

Title: Judgment bias may be explained by shifts in stimulus response curves

Authors: Caroline Strang^{1,2}, Felicity Muth^{1*}

Affiliations

¹Department of Integrative Biology, University of Texas at Austin; Austin, Texas, 78712,
United States.

²School of Behavioural and Social Sciences, Brescia University College; London, Ontario,
Canada, N6G 1H2

*Corresponding author. Email: felicity.muth@austin.utexas.edu

Abstract

Judgment bias, or ‘optimism’ and ‘pessimism’, has been demonstrated across many taxa, yet the cognitive mechanisms underlying this behaviour remain unclear. In an ‘optimism’ paradigm, animals are trained to an association, and, if given a ‘positive’ experience, behave more favourably towards ‘ambiguous’ stimuli. We tested whether this effect could be explained by changes to stimulus response gradients by giving bees a task where their response was tested across a wider gradient of stimuli than typically tested. In line with previous work, we found that bees given a ‘positive’ experience demonstrated judgment bias, being more likely to visit ‘ambiguous’ stimuli. However, bees were also less likely to visit a stimulus on the other side of the S+, and as such had a shifted stimulus response curve, showing a diminished peak shift response. In two follow-up experiments we tested the hypothesis that our ‘positive’ manipulation altered bees’ stimulus response curves via changes to the peak shift response by reducing peak shift in controls. We found that, in support of our hypothesis, elimination of peak shift also eliminated differences between treatments. Our results point towards a cognitive explanation of ‘optimistic’ behaviour in non-human animals and offer a new paradigm for considering emotion-like states.

Keywords: cognitive bias, optimism, peak shift, *Bombus impatiens*, bumblebee, bee.

Introduction

Although difficult to define (1), emotions have been described as ‘suites of cognitive, motivational and physiological changes that are triggered by appraisal of specific classes of environmental situations’ (2). One way that emotional states have been studied is through judgment bias tasks (3, 4) where an individual is trained via differential conditioning to a rewarded stimulus (S+) and unrewarded or punishing stimulus (S-) before undergoing a particular experience (either expected to induce a ‘negative’ or ‘positive’ affective state). The subject is subsequently tested on their response towards the S+, S-, and ‘ambiguous’ stimuli that lie between the two trained stimuli (5). A diminished or enhanced response to the ambiguous stimulus is described as ‘pessimistic’, or ‘optimistic’ bias, respectively (4). The terms ‘pessimistic’ and ‘optimistic’ are operationalized as behavioural responses to ambiguous stimuli and extrapolated to represent emotion-like states (3). Judgment bias tasks have been used most widely in animal welfare studies where they often serve as indicators of animals’ affective states in response to an experience such as a stressful housing condition or environmental enrichment (6, 7).

The first demonstration of judgment bias in a non-human animal was in rats (4), and since then judgment biases have been shown in a number of vertebrates, including European starlings *Sturnus vulgaris* (8, 9), common ravens *Corvus corax* (10), rhesus macaques *Macaca mulatta* (11), and collard peccaries *Pecari tajacu* (12)). Indeed, a meta-analysis in 2020 identified 71 studies in 22 non-human species (6), finding that while most experiments used mammals, judgment biases have been demonstrated across broad taxa. The honeybee was the first invertebrate shown to demonstrate judgment bias, showing a ‘pessimistic’ response to a stressful event (being shaken) (5), see also (13). More recently, bees were also shown to express

positive biases: bumblebees that received an ‘unexpected reward’ took less time to visit an ambiguous visual stimulus, while responses to the S+ and S- remained the same (14).

While the behaviour apparent in judgment bias tasks has been well-documented, and is robust across species and experimental treatments (6), the cognitive mechanisms underlying this behaviour have not been investigated. On the one hand, it has been suggested that judgment bias effects cannot be explained by general sensory or motivational changes because responses to trained cues (S+ and/or S-) often remain the same (3, 5). Others have argued that judgment biases may be explained by what is already known about learning and/or motivational mechanisms (15, 16). Indeed, classic work in experimental psychology tells us that variables such as motivation can alter the shape of stimulus response curves without changing the peak (17, 18). As such, shifts in stimulus response generalization curves may explain how animals may change their response to an ambiguous stimulus, while responses to the S+ and S- remain the same (example in Fig. S1). Similar ideas have also been proposed in (19).

Here we addressed how positive emotion-like states could be explained at a psychophysical level, using bumblebees as a model (20–22). When learning a discrimination, animals typically form a response curve around the S+, generalizing their response to similar stimuli (23), including ambiguous stimuli, and stimuli farther from the S+; hereafter ‘novel’ stimuli (Fig. S1). We refer to these stimuli as ‘ambiguous’ and ‘novel’ in line with existing terminology, while acknowledging that ‘ambiguous stimuli’ are also novel to trained bees, and that the ‘novel’ stimuli may be ambiguous to bees in that they have not been reinforced or unreinforced via training. Based on previous work showing that motivational changes can alter the shape of stimulus response curves (17, 18), we hypothesised that ‘positive’ experiences would broaden the stimulus response curve, thus increasing responsiveness towards the

ambiguous stimuli, but also to novel stimuli (Fig. S1A). Conversely, negative experiences (not tested here) may narrow the range of stimuli that an animal is willing to accept (Fig. S1B). Importantly, such changes would lead to similar responses to the S+ and S-, a key aspect of judgment bias. We initially tested this hypothesis in Experiment 1 using a modified judgment bias task that expanded the range of stimuli typically used. We first trained bees to a S+/S- discrimination, then assigned them to control or experimental groups, the latter of which received a sucrose reward (as in the previous study on positive emotion-like states in bumblebees (24)), before testing them on a range of stimuli in a single probe trial that included the S+, S-, ambiguous and novel stimuli (Fig. 1). Past work on judgment biases has most often used a ‘go/no-go’ design, where the animal is presented with a single stimulus in a probe trial and either responds or withholds their response, while other studies have used an ‘active choice’ design where the animal is presented with a stimulus in a probe trial and must select between two trained responses which correspond to the S+ and S- (6). In both cases, the number of sequential presentations in the probe trial varies greatly among studies, as well as whether stimuli are reinforced or not (6). In our study, we used a design adopted from work on generalization gradients (25, 26), but new in the context of measuring judgment bias, where we presented bees with an array of multiple stimuli simultaneously in order to rapidly assess preference across a stimulus gradient.

In Experiment 1 we found that a ‘positive’ experience of a reward given after learning appeared to alter bees’ stimulus response curves via changes to the peak shift response (see Experiment 1 Results & Discussion). We then tested the hypothesis that an increase in preference for ambiguous stimuli following a ‘positive’ experience is explained by a reduction or elimination of the peak shift response in two follow-up experiments. In Experiment 2, we trained

individuals via absolute conditioning, thus eliminating the inhibitory curve and peak shift. In Experiment 3, we trained individuals to stimuli closer to each other, which may have increased or decreased peak shift (see below). If our hypothesis held, then we predicted that if we eliminated peak shift, we would no longer see differences between treatments. Conversely, if we had increased peak shift, we would expect to see a larger difference between treatments.

