QUANTITATIVE TESTS OF AN ASSOCIATIVE THEORY OF RISK-SENSITIVITY IN HONEYBEES

MARTIN S. SHAPIRO[1], P. A. COUVILLON[2] AND M. E. BITTERMAN[2,*]

[1]*Department of Mathematics and Science, Leeward Community College, 96-045 Ala Ike, Pearl City, HI 96782, USA*
and [2]*Békésy Laboratory of Neurobiology, University of Hawaii, 1993 East-West Road, Honolulu, HI 96822, USA*

*Author for correspondence (e-mail: jeffb@pbrc.hawaii.edu)

Summary

Risk-sensitivity was studied in free-flying honeybees trained individually to choose between two scented targets (A and B) with varying amounts and concentrations of sucrose solution as reward. In the first phase of experiment 1, the animals showed 'risk-aversion,' preferring A, which provided 5 µl of a 40 % sucrose solution on every trial, to B, which provided 30 µl of the same solution once in every six trials (mean amount per trial 5 µl for each alternative). In the second phase, the preference reversed with reversal of the reward assignments. In experiment 2, the consistently rewarded A (5 µl of 40 % sucrose solution per trial) was again preferred, although the inconsistently rewarded B now provided twice the amount of sucrose solution on average (30 µl on two of every six trials, mean amount per trial 10 µl). In experiment 3, with A providing 10 µl of a 15 % sucrose solution on every trial and B providing 10 µl of a 60 % sucrose solution on two of every four trials (mean concentration per trial 30 %), the animals preferred B. In Experiment 4, patterned after experiment 1, similar results were obtained under more natural conditions in which the animals were no longer constrained (as they were in the first three experiments) to go equally often to each alternative. The results of all four experiments were predicted quantitatively and with considerable accuracy by a simple associative theory of discriminative learning in honeybees.

Key words: honeybee, *Apis mellifera*, risk-sensitivity, behaviour, associative learning.

Introduction

What should a foraging animal do when offered a choice between two feeding places providing the same mean amount of food with different variance (say, 1 unit of food on every visit *versus* 3 units on a random third of the visits and nothing on the rest)? The influential energy budget theory of Caraco et al. (Caraco et al., 1980) suggests that if the animal cannot survive on the mean amount of food in prospect, which it is assumed somehow to 'know', it should be risk-prone (that is, prefer the more-variable alternative), while it should be risk-averse (that is, prefer the less-variable alternative) if the prospective mean amount is sufficient for survival. In fact, however, experiments on the amount and quality of reward with a variety of species in a variety of energy states have yielded only marginal evidence that preference changes with energy needs, and even less convincing evidence of risk-proneness under any conditions (Kacelnik and Bateson, 1996). There is, nevertheless, some good evidence for risk-aversion that remains to be understood, and those inclined to 'functional' interpretations of behavior must now consider why it might be advantageous for an animal to prefer consistency. McNamara (McNamara, 1996) does just that.

Perhaps the simplest way to begin to account for risk-aversion in functional terms is on the assumption that the value of a portion of food as defined in terms of net energy gain is nonlinearly related to its amount and quality. Finding bumblebees averse to variability in the amount of a sucrose solution, Harder and Real (Harder and Real, 1987) suggested, for example, that the time and energy expended in probing for sucrose solution may increase disproportionately with amount. If net energy gain is, say, a simple positive growth function of amount, the mean value of two different amounts of food may be substantially less than the value of the mean amount. It has also been suggested in the light of Weber's Law that functions of much the same shape would be expected if rewards were evaluated on the basis of their perceptual properties (Hamm and Shettleworth, 1987); a 20 µl drop of sucrose solution may be perceived by honeybees as less than twice as large as a 10 µl drop, or a 40 % solution as less than twice as sweet as a 20 % solution. The two conceptions are not necessarily incompatible, sharing the implication that high priority should be given in this work to the construction of reward-value functions (Perez and Waddington, 1996).

For a thorough understanding of risk-sensitive behavior, it is necessary in any case to go beyond considerations of adaptive advantage to ask how information about available alternatives is acquired and how the information is translated

into performance. It is easy to agree with Kacelnik and Bateson (Kacelnik and Bateson, 1996) that without answers to such questions, which they classify as 'causal' or 'mechanistic' rather than functional, it is 'virtually impossible' to make exact, falsifiable predictions of behavior (p. 430). An initial step in the direction of causal analysis has been taken by Reboreda and Kacelnik (Reboreda and Kacelnik, 1991), who propose an explanation of risk-aversion for amount of reward in terms of errors of memory. Repeated experience with a given amount is assumed to generate a normal distribution of memories about the true mean whose standard deviation increases with amount in accordance with Weber's Law. In making a choice, an animal is said to sample the distribution of remembered amounts for each alternative, which for the variable alternative is the sum of the distributions for the different amounts and, thus, more likely to yield a memory of an amount smaller than the constant amount. The Reboreda–Kacelnik model has not yet been developed to the point at which it provides exact predictions of performance, although a beginning has been made (Kacelnik and Brito e Abreu, 1998). There is, however, another model, a simple associative model, that already yields such predictions. Developed by Couvillon and Bitterman (Couvillon and Bitterman, 1991) in the course of more traditional work on discriminative learning in honeybees, the model has made it possible for Shapiro (Shapiro, 2000) to simulate, quantitatively and with substantial accuracy, the results of an extensive series of experiments involving variation in both amount and concentration of sucrose solution. The honeybees showed risk-indifference under some conditions and risk-aversion under others.

