantly different from the diagonal (H_0: the slope is not different from 1, df = 62, $t = -0.228$; H_0: the intercept is not different from 0, df = 62, $t = 0.6$). Bees in all experimental conditions visited most of the 20 artificial flowers that were available to them during the experiment. The mean ($\pm$SE) number of different flowers visited was 15.9 $\pm$ 1.5, 14.4 $\pm$ 2.0, 17.6 $\pm$ 1.3, 19.3 $\pm$ 0.2, 17.7 $\pm$ 1.2, 16.4 $\pm$ 1.0, and 20.0 $\pm$ 0.0 for conditions p1p0, p1p1, p5p5, p8p2, p8p4, and p4p2, respectively.

The effects of reward variability, ratios between reward probabilities, and differences between reward probabilities

Treatments p5p5 and p1p1 shared a 1:1 planned reward ratio but differed in reward variability. Bees were rewarded on each flower visit in p1p1, but only in about half of the visits in p5p5. Four out of 10 bees in the p1p1 treatment specialized entirely in one of the flower types, versus only 2 out of 9 in p5p5. However, the variances of the choice proportions did not differ significantly between the two treatments ($F = 1.374$, df = 9,8, $p > .5$).

Statistically significant visit ratios that exceed realized reward ratios (overmatching) occurred both when the probability ratio was 1:4 (treatment p8p2) and when it was 1:2 (treatment p8p4). No overmatching occurred in treatment p4p2, where the reward ratio was 1:2 as well. The difference in reward probability between the two flower types did not account for overmatching either: Overmatching occurred when the difference in reward probability was 0.6 (treatment p8p2) and 0.4 (treatment p8p4), but not in treatment p6p2, where the probability difference was 0.4 as well.

The effects of recent experience on flower-type choice

We calculated the frequency of color shifts after visits to non-rewarding and rewarding flowers at the beginning (visits 1–100) and the end (visits 201–300) of the experiment (Figure 4). The tendency to win-stay-lose-shift was stronger during visits 201–300 than during visits 1–100 and varied markedly among treatments. The variability among treatments was significantly larger during visits 201–300 than during visits 1–100
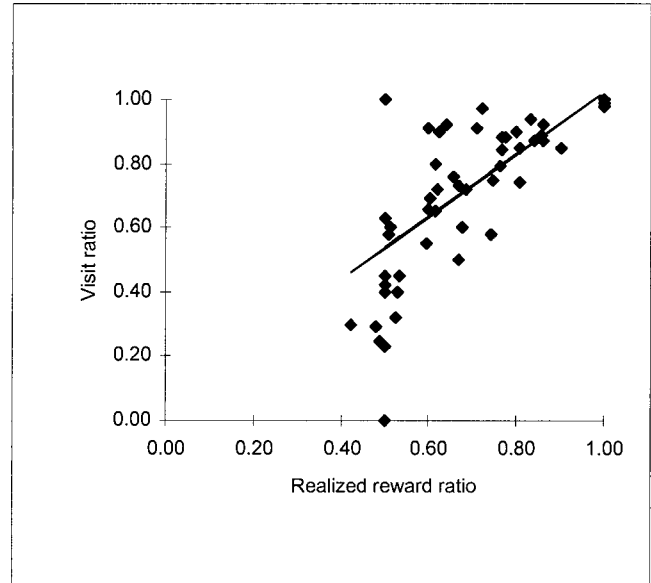
($F = 11.111$, df = 5,5, $p < .01$ for color shifts after non-rewarded visits; $F = 8.704$, df = 5,5, $p < .05$ for color shifts after rewarded visits). The frequency of color-shift flights was reduced after successive rewarded flower visits. Color shifts were more frequent in some of the treatments after a non-rewarded visit than after a rewarded visit. In contrast, color-shift frequencies were not higher after two successive non-rewarded visits than after a single visit to an empty flower (Table 2).