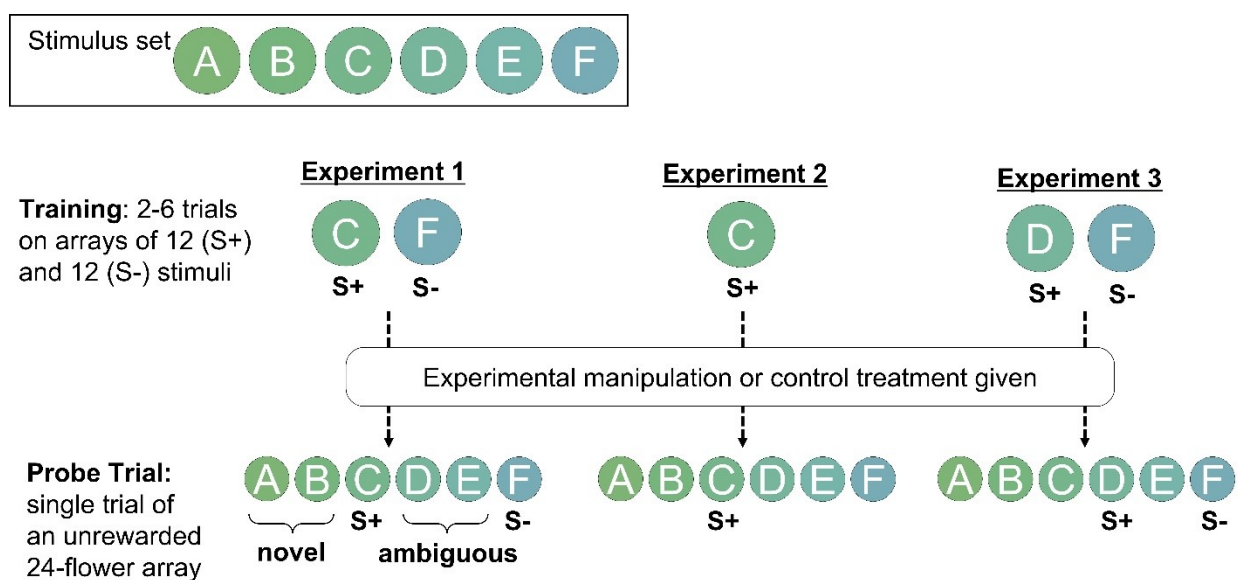


Fig. 1: Diagram of the experimental design and colour stimuli used in the current experiments. Six colour stimuli (indicated by letter; A-F) were used across all experiments; each stimulus was roughly equidistant from the previous in bee colour space (Table S1). In each experiment, bees were first trained to learn a colour association across 2-6 training trials; S+ indicates rewarding flower type; S- indicates unrewarding flower type. Individuals were then either given a reward (5µl of 50% sucrose) in the experimental treatment or no reward in the control treatment. All bees were then tested for their preferences on an unrewarding probe array where they were presented with 24 flowers, with 4 of each colour.

116

117 **Materials and Methods**

118 *General Methods*

119 *Subjects, housing and maintenance*

120 We used worker bumblebees (*Bombus impatiens*) from commercially-produced colonies
121 (Koppert, USA). In each experiment, we used 32 (n=16 control; n=16 experimental) foragers,
122 taken equally from two colonies (i.e. 96 bees total). Colonies were tested sequentially, attached
123 to ‘foraging arenas’ ($l \times w \times h = 122 \times 61 \times 61$ cm) via an ‘entrance tube’ (length = 360cm,
124 diameter = 1.5cm) where individual bees were presented with vertical arrays of 24 artificial
125 flowers. The entrance tube consisted of a ‘walkway’: a 3cm diameter clear plastic tube, attached
126 to a 3cm diameter mesh tube and a ‘holding area’: a 1.5cm diameter Perspex tube (Fig. S2). The
127 holding area contained a small hole, through which we fed bees in the experimental procedure.
128 Both the entrance tube and holding area contained metal ‘gates’ that could be raised and lowered
129 to allow individual bees through the tube and to hold individuals in the holding area. The
130 foraging arena was illuminated by fluorescent room lighting and a 40-Watt LED white light
131 (Commercial Electric, USA) placed directly above the foraging arena. All lights were on a
132 12/12h light/dark schedule.

133 On training days when insufficient sucrose was collected, we pipetted 30% sucrose
134 directly into the colony. We supplied colonies with one tablespoon (approx. 3g) of pollen
135 directly into the colony every two days or as needed (based on visual assessment of colony
136 stores).

137

Training and Testing Stimuli

The shaping, training, and testing arrays consisted of corrugated plastic panels with 24 artificial flower locations creating $10 \times 40\text{cm}$ arrays. The panels were painted a dark grey colour (#2753, BEHR ULTRA™, California, USA). Training and testing arrays consisted of artificial flowers with “corollas” (circular laminated card sized 4cm in diameter). We used 6 coloured stimuli, referred to as A-F, over the course of three experiments. These stimuli were designed to be as close to equidistant as possible in bee colour space (Fig. S3, Table S1). In HSL the hue of the stimuli ranged from 80° (green) to 135° (blue) in 10° increments (i.e. as in (26) and similar to (25)), with the exception of the interval between the 120° to 135° stimuli which was larger to make the stimuli closer to equidistant in bee colour space, with saturation 75% and luminance 150%. Stimuli were printed on an inkjet printer (Epson stylus C88+) on Cotton Fine Art Archival OBA free paper (Pacific Inkjet, USA). We used a Flame UV-VIS spectrometer (Ocean Insight, Florida, USA) to measure reflectance of each stimulus and irradiance in the foraging arena. The reflectance measurements were then analyzed and mapped into bee colour space (27) taking into account the photoreceptor spectral sensitivities of *B. impatiens* (28), using AVICOL, a program for analyzing spectrometric data (29).

In order to avoid biasing bumblebees to particular locations on a given array, we used 6 different training arrays, where flower colour location was pseudorandomized across arrays such that the S+ and S- were represented equally in the top and bottom half of the array. We also pseudorandomized the order that we presented arrays to bees across trials, with all orders presented an equal number of times for control and experimental bees.

For the Probe Trial array, the stimuli were arranged pseudorandomly such that stimuli of the same colour were never directly beside each other and so that there was equal representation

of the S+ and S- stimuli in the top and bottom half of the array, to control for possible location preferences. In Experiment 1, we used four different arrays in the probe phase, which were used an equal number of times across the control and experimental conditions. After one probe array was damaged during the first experiment and we were unable to replace stimuli, we used only three probe arrays in Experiments 2 and 3. In these experiments, the probe arrays were represented roughly equally across treatments and accounted for in statistical analyses.