The model

The first assumption of the Couvillon–Bitterman model is that the attractiveness of a stimulus is given by the strength of its association with the reward, which changes on each rewarded or nonrewarded encounter (trial). The change is described by the linear equation of Bush and Mosteller (Bush and Mosteller, 1951) in the now more familiar notation of Rescorla and Wagner (Rescorla and Wagner, 1972):

$$\Delta V_A = \alpha_A \beta (\lambda - V_A), \tag{1}$$

where V_A is the associative strength of stimulus A at the beginning of a trial, and ΔV_A is the change in V_A on that trial. λ (scaled from 0 to 1) is the maximal associative strength attainable with a given reward and serves as a measure of reward-value; it is taken as 0 when there is no reward. β (scaled from 0 to 1) is the rate of learning (see below) and α_A (scaled from 0 to 1) is the salience (or 'attention-value') of A (see below).

The direction of change is determined by the difference between λ and V_A. When V_A is less than λ, ΔV_A is positive (associative strength increases); when V_A is greater than λ, ΔV_A is negative (associative strength decreases); it may be well to note that associative strength may decline even on a rewarded trial, as when a stimulus is paired with a reward of

low value after repeated pairing with a reward of higher value. Because there has been some confusion in the literature on this point (e.g. Real, 1996), it should be emphasized that the Rescorla–Wagner notation is used only because of its familiarity. The data on compound conditioning in honeybees do not seem to require the assumption of shared associative strength that is the distinctive feature of the Rescorla–Wagner theory (Couvillon et al., 1996; Couvillon et al., 1997); if that assumption were intended, the summed associative strengths of all stimuli present on the trial, rather than V_A alone, would be shown to be subtracted from λ.

The rate of change in V_A is determined by two parameters: α_A (scaled from 0 to 1) is the salience (or 'attention-value') of A, which is taken as 1 for the salient stimuli usually used in the experiments and can, therefore, be ignored. β (scaled from 0 to 1) is the rate of learning, assumed to be characteristic of the species or the behavioral system under study, which may be different on incremental trials, when $\lambda - V_A$ is positive, than on decremental trials, when $\lambda - V_A$ is negative; the incremental rate is written as $U\beta$ and the decremental rate as $D\beta$.

Another important feature of the model is a performance function, which reflects the assumption that choice between two stimuli called, for example, A and B, is determined by their relative associative strength, r. (Without a performance function, only ordinal predictions of experimental outcomes are possible.) The relative strength of A, r_A, is computed as $V_A/(V_A+V_B)$; the relative strength of B, r_B, is $V_B/(V_A+V_B)$, or $1-r_A$. P_A, the probability of choosing A, is assumed to be a power function of r_A, designated by the parameters $K=0.75$ and $s=0.625$, which is plotted for $0.5 \leqslant r \leqslant 1$ in Fig. 1. This function, generated with the scaling equations:

$$P = 0.5 + s(2r - 1)^K \tag{2}$$

for $r \geqslant 0.5$ and

$$P = 0.5 + s(1 - 2r)^K \tag{3}$$

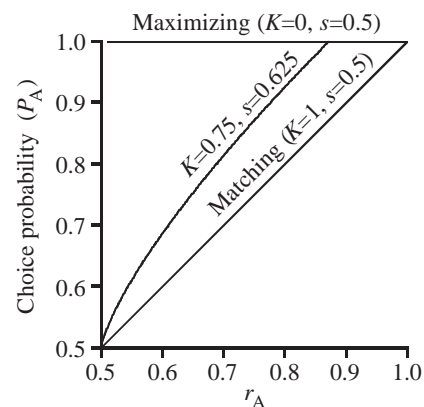for $r \leqslant 0.5$, was selected (Couvillon and Bitterman, 1991) from

Fig. 1. Functions relating predicted probability of choice to relative associative strength (r_A), maximizing ($K=0$, $s=0.5$), matching ($K=1$, $s=0.5$), and the function ($K=0.75$, $s=0.625$) that yielded the best fit to all previous data. K and s are parameters of the model (see equations 2 and 3).

an array of such functions varying in slope and curvature between the matching (K=1, s=0.5) and maximizing (K=0, s=0.5) functions, because it provided the best fit to available data on discriminative learning in honeybees. The matching and maximizing assumptions commonly entertained in the vertebrate literature, but which do not permit such good fits to the honeybee data, are plotted in Fig. 1 for purposes of comparison. (As the plots indicate, the matching assumption is that the probability of choice is given directly by the ratio of associative strengths, and the maximizing assumption is that the alternative of detectably greater strength is always chosen.)

Performance in Shapiro's risk-sensitivity experiments (Shapiro, 2000) could have been predicted (whether correctly or not) in the literal sense of that term if he had been able at the outset to assign a value to λ for each of the various amounts (5, 20 or 30 µl) and concentrations (15, 20 or 40 %) of sucrose solution to be used as reward, but the necessary information did not exist. In all the experiments modeled previously by Couvillon and Bitterman (Couvillon and Bitterman, 1991), the same large reward, feeding to repletion on 50 % sucrose solution (approximately 50 µl), was used, its λ-value being taken simply as 1. Shapiro's inductive alternative was to search factorially for a single set of λ-values with which his data could be accurately simulated. At the same time, $U\beta$ and $D\beta$ were

also varied factorially, because it seemed unsafe to take for granted the learning rates that were adequate to account for the older data but might have been insufficiently constrained by those data. Using the performance function plotted in Fig. 1 (K=0.75, s=0.625), which there was no reason to question, Shapiro found rather good fits to all his data with $U\beta$=0.04, $D\beta$=0.02 and the λ-values that are plotted in Fig. 2. The λ-functions are clearly nonlinear. One of the curves plotted in Fig. 2A shows the hypothetical reward value of feeding to repletion ('fill') for sucrose solutions of different concentration, and a second curve shows the value of 10 µl for solutions of different concentration; a third curve (Fig. 2B) shows the reward value of each of several different amounts for a 40 % solution. The simulations captured very well not only the terminal preference in each of Shapiro's 10 experiments but also the course of its development with increasing experience of the alternatives. The goodness-of-fit of the theory to the data was expressed in terms of the root-mean-square (RMS) deviation of simulated from measured probabilities of choice on each trial in the entire set of experiments (RMS deviation=0.047).

