



Published in final edited form as:

Anim Cogn. 2012 July ; 15(4): 689–698. doi:10.1007/s10071-012-0497-7.

Stereotypic head twirls, but not pacing, are related to a ‘pessimistic’-like judgment bias among captive tufted capuchins (*Cebus apella*)

Ori Pomerantz,

Department of Zoology, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, 69978 Tel Aviv, Israel

Joseph Terkel,

Department of Zoology, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, 69978 Tel Aviv, Israel

Stephen J Suomi, and

Laboratory of Comparative Ethology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health Animal Center, Post Office Box 529, Poolesville, MD 20837, USA

Annika Paukner

Laboratory of Comparative Ethology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health Animal Center, Post Office Box 529, Poolesville, MD 20837, USA

Ori Pomerantz: oripomer@post.tau.ac.il

Abstract

Abnormal stereotypic behaviour is widespread among captive non-human primates and is generally associated with jeopardized well-being. However, attributing the same significance to all of these repetitive, unvarying and apparently functionless behaviours may be misleading, as some behaviours may be better indicators of stress than others. Previous studies have demonstrated that the affective state of the individual can be inferred from its bias in appraising neutral stimuli in its environment. Therefore, in the present study, in order to assess the emotional state of stereotyping individuals, 16 captive tufted capuchins (*Cebus apella*) were tested on a judgment bias paradigm and their faecal corticoid levels were measured in order to assess the intensity of the emotional state. Capuchins with higher levels of stereotypic head twirls exhibited a negative bias while judging ambiguous stimuli and had higher levels of faecal corticoids compared to subjects with lower levels of head twirls. Levels of stereotypic pacing, however, were not correlated with the monkeys' emotional state. This study is the first to reveal a positive correlation between levels of stereotypic behaviour and a ‘pessimistic’-like judgment bias in a non-human primate by employing a recently developed cognitive approach. Combining cognitive tests that evaluate the animals' affective valence (positive or negative) with hormonal measurements that provide information on the strength of the emotional state conduces to a better understanding of the animals' affective state and therefore to their well-being.

Keywords

Stereotypy; Cognitive-bias; Negative affect; Primate welfare

Introduction

The psychological well-being of both humans and animals is affected by intrinsic factors such as motivation, as well as extrinsic factors such as exogenous administrations of glucocorticoids (Lupien et al. 2007; Polivy 1998). A state of bad welfare can have a devastating outcome for the individual. A deprived rearing environment during early human life, for example, may increase the risk of expressing autistic features later on (Rutter et al. 1999), and impoverished captive surroundings may accelerate the progression of disease in animals (Hockly et al. 2002).

In humans, in order to appraise well-being, researchers mainly resort to surveys and patient reports (Diener et al. 2003; Mendl et al. 2010b). With non-human animals however, other, indirect methods must be employed. In order to assess the welfare of their subjects, animal welfare scientists may incorporate an array of techniques such as preference testing (Dawkins 2003), observations of stress-related and abnormal behaviours in different conditions (Pomerantz and Terkel 2009), and the use of autonomic responses as physiological measures of stress (Dawkins 2006). Although valuable, assessing subjective feelings in non-human animals by using indirect techniques poses a number of challenges. Probably, the principal challenge concerns the transformation of objective data into an accurate representation of the individual's subjective feelings.

In order to cope with this challenge, many researchers employ several assessment tools in concert, and new welfare evaluation procedures are also being developed. One such procedure focuses on the role of emotions and mood states, since these are major factors in determining an individual's well-being (Dawkins 2006). It was previously suggested that the mood state of an individual is derived from that individual's perception of positive and negative stimuli in their environment (Mendl et al. 2010b). Interestingly, affective states of individuals were found to create a bias in their processing of information, thereby manipulating memory, attention, decision-making, and judgment (Mendl et al. 2010b). Accordingly, measuring these biases provides a reliable indication of the subject's current affective state. In studies involving animals, it is the manner in which individuals judge ambiguous stimuli that is generally tested. For example, subjects are first trained to distinguish between two stimuli that lie at the ends of a continuous stimulus range: one is associated with an appetitive outcome and the other with an aversive outcome (or a less appetitive one). Once learned, the animals' reactions to an intermediate/ambiguous stimulus within that range are examined. A response to the intermediate stimulus that resembles the one expressed when the stimulus signalling an appetitive outcome is presented suggests that the animal is anticipating a positive outcome and consequently can be defined as showing an 'optimistic'-like judgment bias, while the opposite response indicates an individual expressing a 'pessimistic'-like judgment bias (Matheson et al. 2008). For instance, chicks that were isolated to induce anxiety- and depression-like states took more time to approach ambiguous stimuli than controls, presumably because they associated the former with a greater likelihood of aversive consequences (Salmeto et al. 2011). Starlings housed in sub-optimal conditions (that are assumed to decrease their welfare) were less likely to interpret ambiguous stimuli as signals of a rewarding outcome compared to when they were housed in enriched cages, suggesting that keeping animals in an impoverished environment induces a negative emotional state (Bateson and Matheson 2007).