Pre-training and Shaping

We first ‘pre-trained’ bees to forage in the arena and return to their colony between foraging bouts. To do this, colonies were each given access to a feeder (a 250ml plastic tub containing a white wick) in the foraging arena. The feeder was initially placed at the entrance to the foraging arena; once bees were regularly visiting the feeder we gradually moved it to the back of the arena where the shaping, training and testing arrays would be placed. During pre-training, we restricted colonies’ access to the foraging arena overnight and removed feeders.

In a ‘Shaping’ phase, we trained bees in a step-wise fashion to visit our experimental flowers. To do this, we first placed the ‘shaping array’ at the back of the foraging arena and gave foragers free access to it. This array consisted of 24 equally-spaced (2.5cm) clear Eppendorf tubes, containing cotton wicks soaked in 30% (w/w) sucrose. The wicks were replenished as needed to create a constant sucrose supply in all artificial flowers. Once bees foraged on the array, we gradually pulled the wicks back into the flowers until bees had to fully enter the Eppendorf tubes to access the wicks. This part of shaping typically took 1-2 days and was repeated as needed throughout training and testing. Once bees consistently visited the shaping

array, we replaced the sucrose-soaked wicks with 40µl of sucrose, which required bees to fully enter the artificial flowers to access rewards. We replenished flowers throughout the shaping procedure by pipetting sucrose into the Eppendorf tubes via small holes. This latter part of the shaping phase was repeated each day prior to training and testing. To identify individual bees for use in experiments, we paint-marked foragers with a unique colour combination on their thorax with Posca paint pens (Uni Mitsubishi Pencil, USA) during shaping.

Training

In all experiments, paint-marked, shaped bees were alternately assigned to Experimental and Control treatments. During the training phase, all bees were trained to a S+/S- discrimination (Experiments 1 and 3) or S+ conditioning (Experiment 2). Rewarded flowers always contained 10µl of 30% (w/w) sucrose and unrewarded flowers contained 10µl of water. After training, bees then underwent the experimental treatment before being tested on a set of 6 unrewarded stimuli in a probe trial (Fig. 1). Training phases for each experiment are described in detail below.

Experimental treatment

After training we waited for the bee to leave its colony again to forage, at which point we contained it in a holding area for 2 minutes, during which time bees in the experimental treatment were fed 5µl of 50% (w/w) sucrose (as in (14)); control bees were held but not fed. After 2 minutes, bees were given access to the foraging arena for the probe trial.

203

204 *Probe Trial*

205 During the probe trial, bees were given access to an unrewarded 24-flower array (each flower
206 contained 10µl of water) (Fig. 1) for 10 minutes; after this they were euthanized via freezing.

207

208 *Behavioral coding and inclusion criteria*

209 We filmed all trials and coded bees' choice behaviour in probe trials using Solomon Coder (30).
210 Since flowers were not replenished during learning trials, bees occasionally made visits to
211 previously-emptied flowers; these visits were excluded from the learning criterion calculation
212 because it is not clear if bees would perceive such a visit as a CS+ or CS-. In the probe phase, we
213 defined a choice as the bee landing on a flower and inserting its head or body into the flower. We
214 decided to use the first 20 choices bees made as the response measure in the probe phase to allow
215 for sufficient visits to the 6 stimuli and in line with previous bumblebee cognition experiments
216 (31, 32). Within these 20 choices, bees could visit the same stimulus more than once and it
217 would be counted as a separate choice; we did this to allow a bee to have up to a 100%
218 preference for a given stimulus; otherwise preferences would have been limited to 16.67% (1/6)
219 per stimulus. This decision was informed by observations from similar behavioural experiments
220 that bees will frequently re-visit the same stimulus multiple times in probe phases (personal
221 observation, F.M.). However, we removed bees (n=1) that made more than 50% of their choices
222 to a single location, since this may indicate a strong location rather than stimulus preference. The
223 majority of bees (81/96) made at least 20 choices in the probe phase, yet we did not exclude bees
224 that made fewer than this; all bees included in the final dataset made at least 12 choices. We did

not include all of bees' choices within the 10-minute recorded trial because after 20 unrewarded visits bees may be more likely to shift to random sampling after not receiving rewards.

Previous work on positive judgment bias in bees found that bees given a reward did not differ to controls in their latency to make a choice (24). We confirmed that this was also the case in the current study by extracting the time it took for bees to make their first choice to a stimulus from recorded videos. We found this measure did not vary between treatments for any experiment; statistical methods and results are in the Supplementary Material.

Data Analysis

All statistical analyses were conducted in R ver. 4.1.2 (33). In all experiments, to determine if treatments differed in their stimuli preferences in the probe phase, we asked if bees in the experimental treatment differed in their choices to stimuli compared to control bees. We did this using Linear Mixed Models (LMMs) (package: nlme; function: lme) (34). In all analyses, the response variable was the number of visits to a given stimulus (in the first 20 visits), and the explanatory variables included were: 'Condition' (treatment or control), 'Stimulus' (A, B, C, D, E or F), Probe Array (1-4 in Experiment 1 and 1-3 in Experiments 2 and 3) and 'Bee' nested within 'Colony' as random factors. Bee and Colony were included as random factors in all models because visits to different stimuli were non-independent i.e. a given bee would visit multiple different stimuli. We initially ran a full model including the 3-way interaction between Condition, Stimulus and Probe Array, but then removed non-significant interaction terms ($p > 0.05$) in a step-wise fashion while always keeping main effects in the model (per (35)). We used the anova() function to generate F -values and P -values for model terms. We then used the

package emmeans (36) to carry out Tukey post-hoc comparisons between factor levels of main effects and significant interaction terms. For all models, we checked normality of residuals by plotting them using the ‘resid’ function.

Experiment 1: Standard judgment bias paradigm

Experiment 1 Methods

In each training trial we presented bees with 12 rewarded (S+; stimulus C) and 12 unrewarded (S-; stimulus F) flowers (Fig. 1). Bees were trained over a number of trials (2-6) until they reached a performance criterion of 8/10 correct choices on their first 10 choices within a trial. All training trials were video-recorded, but successful completion of training was assessed by live observation. In each trial bees either depleted the array and then returned to the colony, or if 10 minutes elapsed and the bee had not returned by itself then we returned it to the colony. Individuals that visited fewer than 8 flowers on trial 1 were excluded from the experiment. After training bees were either given experimental or control treatments before proceeding to the test phase (see *General Methods*).

We carried out statistical analysis as described in the *Data Analysis* section of the *General Methods*. One bee (of 32) was excluded from the analysis since she showed extremely strong preferences to a single location (10/13 (77%) of her visits were to the same location). This

resulted in 16 bees in the control treatment and 15 bees in the experimental treatment whose data were included in analyses.

