



Cite this article: Bethell EJ, Koyama NF. 2015
Happy hamsters? Enrichment induces positive
judgement bias for mildly (but not truly)
ambiguous cues to reward and punishment in
Mesocricetus auratus. *R. Soc. open sci.*
2: 140399.
<http://dx.doi.org/10.1098/rsos.140399>

Received: 22 October 2014

Accepted: 3 July 2015

Subject Category:

Psychology and cognitive neuroscience

Subject Areas:

behaviour/cognition/psychology

Keywords:

animal welfare, judgement bias, emotion,
environmental enrichment, psychological
wellbeing, Syrian hamster

Author for correspondence:

Emily J. Bethell

e-mail: e.j.bethell@ljmu.ac.uk

Electronic supplementary material is available
at <http://dx.doi.org/10.1098/rsos.140399> or via
<http://rsos.royalsocietypublishing.org>.

Happy hamsters?

Enrichment induces positive judgement bias for mildly (but not truly) ambiguous cues to reward and punishment in *Mesocricetus auratus*

Emily J. Bethell and Nicola F. Koyama

School of Natural Sciences and Psychology, Liverpool John Moores University,
Byrom Street, Liverpool L3 3AF, UK

Recent developments in the study of animal cognition and emotion have resulted in the ‘judgement bias’ model of animal welfare. Judgement biases describe the way in which changes in affective state are characterized by changes in information processing. In humans, anxiety and depression are characterized by increased expectation of negative events and negative interpretation of ambiguous information. Positive wellbeing is associated with enhanced expectation of positive outcomes and more positive interpretation of ambiguous information. Mood-congruent judgement biases for ambiguous information have been demonstrated in a range of animal species, with large variation in the way tests are administered and in the robustness of analyses. We highlight and address some issues using a laboratory species not previously tested: the Syrian hamster (*Mesocricetus auratus*). Hamsters were tested using a spatial judgement go/no-go task in enriched and unenriched housing. We included a number of controls and additional behavioural tests and applied a robust analytical approach using linear mixed effects models. Hamsters approached the ambiguous cues significantly more often when enriched than unenriched. There was no effect of enrichment on responses to the middle cue. We discuss these findings in light of mechanisms underlying processing cues to reward, punishment and true ambiguity, and the implications for the welfare of laboratory hamsters.

1. Introduction

Accurate assessment of animal emotions is an important goal of welfare science [1,2] and central to the refinement of the use of animals in scientific research [3,4]. In the last decade, methods used to assess the link between affective state and cognitive processes in humans [5,6] have been successfully adapted and applied to animals. The judgement bias model of animal welfare [7–10] argues that emotion states may be measured in animals, as they are in humans, as a function of altered cognitive judgements about the likely rewarding or punishing nature of ambiguous stimuli. The value of judgement bias measures over traditional physiological and behavioural approaches is their potential to distinguish emotional valence (positive versus negative emotions). Given the large number of laboratory rodents used each year, it is important to develop proxy measures of affective state to aid in welfare assessment. Thus far, judgement bias studies on laboratory rodents have focused on rats (e.g. [7]), with two studies on mice [11,12] and none on hamsters.

Judgement bias in animals has been most often measured using a ‘go/no-go’ task in which animals are trained to discriminate between a ‘positive’ and a ‘negative’ cue and then tested on their responses to ambiguous cues [7]. Ambiguous cues possess characteristics intermediate to both the rewarded and non-rewarded stimuli, and it is responses (‘go’ or ‘no-go’) to these that quantify judgement bias. Recent studies with animals including rats [7,13–16], mice [11,12], cats [17], dogs [18–24], dairy calves [25,26], sheep [27–38], chickens [39–42], honeybees [43], rhesus macaques [44], marmosets [45], starlings [46], pigs [47–49], horses [50] and goats [51] demonstrate that emotion-mediated judgement biases may be detected using species-specific variants of the go/no-go task. So far, studies have revealed that manipulations presumed to create a negative state (such as disrupted housing conditions [7] or dehorning in calves [26]) lead to reduced responses (more ‘no-go’s) to the ambiguous probes. This negative shift in judgement bias is presumed to arise from a negative shift in underlying emotion state. Positive manipulations (e.g. addition of environmental enrichment [46]) generally lead to increased responses (more ‘go’s) at one or more of the ambiguous probes. This is presumed to arise from positive shifts in judgement bias and underlying emotion state.

A criticism of the go/no-go approach (discussed more fully in [9]) is that results may not reflect judgements about ambiguity, but instead reveal changes in motivation (due to satiation or learning), arousal or activity (due to treatment effects or task fatigue). A solution to this is to include control trials and a variable reinforcement ratio (VRR) to maintain motivation and reduce learning effects. Triangulating judgement bias with traditional measures of affect (e.g. open field (OF) test and light–dark emergence (LDE) tests) and recording behaviour during testing would allow treatment effects and task fatigue to be considered.

The active choice task is an alternative to the go/no-go task. The response (‘go’ or ‘no-go’) is the same to both learned stimuli, providing a built-in control for arousal effects. The active choice test has been validated with rats [52–64], starlings [65,66], pigs [67], chickens [68], grizzly bear [69] and capuchin monkey [70]. Results from these studies are varied, and several report arousal effects (e.g. [67,68]). Care must be taken to ensure that the perceived affective difference between the two reinforcers is sufficiently great to detect any shift in judgement bias [9]. For example, in the study of Parker *et al.* [64], the ‘positive’ cue signalled two pellets and the ‘negative’ cue signalled one pellet. In this case, repeated training might more rapidly lead to satiation, a decrease in reward value or reduced ability to distinguish between high and low reward. A variant of the active choice task is to use negative reinforcement, e.g. electric shock [52,55–57,59–63], but ethical issues and confounding effects of fear and stress make this method unsuitable for welfare assessment. Results from studies using negative reinforcement, which tends to be pharmacological manipulations conducted in rats, support the judgement bias model of emotion–cognition interaction in non-human animals.

Both approaches are susceptible to a number of limitations. These include: (i) cueing effects (e.g. odour cues from food), (ii) the contribution of other aspects of cognition such as risk-taking behaviour or attention for threat cues, (iii) learning the reward value of the ambiguous probes, (iv) lack of a consistent and robust statistical approach, and (v) a lack of consistency in *a priori* predictions about the direction of change in bias and associated emotion state (for example, increased responding to probes in the ‘stress’ condition is interpreted as indicating relief following the termination of the stressor in Doyle *et al.* [27] and Sanger *et al.* [31]). These limitations are rarely addressed in study design, but failure to do so may lead to misinterpretation of results.

Of the two task designs, the spatial go/no-go task has been the most widely validated and applied across species. Within Mendl *et al.*’s [10] integrative framework, emotional valence is directly associated with approach towards reward (associated with excitement and happiness) and avoidance of (or

inhibition of response to) threat or punishment (associated with fear and anxiety). These underlying associations may facilitate training on the 'go/no-go' task (compared to learning to discriminate high versus low reward in the active choice task), making go/no-go an ideal method to develop a judgement bias task in a new species. Both tasks have been developed using auditory cues (e.g. [48,67]) and visual cues (e.g. [44,70]). In addition, the go/no-go task has been validated with spatial cues (e.g. [7,39,50]), tactile cues (e.g. [53,58]) and odour cues (e.g. [43]). Among rodents, spatial cues [7,13–16] and auditory tones [52,55–57,60–63] have been used successfully with rats, while spatial cues [39,41] and odour cues [12] have been used successfully with mice. Spatial cues are therefore suitable for first development of a task with a new species of rodent.

For the purposes of welfare assessment, the addition of environmental enrichment has been the most common positive manipulation in judgement bias studies. The majority of studies where enrichment was added reported a positive shift in judgement bias. Studies have differed in the type of enrichment added (e.g. increased space, perches, shelters, foraging substrate [44,46]), the duration of treatment (e.g. 2 h in grizzly bears [68], one week in rats [16,53], 10 days in rhesus macaques [44]), whether enrichment is removed (starlings [46]) and whether a switch in housing occurs, for example from enriched to basic and back to enriched (starlings [66] and pigs [48]). There are some exceptions: no judgement bias shift was detected in laying hens [40]; however, this study used a between-subjects design and the individual differences in fearfulness, body condition and neophobia appeared to have a greater impact on performance than the enrichment treatment. Of the four studies that employed a switch in housing conditions [46,48,65,66], three detected a treatment-induced shift in judgement bias [46,48,65]. In starlings, a pessimistic bias was noted following a reduction in enrichment but no change was noted following addition of enrichment [46]; however, a later study using similar enrichment items and increased space found an optimistic bias in enriched cages [65]. Both these studies used a cross-over design with half the subjects experiencing enriched then unenriched housing and the other half unenriched then enriched housing. Douglas *et al.* [48] also used a cross-over design in their study of pigs, with half experiencing the switch enriched-to basic-to enriched and the other half the opposite housing manipulation. Pigs currently in enriched housing interpreted the ambiguous cues more optimistically. Moreover, there was an interaction between current and past environment: pigs that were first in enriched housing reacted more negatively when subsequently housed in a barren environment. In the fourth study involving a switch in housing conditions [66], individual differences in stereotypies were found to predict pessimism. These studies highlight the importance of within-subjects study design, appropriate enrichment to create a shift of sufficient magnitude that is not masked by individual differences, and using a cross-over design to control for order effects.

Here, we aim to validate a judgement bias test for use in hamsters (*Mesocricetus auratus*) and to tackle the above criticisms within spatial judgement bias paradigm [9,13].