If Shapiro had been unable to find a set of parameter values that yielded a good simulation of his data, the validity of the Couvillon–Bitterman model would clearly have been called into question. That he was able to simulate his results with parameter values chosen retrospectively may not seem very impressive until it is appreciated that the set of data is quite extensive (there were 10 experiments in all) and that the number of parameters is relatively small (Uβ, Dβ and λ for the learning function; K and s for the performance function). The reward-value functions are themselves products of the modeling, of course, and it may be well to note that even if such functions could be generated independently, for example in an elaborate set of metabolic or psychophysical experiments, they would provide no more than an ordinal account of the terminal preferences found by Shapiro, with no hint as to absolute magnitudes or the actual behavior of the animals. In any case, the obvious next step in the evaluation of the model was to make some exact new predictions and to test them experimentally, which we have begun here to do in four experiments that are also designed to produce some useful new information on risk-sensitive foraging in honeybees.

New experiments

As McNamara (McNamara, 1996) has recently reminded us, animals are commonly called upon in nature to adjust to fluctuations in available resources, and our purpose in experiment 1 was to examine the flexibility of risk-sensitive behavior in the face of such changes as well as to test the ability of the model to predict the new data. The plan was first to replicate some training with variability in amount of reward in which Shapiro found risk-aversion, preference for a consistent alternative (A) providing 5 µl of a 40 % sucrose solution on every trial to a variable alternative (B) providing 30 µl of the same solution only on every sixth trial (mean amount 5 µl), and then to reverse the treatments of the two alternatives. The
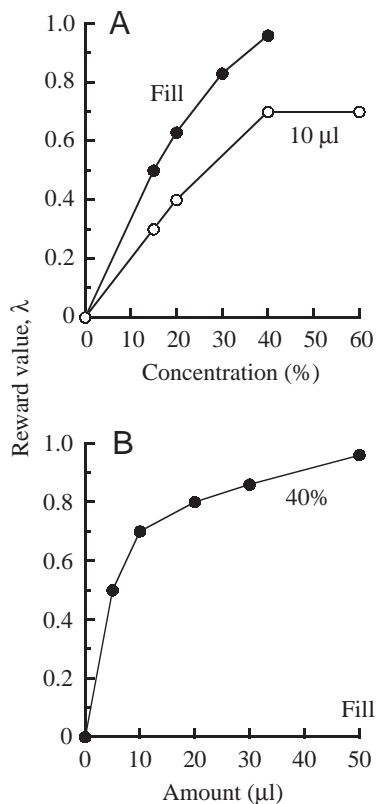

Fig. 2. Reward-value (λ) functions derived in fitting Shapiro's data (Shapiro, 2000). (A) Fill, ●, feeding to repletion; ○, feeding with 10 µl of sucrose solution of increasing concentration. (B) Reward values when fed with increasing amounts of 40 % sucrose solution to repletion at 50 µl.

model predicts not only that the preference established in the initial training will be reversed but also the exact course of reversal.

The convenient strategy commonly employed in work on risk-sensitivity is to offer a choice of alternative feeding places that provide rewards equal in mean value but different in variance. As Stephens (Stephens, 1981) has stressed, however, equality of mean value is extremely unlikely in nature, and in experiments 2 and 3 we studied the behavior of honeybees under conditions in which the alternatives differed in mean value as well as in variance. In experiment 2, A was again rewarded with 5 µl of 40 % sucrose solution on every trial, but now B provided 30 µl of the same solution twice, rather than only once, in every six trials (mean amount 10 µl). The model predicts that the animals will prefer A despite the fact that it yields only half as much sucrose on average as does B. In experiment 3, patterned after a concentration experiment in which Shapiro found risk-aversion with equality of means, A providing 10 µl of a 15 % sucrose solution on every trial and B providing 10 µl of a 60 % sucrose solution once in every four trials (mean concentration 15 %), the treatment of A was the same, but B now provided 10 µl of the 60 % solution twice in every four trials (mean concentration 30 %). Here, in contrast with experiment 2, the model predicts a preference for the variable alternative. Again, in each experiment, the prediction is not only of the terminal preference but also of the exact course of its development.

In experiments 1–3, as in Shapiro's experiments, the training procedure was designed to ensure equal experience with the two targets. The initial choice of the animal in each trial provided a measure of preference, but the trial did not end until the animal had also responded to the other target and taken whatever reward it provided. In experiment 4, which was patterned after experiment 1, the equal-frequency constraint was abandoned, again in the interest of greater compatibility with natural conditions; each trial ended with the first choice made by the animal, which then had no opportunity to explore the alternative presented at that trial. Our interest was in how the animals would adjust in these rather different circumstances and in whether the model was correct in predicting not only that they would continue to be risk-averse but also the exact course of change in their behavior.