Experiment 1 Results and Discussion

Bees in experimental and control treatments differed in how frequently they visited the 6 stimuli, as shown by a significant interaction effect and no main effect of Condition (Stimulus $\times$ Condition: $F_{5, 130} = 2.794, p = 0.0197$; Stimulus = $F_{5, 130} = 34.737, p < 0.0001$; Condition: $F_{1, 25} = 0.00, p = 1.000$; Fig. 2A). In line with previous findings (24), bumblebees given a high-quality reward after learning an association demonstrated judgment bias, showing a higher response rate to the ambiguous stimulus D relative to controls, while responding similarly to the trained C (CS+) and F (CS-) stimuli (Table 1; Fig. 2A). Bees also appeared to show ‘peak shift’; this is typical of S+/S- discrimination training, where the peak of the response curve is not the S+, but rather a novel stimulus biased away from the S- (37), as has previously been described in bees (25, 38). In this case, peak shift was relatively weak but still apparent: bees showed a peak response to not only stimulus C, but also stimulus B shifted away from the S- (no difference in response to stimuli B vs. C in a post-hoc comparison: Control group: $t_{130} = 0.064; p = 0.950$; Experimental group: $t = t_{130} = 0.406; p = 0.685$). However, the strength of this effect was larger for the control than experimental group, with a strong trend towards more bees choosing stimulus B in the control group (Table 1; Fig. 2A). That treatments should show peak shift is in line with previous work with bumblebees (25, 26). Bees were also more likely to visit particular stimuli on specific probe arrays (Stimulus $\times$ Probe array: $F_{15, 130} = 2.010; p = 0.0190$; Probe array: $F_{3, 25} = 0.000; p = 1.000$; for details see Fig. S4, Table S2).

Table 1: Post-hoc differences between Control and Experimental treatments across the 6 experimental stimuli in Experiment 1. Results are averaged across the 4 levels of ‘Probe array’.

Stimulus	Contrast	estimate	SE	df	t-ratio	p-value
A	Control vs. Experimental	0.517	0.757	25	0.683	0.501
B	Control vs. Experimental	1.505	0.757	25	1.988	0.058
C	Control vs. Experimental	1.119	0.757	25	1.478	0.152
D	Control vs. Experimental	-1.630	0.757	25	-2.153	0.041
E	Control vs. Experimental	-1.191	0.757	25	-1.573	0.128
F	Control vs. Experimental	-0.320	0.757	25	-0.422	0.676

One explanation for peak shift is that partially or completely overlapping excitation and inhibition gradients average each other out to form the peak-shifted shape of the response curve (39) (Fig. 3). Since a peak shift response is driven by individuals’ training to the S- (36), our finding that peak shift was lessened by the pretest reward may indicate that the experimental treatment altered the gradient around the S-, for example by narrowing it. To test the hypothesis that the pretest sucrose drove changes to individuals’ responses to ambiguous stimuli via effects on the peak shift response, we conducted Experiment 2. Since inhibitory response curves cannot be easily measured directly, we aimed to eliminate the peak shift response in the control group with the prediction that its elimination should also eliminate the difference between the two experimental groups. To do this, we trained bees via absolute conditioning (S+ only), thus removing the inhibitory response curve completely. Bees then underwent the same experimental manipulation and probe trial as Experiment 1.

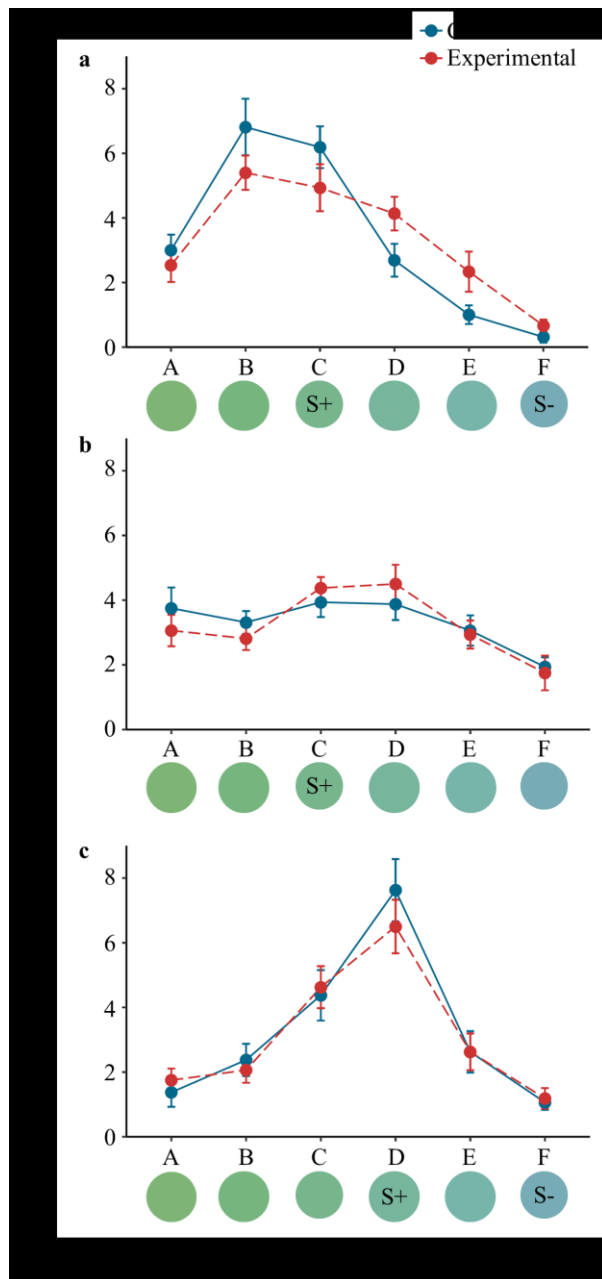


Fig. 2: Results of (a) Experiment 1 (b) Experiment 2 and (c) Experiment 3. Bees' preferences across the 6 stimuli in the unrewarded probe phase. Bees in the experimental treatment (red, dashed line) experienced a reward prior to testing, while control bees (blue, solid line) did not. Colour stimuli are shown on the horizontal axis, with the S+ and S- identified for each experiment. In Experiment 1 where bees were trained via differential conditioning, the

experimental treatment showed an increased response towards ambiguous stimulus D ($p = 0.041$) and a strong trend towards a decreased response towards stimulus B ($p = 0.058$). The experimental group also appeared to show less peak shift relative to controls. In Experiment 2, where we eliminated the peak shift effect by training bees to the S+ only, both treatments generalized broadly and did not differ to each other. In Experiment 3, where we increased the learned response and eliminated peak shift by training bees on stimuli closer together, treatments also did not differ to each other.