To date, most studies that have utilized judgment-bias evaluations first induced a putative affective state in the animals through negative or positive manipulations (Mendl et al. 2009). These studies present a reliable technique by which to evaluate the animal's affective state. However, an individual's well-being is not just a reflection of a momentary condition; rather, it is the product of a more sustained chronic state (Moberg 2001). Moreover, youths with anxiety disorders (a stable trait reflective of a negative bias) present a bias towards threatening stimuli compared to non-anxious subjects (Puliafico and Kendall 2006), which raises the question of whether similarly consistent traits in animals may also exert similar bias.

When chronically exposed to an aversive situation, captive animals may express abnormal stereotypic behaviours (Garner et al. 2003). This group of behaviours is commonly defined as the display of repetitive and unvarying movements with no apparent function or goal (Lanovaz 2011; Garner 2005), and is largely attributed to either incompatible surroundings or a neurological malfunction (Mills 2003). Stereotypies are thus usually associated with impoverished welfare (Mason 1991; Mason et al. 2001; Wiedenmayer 1997). However, such association is not clear-cut, and some researchers attribute a rewarding, stress-reducing effect to stereotypies (Lewis and Bodfish 2007). Accordingly, it has been suggested that within an aversive environment, those individuals that stereotype are likely to be the least compromised (Mason and Latham 2004). Understanding the emotional state of stereotypers by utilizing cognitive-bias assessments may thus provide crucial information about their well-being, and help resolve this dispute.

In the following study, the psychological well-being of a group of 16 laboratory-housed tufted capuchins (*Cebus apella*) was evaluated using observations of the monkeys' natural behaviour in their home cage, together with a judgment-bias paradigm.

In addition to these measurements, the monkeys' faecal corticosterone levels were also analysed. This measurement was applied since the stress reaction includes responses to aversive stimuli that are likely to induce negative emotional states (Koolhaas et al. 1997). Indeed, it was previously reported that the amygdala, which constitutes a central role in the processing of fear and anxiety, influences the HPA axis that mediates the stress response (Dayas et al. 1999). However, autonomic reactions, such as increased heart rate and elevation of corticosteroids, may be similar in response to situations of different emotional quality. For example, the level of cortisol increases after exposure to a threat, but also after physical exercise (Dickerson and Kemeny 2004; Coiro et al. 2011). We thus decided to follow others (Mendl et al. 2009; Paul et al. 2005) in interpreting corticoid levels as an indication of the level of arousal or of the intensity of the emotion that the subject is experiencing, rather than as an absolute representation of an emotional state. In this study, the data from the cognitive-bias tests provided information on the valence of the animals' emotional state (i.e. negative or positive), while the levels of faecal corticosterone indicated the intensity of the emotional experience. This combination of valence and arousal/intensity is conducive to obtaining a high-resolution image of the animals' emotional state.

In accordance with the above, our aims in the present work were (1) to investigate the relationships between sustained individual traits (rather than a temporary environmental manipulation) that are expressed by stereotypic behaviour, and a negative emotional state, reflected in a negative cognitive bias; and (2) to examine the level of arousal (intensity of emotional response) by measuring the levels of faecal corticosterone.

Materials and methods

Animals

Eight male and eight female adult, captive-born and mother-reared, tufted capuchins (*C. apella*) (mean age \pm SD 131 \pm 22 months) were used in this study. Seven monkeys were group-housed in two sets of quad cages (163 \times 163 \times 71 cm, four monkeys in one set and three monkeys in the other), and nine individuals were housed in two indoor runs (6.9 \times 4.1 \times 2.1 m, five individuals in one run and four individuals in the other together with other juvenile and infant animals) at the National Institutes of Health (NIH) Animal Center, Poolesville, MD (see Table 1 for details). Monkeys were provided with a variety of plastic and metal manipulanda. Subjects were fed twice a day with commercial monkey biscuits (Labdiet 5045), and scatter feeds of grains or seeds and enrichment with fresh fruits or nuts were dispensed daily. Water was provided ad libitum. Monkeys were not food-deprived during this study. Monkeys were kept on a 12/12 h light/dark cycle.

This study was approved by the National Institute of Child Health & Human Development (NICHD) Animal Care and Use Committee, #09-015.

Study period

Data were collected between December 2010 and April 2011.

Observations

The subjects' levels of stereotypy and non-stereotypic activity were obtained by observing each monkey five times a week for a 20-min period, using continuous focal sampling (Martin and Bateson 1993). Observations were carried out between 09:00–12:00 and 13:00–16:00, for a total of five morning and five afternoon observations (comprising 200 min of observations per monkey). Stereotypic behaviours (rates of head twirls (HT) and durations of pacing) were recorded on a laptop computer using the JWatcher 1.0 software. Pacing is defined as repetitive walking of the same path in the cage, and HT is defined as repetitive circular movements of the head.

Faecal sampling

In order to measure faecal corticosterone, fresh faeces samples (three per individual) were collected on different days during the morning judgment-bias sessions (9:00–12:00), while monkeys were in individual cages, placed in a 50-ml tube (Thomas Scientific) and kept at -20°C until analysis (RIA). Following Beehner and Whitten 2004, we chose to measure corticosterone and not cortisol as a cumulative measure of stress, since the particular antibody that we used cross-reacts with a number of excreted cortisol metabolites and is therefore considered to be of a broad spectrum. Faecal corticoid analyses were conducted at the Endocrine Laboratory, Smithsonian Conservation Biology Institute, Center for Species Survival, National Zoological Park, Front Royal, USA, according to the following protocol, which was previously validated for use in non-human primates (Wasser et al. 2000): Tubes were freeze-dried, crushed, put into storage tubes, and weighed (0.2 gm) for extraction in 16 mm \times 125 mm glass tubes. Next, 5 ml of 90 % ethanol was added to each sample along with 100 μl corticosterone tracer for recovery. Samples were then capped with rubber stoppers and placed on the multipulse vortexer for 30 min, centrifuged at 2,500 rpm for 15 min, decanted into another set of labelled tubes, and another 5 ml of 90 % ethanol was added to the faecal pellet and vortexed for 30 s. Samples were then centrifuged for 20 min at 2,500 rpm and added to the previously poured-off supernatant. The supernatant was then dried and diluted to 1:500. Finally, samples were run using the MP Biomedicals double-antibody corticosterone ^{125}I RIA kit. Data are presented as nanogram corticosterone per gram dry