Materials and methods
Animals

The subjects were 48 foraging honeybees (*Apis mellifera*), all experimentally naive, from hives situated near the laboratory. They were studied individually, each in a single session 3–4 h long. The 48 animals were assigned at random to three groups of eight subjects, one each in experiments 1–3, and two groups of 12 subjects in experiment 4. Somewhat larger groups were used in experiment 4 (in which the equal-experience constraint was removed) in the interest of obtaining a more reliable sample of performance under the new training conditions.
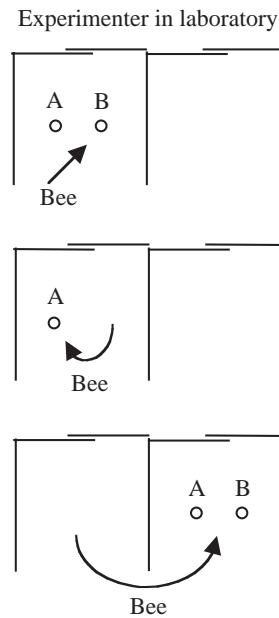

Fig. 3. Diagram of the training situation and the training procedure. In the top panel, the animal chose target B when both targets were presented in the left-hand window. Target B was removed after the animal had consumed the reward it contained, and the trial did not end until the animal had gone to A (middle panel). Target A then was removed, and the animal shuttled to the right-hand window for another choice (bottom panel), and so forth.

Training situation

The training situation, which is shown in Fig. 3, was the same as that used by Shapiro (Shapiro, 2000). It consisted of a pair of immediately adjacent windows (each 58 cm wide, 57 cm high and 55 cm deep) separated by a thin wooden partition around which the animals were required to fly as they shuttled back and forth between the windows on successive trials of each training visit. The training stimuli (called targets) were covered Petri dishes of gray plastic, 5.5 cm in diameter. Drilled in the cover of each dish, 6 mm from its outer circumference, was a circle of eight equally spaced holes, 5 mm in diameter. The dishes contained flattened cotton balls that could be impregnated with the odor of peppermint, geraniol or both. Three sets of targets were used, a peppermint set, a geraniol set and a pretraining set scented with both odors. The covers of the targets used on each visit were washed and exchanged for others in their sets after each visit to randomize extraneous stimuli.

Pretraining procedure

A single animal was selected at random from a group of foragers at a feeding station providing 10–12 % sucrose solution, picked up in a matchbox, carried to the laboratory and released at a large (>100 µl) drop of 40 % sucrose solution on a pretraining target (labeled with the scents both of peppermint and geraniol) centered on the sill of one of the windows (the left-hand window for half the subjects and the right-hand window for the rest). The animal was marked with

a spot of colored lacquer as it fed to repletion, after which it was permitted to leave for the hive, where it deposited the sucrose. Typically, the animal (adapted to 40 % sucrose and now finding the lower concentration at the feeder less acceptable) returned to the training situation after a few minutes, continuing to fly back and forth between the hive and the laboratory as long as sucrose was available there. If the marked animal did not return after its first placement, it was captured again at the feeding station, where it could usually be found, and placed again on the pretraining target. When the animal first returned of its own accord to the pretraining target in the first window, the target, with the animal feeding on it, was picked up and carried to the other window, where the animal continued to feed. On the second return, the pretraining target was again in the alternative window, where again there was feeding to repletion. (After the first return, there was no further handling of the animal.) The pretraining of every animal ended after it had returned twice to each window of its own accord. The experience with both windows facilitated the shuttling between them that was required in the training, which began when the subject made its fifth return to the laboratory.

Training procedure

Experiment 1

On each training visit, the arriving animal found a pair of scented targets (henceforth referred to as A and B), one labeled with peppermint and the other labeled with geraniol, set 10 cm apart in a lateral arrangement on the sill of one of the windows (Fig. 3). Target A always contained a 5 µl drop of 40 % sucrose solution, while B contained a 30 µl drop of 40 % sucrose solution once in every six trials (in quasi-random order). On the other trials, the target contained only water (unacceptable to the animals and distinguishable from the sucrose solution only by taste). The subject chose one of the targets, ingesting the sucrose solution or merely tasting the water (its initial choice was recorded), and then went to the second target (as the first was removed), again ingesting the sucrose solution or merely tasting the water. In the meantime, fresh A and B targets were placed in the adjacent window, to which the subject then shuttled for another trial. This procedure was continued until the subject was replete and returned to its hive. With the crop of the honeybee holding 50 µl on average, the number of choices made on each visit averaged approximately four. On subsequent visits, the subject found either a single target or a pair of targets, depending on the point in the previous visit at which it had left of its own accord for the hive. In this first stage of the training, 24 choice trials were scheduled, which meant that the subject had at least that many experiences with each target. For half the bees, target A was labeled with peppermint and target B with geraniol; for the other half, the odors were reversed. The position of each odor in the choice pair (left or right) was changed in quasi-random order between trials, with the arrangements balanced for each window over the course of training. (The same balancing

procedures were used in all the experiments reported here.) In the second stage of the training, there were 48 choice trials during which the treatments of A and B were reversed; that is, A now became the variable alternative and B the consistent one.

On the visit following the last training visit for each animal, there was a nonrewarded 10 min test (called an extinction test) with a pair of fresh A and B targets, both now containing water. All contacts with each target in successive 30 s intervals were recorded. For half the bees, A was on the left and B was on the right, with the arrangement reversed for the other half. The test took place in either the left-hand or right-hand window (balanced over subjects). During the test, the adjacent window was open and empty.