Simulation results

Simulated steady-state choice proportions (with SDs) of the higher-rewarding flower type were 0.98 $\pm$ 0.03, 0.82 $\pm$ 0.10, 0.75 $\pm$ 0.10, 0.67 $\pm$ 0.11 and 0.67 $\pm$ 0.09 for treatments p1p0, p8p2, p6p2, p8p4, and p4p2, respectively. The simulated choice proportions for the treatments with equally rewarding flower types were 0.50 $\pm$ 0.07 for treatment p1p1 and 0.50 $\pm$ 0.09 for treatment p5p5.

DISCUSSION

When presented with two food sources that differed in reward probabilities, bumblebees in our experiment gradually increased their visit frequency to the higher rewarding food source. Choice frequencies for the two flower types approached stability after 150–200 visits, suggesting that our observations (300 visits) were long enough to study steady-state choices. Generally the bees preferred the feeding option with the higher reward probability at steady state, but did not usually choose it exclusively. The ratios of choices between the two feeding options partly matched the ratios of their reward probabilities. The bees visited most of the artificial flowers in the array in all experimental treatments.

Our experiment extends Fischer et al.'s (1993) investigation of choices among probabilistic food sources in honeybees in the following ways. (1) We worked with a larger number of artificial flowers, smaller volumes of sucrose solution, and a
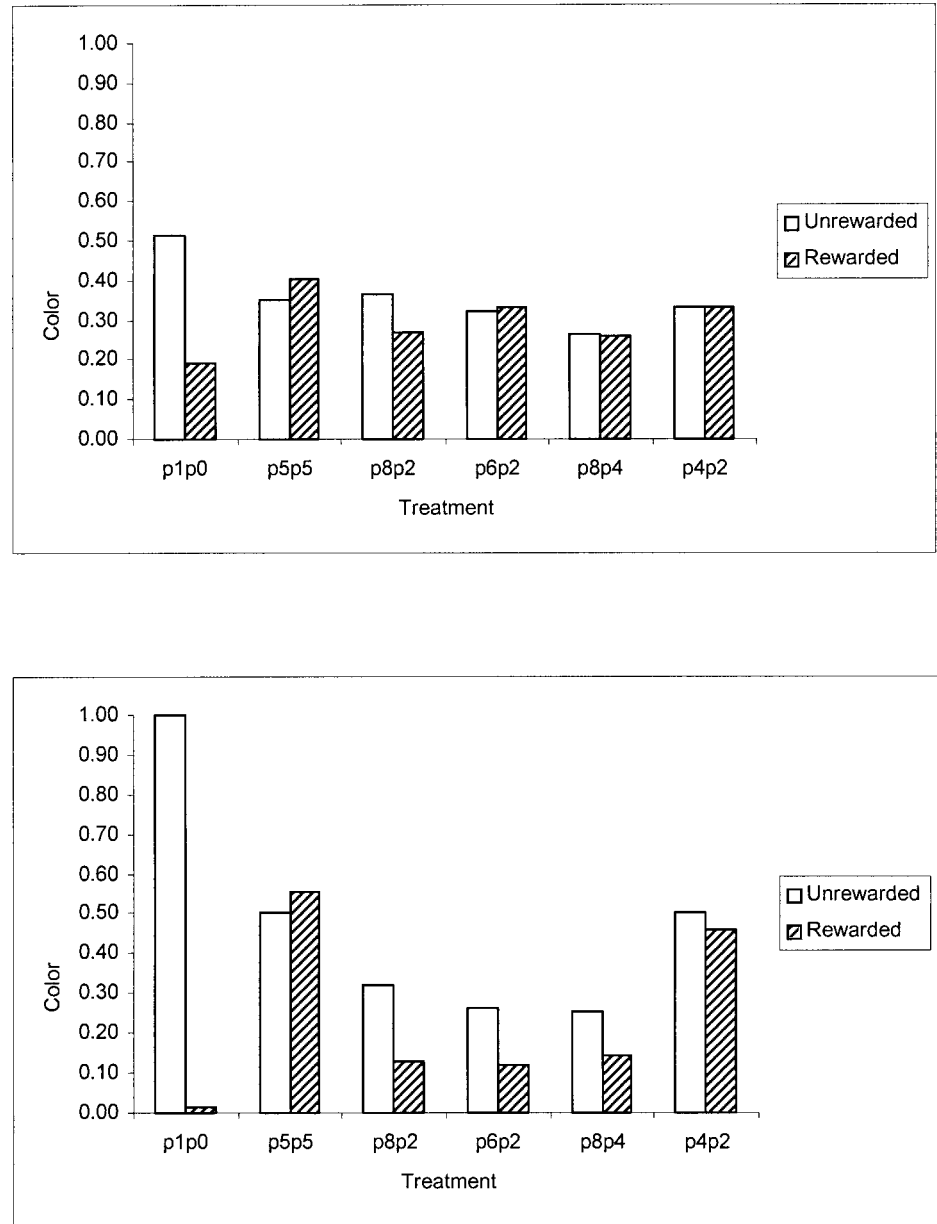
Figure 4
Win-shift and lose-shift frequencies for the seven experimental conditions during visits 1–100 (top) and 201–300 (bottom) of the experiment.