faeces. A mean value was then calculated from the three samples for each monkey, and these means were subsequently correlated with the subjects' levels of stereotypy.

Corticosterone levels and activity

The mean proportion of all active behaviours (defined as instances in which the subject either changed location, engaged in feeding or foraging activity, or in stereotypy) was calculated for each subject from the data obtained during the observations. These means were then correlated with the mean levels of faecal corticosterone in order to determine whether corticosterone levels were linked to the overall activity levels of the monkeys.

Apparatus

The apparatus was positioned in front of the monkey's cage and consisted of a plastic board ($40 \times 40 \times 2$ cm) with two holes (8 cm diameter \times 2 cm deep) covered by opaque sliding caps (one black and one white), which the monkey had to push aside in order to reveal the contents of the hole. Holes were positioned 22 cm apart and 1 cm from the front and side edges (Fig. 1). Two rectangles, each of different length, were used as discriminative stimuli: a 38.1×5 cm striped rectangle (long rectangle) and a 7.6×5 cm striped rectangle (short rectangle). One stimulus (S++) signalled the presence of a high-valued reward in one of the holes, and the other stimulus (s+s-) signalled the presence of a low-valued reward in the other hole. Each subject was randomly allocated to receive either the long or the short rectangle as S++, and reward sides (black or white) were counterbalanced between subjects. The intermediate length stimulus, used as an ambiguous 'probe' (medium probe) for cognitive-bias assessment, comprised a 22.9×5 cm striped rectangle.

Before the first learning session of each subject, the monkeys were introduced to the apparatus while both holes were only partially covered, enabling them to see that one hole was baited and the other was not. Hole lids were positioned in such a manner that the monkeys needed to slide them in order to extract the reward. During those pre-session trials, the monkeys were allowed to explore both holes before the apparatus was removed from their reach, thus learning that only one hole was baited per trial. Following 5–10 introductory trials per individual, the learning sessions commenced.

Cognitive-bias task

Monkeys were presented with a visual conditional discrimination task with differently valued rewards, whereby they had to attend to the length of the rectangle (S++ or s+s) in order to predict which of the two visually distinct lid-covered holes adjacent to it contained a hidden food reward and which was empty (Fig. 1a, b). In S++ trials, the reward was high-value (mix of 1 mm slice of grape, $5 \times 5 \times 5$ mm marshmallow cube, and one 190 mg banana-flavoured pellet, Bio-Serv), while in s+s- trials the reward was low-value ($10 \times 10 \times 10$ mm apple cube). Preferences had been determined in a prior testing that showed that monkeys favoured the mixed reward (grape, marshmallow, and a banana-flavoured pellet) in 97.3 % of trials. In utilizing this design, we sought to ensure that active responses would be given to both the S++ and s+s- stimuli, but that these responses could be differentiated by the subjects' motivation to exploit the reward (as suggested by Brilot et al. 2010). Once the monkeys had learnt this discrimination, the cognitive-bias test involved presenting an intermediate length rectangle between S++ and s+s- without any reward, and recording which of the two holes the subjects chose (Fig. 1c).

The learning stage

The learning stage of the judgment bias task was carried out between 09:00–12:00 and 13:00–16:00. Monkeys were individually caged during the judgment-bias testing by adding

a divider to the quad cages, or allowing the run monkeys to voluntarily enter a tunnel in the runs, comprised of several single cages that could be modified for this purpose. Subjects were required to distinguish between two conditions according to the length of the striped rectangle (S++ or s+s-), and consequently decide which of the two visually distinct lid-covered holes adjacent to it enclosed the preferred reward, and which the less preferred one (in separate trials). In each of the trials [maximum trial duration was 60 s, inter-trial-interval (ITI) was 15 s], only one hole was baited with either the high- or the low-value reward while the other hole remained empty. During the ITIs, the side of the apparatus facing the monkeys was covered with a plastic sheet so that the subjects could not see which hole was being baited for the next trial. Learning sessions consisted of three consecutive blocks: a block of three S++ trials, a block of nine s+s- trials, and a block of 3 S++/s+s- test trials each in random order. The order of S++ and s+s- blocks within a session changed daily. During the learning stage and the cognitive-bias stage, the apparatus was removed from the monkey's reach once a choice was made (either of the baited hole or of the empty hole), such that they were unable to explore both holes in one trial. Each individual received one session/day until reaching a learning criterion that was set at a minimum of 14 'correct' responses (trials in which the monkey chose a baited hole) from 3 consecutive sessions (including at least seven 'correct' s+s- choices) out of a possible 18 test trials. Individuals that had reached the learning criterion continued on to the cognitive-bias test.