Experiment 2

The training procedure was the same as in experiment 1, but here A consistently provided 5 µl of 40 % sucrose solution, while B provided 30 µl of the same solution twice in every six trials (in quasi-random order) and water on the rest (mean amount of the sucrose solution per trial 10 µl). The training of each animal continued until it had had 40 encounters with each of the alternatives. On the visit following the last training visit, there was, as in experiment 1, a 10 min extinction test with targets A and B.

Experiment 3

Here, with the same training procedure as in the preceding experiments, variability in sucrose concentration was studied. While choice of target A was consistently rewarded with 10 µl of a 15 % sucrose solution, choice of target B was rewarded with 10 µl of a 60 % sucrose solution twice in every four trials, with water on the rest (mean sucrose concentration per trial 30 %). As in experiment 2, the training of each animal continued until it had had 40 encounters with each alternative, and on the following visit there was the standard 10 min extinction test with the two targets.

Experiment 4

In the first stage of training, two groups of animals were trained exactly as in the first stage of experiment 1, target A always containing a 5 µl drop of 40 % sucrose solution, and target B containing a 30 µl drop of 40 % sucrose solution once in every six trials (in quasi-random order), with water on the rest. As before, both targets were experienced on each of 24 trials. In the 48 trials of the second stage, the second step in the training procedure as shown in Fig. 3 was omitted. On each trial, the animal experienced only the target chosen initially, the alternative target being removed by the experimenter as the animal sampled the first. For a nonreversal group, the treatments of A and B remained the same as in the first stage, but the treatments were interchanged for a reversal group, target A now becoming the variable alternative and target B the consistent one. On the visit following the last training visit, there was the usual 10 min extinction test with the two targets.

Results

Experiment 1

In Fig. 4, the performance of the animals in the choice training is plotted in terms of the mean proportion of choices of target A (the consistent alternative, providing 5 µl of 40 % sucrose solution on every trial) rather than target B (the variable alternative, providing 30 µl of the same solution on every sixth trial) in successive four-trial blocks. As the curve shows, a preference for target A developed quickly in the first stage of training; the overall proportion of choices of A in that stage was significantly greater than chance, $t(7)=9.01$, $P<0.05$, two-tailed test. (The α-level used throughout is 0.05). In the second stage, with the contingencies of reward reversed, the preference for target A quickly gave way to a preference for target B (now the consistent alternative); the overall proportion of choices of B in the second stage was significantly greater than chance, $t(7)=7.46$. In both stages, that is, the animals showed risk-aversion, a preference for the consistent alternative.

In Fig. 5, the performance of the animals in the extinction test is plotted in terms of the mean cumulative number of responses to each target in successive 30 s intervals of the 10 min test period. These curves, too, show a clear preference for target B, confirming the preference for B demonstrated by the measure of choice at the conclusion of the training. Analysis of variance based on (uncumulated) frequencies of response in four 2.5 min blocks yields a significant stimulus (A *versus* B) effect ($F_{1,7}=245.29$) and a significant stimulus × block interaction ($F_{3,21}=18.81$).

Plotted alongside the actual choice data in Fig. 4 is the prediction from the Couvillon–Bitterman model based on Shapiro's best-fitting parameter values. The prediction

Fig. 5. Extinction test performance in experiment 1 plotted in terms of the mean cumulative number of responses to each target in successive 30 s intervals for targets A and B.

captures the general pattern of the results, but substantially underestimates the speed of reversal. Also plotted in Fig. 4 is a simulation with the same parameter values used by Shapiro except for somewhat larger β-values: $U\beta=0.08$ instead of 0.04, and $D\beta=0.06$ instead of $D\beta=0.02$. The new β-values, which yield a better fit to the data of experiment 1 alone (RMS deviation 0.056 compared with 0.13), do not appreciably change the overall fit of the model to the entire set of Shapiro's data (RMS deviation 0.054 compared with the original 0.047). The new β-values do, however, yield a poorer fit (RMS deviation 0.11 compared with 0.076) to the data of Shapiro's Amt1 experiment, in which the animals preferred an alternative providing 5 µl of 40 % sucrose solution on each trial to one providing 20 µl on every fourth trial. Which of the two experiments provides better estimates of the learning rates? On the grounds that there were substantially fewer training trials in the earlier experiment and that the data of the earlier experiment are more variable (see Fig. 7 in Shapiro, 2000), our decision was to rely on the new data. Accordingly, the new β-values were accepted and were used in predicting the results of the subsequent experiments 2–4. Perhaps it should be emphasized that $U\beta$ and $D\beta$ are not treated as free parameters that are permitted to vary in value capriciously from experiment to experiment. Our estimates of their values, based on the entire set of available data, are simply refined as new data accumulate.

Experiment 2

In Fig. 6, the measured performance of the animals in the choice training is plotted together with the predicted performance in terms of the mean proportion of choices of target A in successive four-trial blocks. The prediction is based here on the new β-values. It is of special interest that a preference for A developed quickly; the animals showed risk-aversion, despite the fact that target A provided on average only half the sucrose (5 µl of a 40 % solution on each trial) provided by B (30 µl of the same solution twice in every six trials). The overall proportion of choices of A was significantly greater than chance, $t(7)=6.82$. Of interest, too, is that the
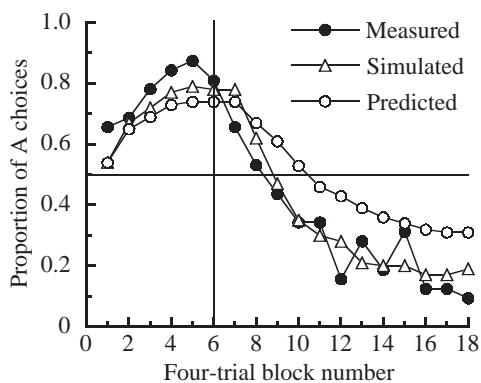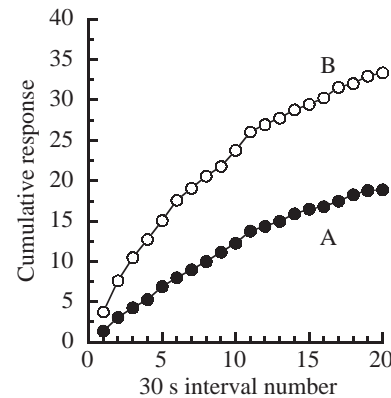