Table 2
Mean frequencies of color shifts after one to four consecutive rewarded and one or two unrewarded flower visits

Treat-ment	Frequency of color shift after n consecutive rewarded visits				Frequency of color shift after n consecutive unrewarded visits	
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 1$	$n = 2$
p5p5	0.47 (155.0)	0.36 (50.2)	0.28 (20.8)		0.46 (141.9)	0.42 (44.7)
p8p2	0.20 (222.8)	0.15 (146.3)	0.11 (104.9)	0.10 (77.7)	0.38 (71.8)	0.34 (74.2)
p6p2	0.24 (155.0)	0.16 (67.2)	0.14 (32.6)		0.31 (142)	0.20 (47.7)
p8p4	0.18 (212.6)	0.13 (135.6)	0.12 (95.7)	0.11 (71.1)	0.25 (84.4)	0.26 (25.1)
p4p2	0.42 (121.3)	0.34 (32.3)			0.43 (175.6)	0.38 (62.5)

Data from visits 1–300 were used for analysis. The number of occurrences of one to four successive rewarded or non-rewarded visits was averaged over all bees in a test condition and is reported in parentheses. Table cells with fewer than 20 mean occurrences are not shown.

larger number of foraging visits per bee. (2) We used a different set of reward probabilities that allowed us to test a prediction arising from an optimal sampling model. Thus, we may try to relate the choice patterns observed in our laboratory experiment to their possible function in sampling of food sources in the bees' natural environment.

The effects of reward variability on the bees' choices

Control conditions p5p5 and p1p1 had similar reward ratios, but the reward schedule of treatment p5p5 was more variable. The existence of reward variability did not affect the bees' choices: Bees in both treatments visited both flower types equally and showed similar individual variability in their choices. This suggests that differences in reward means, rather than differences in variability, probably caused the differences in choice patterns between feeding options in test treatments.

The effects of ratios and differences in reward probabilities on the bees' choices

In agreement with the theoretical prediction (Cohen, 1993), the bees' choices were affected by the difference in mean profitability between the feeding options. The bees specialized more on the higher rewarding flower type when the ratio of mean reward probabilities was 4:1 (0.8:0.2) or 3:1 (0.6:0.2) than when it was 2:1 (0.8:0.4 and 0.4:0.2). This suggests that the extent of specialization on higher rewarding flowers may depend on the ratio of reward probabilities between the two flower types. The bees also specialized more in treatment p8p4 as compared to treatment p4p2. A possible interpretation is that the larger difference in reward probabilities between the flower types (0.4 vs. 0.2) also promoted exclusiveness. In addition, bees may have specialized more in condition p8p4 because the average reward probability of the whole patch was higher than in condition p4p2.