Cognitive-bias test

In the next stage, the monkeys were given one 20-trial session per day for five consecutive days (randomized nine reinforced S++ trials, nine reinforced s+s- trials, and two *non* reinforced trials with the intermediate probe presented). During these sessions, the number of S++ like choices (i.e. choice of the hole that was baited with the preferred reward during the learning stage) when the medium rectangle was presented were measured. Trial maximum duration and ITIs were the same as during the learning sessions. Responses to the medium probe were used to evaluate the monkeys' affective state. A s+s- like response (i.e. choice of the hole that was baited with the less preferred reward during the learning stage) would imply a more 'pessimistic'-like judgment, and a S++ like response, a more 'optimistic' one.

Statistical analysis

Data were statistically analysed using SPSS 18; graphs were created using SigmaPlot 12.0; Fig. 1 was created using Adobe Photoshop CS3. Parametric statistics (repeated-measures ANOVA, Pearson correlation) were used when assumptions for parametric analysis were met. Non-parametric statistics (Friedman's test, Spearman's correlation) were employed when parametric assumptions were violated. Alpha was set at 0.05.

Results

Assessment of behavioural stereotypes

Stereotypic HT ranged from 0.01 to 7.17 occurrences/min with a mean \pm SD of 2.28 ± 2.44 , while proportions of stereotypic pacing (% of time) ranged from 0.00 to 30.75 %, with a mean \pm SD of 8.94 ± 10.65 %. Both these stereotypic behaviours were found to be significantly positively correlated ($n = 16$, $r_s = 0.72$, $p = 0.002$) (Table 1; Fig. 2).

Judgment-bias task

Twelve out of 16 monkeys reached the testing stage following 14.42 ± 8.44 sessions (mean \pm SD). Only one subject out of the four that did not reach the learning criteria showed a relatively high proportion of pacing (22.9 % of observation time), and high rates of HT (4.73

occurrences/min), while the other three had comparatively low values of pacing (2.2, 0, 2 %) and low rates of HT (0.63, 0.01, 0.12 occurrences/min) (see Table 1). For those individuals that did reach the testing stage, rectangle length (S++, medium, and s+s-) affected the probability of choosing the hole associated with the preferred reward, reflecting a generalization response (Friedmans' test: $\chi^2(2) = 19.50, p < 0.01$). The probability of choosing the hole associated with the preferred reward for the S++, medium, and s+s- stimuli was $89 \pm 10 \%$, $46 \pm 21 \%$ and $35 \pm 18 \%$, respectively (Fig. 3). In the S++ versus s+s- trial, there was a statistically significant reduction in the probability of choosing the hole associated with the preferred reward (Wilcoxon signed-rank tests with Bonferroni correction applied, $Z = -3.06, p = 0.002$), and between S++ and the medium probe ($Z = -3.06, p = 0.002$). In addition, there was a reduction, albeit not significant ($Z = -1.57, p = 0.11$), in the probability of choosing the preferred reward in the medium probe trials compared to the s+s- trials.

Head twirls (HT)

With regard to the link between stereotypies and judgment bias, there was a significant negative correlation between rates of HT and the probability of choosing the hole associated with the preferred reward when the medium probe was presented ($r_s = -0.59, p = 0.04$, Table 1; Fig. 4). In addition, assessing the connection between stereotypic behaviours and physiological indicators of emotional arousal revealed that rates of HT were positively correlated with levels of faecal corticosterone ($r_p = 0.66, p = 0.02$) (Table 1; Fig. 5). Out of the 16 individuals, four females that were pregnant during the study were excluded from hormonal analyses since pregnancy has been previously reported to influence the levels of these hormones (Bell et al. 1991).

Pacing

Unlike rates of HT, the proportions of stereotypic pacing were not significantly correlated with the probability of choosing the hole associated with the preferred reward when the medium probe was presented ($r_s = -0.10, p = 0.76$). However, similar to rates of head twirls, there was a significant correlation between proportion of ($r_s = 0.74, p = 0.006$, Table 1; Fig. 6).

There was no significant correlation between the proportion of time the monkeys engaged in overall active behaviour (including stereotypy) and the levels of faecal corticosterone ($r_s = 0.32, p = 0.31$) (see Table 1).

Discussion

This study reports for the first time a positive correlation between a particular form of stereotypic behaviour and a 'pessimistic'-like judgment in a non-human primate. 'Pessimistic'-like judgment has been previously defined as the inclination of a subject to anticipate relatively negative outcomes in response to affectively ambiguous stimuli (Matheson et al. 2008; Mendl et al. 2009). Indeed, tufted capuchins displaying higher levels of stereotypic head twirls (HT) displayed a lower probability to choose the hole associated with the preferred reward when presented with the medium probe compared to individuals with lower levels of this abnormal repetitive behaviour. The present findings suggest that individuals expressing higher levels of HT anticipated the less preferred reward when presented with an ambiguous stimulus and can therefore be considered to have a more 'pessimistic'-like judgment; while those with low (or no) HT levels tended to interpret the ambiguous probe as a precursor of the preferred reward and can therefore be considered to have a more 'optimistic'-like judgment. These results are in agreement with those reported

by Brilot et al. (2010), wherein European starlings that exhibited stereotypic somersaults were found to interpret ambiguous stimuli in a more 'pessimistic' manner.