Fig. 4. Measured (filled circles), predicted (open circles) and subsequently simulated (open triangles) acquisition performance in experiment 1 plotted in terms of the mean proportion of choices of target A in successive four-trial blocks. The prediction was based on the estimated learning rates ($U\beta=0.04$ and $D\beta=0.02$) that yielded the best fits to the data of Shapiro (Shapiro, 2000). The *post-hoc* simulation was based on somewhat larger values ($U\beta=0.08$ and $D\beta=0.06$). The vertical line separates the first and second stages of training. $U\beta$, incremental rate of learning; $D\beta$, decremental rate of learning.
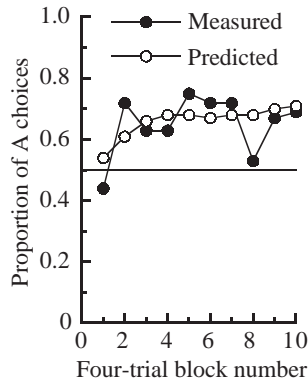
Fig. 6. Measured (filled circles) and predicted (open circles) acquisition performance in experiment 2 plotted in terms of the mean proportion of choices of target A in successive four-trial blocks. The new β-values suggested by the results of experiment 1 (Uβ=0.08 and Dβ=0.06) were used in the prediction.

choice data were predicted with considerable accuracy by the model.

In Fig. 7, the extinction test performance of these animals is plotted in terms of the mean cumulative number of responses to each target in successive 30 s intervals of the 10 min period. Here again, in agreement with the terminal choice data, we see a clear preference for A. Analysis of variance based on (uncumulated) frequencies of response in four 2.5 min blocks yields a significant stimulus (A *versus* B) effect ($F_{1,7}$=19.30) and a significant stimulus × block interaction ($F_{3,21}$=6.14).

Experiment 3

In Fig. 8, the performance of the animals in the choice training is plotted together with the predicted performance in terms of the mean proportion of choices of target A in successive four-trial blocks. The prediction is again based on the β-values as modified in accordance with the results of experiment 1. The curves show the development of a clear preference for target B (the variable alternative, providing 10 μl of a 60 % sucrose solution twice in every four trials) over
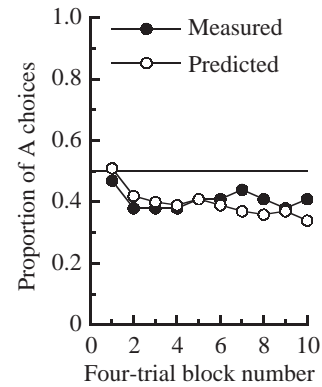
A (the consistent alternative, providing 10 μl of 15 % solution on every trial); the overall proportion of choices of B was significantly greater than chance, $t(7)$=2.49. The curves also show that the course of development was predicted with considerable accuracy by the model.

In Fig. 9, the extinction test performance of these animals is plotted as usual in terms of the mean cumulative number of responses to each target in successive 30 s intervals of the 10 min period. Here again, in agreement with the terminal choice data, there was a clear preference for target B. Analysis of variance based on (uncumulated) frequencies of response in four 2.5 min blocks yields a significant stimulus (A *versus* B) effect ($F_{1,7}$=10.76), but no significant stimulus × block interaction ($F_{3,21}$=1.53).

Experiment 4

In Fig. 10, the performance of the animals in the choice training is plotted in terms of the mean proportion of choices of target A in successive four-trial blocks. As in experiment 1, a preference for A (the consistent alternative, providing 5 μl of a 40 % sucrose solution) over B (the variable alternative, providing 30 μl of the same solution on every sixth trial)

Fig. 8. Measured (filled circles) and predicted (open circles) acquisition performance in experiment 3 plotted in terms of the mean proportion of choices of A in successive four-trial blocks. The new β-values suggested by the results of experiment 1 (Uβ=0.08 and Dβ=0.06) were used in the prediction.
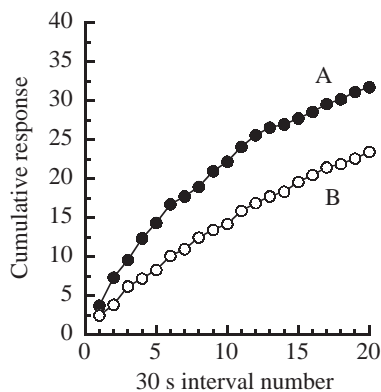

Fig. 7. Extinction test performance in experiment 2 plotted in terms of the mean cumulative number of responses to each target in successive 30 s intervals for targets A and B.
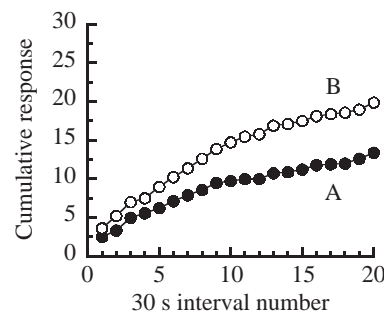