We measured judgement bias using a balanced, cross-over design [48] using changes in enrichment to validate the task. Hamsters were initially housed in standard laboratory cages with basic enrichment and then received highly enriched housing (E) first or experienced removal of some of the basic enrichment (R). Enrichment devices were selected according to published data from hamster preference tests and natural behaviour [71–73]. A week later hamsters moved to the opposite housing conditions allowing comparison of the two groups (ER or RE). In response to our critique above we proceeded as follows. (i) We excluded olfactory cueing effects by introducing a VRR during training. (ii) We statistically controlled for confounding effects of arousal, by including day and block of testing in our model (see below), and controlling for other aspects of cognition by triangulating data with traditional tests of affect in rodents such as OF and LDE [74], neophobia [75] tests and recording in-trial behaviours traditionally used to measure emotion state or arousal in laboratory rodents (e.g. [7,16,73–77]). (iii) To reduce the likelihood of hamsters learning that probes were never rewarded we used a VRR during training, and included unreinforced control trials at the positive and negative stimulus locations during testing. (iv) We applied a robust information-theory approach (multi-model inference) using linear mixed effects models [78–81] to assess the relative contribution of different variables to responses on the task. Generalized linear mixed models incorporate random effects, test fixed effects using generalized least squares results in more powerful tests, handle non-normal data by using link functions and exponential family distributions (e.g. normal, Poisson or binomial), and are able to handle missing data and unequal spacing over time [78,79]. Linear mixed effects models are beginning to appear in judgement bias studies (e.g. [22,51,82]) and provide a robust and flexible approach to analysis that will aid interpretation of output in light of potential confounding factors. (v) We predicted that the addition of enrichment would be associated with a positive affective state and corresponding optimistic judgement bias, while the removal of enrichment would be associated with a more negative affective state and judgement bias.

2. Material and methods

2.1. Study animals and housing

Thirty captive-born male Syrian hamsters (*Mesocricetus auratus*) took part in the study. Hamsters were weaned at 21 days and retained in eight same-sex sibling groups throughout the study. Hamsters were 8 weeks old at the start of the study and 13 weeks old by the end of the study. The study was conducted with one cohort of hamsters in 2011 ($n = 17$) and a second cohort in 2013 ($n = 13$). In both years, sibling groups were housed in the same style of caging ($56 \times 38 \times 22$ cm) in the same air-conditioned room kept at a constant 20–24°C on a 12 L:12 D cycle, with humidity maintained between 45 and 70%. Hamsters had *ad libitum* access to water and dry rodent pellets (Eurodent Diet 22% 5LF5, PMI Nutrition International Plc) and were handled twice per week from weaning for routine cage cleaning and health inspections by care staff. All animals that completed the study were assessed by care staff as healthy throughout.

2.2. Apparatus

Training and testing sessions were conducted in a high-sided blue opaque plastic arena ($100 \times 80 \times 85$ cm) in the housing room during the first hours of the active (dark) phase using low level light. The arena had five holes drilled along one wall, with attachments on the outside so that on each trial a drinker could be attached at any location with only the spout protruding into the arena. Three drinkers were used during the experiment. Two drinkers were used for training and maintenance during testing; one drinker contained 0.3 M sugar water and one drinker contained 0.01 M quinine hydrochloride (QHCl) solution. A third drinker was used for experimental trials ('control' and 'probe' trials) and was kept empty throughout.

2.3. Habituation

The timeline for the study is shown in [figure 1](#). Initially, all animals were familiarized with the arena. One familiarization session was conducted on each of four days (Tuesday–Friday) for each sibling group. The sibling group was placed in the arena with scattered food and a bottle containing sugar water attached at the rewarded location (left or right) for that group. Animals were monitored for 5 min to ensure all animals had drunk from the drinker. At the end of each daily session, the number of faecal pellets left in the arena was counted.

2.4. Training

2.4.1. Phase A (weeks 1 and 2: days 1–10)

We employed a spatial judgement task used previously ([13], [figure 2](#)). In brief, hamsters were trained to approach a drinker at one location (left or right) to obtain a reward ('go' towards sugar water), and to refrain from approaching a drinker at the other location (right or left) to avoid an aversive liquid ('no go' towards QHCl). A trial began when the hamster was placed at a start point at the centre back of the arena. The trial ended when the hamster approached the drinker, or after a predetermined number of seconds, whichever occurred sooner. On the first day of Phase A, each hamster took part in 8×120 s trials. As hamsters learned the task and became faster to approach, sessions were sequentially adjusted to 8×60 s trials and then 10×30 s trials. This adjustment was made according to individual response latencies so that trial duration was reduced quickly for fast-responding hamsters, and more slowly for hamsters that needed the extra time to approach. If the hamster approached to within 3 mm of the drinker in the predetermined trial duration this was scored as a 'go'. If the hamster did not approach the drinker in the predetermined trial duration this was scored as a 'no-go'. An equal number of sugar and QHCl trials were run within each daily training session for each hamster. The first and last trials were always sugar trials so that each session began and ended with a positive association with the arena (same throughout all subsequent phases). A hamster was removed once he had finished drinking. Criterion for learning the discrimination task was average latency to approach the sugar drinker being shorter than average latency to approach the QHCl drinker on each of three consecutive days.

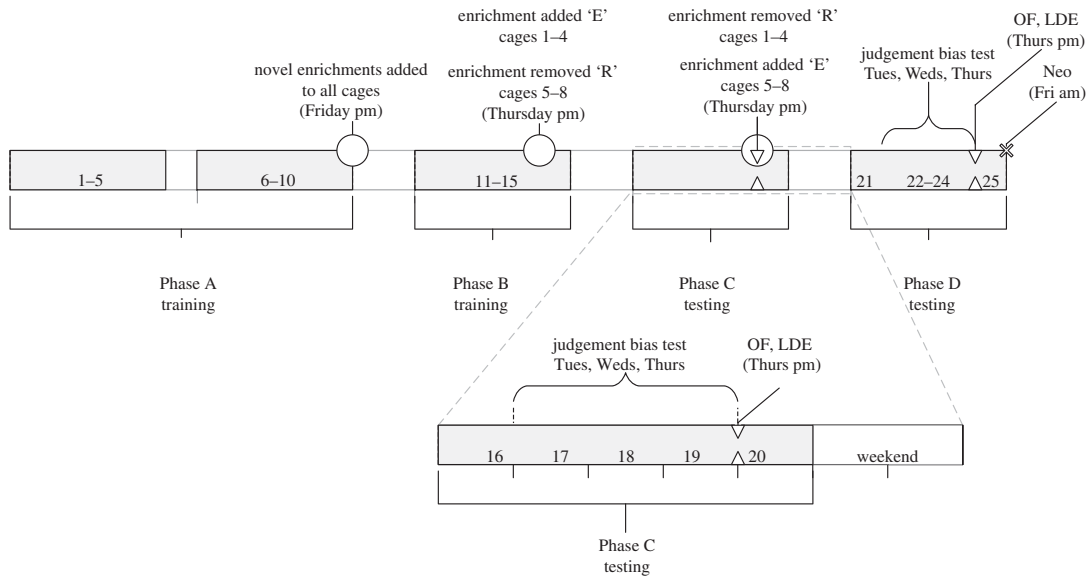


Figure 1. Timeline for the study, conducted over five weeks. Day numbers are shown: discrimination training (days 1–15); judgement bias testing (days 17–19 and 22–24; maintenance trials) days 16, 20 and 21). ○, Enrichment manipulation; additional behavioural tests—△, OF and LDE tests (days 19 and 24) and ×, neophobia test (Neo) (day 25).

2.4.2. Phase B (week 3; training days 11–15)

Training continued as for Phase A with the introduction of control trials to create a 60% variable reinforcement ratio. A control trial was a trial on which the empty drinker was presented at either of the two trained (sugar/QHCl) locations. These trials were included to ensure animals were responding to drinker location and not to odour cues. A daily training session in Phase B comprised 16×30 s trials divided into two blocks (table 1). Block 1 contained nine trials: five reinforcement trials (2 sugar : 2 QHCl, plus one) pseudorandomized with four control trials ($2C+ : 2C-$). Block 2 comprised a total of seven trials: three reinforcement trials (1 sugar : 1 QHCl, plus one) pseudorandomized with four control trials ($2C+ : 2C-$). Care was taken to remove the hamster as soon as he reached the empty drinker on control trials to reduce the possibility of animals becoming frustrated if left to attempt to drink from an empty drinker.

2.5. Judgement bias experiment

2.5.1. Phases C and D (weeks 4 and 5; days 16–25)

The judgement bias testing phases followed the design used in Phase B with the addition of experimental probe trials on the Tuesday, Wednesday and Thursday of each week (Phase C testing days: 17–19; Phase D testing days: 22–24). An experimental probe trial was a trial on which the empty drinker was presented at one of the three intermediate locations. A daily testing session in Phase C comprised 16×30 s trials divided into two blocks. Block 1 comprised four reinforcement trials (2 sugar : 2 QHCl, pseudorandomized with two control trials ($C+$, $C-$) and three probe trials ($P+$, P_{mid} and $P-$). Block 2 comprised two control trials ($C+$, $C-$), pseudorandomized with three probe trials ($P+$, P_{mid} and $P-$), and two reinforcement trials ($Q-$, $S+$). Hamsters therefore had six experimental probe trials (two at each location: $P+$, P_{mid} and $P-$) in each daily session. Hamsters were removed as soon as they reached the empty drinker on all probe and control trials. Phase C comprised three testing days which generated six data points per probe location per hamster. On the Monday and Friday (days 16 and 20), hamsters received sessions as in Phase B. Phase D (week 5) was an exact repetition of Phase C, except that housing treatments were switched as explained below.