Although the results show that capuchins with higher levels of HT also displayed higher levels of stereotypic pacing, there was no significant correlation between stereotypic pacing and the probability to choose the hole associated with the preferred reward when the medium length rectangle was presented. This discrepancy suggests that stereotypic pacing may be a weaker marker of a negative mood state than HT, since individuals that expressed higher levels of stereotypic pacing did not process the ambiguous stimulus in a more negative manner than monkeys with low levels of that behaviour whereas subjects with higher rates of HT did judge this intermediate cue as a sign predicting a less preferred outcome. This finding is of great interest, since researchers tend to clump various categories of abnormal repetitive behaviours (e.g. thumb-sucking, rocking, self-biting) under one umbrella of stereotypic behaviour or to separate them according to their appearance (e.g. whole body stereotypy, self-directed stereotypy; Mason and Rushen 2006) whereas here it is suggested that certain behaviours should be separated according to their relevance to the mood state of the individuals, thus adding a specific affective characterization to a specific behaviour. A deeper understanding of the affective state that is related to each abnormal repetitive behaviour can potentially contribute to the treatment of those individuals that express them.

In zoo-housed carnivores, levels of pacing were found to be correlated with the habitat-range size, and the daily distance travelled in the wild (Clubb and Mason 2003). Therefore, although regarded as a stereotypy, pacing may act as a way for the animal to expend energy rather than reflecting specific affective states. From the above findings, it would seem that combining several stereotypic behaviours into one behavioural aggregate is probably misleading from a welfare point of view.

In order to further investigate the mood state of stereotyping capuchins, we incorporated hormonal measurements of extensively used 'stress'-indicators. It was previously suggested that stress is a behavioural and physiological accumulation of reactions to aversive stimuli that most likely induces negative emotions (Paul et al. 2005). Therefore, it was predicted that both types of measurement (judgment bias and faecal corticosterone) would be significantly correlated. Such correlation would increase the robustness of the findings and strengthen the ensuing conclusions.

Indeed, levels of faecal corticosteroids, used as a measure of stress in many studies (Keay et al. 2006; Weingrill et al. 2011; Beehner and Whitten 2004), were found to be positively correlated with both types of stereotypic behaviours. Furthermore, it is becoming evident that the interpretation of physiological indicators alone may be misleading, since at least some measurements (e.g. corticoids and heart rate) are probably better markers of arousal or of the magnitude of the emotional reaction, rather than of the valence (positive or negative) of the affective state (Leknes and Tracey 2008; Lewis and Bodfish 2007; Paul et al. 2005). Combining the valence with the level of arousal (intensity of the emotional response) provides a richer image of the individual's affective state and may contribute to a better analysis of that individual's welfare.

Although activity levels may also influence the levels of corticosteroids (Coiro et al. 2011), our analysis found no connections here between these two factors. This finding suggests that the significant correlation found between HT levels and corticosterone indeed reflects the higher intensity of a negative emotional state experienced by high stereotypers.

In contrast, the correlation found between percentage of time spent pacing and corticosterone levels is probably not indicative of stress levels, since a state of stress is likely

to be accompanied by a negative affective state (Koolhaas et al. 1997), which our analysis failed to reveal with regard to pacing. However, although a significant correlation was found between corticosterone levels and the two types of stereotypic behaviours, these results should be interpreted with caution since only three faecal samples were collected for each subject. Thus, in order to validate the current findings, in future studies more samples should be obtained for each subject.

Thus, integrating the above two types of measurements (judgment bias and faecal corticosterone) yielded a detailed representation of the monkeys' emotional state, revealing that individuals with high rates of HT are not only in a negatively valenced state, but additionally experience (not necessarily consciously) a high intensity of that emotion.

It has been previously suggested that using stressed captive animals in scientific research can yield distorted data, affected by factors such as the animals' fear, anxiety and depression, thus impairing the credibility of the results (Reinhardt 2002). In particular, it was suggested that in studies related to cognition, the manner by which animals handle the stress of captivity may influence their emotional state and therefore introduce unaccounted-for factors in cognitive tasks (Brilot et al. 2010). Others have claimed that stereotypic behaviours may stem from abnormal brain function, and therefore, incorporating stereotypers in an experiment could potentially invalidate the data (Garner 2006). The current study shows that tufted capuchins with high rates of HT are in a more negative affective state than their non-stereotyping conspecifics. It can thus be suggested that using capuchins that express HT should either be considered with caution, in order to avoid the potential influence of unaccounted-for parameters that might put the research findings into question, or the influence of their affective state should also be evaluated. Further research is needed to reveal whether stereotyping subjects increase the variability of data gathered in psychology and neuroscience research.

Finally, there is a wealth of literature regarding the connection between abnormal stereotypic behaviour and the welfare of those that express it (Clubb and Mason 2003; Garner 2005; Mason and Rushen 2006; Mason 1991; Mason et al. 2001; Mason and Latham 2004; Quirke and O'Riordan 2011; Wurbel 2001; Stewart et al. 2011). From these studies, it is clear that impoverished housing conditions and chronic exposure to aversive stimuli are strongly linked to the expression of stereotypies. Despite the argument by some researchers that, within a group of captive animals, those that stereotype experience a better welfare state than non-stereotyping animals (possibly through some sort of stress-coping mechanism; Mason and Latham 2004), the results of the present study suggest otherwise: animals with high levels of HT were found to judge an ambiguous stimulus in a more negative manner than their counterparts and to be more negatively emotionally aroused.