The steady-state proportion of visits to the high-rewarding flower type reached 100% in treatment p1p0. This can be interpreted as no sampling at all at the experienced phase. This finding is also consistent with our working hypothesis because treatment p1p0 had the largest difference in profitability between both flower types. Identifying the better flower type was probably easy for the bees in this treatment. This presumably reduced the benefits of continued sampling of the non-rewarding flowers. Treatment p1p0 also differed from the other test treatments in that it offered constant rather than probabilistic reward (i.e., there was no reward variability). This lack of variability may have made it even easier for the bees to identify the higher rewarding flower type and may have further reduced their benefit from sampling the alternative. However, our discussion of the effects of reward variability (see previous section) renders this interpretation less likely.

In an earlier experiment, we allowed bumblebees to forage on three types of flowers that were always rewarding for 150 visits and then exposed them to an additional, non-rewarding flower type. The bees visited the non-rewarding flowers regularly (Keasar, 2000). This finding is different from the present result. The design of the present experiment is much simpler, however, because it contains only two flower types and no change in conditions throughout the observation session. Possibly, the extreme simplicity of foraging conditions in the p1p0 treatment allowed for exclusiveness on one food source.

Choices between probabilistic rewards and the matching law

Our results superficially resemble the findings of many experiments on matching behavior performed on humans and other vertebrates (Herrnstein, 1970; Heyman, 1979; review by Herrnstein, 1997). These experiments show that the rate of response to a behavioral option is roughly linearly proportional to the option's relative reward rate. Similarly, we show that the choice proportions of individual bees are linearly related to the ratio of rewards they had experienced (Figure 3). However, the linear relationship explains only 54% of the variation in the bees' choices, suggesting that linear matching is an incomplete description of the bees' behavior. Tests of homogeneity indicate a significant variation in choice proportions among bees in most experimental conditions. This suggests that some individuals specialized on the higher-probability option more than expected by linear matching, while others specialized less. Moreover, the average choice ratios of all bees within a treatment deviate significantly from the average realized reward ratio in three out of seven experimental conditions (Figure 2). We therefore cannot conclude that the bees matched their choices to the realized reward probabilities.

Matching experiments commonly assign variable-interval reward schedules (concurrent VI–VI) to the various behavioral options. In simple VI–VI schedules, matching leads to maximization of reward intake rates (Houston, 1983; Houston et al., 1982; Staddon, 1983). Partial matching was also observed in experiments on honeybees that used modified VI–VI schedules (Greggers and Maulhagen, 1997; Greggers and Menzel, 1993). Our results (as well as Fischer et al.'s [1993] findings) show that incomplete exclusiveness also occurs with probabilistic rewards (concurrent VR-VR schedule). Similar results were obtained in sparrows (Gray, 1994). This incomplete exclusiveness clearly does not maximize reward intake rates in our experiment. It is also not predicted by the matching law (Herrnstein, 1982).

We now turn to examine proximate mechanisms that may account for the observed choice behavior.

Possible choice mechanisms

Associative learning—the Rescorla-Wagner learning rule
Simulations of the Rescorla-Wagner difference learning rule predict visit ratios that resemble the expected reward ratios in each experimental condition. As in the experiment, the simulated bees completely specialized on the rewarding flower type in condition p1p0. Their choices were indifferent in simulated conditions p1p1 and p5p5; and they showed incomplete exclusiveness in the remaining simulated conditions. Thus, the simulation reproduced the main trends of the empirical results, in agreement with previous work (Fischer et al., 1993; Greggers and Menzel, 1993). On the other hand, it failed to predict the deviations between visit ratios and expected reward ratios that were observed in the experiment. These deviations were especially large in treatment p6p2. In spite of these deviations, the Rescorla-Wagner model may be considered a successful rough predictor of the bees' choices. A reasonable agreement between model predictions and results does not prove that the model is mechanistically valid. In particular, the memory mechanisms of bees are now known to be much more complex than described by the Rescorla-Wagner rule (Menzel, 1999). We do not claim, therefore, that the Rescorla-Wagner rule provides a mechanistic description of the learning underlying the bees' choices. Rather, we wish to test whether the experimental results are compatible with decision-making rules that may be simpler, such as the ϵ-sampling rule and the failures rule. The simplest model that accurately describes the bees' choices can be used to predict choice behavior in new situations.