Fig. 9. Extinction test performance in experiment 3 plotted in terms of the mean cumulative number of responses to each target in successive 30 s intervals for targets A and B.
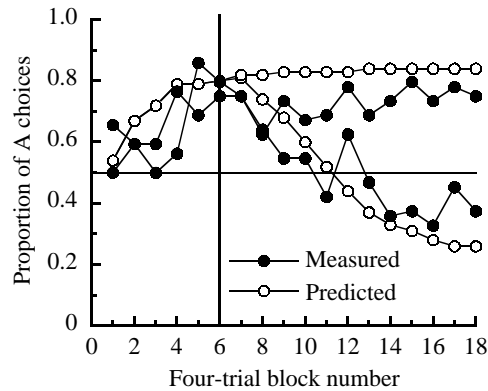
Fig. 10. Measured (filled circles) and predicted (open circles) acquisition performance in experiment 4 plotted in terms of the mean proportion of choices of A in successive four-trial blocks. The new β-values suggested by the results of experiment 1 ($U\beta$=0.08 and $D\beta$=0.06) were used in the prediction. The top pair of plots are for the nonreversal group, the lower pair for the reversal group. The vertical line separates the first and second stages of the training.

developed quickly in the first stage of training. The overall proportion of choices of A in that stage was significantly greater than chance; $t(11)$=5.79 for the nonreversal group and 6.85 for the reversal group. In the second stage, the preference for A remained about the same in the nonreversal group, for which the contingencies of reward remained the same, but quickly gave way to a preference for target B in the reversal group (persistent risk-aversion). Analysis of variance based on the data for the second stage yields a significant group effect ($F_{1,22}$=24.24) and a significant group $\times$ four-trial block
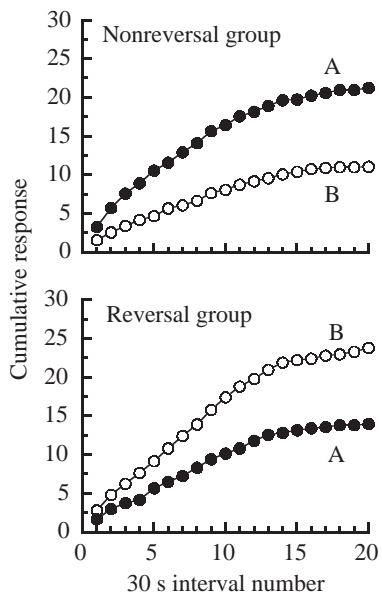

Fig. 11. Extinction test performance of the nonreversal (top) and reversal (bottom) groups in experiment 4 plotted in terms of the mean cumulative number of responses to each target in successive 30 s intervals for targets A and B.

interaction ($F_{11,242}$=2.05), and the correspondence between the predicted and measured results continues to be good. Comparison with Fig. 4 shows that the theory predicted slower reversal of preference in this experiment than in experiment 1, which was confirmed by the data ($F_{1,18}$=18.23). The explanation, of course, is that there was much less experience with the stimuli in the reversal training of experiment 4; on each trial, the animal was permitted to visit only one target of each pair.

In Fig. 11, the performance of the nonreversal and reversal groups of animals in the extinction test is plotted in terms of the mean cumulative number of responses to each target in successive 30 s intervals of the 10 min period. These curves, too, show a preference for target A in the nonreversal group and a preference for target B in the reversal group, confirming the preferences demonstrated by the choice measures at the conclusion of the training. Analysis of variance based on (uncumulated) frequencies of response in four 2.5 min blocks yields, for each group, a significant stimulus (A *versus* B) effect ($F_{1,11}$=22.16 for the nonreversal group and $F_{1,11}$=23.56 for the reversal group) and a significant stimulus $\times$ block interaction ($F_{3,33}$=9.13) for the nonreversal group but not for the reversal group ($F_{3,33}$=1.93).

Discussion

The rather simple model of discriminative learning developed by Couvillon and Bitterman (Couvillon and Bitterman, 1991) and enriched with reward-value functions generated by Shapiro (Shapiro, 2000) provides an accurate account of choices made by foraging honeybees under a considerable variety of conditions. The first postulate of the model is that the attractiveness of a feeding place is based simply on the strength of its association with reward, which grows as a function of the value of the reward. It may be well to note that there is no place here for the assumption of learning 'about' reward, which has figured prominently in the vertebrate literature (Bitterman, 2000). There is no need to assume that honeybees remember the specific properties of the various rewards they encounter; associations of identical strength may be generated by very different schedules of reward. A second postulate of the model is that choice is determined by relative associative strength. For exact rather than merely ordinal predictions of performance, it is necessary to assign numerical values to the parameters of the equations that express these postulates, and the goodness-of-fit to the substantial array of data currently available engenders confidence in our estimates of those parameter values and in the model itself.

Our prediction of the results of experiment 1 on the basis of parameter values that yielded good fits to Shapiro's data were not entirely accurate, capturing the qualitative features of the new results while underestimating the speed with which the preference acquired in the first stage of the training was reversed in the second stage, but the discrepancy could be corrected by moderate increments in the hypothetical learning rates (the β-values). Because there is nothing in the model to

suggest that the learning rates should have been different in experiment 1 from those in earlier experiments (the learning rates are assumed to be fixed for the species and the behavioral system under investigation), it was necessary to ask whether the new rates would impair the fit of the model to Shapiro's data, and it turned out that they did not do so appreciably. Furthermore, the RMS deviation of simulated from measured values for the entire set of data (data from Shapiro, 2000, and those of the four present experiments) proved to be smaller for the new rates than for the old (0.069 compared with 0.086). What now seems to have been an underestimation of the β-values by Shapiro is attributable to some unusual variability in the acquisition data of one of his experiments.