While the expression of abnormal behaviours alone may reflect 'scars' induced by past experiences (Mason and Latham 2004), or damage to an underlying neurological mechanism (Garner et al. 2003), testing the connection between stereotypies and judgment bias provides a representation of the individual's current affective state. Since the subject's affective state is inherently related to its well-being (Dawkins 2006), employing a method that measures both the valence of the affective state and the subject's level of arousal enables us to portray the subjective condition of the animal's well-being. This rather simple method has yielded similar results in a number of different animal species across several orders (Brydges et al. 2011; Mendl et al. 2010a; Matheson et al. 2008) [perhaps through a shared adaptive mechanism in which the negative mood state of animals living in a threatening environment increases their fitness by 'guiding' them to avoid ambiguous stimuli that are very likely to jeopardize them (Mendl et al. 2010b)], thus reinforcing the reliability and validity of this technique. Head-twirling subjects are more likely to be in a negative emotional state and

therefore require special attention from the researchers and care-givers responsible for their well-being, as reflected in close monitoring of their welfare state and greater emphasis on appetitive stimuli within the individuals' environment.

Finally, similar to assessment of a 'pessimistic'-like judgment bias, the paradigm employed here also allows for the evaluation of an 'optimistic'-like judgment bias, rendering it a practical welfare assessment tool. Indeed, Matheson et al. (2008) reported that European starlings housed in larger enriched cages showed significantly increased 'optimism' than when housed in smaller standard cages. Accordingly, it is predicted that an 'optimistic'-like judgment bias will be correlated with behaviours related to positive emotions such as social grooming. This issue should be determined in future research.

In conclusion, the present study shows for the first time that head-twirling, but not stereotypic pacing, is a reliable indicator of the well-being of tufted capuchins. By utilizing a novel approach combining cognitive processes with hormonal analyses, we now have a greater possibility of understanding the inner emotional world of captive animals and of harnessing that understanding to the provision of better care for those animals.

Acknowledgments

We wish to thank Ernest Davis, Seth Bower On Pomerantz, and Naomi Paz for her help in preparing and editing this manuscript.

References

- Bateson M, Matheson SM. Performance on a categorisation task suggests that removal of environmental enrichment induces 'pessimism' in captive European starlings (*Sturnus vulgaris*). *Anim Welf*. 2007; 16(Supplement 1):33–36.
- Beehner JC, Whitten PL. Modifications of a field method for fecal steroid analysis in baboons. *Physiol Behav*. 2004; 82(2–3):269–277. [PubMed: 15276788]
- Bell ME, Wood CE, Keller-Wood M, Kane C, Kluwe C, Manlove E, Taranovich C, Johnson J. Influence of reproductive state on pituitary-adrenal activity in the ewe. *Domest Anim Endocrinol*. 1991; 8(2):245–254. [PubMed: 1649030]
- Brilot B, Asher L, Bateson M. Stereotyping starlings are more "pessimistic". *Anim Cogn*. 2010; 13(5): 721–731.10.1007/s10071-010-0323-z [PubMed: 20464439]
- Brydges NM, Leach M, Nicol K, Wright R, Bateson M. Environmental enrichment induces optimistic cognitive bias in rats. *Anim Behav*. 2011; 81(1):169–175.
- Clubb R, Mason G. Animal welfare: captivity effects on wide-ranging carnivores. *Nature*. 2003; 425(6957):473–474. [PubMed: 14523435]
- Coiro V, Volpi R, Casti A, Maffei ML, Stella A, Volta E, Chiodera P. Naloxone decreases the inhibitory effect of alprazolam on the release of adrenocorticotropin/cortisol induced by physical exercise in man. *Br J Clin Pharmacol*. 2011; 71(6):951–955.10.1111/j.1365-2125.2010.03900.x [PubMed: 21564163]
- Dawkins MS. Behaviour as a tool in the assessment of animal welfare. *Zoology*. 2003; 106(4):383–387. [PubMed: 16351922]
- Dawkins MS. A user's guide to animal welfare science. *Trends Ecol Evol*. 2006; 21(2):77–82. [PubMed: 16701478]
- Dayas CV, Buller KM, Day TA. Neuroendocrine responses to an emotional stressor: evidence for involvement of the medial but not the central amygdala. *Eur J Neurosci*. 1999; 11(7):2312–2322.10.1046/j.1460-9568.1999.00645.x [PubMed: 10383620]
- Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull*. 2004; 130(3):355–391. [PubMed: 15122924]