2.6. Housing treatments

At the start of the study hamsters were housed in their home cages provisioned with a thin layer of aspen chip substrate and lignocel bedding material (both from IPS, London, UK), a free-standing barred running wheel (diameter 15 cm) and two spirally wound cardboard tubes (20×10 cm). In the

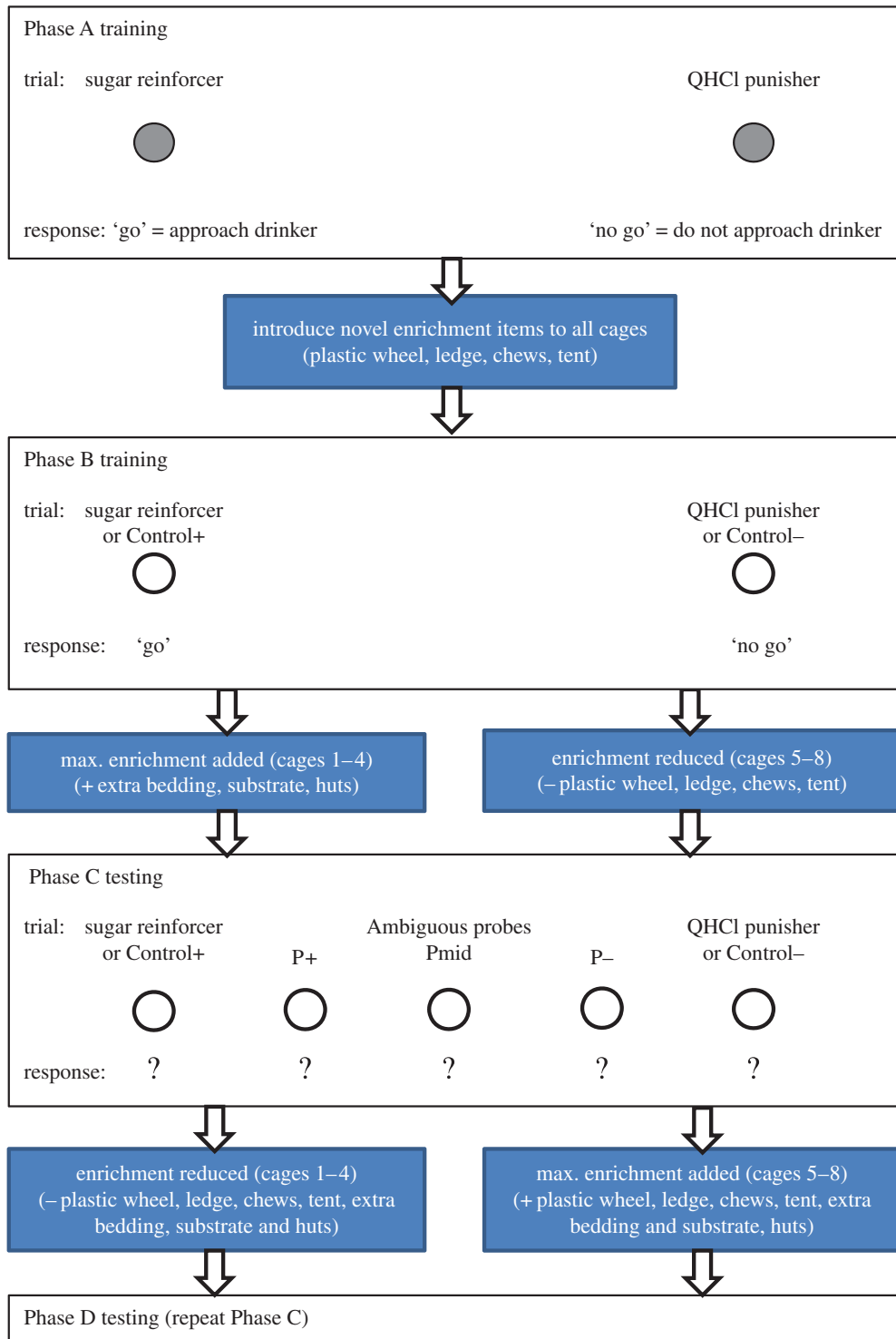


Figure 2. The judgement bias task. During training (Phase A) hamsters learned to approach the sugar drinker ('go') and not to approach ('no go') the QHCl drinker. At the end of Phase A, all animals received a basic provision of novel enrichment items. Training on the go/no-go task was maintained in Phase B with the inclusion of the empty drinker on some trials to create a 50% VRR (empty drinker trials: 'Control +' and 'Control -'). Towards the end of Phase B half of the sibling groups received the maximum complement of enrichment, while the novel enrichment items were removed from the other half of the sibling groups. Animals were tested on the judgement bias task on the following week (Phase C) with the presentation of the empty drinker at all five locations, including the three ambiguous intermediate probe locations. Enrichments were then switched and animals tested again in Phase D.

Table 1. The blocked design for training in Phase B and testing in Phases C and D. Each daily session began (block 1: trials 1 and 2) with a rewarding sugar trial followed by a QHCl trial to remind hamsters of the task contingencies. Each daily session ended (block 2: trial 16) with a sugar trial to end on a positive association and encourage maintained performance over subsequent days. Order of all other trials was pseudorandomized within each block. S+ = Sugar; Q- = QHCl; C+ = Control at sugar location; C- = Control at QHCl location; P+ = ambiguous probe closest to the sugar location; Pmid = middle probe; P- = ambiguous probe closest to the QHCl location. Trial numbers are shown in parentheses.

Phase	block			
	1		2	
B (trial n)	S+, Q- (1, 2)	2C+, 2C-, S+, Q- (3–9)	2C+, 2C-, S+, Q- (10–15)	S+ (16)
C/D (trial n)	S+, Q- (1, 2)	C+, C-, P+, Pmid, P- S+, Q- (3–9)	C+, C-, P+, Pmid, P-, Q- (10–15)	S+ (16)

current study, emotion state was manipulated by adding or reducing enrichment. Choice of enrichment devices was informed by published work showing that their presence leads to changes in behaviour assumed to indicate preferences and improved welfare in hamsters [71–73]. Additional enrichment items comprised deeper aspen substrate and extra nesting material [71], two coloured transparent plastic huts (10 × 12 cm), a suspended hamster tent (15 × 12 cm), four hamster gnaw sticks and a wooden ledge (18 × 13 cm: all of which increase opportunity for natural and exploratory behaviour [72,73,77]). The metal barred wheel was replaced with a larger solid-floor plastic silent running ball-bearing wheel (16.5 cm, Silent Spinner Regular [83]).

Enrichment items that hamsters had not encountered previously (hamster gnaw sticks, tent, ledge and plastic running wheel) were added to all home cages on the Friday afternoon of Phase A (day 10 after training: figure 1). Animals were left to habituate to the novel enrichment devices over the weekend and throughout the next week (Phase B on the following Monday–Friday). Introducing novel enrichment items during training should reduce potentially confounding context-dependent learning effects on animals' performance in the two treatments during later testing sessions [84]. After training on the Thursday of Phase B, the maximum complement of enrichment items was added to half of the cages (Treatment E: 'enrichment added') while the novel enrichment items were removed from the other half of the cages (Treatment R: 'enrichment removed'). Discrimination training was then maintained on the Friday and the following Monday. This ensured all animals had undergone some form of 'change' in housing environment prior to the start of testing in Phase C). Testing in Phase C was conducted on three days (Tuesday–Thursday), which corresponded to the fifth, sixth and seventh days of having a full complement of enrichment devices added to the home cage for four sibling groups, and the fifth, sixth and seventh days following reduction in enrichment from the home cage for four sibling groups. Following the third day of testing in Phase C (Thursday), housing treatments were switched so that the hamsters receiving Treatment E now received Treatment R, and vice versa. Maintenance on the task following the procedure described for Phase C was conducted on the Friday and following Monday. Hamsters were then tested again in Phase D on the following Tuesday–Thursday.

2.7. Behavioural measures

Behaviour during experimental trials (probe and control trials on the Tuesday–Thursday in Phases C and D) was recorded to explore possible treatment-induced differences in arousal and stress-related activity. Arousal was assessed by calculating a 'locomotion score': the arena floor was marked out as a grid of 12 squares, and the number of squares the hamsters entered prior to reaching the drinker (or until the end of the 30 s trial on 'no-go' trials) was recorded. General stress-related activity was assessed by calculating an 'anxiety score': behaviours considered to indicate stress or anxiety in hamsters (autogrooming, attempting to climb arena walls to 'escape', freeze, stereotypy, rearing and sniffing [7,74,76,77]) were recorded using 1/0 scoring [85] (range 0–6) during each trial.

OF, LDE and neophobia tests were conducted during Phase C and D (figure 1). Each hamster took part in one 2 min trial on each of an OF and an LDE test on the Thursday afternoon after judgement bias testing in each phase. The OF apparatus measured 1 m². Total number of squares entered and time spent in the middle square were recorded. In the LD emergence test, the hamster was placed in a wooden hut (15 cm³) and time taken to emerge into a well-lit exterior was recorded. Neophobia was tested on the Friday of Phase D. Each hamster took part in nine trials in the arena, during which the time to approach

the sugar drinker when a novel object was next to the drinker ($n = 4$ trials) or not present ($n = 5$ trials) was recorded. A different novel object was used on each trial and objects varied in shape and colour and were unfamiliar to the hamsters.

2.8. Data analysis

Each hamster took part in a total of 60 experimental trials (30 in each housing treatment): 36 probe trials (12 at each of the three probe locations) and 24 control trials (12 at each of the C+ and C− locations). These were dispersed amongst a total of 36 reinforcement trials (Sugar and QHCl). A criterion of more than 50% correct responses ('go' to the C+ and 'no-go' to the C−) on control trials across testing sessions was set for each hamster to be included in the analyses. We used a generous criterion for inclusion in the analyses since we wanted to explore the potential influence of a range of factors on performance in the judgement bias task. We did not want to bias findings by selectively removing hamsters who were perceived to be poor performers if changes in performance reflected meaningful responses to the enrichment treatments. However, we wanted to remove hamsters who stopped responding on the task altogether.