Rules of thumb—the ϵ-sampling rule and the failures rule

Table 2 shows that the frequency of flower-type shifts differs among treatments, and decreases with consecutive rewarded visits at the same flower types. Both of these trends are not compatible with basic assumptions of the ϵ-sampling rule, which hypothesizes a fixed probability of switching flower type, regardless of previous rewards. Neither is the failures rule supported by the experimental data, because the tendency to shift flower type is similar after one rewarded visit, one unrewarded visit or two consecutive unrewarded visits (Table 2). Moreover, differences between treatments in the tendency to win-stay-lose-shift developed over the course of the experiment: The frequency of color-shifts in inexperienced bees ranged 0.3–0.4 in most experimental treatments, regardless of the most recent reward obtained (Figure 4a). A different pattern is evident during the last 100 visits of the experiment. Bees were more likely to shift color after visiting an empty flower than after a visit to a rewarding flower in most treatments, and they were more likely to shift in conditions p5p5 and p4p2 than in conditions p8p2, p6p2 and p8p4 (Figure 4b). These experience-related changes suggest that the differences in foraging history between treatments play an important role in the decision-making of experienced bees. For example, the fact that start conditions (all 20 flowers filled) were identical for all treatments (except p1p0) may have decreased the variability between the choices of bees from different treatments during visits 1–100. In other words, simple rules of thumb, such as the ϵ-sampling rule, the failures rule, or other variants that ignore long-term experience, do not suffice to explain the bees' choices in the present experiment.

It is interesting to note that treatment p1p0 differs from other treatments in the frequencies of color shifts as well as in the overall foraging choice pattern. Unlike other treatments, bees in treatment p1p0 used a clear win-stay-lose-shift strategy already during visits 1–100. During visits 201–300, they hardly shifted between flower types at all. This may be a further indication that condition p1p0 was an easy learning task for the bees.

The adaptive value of the bees' foraging choices

The artificial flowers in the experiment refilled instantly and probabilistically. The optimal steady-state foraging strategy would have been, therefore, to visit a single flower (or to shuttle between two neighboring flowers) of the higher rewarding type. The bees, however, visited most of the artificial flowers in the array and foraged on both flower types (except in treatment p1p0). This strategy certainly did not maximize their food intake in the experiment. Is it adaptive under field conditions?

Flowers in nature produce nectar slowly and gradually. Therefore, foragers that revisit the same flowers too frequently will find them empty of reward. This may have selected for foraging on many flowers within each patch, as we observed in our arrays of artificial flowers.

The incomplete exclusiveness on the higher-rewarding feeding option in our experiment resembles the majoring-minoring strategy that bees use in natural foraging situations. Incomplete exclusiveness may be adaptive under field conditions for two reasons: First, it may serve as a sampling strategy, which allows bees to track changes in the quality and availability of food sources. Such changes can result from the depletion of food sources by competitors or from changes in nectar and pollen production by flowers over time. Under such conditions, individuals are expected to increase their food intake rates by sampling a few available feeding options occasionally (Heinrich, 1979; Keasar, 2000). This interpretation is indirectly supported by the agreement between our results and the predictions of an optimal sampling model. Second, complete exclusiveness under natural conditions may incur high travel costs because it requires bees to fly between flowers of one species only and ignore any other species they encounter on their way. Incomplete exclusiveness should be adaptive if travel costs to high-quality, distant food sources are higher than the energy losses incurred by occasionally visiting low-quality food sources that are abundant on the way (Thomson et al., 1987).