- Diener E, Oishi S, Lucas RE. Personality, culture, and subjective well-being: emotional and cognitive evaluations of life. *Annu Rev Psychol.* 2003; 54(1):403–425.10.1146/annurev.psych.54.10160.1.145056 [PubMed: 12172000]
- Garner JP. Stereotypies and other abnormal repetitive behaviors: potential impact on validity, reliability, and replicability of scientific outcomes. *ILAR.* 2005; 46(2):106–117.
- Garner, JP. Perseveration and stereotypy—system-level insights from clinical psychology. In: Mason, GJ.; Rushen, J., editors. *Stereotypic animal behaviour: fundamentals and applications to welfare.* 2. CABI; Wallingford: 2006. p. 121-152.
- Garner JP, Meehan CL, Mench JA. Stereotypies in caged parrots, schizophrenia and autism: evidence for a common mechanism. *Behav Brain Res.* 2003; 145(1–2):125–134. [PubMed: 14529811]
- Hockly E, Cordery PM, Woodman B, Mahal A, Van Dellen A, Blakemore C, Lewis CM, Hannan AJ, Bates GP. Environmental enrichment slows disease progression in R6/2 Huntington's disease mice. *Ann Neurol.* 2002; 51(2):235–242.10.1002/ana.10094 [PubMed: 11835380]
- Keay JM, Singh J, Gaunt MC, Kaur T. Fecal glucocorticoids and their metabolites as indicators of stress in various mammalian species: a literature review. *J Zoo Wildl Med.* 2006; 37(3):234–244.10.1638/05-050.1 [PubMed: 17319120]
- Koolhaas JM, Meerlo P, de Boer SF, Strubbe JH, Bohus B. The temporal dynamics of the stress response. *Neurosci Biobehav Rev.* 1997; 21:775–782. [PubMed: 9415902]
- Lanovaz MJ. Towards a comprehensive model of stereotypy: integrating operant and neurobiological interpretations. *Res Dev Disabil.* 2011; 32(2):447–455. [PubMed: 21236636]
- Leknes S, Tracey I. A common neurobiology for pain and pleasure. *Nat Rev Neurosci.* 2008; 9(4): 314–320. [PubMed: 18354400]
- Lewis MH, Bodfish JW. Repetitive behavior disorders in autism. *Ment Retard Dev Disabil Res Rev.* 2007; 4:80–89.
- Lupien SJ, Maheu F, Tu M, Fiocco A, Schramek TE. The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain Cogn.* 2007; 65(3):209–237. [PubMed: 17466428]
- Martin, P.; Bateson, P. *Measuring behaviour.* 2. Cambridge University Press; Cambridge: 1993.
- Mason G. Stereotypies and suffering. *Behav Process.* 1991; 25(2–3):103–115.
- Mason GJ, Latham NR. Can't stop, won't stop: is stereotypy a reliable animal welfare indicator? *Anim Welf.* 2004; 13:57–69.
- Mason, G.; Rushen, J. *Stereotypic animal behaviour: fundamentals and applications to welfare.* 2. CABI Pub; Wallingford: 2006.
- Mason GJ, Cooper J, Clarebrough C. Frustrations of fur-farmed mink. *Nature.* 2001; 410(6824):35–36. [PubMed: 11242031]
- Matheson SM, Asher L, Bateson M. Larger, enriched cages are associated with 'optimistic' response biases in captive European starlings (*Sturnus vulgaris*). *Appl Anim Behav Sci.* 2008; 109(2–4): 374–383.
- Mendl M, Burman OHP, Parker RMA, Paul ES. Cognitive bias as an indicator of animal emotion and welfare: emerging evidence and underlying mechanisms. *Appl Anim Behav Sci.* 2009; 118(3): 161–181.
- Mendl M, Brooks J, Basse C, Burman O, Paul E, Blackwell E, Casey R. Dogs showing separation-related behaviour exhibit a 'pessimistic' cognitive bias. *Curr Biol.* 2010a; 20(19):R839–R840. [PubMed: 20937467]
- Mendl M, Burman OHP, Paul ES. An integrative and functional framework for the study of animal emotion and mood. *Proc R Soc B Biol Sci.* 2010b; 277(1696):2895–2904.10.1098/rspb.2010.0303
- Mills DS. Medical paradigms for the study of problem behaviour: a critical review. *Appl Anim Behav Sci.* 2003; 81(3):265–277.
- Moberg, GP. Biological response to stress: implications for animal welfare. In: Moberg, GP.; Joy, AM., editors. *The biology of animal stress: basic principles and implications for animal welfare.* CABI; New York: 2001. p. 1-22.
- Paul ES, Harding EJ, Mendl M. Measuring emotional processes in animals: the utility of a cognitive approach. *Neurosci Biobehav Rev.* 2005; 29(3):469–491. [PubMed: 15820551]