2.8.1. Model development

All recorded explanatory variables were considered for inclusion in the maximal model. These were: treatment (E, R), trial type (C+, P+, Pmid, P−, C−), testing block (blocks 1 and 2), year of testing (2011, 2013), day of testing (Tuesday, Wednesday or Thursday), order of testing (E first, R first), trial number (trial 1–16), experimenter ID, hamster ID (24 hamsters), sibling group (eight groups) and speed to learn the discrimination task (3–16 days). Preliminary assessment for collinearity between these variables was conducted in the statistical package R v. 3.0.1 [86]. Where two or more variables covaried, these were serially entered into a series of generalized linear models (GLMs) and the variable resulting in the GLM with the lowest AIC_c value was retained [80,81].

Preliminary analyses revealed that testing block correlated with trial number ($r = 0.63$); block had the lowest AIC_c and was retained. Week of testing correlated strongly with day of testing ($r > 0.9$); day was retained. Sibling group, year in which the data were collected, experimenter ID and speed to learn the discrimination task were correlated with each other (all $r_s > 0.36$); sibling group was retained. Order of testing was correlated with sibling group and hamster ID (as siblings were housed together and therefore underwent the same housing treatment together: $r_s > 0.55$). Because previous studies have found an effect of order on responses during judgement bias tests, we conducted a glm (order \times treatment) to explore possible order effects. This revealed a significant order \times treatment interaction (estimate = 0.512, error = 0.23, $z = 2.19$, $p = 0.029$). To control for this effect in the final model order of testing was retained with sibling group and male ID as nested random factors.

A maximal model was defined using the glmer package in R [86]. Go/no-go responses were the dependent binary variable. Treatment, trial type and block were the fixed factors; day of testing was included as a random factor and order of testing, sibling group and male ID were included as nested random factors. Treatment, trial type and block were also included as interaction terms:

$$\text{glmer}(\text{NoGo} \sim \text{Treatment} \times \text{Trial Type} \times \text{Block} + (1|\text{Day Test}) + (1|\text{Order/Sibgrp/Male ID}), \\ \text{family} = \text{binomial}).$$

These explanatory variables were used to create a total of 15 candidate models all with a logit link function and a binomial error distribution (electronic supplementary material, table S1). Models were compared and selected using the AIC_c, with the lowest AIC_c value indicating the best model fit. All models within 2 AIC_c of the best model were retained as the final subset of models [80,81].

To test for an effect of fixed factors featured in the best-fit model(s) at each of the probe and control locations separately, test models were created for each probe and control location (with random factors held constant), and with each fixed factor from the best model(s) included separately. Test models were then tested against the null model for each probe and control location using the anova function in R [86]. Criterion for significance was $p < 0.05$.

An example of the null model used with data for trial location separately is as follows:

$$\text{NoGo} \sim (1|\text{Day Test}) + (1|\text{Order/Sibgrp/Male ID}).$$

2.9. Behavioural measures

Behavioural data for hamsters while in the arena were analysed separately as they were collected on a subset of the experimental trials. For each hamster, the locomotion score was divided by the trial duration (latency to approach the drinker, or 30 s in the case of no-gos) to give a locomotion index for each trial. For behaviour in the arena, anxiety score was divided by the trial duration. Data were tested for normality and sqrt transformed where necessary. Locomotion and anxiety indices were then entered separately as dependent variables in the 15 candidate models and the best-fit models selected as described above.

LD emergence, OF and neophobia test data were also analysed separately as these data were collected at the end of each treatment week (Phases C and D—Phase D only for neophobia data). Owing to the small number of trials, these data were analysed using non-parametric Wilcoxon tests. Neophobia data were tested following a model selection procedure as outlined above using the lmer function in R [87]. Time to approach the drinker was the dependent variable (since hamsters approached the drinker on all novelty trials and most sugar trials), with fixed factors treatment (E or R) and trial type (sugar or sugar + novel) with nested random factors sibling group and hamster ID.

3. Results

3.1. Habituation

Animals defecated in the arena on the first day of familiarization, but not thereafter. Hamsters readily approached and drank from the drinker within each 5 min session.

3.2. Learning the discrimination task (days 1–16)

Performance data for hamsters during Phases A and B are given in the electronic supplementary material, table S2 and table S3. On average, hamsters reached criterion for learning the discrimination task in 8 days (79 trials; range 3–16 days). Eighteen hamsters reached criterion for learning the discrimination task before the end of Phase A (day 10). Five hamsters reached criterion before the end of Phase B (day 15). One hamster reached criterion on the first day in Phase C (day 16).

3.3. Performance on control trials in Phase B

Group-level accuracy for discrimination and control trials during phase B are given in the electronic supplementary material, table S3. Hamsters correctly approached the sugar location on 84% of trials and the sugar control (C+) on 71% of trials. They correctly did not approach the QHCl location on 64% of trials and the QHCl control (C–) on 69% of trials.

Maintenance of performance on the discrimination task during testing in Phases C and D: of the 30 hamsters that took part in testing, 24 hamsters ($n = 15$ from 2011 and $n = 9$ from 2013) from eight sibling groups performed to criterion on control trials during the judgement bias test in Phases C and D and were included in the final analysis. Performance data for these 24 are given in the electronic supplementary material, table S4. Of the six hamsters that were not included, one was removed from the study due to health reasons, and the remaining five failed to approach the C+ during the enrichment treatment ($n = 3$) or during both treatments ($n = 2$).

3.4. Model selection

The factors treatment, trial type and block of testing appeared in the best-fitting models, explaining 84% of the variance in the data (table 2).

3.5. Localizing significant effects

3.5.1. Probe trials

Anovas comparing the effect of treatment against the null model were conducted for each probe separately (table 3). These revealed a strong effect of treatment at the two outer probes, with no effect at

Table 2. AIC_c-ranked candidate model set showing the relative importance of fixed effects (treatment, trial type, block) and random effects (order of testing, sibling group, hamster ID and day) in explaining tendency to approach drinkers of known and uncertain reward and punishment value.

model fixed effects	random effects (/nested)	d.f.	log likelihood	AIC	delta	weight
treatment, trial × block	(order/sibling grp/ID), day	15	−772.932	1576.2	0.00	0.622
treatment, trial	(order/sibling grp/ID), day	11	−778.072	1578.3	2.11	0.216

Table 3. Comparison of test models (fixed factor of either treatment or block) with the null model for each of the three probes, two control locations and the reinforcing sugar or QHCl trials. Values for the null model are given in parentheses. Test models which were shown by ANOVA to deviate significantly from the null are indicated by asterisks.

trial	intercept	d.f.	log likelihood	AIC _c	delta	weight
probe (+)						
treatment	−1.268	6	−143.38	299.1*	0.00	0.589
block	−2.743	6	−143.98	300.3*	1.19	0.325
null	(−0.879)	(5)	(−146.35)	(302.9)	(3.84)	(0.09)
probe (mid)						
treatment	−0.512	6	−166.48	345.3	6.37	0.036
block	−2.381	6	−163.30	338.9*	0.00	0.870
null	(−0.453)	(5)	(−166.58)	(343.4)	(4.46)	(0.094)
probe (−)						
treatment	−0.508	6	−173.17	358.7*	0.00	0.745
block	0.244	6	−175.46	363.3	4.59	0.075
null	(−0.202)	(5)	(−175.64)	(361.5)	(2.85)	(0.18)
control (+)						
treatment	−1.707	6	−129.28	270.9	0.25	0.40
block	−1.477	6	−130.20	272.7	2.10	0.16
null	(−1.484)	(5)	(−130.20)	(270.6)	(0.00)	(0.45)
control (−)						
treatment	0.571	6	−161.51	335.4	0.71	0.29
block	−0.130	6	161.49	335.3	0.66	0.30
null	(0.732)	(5)	(−162.20)	(334.6)	(0.00)	(0.41)
reinforcer (S+)						
treatment	−2.807	6	−138.62	289.4	9.32	0.00
block	−3.990	6	−133.96	280.1**	0.00	0.97
null	(−2.695)	(5)	(−138.84)	(287.8)	(7.70)	(0.02)
punisher (QHCl)						
treatment	0.509	6	−226.53	465.3	1.38	0.23
block	0.246	6	−226.25	464.7	0.82	0.31
null	(0.602)	(5)	(−226.88)	(463.9)	(0.00)	(0.46)

* $p < 0.05$; ** $p < 0.01$.

the middle probe: hamsters approached the two outer probes more often in treatment E than R (figure 3). Comparison of the effect of block against the null model for each probe separately revealed a strong effect of block on responses to the Pmid and P+, but not the P−: hamsters made fewer approaches to the Pmid and P+ in the second compared to the first block.