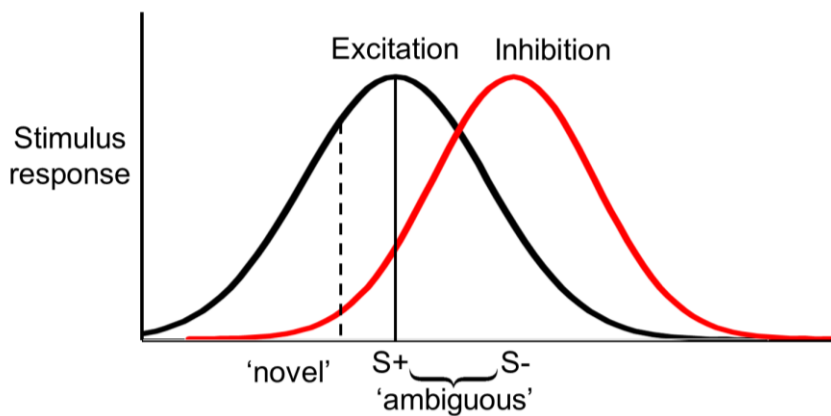


Fig. 3: Diagram showing hypothetical excitation (black) and inhibition (red) curves generated by differential training to a S+ and S-. An averaging of these overlapping curves may cause peak shift (39). On the above diagram, the difference between the excitation and inhibition curves is larger at a value shifted away from the S+ (dashed line) than at the S+ (solid line), and thus there may be a higher response to stimuli at this point.

Experiment 2: Absolute conditioning with the S+ only

In this experiment we tested the hypothesis that our results from Experiment 1 were due to a reduction or elimination of the peak shift response. To do this, we trained bees via absolute conditioning; without a S- no inhibitory curve would be generated and thus no peak shift response. We predicted that if our hypothesis held, the differences between the treatments would be eliminated when peak shift was eliminated.

Experiment 2 Methods

Bees exclusively experienced the S+ (C) stimuli during training (Fig. 1); S- stimuli were removed and replaced with black rubber stoppers. All bees were given 3 training trials; this number was chosen since it was the most frequent number of trials needed for bees to reach the learning criterion in Experiment 1 and thus was the closest match possible to the number of rewarded experiences with the S+ that bees would have had in that experiment; a similar calculation was used in (25). Once bees completed training and returned to the colony they were assigned to experimental or control treatments and were given probe trials as in Experiment 1.

We carried out statistical analysis as described in the *Data Analysis* section of the *General Methods*. 4 bees made fewer than 20 visits (control n=2; experimental n=2); these varied

from 14 to 19 visits and were included, meaning a final sample size of 16 experimental and 16 control bees included in data analyses.

Experiment 2 Results and Discussion

Without the S-, bees generalized broadly across stimuli, widening the stimulus response curve (Fig. 2B). In support of our hypothesis, control and experimental groups did not differ from each other (Condition: $F_{1,27} = 0.0737$, $p = 0.788$). However, while bees generally chose the S+ above other stimuli (Stimulus: $F_{5,155} = 0.7011$, $p < 0.001$), learning was relatively weak, with bees choosing the S+ stimulus above B, E and F stimuli (significant effect or trend with post-hoc pairwise comparisons), but not above stimuli A or D (see Table 2). In this experiment the probe array used did not affect bees' choice of stimuli (Probe array: $F_{2,27} = 0.0784$, $p = 0.925$).

Table 2: Post-hoc differences between the S+ (C) stimulus and 5 other stimuli in Experiment 2. Results are averaged across the two levels of the factor 'Condition' and four levels of the factor 'Probe array'.

Contrast	estimate	SE	df	t-ratio	p-value
C - A	-0.750	0.463	155	-1.62	0.353
C - B	-1.094	0.463	155	-2.362	0.080
C - D	0.031	0.463	155	0.067	1.000
C - E	-1.156	0.463	155	-2.497	0.058
C - F	-2.313	0.463	155	-4.994	<0.001

While the results of this experiment supported our hypothesis, we only saw weak learning in both treatments. That absolute conditioning only led to weak learning is in line with previous work showing that differential conditioning is necessary for fine-scale colour discrimination (40). However, other work addressing bumblebee generalization gradients found equivalent learning when bees were trained via absolute conditioning using the average number of training sessions taken by groups that received differential conditioning (25); differences between that study and our own are likely due to subtle differences in training regime or stimulus discriminability.

To tackle this limitation, in Experiment 3 we aimed to determine if our hypothesis held when bees showed a stronger learned response. To increase learning while manipulating the peak shift response, we decreased the distance between our S+ and S- stimuli. While decreasing the distance between stimuli can increase the peak shift if curves overlap more with each other (41), training animals to stimuli closer together (i.e. requiring finer discriminability) typically reduces stimulus generalization (40, 41) and could also reduce or eliminate peak shift if the excitation and inhibition curves narrowed to the point of not overlapping. If our ‘peak shift hypothesis’ as an explanation for judgement bias effects held, we predicted that if peak shift increased, we would increase the difference between our treatments, whereas if peak shift decreased, we would decrease the difference between our treatments.

Experiment 3: Standard JBT with reduced distance between the trained stimuli

Experiment 3 Methods

Bees underwent the same training and testing procedures in Experiment 3 as for Experiment 1, with the exception that the rewarded stimulus (S+) was Stimulus D and the unrewarded stimulus (S-) was Stimulus F.

We carried out statistical analysis as described in the *Data Analysis* section of the *General Methods*. 10 bees made fewer than 20 visits (control n=4; experimental n=6); these varied from 12 to 19 visits and were included in the final data analysis.

Experiment 3 Results and Discussion

When we trained bees to a S+ and S- closer together, it narrowed the stimulus response curve, causing a strong learned response to the S+ and no peak shift (Fig. 2C, Table 3). Accordingly, bees visited the S+ (D) stimulus more than all other stimuli (Stimulus: $F_{5, 145} = 29.629$, $p < 0.0001$; Table 3). In support of our hypothesis, once peak shift was eliminated, there was no difference between experimental and control treatments in the stimuli they visited (Condition: $F_{1, 27} = 0.117$, $p = 0.735$). As in Experiment 1, we found that certain stimuli were preferred on particular probe arrays (Stimulus $\times$ Probe array: $F_{10, 145} = 1.910$, $p = 0.0482$; Probe array: $F_{2, 27} = 0.084$; $p = 0.920$); Fig. S5, Table S3), likely due to location preferences.

Table 3: Post-hoc differences between the S+ (D) stimulus and 5 other stimuli in Experiment 3. Results are averaged across the two levels of the factor ‘Condition’ and four levels of the factor ‘Probe array’.

Contrast	estimate	SE	df	t-ratio	p-value
D - A	-5.67	0.595	145	-9.532	< 0.0001
D - B	-4.99	0.595	145	-8.384	< 0.0001
D - C	-2.95	0.595	145	-4.951	< 0.0001
D - E	-4.62	0.595	145	-7.756	< 0.0001
D - F	-6.06	0.595	145	-10.187	< 0.0001

General Discussion

The discovery of emotion-like states in invertebrates raises questions about the evolutionary origin of emotions, as well as having implications for animal welfare. A cognitive framework for the study of judgment bias may help explain discrepancies between previous studies and offer a new perspective on inter-specific comparisons. By addressing bees’ generalization curves after receiving a reward, we found that rewarded individuals showed an increased response to an ambiguous stimulus, but a decreased response to a novel stimulus. As such, this group showed a lessened peak shift response across a gradient of stimuli. When we manipulated the training procedure to eliminate peak shift, we no longer found a difference between experimental and control treatments in their response to any stimuli.