Beyond considering the functional significance of incomplete exclusiveness, it is also important to think about the adaptive value of the decision rules underlying it. Different decision rules can lead to incomplete exclusiveness, but foraging bees use only a subset of these rules. Bees have presumably evolved to use those choice rules that are most adaptive and not to use others. We can therefore use our analysis of possible choice mechanisms to speculate on their adaptive significance.

We cannot determine exactly what choice rules the bees used in our experiments. Nevertheless, two main patterns emerge from our examination of possible rules. First, bees do not base their choices only on their very last foraging visits. This implies that it is advantageous for them to incorporate more distant past experience into decision making. This could help bees detect trends or regularities in their foraging environment (e.g., gradual depletion of a food source, gradual change in weather conditions), possibly improving their ability to respond to such changes. The ϵ-sampling rule, the failures rule, and other possible rules of thumb that we did not examine assume a selective advantage to choice rules that require a short recall and simple computation. This is because bees (and other small invertebrates) may be constrained by their data processing capabilities (Real, 1991; Thuijsman et al., 1995). Our data suggest, however, that bees are nevertheless selected to rely on more distant experience, in spite of possible neural limitations. We do not know the exact form of the bees' choice rule or how complex it is to implement. This may be a case where an optimal decision rule is "difficult for the biologist to determine, but simple enough for an animal to use" (Green, 1987: 287).

Second, bees use choice rules that lead to incomplete exclusiveness both under field conditions and in laboratory two-armed bandit situations. This suggests that they were not selected to use different choice rules for different foraging situations (as, for example, humans would do). Possibly, most foraging situations that bees encounter in nature favor incomplete exclusiveness, and may not select for a wide repertoire of decision rules.

Directions for further study

Our experiment reproduces some of the features of natural foraging situations that involve sampling, but the generality of the results is limited in two important aspects. First, the bees were only offered two flower types, so that leaving one of them automatically meant choosing the other. This limitation has been addressed in sampling experiments that include more than two food sources (Keasar, 2000, unpublished data). Second, only one parameter, reward probability, was varied in the experiment. Different combinations of environmental stimuli, such as the presence of landmarks (Greggers and Maulshagen, 1997) and bee genotype (Page et al., 1998) should be included in further studies.

REFERENCES

Auer A, Cesa-Bianchi N, Freund Y, Shapire RE, 1995. Gambling in a rigged casino: the adversarial multi-armed bandit problem. IEEE Proceedings, Foundations of Computer Science 1995. Los Alamitos, California: IEEE Computer Society Press; 322–331.

Bush RR, Mosteller F, 1951. A mathematical model for simple learning. Psych Rev 58:313–323.

Cohen D, 1993. The equilibrium distribution of optimal search and sampling effort of foraging animals in patchy environments. Lecture Notes Biomath 98:8–25.

Estabrook GF, Jespersen DC, 1974. Strategy for a predator encountering a model-mimic system. Am Nat 108:443–457.

Fischer ME, Couvillon PA, Bitterman ME, 1993. Choice in honeybees as a function of the probability of reward. Anim Learn Behav 21:187–195.

Giurfa M, 1993. The repellent scent-mark of the honeybee *Apis mellifera* and its role as communication cue during foraging. Insect Soc 40:59–78.

Goulson D, Hawson SA, Stout JC, 1998. Foraging bumblebees avoid flowers already visited by conspecifics or by other bumblebee species. Anim Behav 55:199–206.

Gray RD, 1994. Sparrows, matching and the ideal free distribution—can biological and psychological approaches be synthesized? Anim Behav 48:411–423.

Green RF, 1987. Stochastic models of optimal foraging. In: Foraging behavior (Kamil AC, Krebs JR, Pulliam HR, eds). New York: Plenum Press; 273–302.