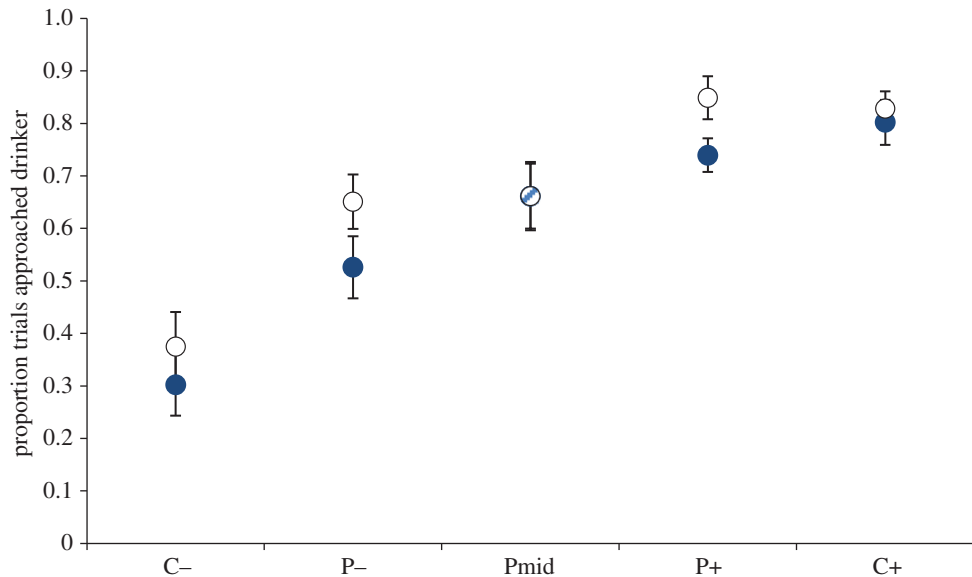


Figure 3. Mean proportion (\pm s.e.) of approaches to each of the probe and control locations when enrichment had been added (open circles) or reduced (closed circles). Hamsters approached the outer probes (P— and P+) significantly more often following addition of enrichment compared to when enrichment had been reduced. Hamsters approached the middle probe equally often in both treatments. CQHCl control; P— probe closest to the QHCl location; Pmid middle probe; P+ probe closest to the sugar location; C+ sugar control.

Table 4. Best-fit models for locomotion and anxiety indices for hamsters during test trials in the arena. All models within 2AIC_c of the best model are shown for each behavioural index.

model fixed effects	intercept	d.f.	log likelihood	AIC	delta	weight
locomotion index						
trial	0.347	12	−188.11	400.4**	0.00	0.762
null	(0.443)	(6)	(−255.45)	(523.0)	(122.51)	(0.000)
anxiety index						
block	0.2732	7	712.43	−1410.8	0.00	0.412
null	(0.3398)	(6)	(711.01)	(−1410.0)	(0.83)	(0.273)
block + treatment	0.2828	8	712.65	−1409.2	1.58	0.187

** $p < 0.01$.

3.5.2. Control and reinforcement trials

There was no effect of treatment or block at either of the two control locations (or on the QHCl trials: table 3). There was a significant effect of block on approaches to the sugar drinker with hamsters approaching less often in the second block of testing compared to the first. Comparison of bottle weights revealed hamsters drank equivalent amounts of sugar water in the two treatments ($t_{15} = 1.56$, $p = 0.14$).

3.6. Behavioural measures

Trial type explained more than 76% of the variance in locomotion in the arena (table 4). Hamsters locomoted more per second on sugar, C+ and P+ trials compared to the C— baseline (table 5). For anxiety index, no model explained the variance any better than the null model.

Non-parametric tests revealed no difference between treatments in performance in the OF (n squares entered: $W = 105$, $p = 0.50$; time in centre: $W = 129$, $p = 0.95$) or LDE ($W = 87$, $p = 0.12$). Model selection performed on data from neophobia tests revealed that trial type explained 70% of the variance in the data (all other models greater than 2 AIC_c points from best model). Hamsters were faster to approach the sugar drinker when there was a novel object situated next to it than when the sugar drinker alone was presented (figure 4). This was not affected by housing treatment.

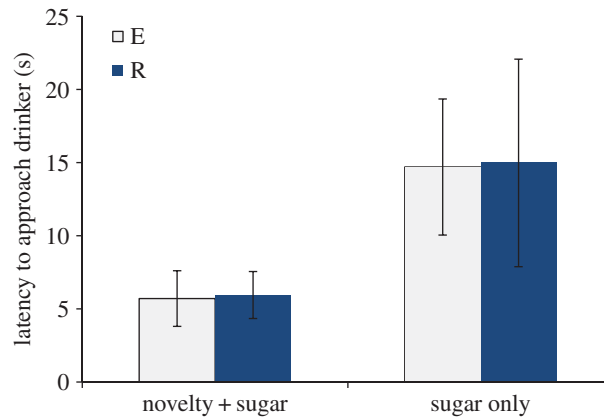


Figure 4. Mean latency (\pm s.e.) for hamsters to approach the sugar drinker with and without a novel object present. E, Enrichment added; R, Enrichment removed.

Table 5. The linear relationship between trial type and locomotion in the arena. Hamsters locomoted most per second on sugar trials and least on QHCl trials. Values calculated using C— used as reference.

	locomotion	estimate	s.e.	t-value
	intercept	0.35	0.09	3.91
	sugar	0.22	0.02	8.17
	C+	0.21	0.02	7.95
	P+	0.18	0.02	6.74
	Pmid	0.06	0.02	2.31
	P—	0.02	0.02	0.67
	QHCl	−0.01	0.02	−0.50

4. Discussion

We developed a spatial judgement bias test for use with hamsters, using short-term changes in enrichment to validate the task. Firstly, hamsters learned to discriminate between drinkers at two locations: they were more likely to approach the positive location and less likely to approach the negative location. Therefore, the spatial judgement bias task is suitable for further development with hamsters. Secondly, hamsters were more likely to approach an empty drinker at ambiguous locations when enrichment had been added to their cage for the previous week, than when it had been removed. Hamsters approached more often as the probe neared the positive location. Importantly, there was no effect of treatment on approaches at the extreme cue locations. This difference in response to ambiguity is consistent with the judgement bias model of emotion–cognition interaction previously reported for a range of species [82]. This is the first study to report evidence for emotion-mediated judgement bias in a species of hamster.

Our unique pattern of findings has implications for continued refinement in the design of the judgement bias task by: (i) using a VRR during training to remove cueing effects of reinforcers; (ii) statistically controlling for potentially confounding effects by including additional factors (e.g. time, order of testing) as random effects in mixed models, and using behavioural measures to distinguish valenced judgements from arousal, risk-taking or perceptual factors; (iii) including a VRR during training, and control trials during testing to reduce rate of learning that probes were never rewarded; and (iv) using robust statistical approaches such as mixed effects models to control for potentially confounding factors in studies that often have small sample sizes. Our findings also have theoretical implications for: the value of cognitive measures of emotion over behavioural and experimental measures alone; the study of positive emotions; and further understanding the mechanisms underlying responses to ambiguous cues to reward and punishment.

4.1. Control trials and behavioural measures distinguish judgement bias from alternative explanations

Increased responding to two of the probes may be interpreted as valenced judgements of likelihood of reward or punishment because hamsters showed no difference between treatments in (i) proportion of approaches on control trials at the sugar drinker location and (ii) proportion of approaches on control trials at the QHCl location. If changes in responses were due to a generalized increase in arousal (or a shift in perception of the empty drinker, or risk-taking behaviour), then we would expect to see greater or fewer responses on the control trials as well as at the probes. Behavioural indices of arousal (locomotion and activity) collected during trials further support this interpretation. Hamsters showed no difference between treatments in rate of locomotion or anxiety-related behaviour during trials. Control trials and additional behavioural measures are therefore an effective way of controlling and accounting for possible arousal effects on responses.

4.2. Linear mixed effects models with random and nested factors provide robust analysis of data

The information-theoretic approach using generalized linear mixed models (GLMMs) provided a flexible and highly informative means of exploring our data. By testing the effect of fixed factors (here treatment and block) at each drinker location separately, we were able to tease apart the effects of different factors on responses at each probe. An advantage of the multi-model inference approach is that it provides one or more 'best-fit' models with information on the relative fit of these models to the spread of data. Using a hierarchical nested approach, we were able to control for a range of additional factors including individual differences in rate of learning, relatedness within sibling groups, order of testing, as well as for changes in performance over testing days. Using mixed effects models prevents problems of interpretation when there are temporal effects on performance (for example, in Scollo *et al.* [49] a shift in judgement bias was reported in pigs on the third day of testing, but no difference was found on previous days or when a mean across days was considered).

The results we report here meet the performance criteria suggested elsewhere for applying mixed models to judgement bias data [82], but we argue the model selection approach we present provides a more informative and intuitive analysis than that proposed by Gyga [82]. By testing our best-fit models against the null model at each probe location, we were able to identify the location of mood effects on responses to the probes. Probes are categorically different from the control locations (about which we have different predictions). The approach we advocate may be more useful for distinguishing the location of effects when identifying different emotion states, or processes involved in approach and avoidance (discussed below). Furthermore, a unified statistical approach to judgement bias studies would allow better comparison between studies and meta-analyses.

4.3. Learning and motivation may influence responses to ambiguous probes

In some previous studies, it has not been possible to rule out possible learning or motivational effects on responses to probes, especially if the effect of treatment is small [22,67,69], animals took part in a relatively large number of trials (more than 50 trials in some studies [26,44,62,65]) or where changes in responses at the learned locations were also detected (e.g. [7,30,32,39,41,43]). To reduce the likelihood, or speed, of learning that the probes were never rewarded, we used variable reinforcement during training. We then controlled for learning effects in the analyses by including day of testing as a random factor in our models. Block featured in the best model only in interaction with trial type. The main effect of block was found at the S+, P+ and Pmid (hamsters approached less often in the second of the two experimental blocks), but not at the other locations. There is therefore a possibility that learning may affect responses to ambiguous cues to reward, or that learning occurred across all locations but the lack of reduction in responses in the second block for either of the P−, C− or QHCl trials was concealed by a floor effect in the data. Alternatively, reduced responding to locations at the positive end may reflect reduced sugar motivation. In either case, any changes due to learning or motivation did not differ between treatments: hamsters drank equal amounts of sugar water during the two treatments, and there was no interaction of block with treatment. For hamsters, at least, it may therefore not be suitable to run more than one or two trials per probe location in a daily testing session. Including a measure of time (here, block and day) in analyses can provide a way of examining potential changes in responding over trials. This may be useful for refining protocols to optimize aspects of study design such as number and spacing of trials within and across days.