Taken together, our experiments support the hypothesis that behaviour observed towards ambiguous stimuli during ‘optimistic’ judgment bias protocols can be explained by changes to the peak shift effect of stimulus response curves, presumably through a change to the inhibitory response curve around the S-. Since we did not measure the S- curve directly we cannot be

certain as to how it changed, however, one possibility is that the experimental treatment in Experiment 1 narrowed it; this would explain the reduction of peak shift while response to the S- remained unchanged. This explanation would be consistent with our manipulation in Experiment 3, which we believe also narrowed the inhibitory generalization curve. Our hypothesis is also consistent with work showing that a heightened response to ambiguous stimuli is caused by changes in the dopaminergic system (14), which has been associated with inhibitory conditioning in honeybees: blocking dopaminergic receptors suppresses aversive learning (42) and abolishes judgment bias effects (24).

One explanation for previously-reported judgment bias effects in bees (24) is that a high-quality sucrose reward after training may increase foraging motivation in general, and in doing so increase responsiveness to ambiguous stimuli (15), as well as novel stimuli (as tested in our original hypothesis in Experiment 1). Indeed, high-quality sucrose generally motivates bees' foraging: experimental addition of sucrose into a colony can trigger foraging, with higher-concentration sucrose causing greater foraging activity (43). However, our findings do not support this explanation, since this would have led to a higher response to the similar novel stimulus B in Experiment 1 and stimuli similar to the S+s in Experiments 2 and 3. Finally, we also found that another potential measure of foraging motivation, the bees' time to the visit the array, did not differ between experimental and control treatments in any of our experiments. Another possibility for an increase in visitation to ambiguous stimuli following a high-quality sucrose reward could be increased sampling. Receiving a high-quality food reward, either in the colony, or in this case, on the outside of the colony, may give individuals information that there is higher-quality food in their environment (44, 45). We would expect this to lead to greater search on 'novel' flowers that were not previously associated with a particular reward value,

since bees form specific expectations of reward quality based on the associated stimuli (46, 47). However, this would not explain our results either: bees that experienced the higher-quality sucrose reward did not visit novel stimuli A and B more in Experiment 1, nor any novel stimuli more in Experiments 2 and 3.

Judgment bias is typically assessed by offering an animal the two trained stimuli and ambiguous stimuli between these; our method and framework of offering a wider range of stimuli to determine stimulus generalization curves may be a useful means to address why and how different types of experience alter animals' stimulus perception. Judgment biases are likely a result of several different processes (rev. (3), section 4). For example, factors known to affect discrimination training (the distance between the trained S+ and S- stimuli, salience of stimuli etc. (48)) will determine generalization curves post initial training. These curves may then be altered by affective state. A better understanding of how particular experiences change generalization curves (e.g. chronic vs. acute stress, the context of the reward or punishment, the timing of the experience relative to training) may shed light on discrepancies in results from judgment bias tasks (6) and provide a framework for generating predictions in future work. Existing knowledge of how affective states affect learning and generalization may be useful in this regard. For example, classic work in experimental psychology shows that pigeons that are hungrier (i.e. starved) have broader stimulus generalization gradients (17). If this is a general phenomenon, it could explain why chronically food-restricted sheep *Ovis aries* behaved more 'optimistically' on a judgment bias task (49) and why dogs *Canis familiaris* fed prior to testing had seemingly 'pessimistic' judgments, in the opposite directions to what was expected based on the animal's presumed emotional state (50). These examples highlight how, without a mechanistic framework to understand judgment bias, we risk misinterpreting animals' emotional

states. Changes to generalization curves may also explain why in some studies the response to the S- changes in addition to the response to ambiguous stimuli: if a negative experience increases the peak of the inhibitory curve as well as broadening it, the experimental group would respond more to the S- in addition to ambiguous stimuli, as has been found for honeybees stressed via shaking (5) and sheep given chronic stress in an agricultural setting (51). Indeed, stressed states can both lead to higher responsiveness to negative stimuli and overgeneralization of responses (reviewed in: (52, 53)). The timing of the affective state manipulation may also affect the learned response; many judgment bias studies manipulate state during learning (rev. 6), rather than post-learning as in the current study, and the state an animal is in when learning can affect how it perceives the relative value of the reward (e.g. 47) and thus determine learning and generalization curves. Another consideration in predicting how a generalization curve may change in response to an experimental manipulation is the type of stimulus used in differential conditioning. In the current study, we focused on colour, a ‘rearrangement dimension’ (41), where the generalization gradient forms a Gaussian curve. Other stimuli follow an ‘intensity dimension’, where the same receptors are stimulated to a different extent, for example light intensity or tone frequency (41). In these cases, animals may not form a Gaussian curve when generalizing, but instead always respond more to higher intensities of the stimulus, and as such show a different peak shift response (41, 55). Finally, Spence’s model of peak shift (39) that we assume here (Fig. 3) is only one explanation for how peak shift occurs. Going forward, other models could be considered in terms of how we would expect animals’ affective state to alter stimulus generalization (41, 56).

In addition to expanding the range of stimuli used, we also adopted a different protocol to one that is typically used in judgment bias tasks in the probe trial phase of our experiment.

Previous research on judgment biases has been conducted using a variety of methods, that vary across all phases of the experiment, including discrimination training, experimental manipulation and the probe trial (6), all of which may affect interpretation of differences in results across studies (57). In regards to the probe trial, previous work has either presented animals with a single stimulus at a time and measured propensity or latency to perform an action ('go/no-go' design), or measured 'active choice' between two options, as a means of controlling for motivation, since a response is necessary for any choice made (3, 8, 9, 57, 58). Furthermore, in some experiments (as in the original experiment investigating judgment bias in bumblebees (14)), the experimental manipulation is given before each stimulus presentation in the probe trial, whereas in other experiments, multiple stimuli are presented sequentially in the probe phase, without repeating the experimental manipulation between exposures (e.g. (11, 59)). In our experiment, bees visited 20 unrewarding stimuli in succession; we designed our experiment this way in line with previous work on generalization gradients (25, 26). However, a clear next step would be to determine if our results hold in other paradigms, for example if stimuli are presented sequentially rather than simultaneously, and in go/no-go designs.

In conclusion, our results imply that good or bad experiences may alter how animals classify a range of stimuli in their worlds, and not only the 'ambiguous'. Thus, 'optimism' and 'pessimism' may not simply be a change in animals' perception of the uncertain, but rather a shift in how many stimuli are perceived. Moving forward, our suggestion that changes to peak shift underlie positive judgment bias provides a framework for advancing the comparative study of emotions. It also serves to reinvigorate investigation of how animals' experiences can influence psychophysical aspects of learning tasks. Given the central role of judgment bias tasks in topics of animal emotion and welfare, it is essential that we understand the underlying

cognitive mechanisms to avoid superficial intra- and inter-specific comparisons that may lead to inaccurate equivalence.