The results of the present experiments should be of some interest to students of foraging apart from their bearing on the validity of the Couvillon–Bitterman analysis. Experiments 1 and 4, which provide further evidence that honeybees are sensitive to variability in reward, suggest that they can adjust rather quickly to changes in the distribution of available resources (see also the more conventional reversal experiments of Couvillon and Bitterman, 1986). Experiments 2 and 3 point again to the importance of the question about the source of reward-value. The results of experiment 3 are immediately compatible with the traditional assumption that foraging choices are designed to maximize net energy gain, but those of experiment 2 are more problematic. To account for them in the same terms, it would be necessary to show that the net energy gain from two 30 μl drops (60 μl in total) of 40 % sucrose solution found in the course of six flower-visits is less than that from six 5 μl drops (30 μl) of the same solution found in the course of the same number of visits, which is not a simple task. The same problem is posed by a recent experiment on conditioned proboscis-extension in harnessed honeybees (Shafir et al., 1999), which oriented somewhat more readily to an odor paired consistently with 0.4 μl of a 1.5 mol l^{-1} (approximately 33 %) sucrose solution than to an odor paired with 2.4 μl of the same solution on only one of every four trials (mean amount 0.6 μl). Nor does it seem that it would be any easier to find independent support for an account of such results in purely perceptual terms.

It is to be hoped that meaningful answers to questions about reward value and its sources will be provided by experiments with other amounts and concentrations designed to refine the existing reward-value functions and to develop new ones for other amounts and concentrations. Experiments with targets consistently providing drops of sucrose solution that differ both in amount and in concentration might be especially informative. It should be interesting, too, as Shapiro (Shapiro, 2000) has proposed, to perform some analogous experiments with variability in delay rather than in amount and concentration of reward. The model suggests that honeybees would show risk-proneness such as that often found in delay experiments with a variety of other animals (Kacelnik and Bateson, 1996).

References

Bitterman, M. E. (2000). Cognitive evolution: A psychological perspective. In *Evolution of Cognition* (ed. C. Heyes and L. Huber), pp. 61–79. Cambridge, MA: MIT Press.

Bush, R. R. and Mosteller, F. C. (1951). A mathematical model for simple learning. *Psychol. Rev.* **58**, 313–323.

Caraco, T., Martindale, S. and Whittam, T. S. (1980). An empirical demonstration of risk sensitive foraging preferences. *Anim. Behav.* **28**, 338–345.

Couvillon, P. A., Arakaki, L. and Bitterman, M. E. (1997). Intramodal blocking in honeybees. *Animal Learn. Behav.* **25**, 277–282.

Couvillon, P. A. and Bitterman, M. E. (1986). Performance of honeybees in reversal and ambiguous-cue problems: Tests of a choice model. *Animal Learn. Behav.* **14**, 225–231.

Couvillon, P. A. and Bitterman, M. E. (1991). How honeybees make choices. In *The Behaviour and Physiology of Bees* (ed. J. L. Goodman and R. C. Fischer), pp. 116–130. Wallingford, UK: CAB International.

Couvillon, P. A., Mateo, E. T. and Bitterman, M. E. (1996). Reward and learning in honeybees: Analysis of an overshadowing effect. *Animal Learn. Behav.* **24**, 19–27.

Hamm, S. L. and Shettleworth, S. J. (1987). Risk-aversion in pigeons. *J. Exp. Psychol. Animal Behav. Proc.* **13**, 376–383.

Harder, L. D. and Real, L. A. (1987). Why are bumble bees risk averse? *Ecology* **68**, 1104–1108.

Kacelnik, A. and Bateson, M. (1996). Risk-theories – The effect of variance on foraging decisions. *Am. Zool.* **36**, 402–434.

Kacelnik, A. and Brito e Abreu, F. (1998). Risky choice and Weber's Law. *J. Theor. Biol.* **194**, 289–298.

McNamara, J. M. (1996). Risk-prone behavior under rules which have evolved in a changing environment. *Am. Zool.* **36**, 484–495.

Perez, S. M. and Waddington, K. D. (1996). Carpenter bee (*Xylocopa micans*) risk indifference and a review of nectarivore risk sensitivity studies. *Am. Zool.* **36**, 435–446.

Real, L. A. (1996). Paradox, performance and the architecture of decision-making in animals. *Am. Zool.* **36**, 518–529.

Reboreda, J. C. and Kacelnik, A. (1991). Risk sensitivity in starlings: Variability in food amount and food delay. *Behav. Ecol.* **2**, 301–308.

Rescorla, R. A. and Wagner, A. R. (1972). A theory of classical conditioning: Variation in the effectiveness of reinforcement and nonreinforcement. In *Conditioning II: Current Research and Theory* (ed. A. H. Black and W. F. Prokasy), pp. 64–99. New York: Appleton-Century-Crofts.

Shafir, S., Wiegmann, D. D., Smith, B. H. and Real, L. A. (1999). Risk-sensitive foraging: choice behaviour of honeybees in response to variability in volume of reward. *Animal Behav*. **57**, 1055–1061.

Shapiro, M. S. (2000). Quantitative analysis of risk-sensitivity in honeybees (*Apis mellifera*) with variability in concentration and amount of reward. *J. Exp. Psychol. Animal Behav. Proc.* **26**, 196–205.

Stephens, D. W. (1981). The logic of risk-sensitive foraging preferences. *Animal Behav.* **29**, 628–629.