4.4. Not all probes are equal: the spatial judgement task may reveal qualitatively different emotion–cognition interactions

Our study presents a unique pattern of findings: a shift in bias at the two outer probes with no evidence of a bias at the middle probe. Changes in enrichment explained 59% and 76% of the variance in responses to the P+ and P−, respectively, yet treatment failed to explain variance at the middle probe. There are several possible reasons for the lack of bias at the middle probe: the middle probe block explained 87% of the variance which may suggest a greater influence of learning effects at this (perhaps more highly) ambiguous location; the middle probe, being truly ambiguous, may have induced a greater level of approach–avoidance conflict and anxiety which masked any other emotion effects arising from the housing treatments; or hamsters may process mildly ambiguous cues to punishment (P−) or reward (P+) differently from ambiguous cues with no partial negative or positive association.

Responses at the two outer probes have been proposed to distinguish ‘anxiety’ and ‘depression’ in animal models [9,39,41,62]. Humans suffering from anxiety have an increased expectation of negative events, while people suffering from depression have both an increased expectation of negative events and a reduced expectation of positive events [88]. In the animal judgement bias task, responses to the rewarded stimulus and adjacent probe are considered to tap into brain mechanisms associated with seeking, approaching and gaining reward; responses to the punished stimulus are considered to tap into brain mechanisms associated with avoidance of threat. Reinforcement sensitivity theory [89] identifies that emotional responses to stimuli arise from the interaction of three motivational systems sensitive to reward (dopaminergic and associated systems relating to positive emotions and approach), punishment (amygdala, anterior cingulate and serotonergic systems relating to negative emotions and avoidance) and reward–punishment conflict [90,91]. Different emotional and cognitive processes may underlie responses to the ambiguous probes, reflecting differences in degree of ambiguity, approach, avoidance and approach–avoidance anxiety. Studies involving genetic models [12,16,55] or pharmaceutical manipulation of brain systems [35,36,41,55,59,62] may elucidate the mechanisms driving responses on the judgement bias test. At the least, we suggest several probes should be included in the test when used with hamsters (cf. for example, single ambiguous probe used with rats [58,59,63] and pigs [48]), and each probe should be examined separately in analyses (cf. [36]): in the example we present here, this was done by testing against the null model at each probe location.

4.5. Optimistic judgement bias in hamsters?

Mapping shifts in judgement bias on to emotional ‘states’ [9] remains a challenge. Hamsters responded to all probes at least as often as expected by chance in both treatments; the most parsimonious explanation is therefore that the addition of extra cage enrichment leads to a more ‘optimistic’ judgement bias in the hamsters. Removal of the additional enrichment items (back to standard enrichment condition which hamsters were used to) resulted in a negative shift (possibly back to the standard baseline?). It is difficult to interpret animals as categorically ‘optimistic’ or ‘pessimistic’ as there is no reference baseline against which to assess this.

4.6. Judgement bias tests may be more sensitive to changes in emotion than traditional behavioural tests

Finally, while the judgement bias task revealed a change in responses to ambiguous cues to possible reward and possible punishment, we found no differences in responses between treatments on the additional standard behavioural tests. Hamsters performed similarly on the OF, LDE and neophobia tests in both housing treatments. In the latter test, hamsters were equally interested in investigating the novel objects in both treatments and were much faster on trials when a novel object was present than for sugar drinker only trials. There was therefore no difference in exploratory tendency or fear of new stimuli (e.g. a drinker spout appearing at a new location) between treatments. These findings lend weight to the argument that the judgement bias test may detect shifts in emotion state not discernable from, or apparently contra to, traditional behavioural measures alone ([7,13,27,29,34,51,70], cf. [24,40,49,66]); may detect shifts earlier (the housing treatments were one week in duration and therefore relatively short-term); or may detect these shifts at different points in the time course of emotional response (e.g. [32,51,58]).

This study adds to the rapidly expanding body of literature which supports the judgement bias task as a means of detecting affect-mediated changes in cognition in non-human animals. This literature shows that shifts in judgement bias may be detected following a range of affect manipulations in a range of species, now including hamsters; and that cognitive measures may detect shifts in emotion not apparent from behavioural observation or traditional behavioural tests of emotion alone. The unique pattern of results reported in this paper adds weight to suggestions that different underlying brain systems may contribute to differences in responses to ambiguous cues to reward or punishment that may have value in distinguishing between different, but similarly valenced, emotional states. We cannot say whether the hamsters in our study *felt* happy in their enriched housing, but the changes in cognitive processing of ambiguous cues certainly suggests enriched hamsters became more optimistic about the likelihood of future reward when faced with uncertain information. Judgement bias tasks present a unique and valuable approach to assessing emotion in laboratory rodents, including hamsters. Future development of these approaches (for example, incorporating automated systems, exploring long-term and developmental effects of captive animal husbandry) should lead to improved welfare assessment across species.

Ethics. This research was conducted at Liverpool John Moores University and conformed to the current UK Home Office Guidelines and Codes of Practice for housing and husbandry of laboratory animals: <http://www.homeoffice.gov.uk/science-research/animal-research/> and the Association for the Study of Animal Behaviour's guidelines on the use of animals in research: <http://asab.nottingham.ac.uk/ethics/guidelines.php>. All procedures were approved by the local ethics committee prior to the start of the study.

Data accessibility. The datasets and R code supporting this article have been uploaded as part of the electronic supplementary material, tables S1–S5. R is a freely available statistical software package and is available to download at: <http://www.r-project.org/>.

Authors' contributions. E.J.B. and N.F.K. jointly designed the study, participated in data analysis, carried out the statistical analyses and drafted the manuscript. Both authors give final approval for publication.

Competing interests. We have no competing interests.

Funding. Data collection was part-funded by a small grant from Liverpool John Moores University, School of Natural Sciences and Psychology. Training in R programming was funded by the Liverpool John Moores University Research Office.

Acknowledgements. We would like to thank Steve Broadfoot and Roy Williams for care of the animals; Alice Maher, Katie Garland, Samantha Noonan and Charlotte Buckley for assistance with data collection; Kevin Arbuckle, Hazel Nichols and Richard Brown for statistical advice. We are grateful to two anonymous reviewers for their helpful feedback.

References

- Dawkins MS. 1990 From an animal's point of view: motivation, fitness, and animal welfare. *Behav. Brain Sci.* **13**, 1–61. (doi:10.1017/S0140525X00077104)
- Mendl M, Paul ES. 2004 Consciousness, emotion and animal welfare: insights from cognitive science. *Anim. Welf.* **13**, 17–25.
- Russell WMS, Burch RL. 1959 *The principles of humane experimental technique*. See <http://web.archive.org/web/20070403225227>.
- National Centre for the Replacement, Refinement & Reduction of Animals in Research. 2014 See <http://www.nc3rs.org.uk/>
- Mogg K, Bradbury KE, Bradley BP. 2006 Interpretation of ambiguous information in clinical depression. *Behav. Res. Ther.* **44**, 1411–1419. (doi:10.1016/j.brat.2005.10.008)
- Richards A, French CC, Calder AJ, Webb B, Fox R, Young AW. 2002 Anxiety-related bias in the classification of emotionally ambiguous facial expressions. *Emotion* **2**, 273–287. (doi:10.1037/1528-3542.2.3.273)
- Harding EJ, Paul ES, Mendl M. 2004 Animal behaviour: judgement bias and affective state. *Nature* **427**, 312. (doi:10.1038/427312a)
- Paul E, Harding EJ, Mendl M. 2005 Measuring emotional processes in animals: the utility of a cognitive approach. *Neurosci. Biobehav. Rev.* **29**, 469–491. (doi:10.1016/j.neubiorev.2005.01.002)
- Mendl M, Burman OHP, Parker RMA, Paul ES. 2009 Judgement biases an indicator of animal emotion and welfare: emerging evidence and underlying mechanisms. *Appl. Anim. Behav. Sci.* **118**, 161–181. (doi:10.1016/j.applanim.2009.02.023)
- Mendl M, Burman OHP, Paul ES. 2010 An integrative and functional framework for the study of animal emotion and mood. *Proc. R. Soc. B* **277**, 2895–2904. (doi:10.1098/rspb.2010.0303)
- Kloke V, Schreiber RS, Bodden C, Möllers J, Ruhmann H, Kaiser S, Lesch KP, Sachser N, Levejohann L. 2014 Hope for the best or prepare for the worst? Towards a spatial cognitive bias test for mice. *PLoS ONE* **9**, e105431. (doi:10.1371/journal.pone.0105431)
- Boleij H, Klooster J, Lavrijsen M, Kirchhoff S, Arndt SS, Ohl F. 2012 A test to identify judgement bias in mice. *Behav. Brain Res.* **233**, 45–54. (doi:10.1016/j.bbr.2012.04.039)
- Burman OHP, Parker R, Paul ES, Mendl M. 2008 A spatial judgement task to determine background emotional state in laboratory rats, *Rattus norvegicus*. *Anim. Behav.* **76**, 801–809. (doi:10.1016/j.anbehav.2008.02.014)
- Burman OHP, Parker RMA, Paul ES, Mendl MT. 2009 Anxiety-induced judgement bias in non-human animals. *Physiol. Behav.* **98**, 345–350. (doi:10.1016/j.physbeh.2009.06.012)
- Wheeler RR, Swan MP, Hickman DL. 2014 Effect of multilevel laboratory rat caging system on the well-being of the singly-housed Sprague Dawley rat. *Lab. Anim.* **49**, 10–19. (doi:10.1177/0023677214547404)
- Richter SH, Schick A, Hoyer C, Lankisch K, Gass P, Vollmayr B. 2012 A glass full of optimism: enrichment effects on judgement bias in a rat model of depression. *Cogn. Affect. Behav. Neurosci.* **12**, 527–542. (doi:10.3758/s13415-012-0101-2)
- Tami G, Torre C, Compagnucci M, Manteca X. 2011 Interpretation of ambiguous spatial stimuli in cats. *Anim. Welf.* **20**, 185–189.
- Starling MJ, Branson N, Cody D, Starling TR, McGreevy PD. 2014 Canine sense and sensibility: tipping points and response latency variability as an optimism index in a canine judgement bias assessment. *PLoS ONE* **9**, e107794. (doi:10.1371/journal.pone.0107794)
- Kis A, Hernadi A, Kanizsar O, Gacsi M, Topal J. 2015 Oxytocin induces positive expectations about ambivalent stimuli (cognitive bias) in dogs. *Horm. Behav.* **69**, 1–7. (doi:10.1016/j.yhbeh.2014.12.004)
- Burman O, McGowan R, Mendl M, Norling Y, Paul E, Rehn T, Keeling L. 2011 Using judgement bias to measure positive affective state in dogs. *Appl. Anim. Behav. Sci.* **132**, 160–168. (doi:10.1016/j.applanim.2011.04.001)