Data accessibility. Data are provided as Supplementary Material and will be uploaded to Dryad upon acceptance of the manuscript.

Competing interests. The authors have no competing interests.

Funding. Funding was provided by NSF award number IOS-2028613.

Acknowledgements. We would like to thank Michael Domjan and the Muth Lab for valuable feedback on earlier versions of this manuscript. We would also like to thank Alexandra Urza and Daniel Katz who offered advice on statistical analyses.

References

1. J. LeDoux, Rethinking the Emotional Brain. *Neuron*. **73**, 653–676 (2012).
2. D. Nettle, M. Bateson, The Evolutionary Origins of Mood and Its Disorders. *Curr. Biol.* **22**, R712–R721 (2012).

- 525 3. M. Mendl, O. H. P. Burman, R. M. A. Parker, E. S. Paul, Cognitive bias as an indicator of
526 animal emotion and welfare: Emerging evidence and underlying mechanisms. *Appl. Anim.*
527 *Behav. Sci.* **118**, 161–181 (2009).
- 528 4. E. J. Harding, E. S. Paul, M. Mendl, Cognitive bias and affective state. *Nature*. **427**, 312–
529 312 (2004).
- 530 5. M. Bateson, S. Desire, S. E. Gartside, G. A. Wright, Agitated Honeybees Exhibit
531 Pessimistic Cognitive Biases. *Curr. Biol.* **21**, 1070–1073 (2011).
- 532 6. M. Lagisz, J. Zidar, S. Nakagawa, V. Neville, E. Sorato, E. S. Paul, M. Bateson, M.
533 Mendl, H. Løvlie, Optimism, pessimism and judgement bias in animals: A systematic
534 review and meta-analysis. *Neurosci. Biobehav. Rev.* **118**, 3–17 (2020).
- 535 7. E. J. Bethell, A “How-To” Guide for Designing Judgment Bias Studies to Assess Captive
536 Animal Welfare. *J. Appl. Anim. Welf. Sci.* **18**, S18–S42 (2015).
- 537 8. B. O. Brilot, L. Asher, M. Bateson, Stereotyping starlings are more ‘pessimistic.’ *Anim.*
538 *Cogn.* **13**, 721–731 (2010).
- 539 9. S. M. Matheson, L. Asher, M. Bateson, Larger, enriched cages are associated with
540 ‘optimistic’ response biases in captive European starlings (*Sturnus vulgaris*). *Appl. Anim.*
541 *Behav. Sci.* **109**, 374–383 (2008).
- 542 10. J. E. C. Adriaense, J. S. Martin, M. Schiestl, C. Lamm, T. Bugnyar, Negative emotional
543 contagion and cognitive bias in common ravens (*Corvus corax*). *Proc. Natl. Acad. Sci.*
544 **116**, 11547–11552 (2019).
- 545 11. E. Bethell, A. Holmes, A. Maclarnon, S. Semple, Cognitive bias in a non-human primate:

- 546 husbandry procedures influence cognitive indicators of psychological well-being in
547 captive rhesus macaques. *Anim. Welf.* **21**, 185–195 (2012).
- 548 12. F. R. M. Oliveira, S. L. G. Nogueira-Filho, M. B. C. Sousa, C. T. S. Dias, M. Mendl, S. S.
549 C. Nogueira, Measurement of cognitive bias and cortisol levels to evaluate the effects of
550 space restriction on captive collared peccary (Mammalia, Tayassuidae). *Appl. Anim.*
551 *Behav. Sci.* **181**, 76–82 (2016).
- 552 13. H. Schlüns, H. Welling, J. R. Federici, L. Lewejohann, The glass is not yet half empty:
553 agitation but not *Varroa* treatment causes cognitive bias in honey bees. *Anim. Cogn.* **20**,
554 233–241 (2017).
- 555 14. C. Solvi, L. Baciadonna, L. Chittka, Unexpected rewards induce dopamine-dependent
556 positive emotion-like state changes in bumblebees. *Science*. **353**, 1529–1531 (2016).
- 557 15. D. Baracchi, M. Lihoreau, M. Giurfa, Do Insects Have Emotions? Some Insights from
558 Bumble Bees. *Front. Behav. Neurosci.* **11**, 157 (2017).
- 559 16. C. M. S. Plowright, Bumblebees at work in an emotion-like state. *Learn. Behav.* **45**, 207–
560 208 (2017).
- 561 17. D. R. Thomas, R. A. King, Stimulus generalization as a function of level of motivation. *J.*
562 *Exp. Psychol.* **57**, 323–328 (1959).
- 563 18. A. D. Lotfizadeh, T. L. Edwards, R. Redner, A. Poling, Motivating operations affect
564 stimulus control: A largely overlooked phenomenon in discrimination learning. *Behav.*
565 *Anal.* **35**, 89–100 (2012).
- 566 19. S. Roelofs, H. Boleij, R. E. Nordquist, F. J. van der Staay, Making Decisions under

- 567 Ambiguity: Judgment Bias Tasks for Assessing Emotional State in Animals. *Front.*
568 *Behav. Neurosci.* **10**, 119 (2016).
- 569 20. E. Leadbeater, E. H. Dawson, A social insect perspective on the evolution of social
570 learning mechanisms. *Proc. Natl. Acad. Sci.* **114**, 7838–7845 (2017).
- 571 21. H. Di Maboudi, A. B. Barron, S. Li, M. Honkanen, O. J. Loukola, F. Peng, W. Li, J. A. R.
572 Marshall, A. Cope, E. Vasilaki, C. Solvi, Non-numerical strategies used by bees to solve
573 numerical cognition tasks. *Proc. R. Soc. B Biol. Sci.* **288** (2021),
574 doi:10.1098/rspb.2020.2711.
- 575 22. L. Chittka, S. J. Rossiter, P. Skorupski, C. Fernando, What is comparable in comparative
576 cognition? *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2677–2685 (2012).
- 577 23. S. J. Shettleworth, *Cognition, Evolution, and Behavior* (Oxford University Press, ed. 2nd,
578 2010).
- 579 24. C. J. Perry, L. Baciadonna, L. Chittka, Unexpected rewards induce dopamine-dependent
580 positive emotion-like state changes in bumblebees. *Science.* **353**, 1529–1531 (2016).
- 581 25. S. K. Lynn, J. Cnaani, D. R. Papaj, Peak shift discrimination learning as a mechanism of
582 signal evolution. *Evolution.* **59**, 1300 (2005).
- 583 26. A. S. Leonard, A. Dornhaus, D. R. Papaj, Flowers help bees cope with uncertainty: signal
584 detection and the function of floral complexity. *J. Exp. Biol.* **214**, 113–121 (2011).
- 585 27. L. Chittka, The colour hexagon: A chromaticity diagram based on photoreceptor
586 excitations as a generalized representation of colour opponency. *J. Comp. Physiol. A.* **170**,
587 533–543 (1992).