Greggers U, Maulshagen J, 1997. Matching behavior of honeybees in a multiple-choice situation: the differential effect of environmental stimuli on the choice process. Anim Learn Behav 25:458–472.

Greggers U, Menzel R, 1993. Memory dynamics and foraging strategies of honeybees. Behav Ecol Sociobiol 32:17–29.

Heinrich B, 1979. ''Majoring'' and ''minoring'' by foraging bumblebees, *Bombus vagans*: an experimental analysis. Ecology 60:245–255.

Herrnstein RJ, 1970. On the law of effect. J Exp Anal Behav 13:243–266.

Herrnstein RJ, 1982. Melioration as behavioral dynamism. In: Quantitative analyses of behavior, vol. 2: matching and maximizing accounts (Commons ML, Herrnstein RJ, Rachlin H, eds). Cambridge, Massachusetts: Ballinger; 433–458.

Herrnstein RJ, 1997. The matching law (Rachlin H, Laibson DI, eds). Cambridge, Massachusetts: Harvard University Press.

Heyman GM, 1979. A Markov model description of changeover probabilities on concurrent variable-interval schedules. J Exp Anal Behav 31:41–51.

Houston AI, 1983. Optimality theory and matching. Behav Anal Lett 3:1–15.

Houston AI, Kacelnik A, McNamara JM, 1982. Some learning rules for acquiring information. In: Functional ontogeny (McFarland DJ, ed). Pitman; 140–191.

Keasar T, 2000. The spatial distribution of non-rewarding artificial flowers affects pollinator attraction. Anim Behav 60:639–646.

Keasar T, Motro U, Shur Y, Shmida A, 1996. Overnight memory retention of foraging bumblebees in imperfect. Anim Behav 52:95–104.

McNamara JM, Houston AI, 1985a. Optimal foraging and learning. J Theor Biol 117:231–249.

McNamara JM, Houston AI, 1985b. A simple model of information use in the exploitation of patchily distributed food. Anim Behav 33:553–560.

Menzel R, 1999. Memory dynamics in the honeybee. J Comp Physiol A 185:323–340.

Milano N, 1994. Simple decision rules for foraging bees (MSc thesis). Jerusalem: The Hebrew University.

Ney-Nifle M, Keasar T, Shmida A, 2001. Location and color learning in bumblebees in a two-phase conditioning experiment. J Insect Behav 14:697–711.

Page RE, Erber J, Fondrk MK, 1998. The effect of genotype on response thresholds to sucrose and foraging behavior of honey bees (*Apis mellifera* L.). J Comp Physiol A 182:489–500.

Real LA, 1991. Animal choice behavior and the evolution of cognitive architecture. Science 253:980–986.

Real LA, 1992. Information processing and evolutionary ecology of cognitive architecture. Am Nat 140:S108–S145.

Rescorla RA, Wagner AR, 1972. A theory of classical conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In: Classical conditioning II: Current research and theory (Black AH, Prokasy WF, eds). New York: Appleton-Century-Crofts; 64–99.

Shettleworth SJ, Krebs JR, Stephens DW, Gibbon J, 1988. Tracking a fluctuating environment: a study of sampling. Anim Behav 36:87–105.

Sokal RR, Rohlf FJ, 1981. Biometry. New York: Freeman.

Staddon JER, 1983. Adaptive behaviour and learning. Cambridge: Cambridge University Press.

Stephens D, 1987. On economically tracking a variable environment. Theor Popul Biol 32:15–25.

Stephens DW, Krebs JR, 1986. Foraging theory. Princeton, New Jersey: Princeton University Press.

Thomson, JD, Peterson SC, Harder LD, 1987. Response of traplining bumble bees to competition experiments: shifts in feeding location and efficiency. Oecologia 71:295–300.

Thuijsman F, Peleg B, Amitai M, Shmida A, 1995. Automata, matching and foraging behavior of bees. J Theor Biol 175:305–316.