21. Mendl M, Brooks J, Basse C, Burman O, Paul E, Blackwell E, Casey R. 2010 Dogs showing separation-related behaviour exhibit a pessimistic judgement bias. *Curr. Biol.* **20**, R839–R840. (doi:10.1016/j.cub.2010.08.030)
22. Müller CA, Riemer S, Rosam CM, Schöbwendler J, Range F, Huber L. 2012 Brief owner absence does not induce negative judgement bias in pet dogs. *Anim. Cogn.* **15**, 1031–1035. (doi:10.1007/s10071-012-0526-6)
23. Titulaer M, Blackwell EJ, Mendl M, Casey RA. 2013 Cross sectional study comparing behavioural, cognitive and physiological indicators of welfare between short and long term kennelled domestic dogs. *Appl. Anim. Behav. Sci.* **147**, 149–158. (doi:10.1016/j.applanim.2013.05.001)
24. Walker JK, Waran NK, Phillips JJC. 2014 The effect of conspecific removal on the behaviour and physiology of pair-housed shelter dogs. *Appl. Anim. Behav. Sci.* **158**, 46–56. (doi:10.1016/j.applanim.2014.06.010)
25. Daros RR, Costa JHC, von Keyserlingk MAG, Hötzel MJ, Weary DM. 2014 Separation from the dam causes negative judgement bias in dairy calves. *PLoS ONE* **9**, e98429. (doi:10.1371/journal.pone.0098429)
26. Neave HW, Daros RR, Costa JHC, von Keyserlingk MAG, Weary DM. 2013 Pain and pessimism: dairy calves exhibit negative judgement bias following hot-iron disbudding. *PLoS ONE* **8**, e80556. (doi:10.1371/journal.pone.0080556)
27. Doyle RE, Fisher AD, Hinch GN, Boissy A, Lee C. 2010 Release from restraint generates a positive judgement bias in sheep. *Appl. Anim. Behav. Sci.* **122**, 28–34. (doi:10.1016/j.applanim.2009.11.003)
28. Doyle RE, Vidal S, Hinch GN, Fisher AD, Boissy A, Lee C. 2010 The effect of repeated testing on judgement biases in sheep. *Behav. Process.* **83**, 349–352. (doi:10.1016/j.beproc.2010.01.019)
29. Doyle RE, Hinch GN, Fisher AD, Boissy A, Henshall JM, Lee C. 2011 Administration of serotonin inhibitor p-chlorophenylalanine induces pessimistic-like judgement bias in sheep. *Psychoneuroendocrinology* **36**, 279–288. (doi:10.1016/j.psyneuen.2010.07.018)
30. Doyle RE, Lee C, Deiss V, Fisher AD, Hinch GN, Boissy A. 2011 measuring judgement bias and emotional reactivity in sheep following long-term exposure to unpredictable and aversive events. *Physiol. Behav.* **102**, 503–510. (doi:10.1016/j.physbeh.2011.01.001)
31. Sanger ME, Doyle RE, Hinch GN, Lee C. 2011 Sheep exhibit a positive judgement bias and stress-induced hyperthermia following shearing. *Appl. Anim. Behav. Sci.* **131**, 94–103. (doi:10.1016/j.applanim.2011.02.001)
32. Destrez A, Deiss V, Belzung C, Lee C, Boissy A. 2012 Does reduction of fearfulness tend to reduce pessimistic-like judgment in lambs? *Appl. Anim. Behav. Sci.* **139**, 233–241. (doi:10.1016/j.applanim.2012.04.006)
33. Destrez A, Deiss V, Lévy F, Calandreau L, Lee C, Chaillou E, Boissy A. 2013 Chronic stress induces pessimistic-like judgment and learning deficits in sheep. *Appl. Anim. Behav. Sci.* **148**, 28–36. (doi:10.1016/j.applanim.2013.07.016)
34. Destrez A, Deiss V, Leterrier C, Calandreau L, Boissy A. 2014 Repeated exposure to positive events induces optimistic-like judgment and enhances fearfulness in chronically stressed sheep. *Appl. Anim. Behav. Sci.* **154**, 30–38. (doi:10.1016/j.applanim.2014.01.005)
35. Verbeek E, Ferguson D, Lee C. 2014 Are hungry sheep more pessimistic? The effects of food restriction on judgement bias and the involvement of ghrelin in its regulation. *Physiol. Behav.* **123**, 67–75. (doi:10.1016/j.physbeh.2013.09.017)
36. Verbeek E, Ferguson D, Quinquet de Monjour P, Lee C. 2014 Generating positive affective states in sheep: the influence of food rewards and opioid administration. *Appl. Anim. Behav. Sci.* **154**, 39–47. (doi:10.1016/j.applanim.2014.02.011)
37. Vögeli S, Lutz J, Wolf M, Wechsler B, Gygas L. 2014 Stimulus valence, not housing conditions, affects behavioural and brain reactions in sheep. *Behav. Brain Res.* **267**, 144–155. (doi:10.1016/j.bbr.2014.03.036)
38. Guldemann K, Vögeli S, Wolf M, Wechsler B, Gygas L. 2015 Frontal brain deactivation during a non-verbal cognitive judgement bias test in sheep. *Brain Cogn.* **93**, 35–41. (doi:10.1016/j.bandc.2014.11.004)
39. Salmeto AL, Hymel KA, Carpenter EC, Brilot BO, Bateson M, Sufka KJ. 2011 Judgement bias in the chick anxiety–depression model. *Brain Res.* **1373**, 124–130. (doi:10.1016/j.brainres.2010.12.007)
40. Wichman A, Keeling LJ, Forkman B. 2012 Judgement bias and anticipatory behaviour of laying hens housed in basic and enriched pens. *Appl. Anim. Behav. Sci.* **140**, 62–69. (doi:10.1016/j.applanim.2012.05.006)
41. Hymel KA, Sufka KJ. 2012 Pharmacological reversal of judgement bias in the chick anxiety–depression model. *Neuropharmacology* **62**, 161–166. (doi:10.1016/j.neuropharm.2011.06.009)
42. Seehuus B, Mendl M, Keeling LJ, Blokhuis H. 2013 Disrupting motivational sequences in chicks: are there affective consequences? *Appl. Anim. Behav. Sci.* **148**, 85–92. (doi:10.1016/j.applanim.2013.07.008)
43. Bateson M, Desire S, Gartside SE, Wright GA. 2011 Agitated honey-bees exhibit pessimistic judgement biases. *Curr. Biol.* **21**, 1070–1073. (doi:10.1016/j.cub.2011.05.017)
44. Bethell EJ, Holmes A, MacLarnon A, Semple S. 2012 Judgement bias in a non-human primate: husbandry procedures influence cognitive indicators of psychological well-being in captive rhesus macaques. *Anim. Welf.* **21**, 185–195. (doi:10.7120/09627286.21.2.185)
45. Gordon DJ, Rogers LJ. 2015 Cognitive bias, hand preference and welfare of common marmosets. *Behav. Brain Res.* **287**, 100–108. (doi:10.1016/j.bbr.2015.03.037)
46. Bateson M, Matheson SM. 2007 Performance on a categorisation task suggests that removal of environmental enrichment induces 'pessimism' in captive European starlings (*Sturnus vulgaris*). *Anim. Welf.* **16**, S33–S36.
47. D'upjan S, Ramp C, Kanitz E, Tuchscherer A, Puppe B. 2013 A design for studies on judgement bias in the domestic pig. *J. Vet. Behav.* **8**, 485–489. (doi:10.1016/j.jveb.2013.05.007)
48. Douglas C, Bateson M, Walsh C, Bédoué A, Edwards SA. 2012 Environmental enrichment induces optimistic judgement biases in pigs. *Appl. Anim. Behav. Sci.* **139**, 65–73. (doi:10.1016/j.applanim.2012.02.018)
49. Scollo A, Gottardo F, Contiero B, Edwards SA. 2014 Does stocking density modify affective state in pigs as assessed by judgement bias, behavioural and physiological parameters? *Appl. Anim. Behav. Sci.* **153**, 26–35. (doi:10.1016/j.applanim.2014.01.006)
50. Briefer-Freymond S, Briefer EF, Zollinger A, Gindrat-von Allmen Y, Wyss C, Bachmann I. 2014 Behaviour of horses in a judgment bias test associated with positive or negative reinforcement. *Appl. Anim. Behav. Sci.* **158**, 34–45. (doi:10.1016/j.applanim.2014.06.006)
51. Briefer EF, McElligott AG. 2013 Rescued goats at a sanctuary display positive mood after former neglect. *Appl. Anim. Behav. Sci.* **146**, 45–55. (doi:10.1016/j.applanim.2013.03.007)
52. Rygula R, Szczec E, Papciak J, Nikiforuk A, Popik P. 2014 The effects of cocaine and mazindol on the cognitive judgement bias of rats in the ambiguous-cue interpretation paradigm. *Behav. Brain Res.* **270**, 206–212. (doi:10.1016/j.bbr.2014.05.026)
53. Brydges NM, Leach M, Nicol K, Wright R, Bateson M. 2011 Environmental enrichment induces optimistic judgement bias in rats. *Anim. Behav.* **81**, 169–175. (doi:10.1016/j.anbehav.2010.09.030)
54. Brydges NM, Hall L, Nicolson R, Holmes MC, Hall J. 2012 The effects of juvenile stress on anxiety, judgement bias and decision making in adulthood: a rat model. *PLoS ONE* **7**, e48143.
55. Enkel T, Gholizadeh D, von Bohlen und Halbach O, Sanchis-Segura C, Hurlmann R, Spanagel R, Gass P, Vollmayr B. 2010 Ambiguous-cue interpretation is biased under stress- and depression-like states in rats. *Neuropsychopharmacology* **35**, 1008–1015. (doi:10.1038/npp.2009.204)
56. Rygula R, Pluta H, Popik P. 2012 Laughing rats are optimistic. *PLoS ONE* **7**, e51959. (doi:10.1371/journal.pone.0051959)
57. Rygula R, Papciak J, Popik P. 2013 Trait pessimism predicts vulnerability to stress-induced anhedonia in rats. *Neuropsychopharmacology* **38**, 2188–2196. (doi:10.1038/npp.2013.116)
58. Chaby LE, Cavigelli SA, White A, Wang K, Braithwaite VA. 2013 Long-term changes in judgement bias and coping response as a result of chronic unpredictable stress during adolescence. *Front. Hum. Neurosci.* **7**, 328. (doi:10.3389/fnhum.2013.00328)
59. Rygula R, Papciak J, Popik P. 2014 The effects of acute pharmacological stimulation of the 5-HT_{1A} and DA systems on the cognitive judgement bias of rats in the ambiguous-cue interpretation paradigm. *Eur. Neuropsychopharmacol.* **24**, 1103–1111. (doi:10.1016/j.euroneuro.2014.01.012)
60. Rygula R, Golebiowska J, Kregiel J, Kubik J, Popik P. 2015 Effects of optimism on motivation in rats. *Front. Behav. Neurosci.* **9**, 32. (doi:10.3389/fnbeh.2015.00032)
61. Rygula R, Szczec E, Kregiel J, Golebiowska J, Kubik J, Popik P. 2015 Cognitive judgement bias in the psychostimulant-induced model of mania in rats. *Psychopharmacology* **232**, 651–660. (doi:10.1007/s00213-014-3707-y)
62. Anderson M, Munafò MR, Robinson ESJ. 2013 Investigating the psychopharmacology of cognitive affective bias in rats using an affective tone discrimination task. *Psychopharmacology* **226**, 601–613. (doi:10.1007/s00213-012-2932-5)