- 588 28. P. Skorupski, L. Chittka, Photoreceptor spectral sensitivity in the bumblebee, *Bombus*
589 *impatiens* (Hymenoptera: Apidae). *PLoS One*. **5**, e12049 (2010).
- 590 29. D. Gomez, AVICOL, a program to analyse spectrometric data.
591 <http://sites.google.com/site/avicolprogram/> (2006).
- 592 30. A. Péter, Solomon Coder, software for behavioral coding. solomon@andraspeter.com
593 (2011).
- 594 31. F. Muth, J. S. Francis, A. S. Leonard, Modality-specific impairment of learning by a
595 neonicotinoid pesticide. *Biol. Lett.* **15**, 20190359 (2019).
- 596 32. A. G. Dyer, M. G. P. Rosa, D. H. Reser, Honeybees can recognise images of complex
597 natural scenes for use as potential landmarks. *J. Exp. Biol.* **211**, 1180–1186 (2008).
- 598 33. R Core Team: A language and environment for statistical computing. R Foundation for
599 Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>. (2020).
- 600 34. J. Pinheiro, D. Bates, S., R Core Team, nlme: Linear and Nonlinear Mixed Effects
601 Models_. R package version 3.1-128, <URL: <http://CRAN.R-project.org/package=nlme>>
602 (2020).
- 603 35. N. Colegrave, G. D. Ruxton, Statistical model specification and power: recommendations
604 on the use of test-qualified pooling in analysis of experimental data. *Proc. R. Soc. B Biol.*
605 *Sci.* **284**, 20161850 (2017).
- 606 36. R. Lenth, Emmeans: Estimated marginal means, aka least-squares means. *R Packag.*
607 *version 1.0.* <https://CRAN.R-project.org/package=emmeans> (2017).
- 608 37. H. M. Hanson, Effects of discrimination training on stimulus generalization. *J. Exp.*

- 609 *Psychol.* **58**, 321–334 (1959).
- 610 38. S. C. Andrew, C. J. Perry, A. B. Barron, K. Berthon, V. Peralta, K. Cheng, Peak shift in
611 honey bee olfactory learning. *Anim. Cogn.* **17**, 1177–1186 (2014).
- 612 39. K. W. Spence, The differential response in animals to stimuli varying within a single
613 dimension. *Psychol. Rev.* **44**, 430–444 (1937).
- 614 40. A. G. Dyer, L. Chittka, Fine colour discrimination requires differential conditioning in
615 bumblebees. *Naturwissenschaften.* **91**, 224–227 (2004).
- 616 41. S. Ghirlanda, M. Enquist, A century of generalization. *Anim. Behav.* **66**, 15–36 (2003).
- 617 42. V. Vergoz, E. Roussel, J.-C. Sandoz, M. Giurfa, Aversive Learning in Honeybees
618 Revealed by the Olfactory Conditioning of the Sting Extension Reflex. *PLoS One.* **2**, e288
619 (2007).
- 620 43. A. Dornhaus, L. Chittka, Information flow and regulation of foraging activity in bumble
621 bees (*Bombus spp.*). *Apidologie.* **35**, 183–192 (2004).
- 622 44. J. M. Townsend-Mehler, F. C. Dyer, K. Maida, Deciding when to explore and when to
623 persist: a comparison of honeybees and bumblebees in their response to downshifts in
624 reward. *Behav. Ecol. Sociobiol.* **65**, 305–312 (2011).
- 625 45. J. M. Townsend-Mehler, F. C. Dyer, An integrated look at decision-making in bees as
626 they abandon a depleted food source. *Behav. Ecol. Sociobiol.* **66**, 275–286 (2012).
- 627 46. C. T. Hemingway, F. Muth, Label-based expectations affect incentive contrast effects in
628 bumblebees. *Biol. Lett.* **18** (2022), doi:10.1098/rsbl.2021.0549.
- 629 47. M. J. Boisvert, A. J. Veal, D. F. Sherry, Floral reward production is timed by an insect

- 630 pollinator. *Proc. R. Soc. London B Biol. Sci.* **274** (2007).
- 631 48. M. Domjan, *The Principles of Learning and Behavior* (Wadsworth, 5th editio., 2003).
- 632 49. E. Verbeek, D. Ferguson, C. Lee, Are hungry sheep more pessimistic? The effects of food
633 restriction on cognitive bias and the involvement of ghrelin in its regulation. *Physiol.*
634 *Behav.* **123**, 67–75 (2014).
- 635 50. O. Burman, R. McGowan, M. Mendl, Y. Norling, E. Paul, T. Rehn, L. Keeling, Using
636 judgement bias to measure positive affective state in dogs. *Appl. Anim. Behav. Sci.* **132**,
637 160–168 (2011).
- 638 51. A. Destrez, V. Deiss, F. Lévy, L. Calandreau, C. Lee, E. Chaillou-Sagon, A. Boissy,
639 Chronic stress induces pessimistic-like judgment and learning deficits in sheep. *Appl.*
640 *Anim. Behav. Sci.* **148**, 28–36 (2013).
- 641 52. J. E. Dunsmoor, R. Paz, Fear Generalization and Anxiety: Behavioral and Neural
642 Mechanisms. *Biol. Psychiatry.* **78**, 336–343 (2015).
- 643 53. D. Struyf, J. Zaman, B. Vervliet, I. Van Diest, Perceptual discrimination in fear
644 generalization: Mechanistic and clinical implications. *Neurosci. Biobehav. Rev.* **59**, 201–
645 207 (2015).
- 646 54. L. Pompilio, A. Kacelnik, S. T. Behmer, State-dependent learned valuation drives choice
647 in an invertebrate. *Science.* **311**, 1613–5 (2006).
- 648 55. C. Tencate, C. Rowe, Biases in signal evolution: learning makes a difference. *Trends*
649 *Ecol. Evol.* **22**, 380–387 (2007).
- 650 56. D. R. Thomas, K. Mood, S. Morrison, E. Wiertelak, Peak shift revisited: A test of

alternative interpretations. *J. Exp. Psychol. Anim. Behav. Process.* **17**, 130–140 (1991).

57. A. L. Whittaker, T. H. Barker, A consideration of the role of biology and test design as confounding factors in judgement bias tests. *Appl. Anim. Behav. Sci.* **232**, 105126 (2020).

58. T. Enkel, D. Gholizadeh, O. von Bohlen und Halbach, C. Sanchis-Segura, R. Hurlemann, R. Spanagel, P. Gass, B. Vollmayr, Ambiguous-Cue Interpretation is Biased Under Stress- and Depression-Like States in Rats. *Neuropsychopharmacology*. **35**, 1008–1015 (2010).

59. J. Zidar, I. Campderrich, E. Jansson, A. Wichman, S. Winberg, L. Keeling, H. Løvlie, Environmental complexity buffers against stress-induced negative judgement bias in female chickens. *Sci. Rep.* **8**, 5404 (2018).