63. Papić J, Popik P, Fuchs E, Rygula R. 2013 Chronic psychosocial stress makes rats more 'pessimistic' in the ambiguous-cue interpretation paradigm. *Behav. Brain Res.* **256**, 305–310. (doi:10.1016/j.bbr.2013.08.036)
64. Parker RMA, Paul ES, Burman OHP, Browne WJ, Mendl M. 2014 Housing conditions affect rat responses to two types of ambiguity in a reward–reward discrimination judgement bias task. *Behav. Brain Sci.* **274**, 73–83. (doi:10.1016/j.bbr.2014.07.048)
65. Matheson SM, Asher L, Bateson M. 2008 Larger, enriched cages are associated with optimistic response biases in captive European starlings (*Sturnus vulgaris*). *Appl. Anim. Behav. Sci.* **109**, 374–383. (doi:10.1016/j.applanim.2007.03.007)
66. Brilot B, Asher L, Bateson M. 2010 Stereotyping starlings are more pessimistic. *Anim. Cogn.* **13**, 721–731. (doi:10.1007/s10071-010-0323-z)
67. Murphy E, Nordquist RE, van der Staay FJ. 2013 responses of conventional pigs and Gottingen miniature pigs in an active choice judgement bias task. *Appl. Anim. Behav. Sci.* **148**, 64–67. (doi:10.1016/j.applanim.2013.07.011)
68. Hernandez CE, Hinch GE, Lea J, Ferguson D, Lee C. 2015 Acute stress enhances sensitivity to a highly attractive food reward without affecting judgement bias in laying hens. *Appl. Anim. Behav. Sci.* **163**, 135–143. (doi:10.1016/j.applanim.2014.12.002)
69. Keen HA, Nelson OL, Robbins CT, Evans M, Shepherdson DJ, Newberry RC. 2014 Validation of a novel judgement bias task based on difference in quantity of reinforcement for assessing environmental enrichment. *Anim. Cogn.* **17**, 529–541. (doi:10.1007/s10071-013-0684-1)
70. Pomerantz Q, Terkel J, Suomi SJ, Paukner A. 2012 Stereotypic head twirls, but not pacing, are related to a 'pessimistic'-like judgment bias among captive tufted capuchins (*Cebus apella*). *Anim. Cogn.* **15**, 689–698. (doi:10.1007/s10071-012-0497-7)
71. Lanteigne M, Reeb SG. 2006 Preference for bedding material in Syrian hamsters. *Lab. Anim.* **40**, 410–418. (doi:10.1258/002367706778476424)
72. Veillette M, Reeb SG. 2011 Shelter choice by Syrian hamsters (*Mesocricetus auratus*) in the laboratory. *Anim. Welf.* **20**, 603–611.
73. Sørensen DB, Krohn T, Hansen HN, Ottesen JL, Hansen AK. 2005 An ethological approach to housing requirements of golden hamsters, Mongolian gerbils and fat sand rats in the laboratory—a review. *Appl. Anim. Behav. Sci.* **94**, 181–195. (doi:10.1016/j.applanim.2005.02.004)
74. Swanson HH. 1966 Sex differences in behaviour of hamsters in open field and emergence tests: effects of pre- and post-pubertal gonadectomy. *Anim. Behav.* **14**, 522–529. (doi:10.1016/S0003-3472(66)80054-4)
75. Belzung C, Le Pape G. 1994 Comparison of different behavioral test situations used in psychopharmacology for measurement of anxiety. *Physiol. Behav.* **56**, 623–628. (doi:10.1016/0031-9384(94)90311-5)
76. Shettleworth SJ. 1975 Reinforcement and the organization of behaviour in Golden Hamsters: hunger, environment, and food reinforcement. *J. Exp. Psychol. Anim. Behav. Proc.* **104**, 56–87. (doi:10.1037/0097-7403.1.1.56)
77. Ziomek J, Zgrabczynska E, Poradzisz A. 2009 the behaviour of the common hamster (*Cricetus cricetus*) under zoo conditions. *Der Zoologische Garten.* **78**, 221–231. (doi:10.1016/j.zoolgart.2009.08.006)
78. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Steebns MHH, White JS. 2008 Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* **24**, 127–135. (doi:10.1016/j.tree.2008.10.008)
79. Wang Z, Goonewardene LA. 2004 The use of mixed models in the analysis of animal experiments with repeated measures data. *Can. J. Anim. Sci.* **84**, 1–11. (doi:10.4141/A03-123)
80. Thomas R, Vaughn I, Lello J. 2013 *Data analysis with R statistical software: a guide for scientists*. Newport, UK: Eco-explore.
81. Crawley MJ. 2007 *The R book*. Chichester, UK: Wiley-Blackwell.
82. Gyax L. 2014 The A to Z of statistics for testing cognitive judgement bias. *Anim. Behav.* **95**, 59–69. (doi:10.1016/j.anbehav.2014.06.013)
83. Reeb SG, St Onga P. 2005 Running wheel choice by Syrian hamsters. *Lab. Anim.* **39**, 442–451. (doi:10.1258/002367705774286493)
84. Mendl M. 1999 Performing under pressure: stress and cognitive function. *Appl. Anim. Behav. Sci.* **65**, 221–244. (doi:10.1016/S0168-1591(99)00088-X)
85. Martin P, Bateson P. 1998 *Measuring behaviour: an introductory guide*. Cambridge, UK: Cambridge University Press.
86. R Core Team. 2013 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
87. Bates D, Maechler M, Bolker B, Walker S. 2013 *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.0–5. See <http://CRAN.R-project.org/package=lme4>.
88. Eysenck M, Payne S, Santos R. 2006 Anxiety and depression past, present and future events. *Cog. Emot.* **20**, 274–294. (doi:10.1080/026999930500220066)
89. Gray JA 1971 *The psychology of fear and stress*. London, UK: World University Library.
90. McNaughton N, Corr PJ. 2004 A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neurosci. Biobehav. Rev.* **28**, 285–305. (doi:10.1016/j.neubiorev.2004.03.005)
91. Panksepp J. 1998 *Affective neuroscience. The foundations of human and animal emotion*. New York, NY: Oxford University Press.