- Polivy J. The effects of behavioral inhibition: integrating internal cues, cognition, behavior, and affect. *Psychol Inq.* 1998; 9(3):181–204.10.1207/s15327965pli0903_1
- Pomerantz O, Terkel J. Effects of positive reinforcement training techniques on the psychological welfare of zoo-housed chimpanzees (*Pan troglodytes*). *Am J Primatol.* 2009; 71(8):687–695.10.1002/ajp.20703 [PubMed: 19434627]
- Puliafico A, Kendall P. Threat-related attentional bias in anxious youth: a review. *Clin Child Fam Psychol Rev.* 2006; 9(3):162–180.10.1007/s10567-006-0009-x [PubMed: 17053961]
- Quirke T, O’Riordan RM. The effect of a randomised enrichment treatment schedule on the behaviour of cheetahs (*Acinonyx jubatus*). *Appl Anim Behav Sci.* 2011; 135(1–2):103–109.
- Reinhardt, V. Comfortable quarters for laboratory animals. 9. Animal Welfare Institute; Washington DC: 2002.
- Rutter M, Andersen-Wood L, Beckett C, Bredenkamp D, Castle J, Groothues C, Kreppner J, Keaveney L, Lord C, O’Connor TG. Quasi-autistic patterns following severe early global privation. *J Child Psychol Psychiatry Allied Discipl.* 1999; 40:537–549.
- Salmeto AL, Hymel KA, Carpenter EC, Brilot BO, Bateson M, Sufka KJ. Cognitive bias in the chick anxiety-depression model. *Brain Res.* 2011; 1373:124–130. [PubMed: 21156165]
- Stewart CL, Boyle LA, O’Connell NE. The effect of increasing dietary fibre and the provision of straw racks on the welfare of sows housed in small static groups. *Anim Welf.* 2011; 20(4):633–640.
- Wasser SK, Hunt KE, Brown JL, Cooper K, Crockett CM, Bechert U, Millspaugh JJ, Larson S, Monfort SL. A generalized fecal glucocorticoid assay for use in a diverse array of nondomestic mammalian and avian species. *Gen Comp Endocrinol.* 2000; 120(3):260–275. [PubMed: 11121291]
- Weingrill T, Willems EP, Zimmermann N, Steinmetz H, Heistermann M. Species-specific patterns in fecal glucocorticoid and androgen levels in zoo-living orangutans (*Pongo* spp.). *Gen Comp Endocrinol.* 2011; 172(3):446–457. [PubMed: 21524655]
- Wiedenmayer C. Causation of the ontogenetic development of stereotypic digging in gerbils. *Anim Behav.* 1997; 53(3):461–470.
- Wurbel H. Ideal homes? Housing effects on rodent brain and behaviour. *Trends Neurosci.* 2001; 24(4): 207–211. [PubMed: 11250003]

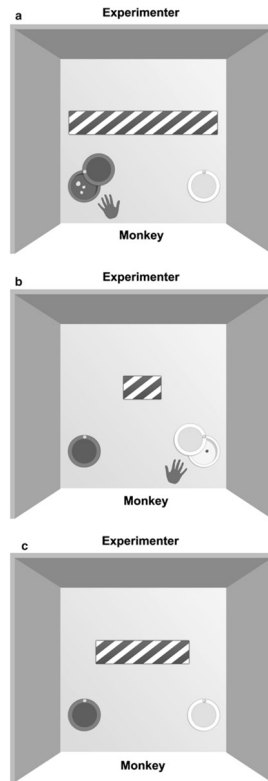


Fig. 1.

Top view of the apparatus. During the learning stage, the monkeys learned to discriminate between the two situations depicted in **(a)** and **(b)**: **a** for example, in trials when the subjects were presented with the long striped rectangle they had to slide the *black cap* in order to reveal the hole and gain the preferred reward, while the other hole covered by the white cap remained empty; **b** in trials when the monkeys were presented with the short striped rectangle, they needed to choose the *white-capped* hole in order to obtain the less preferred reward, while the other hole remained empty; **c** during the cognitive-bias tasks, monkeys were given trials similar to those in the learning stage, in addition to test trials in which subjects were presented with an intermediate length striped rectangle (the medium probe), and their responses (to the *black-* or *white-*capped holes) were recorded. For additional details see ‘Materials and methods’

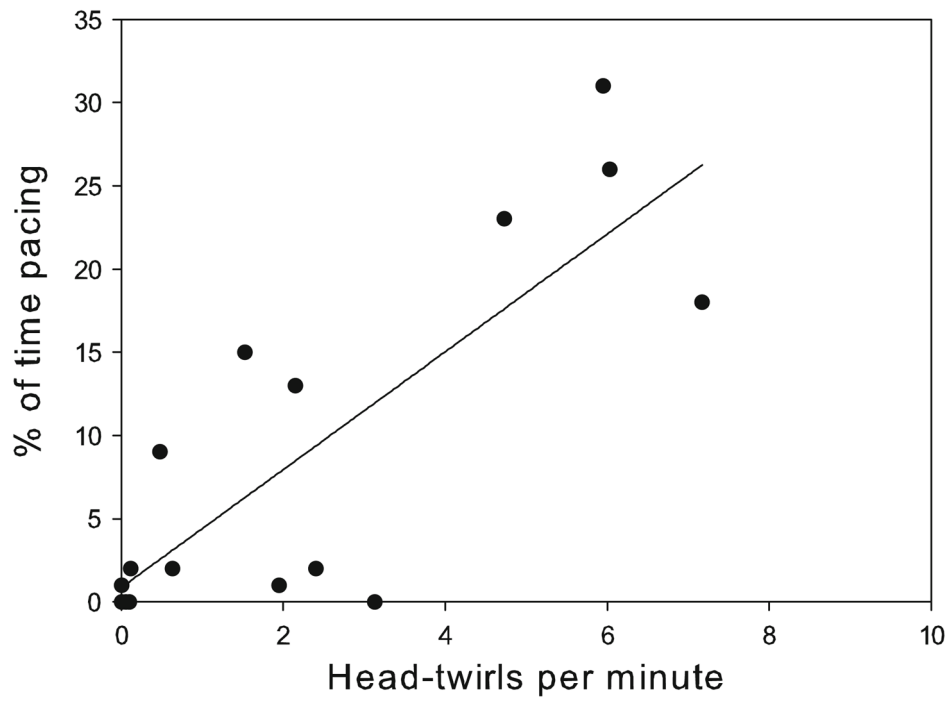


Fig. 2.
Correlation between the two types of stereotypic behaviour: head twirls and pacing, $n = 16$

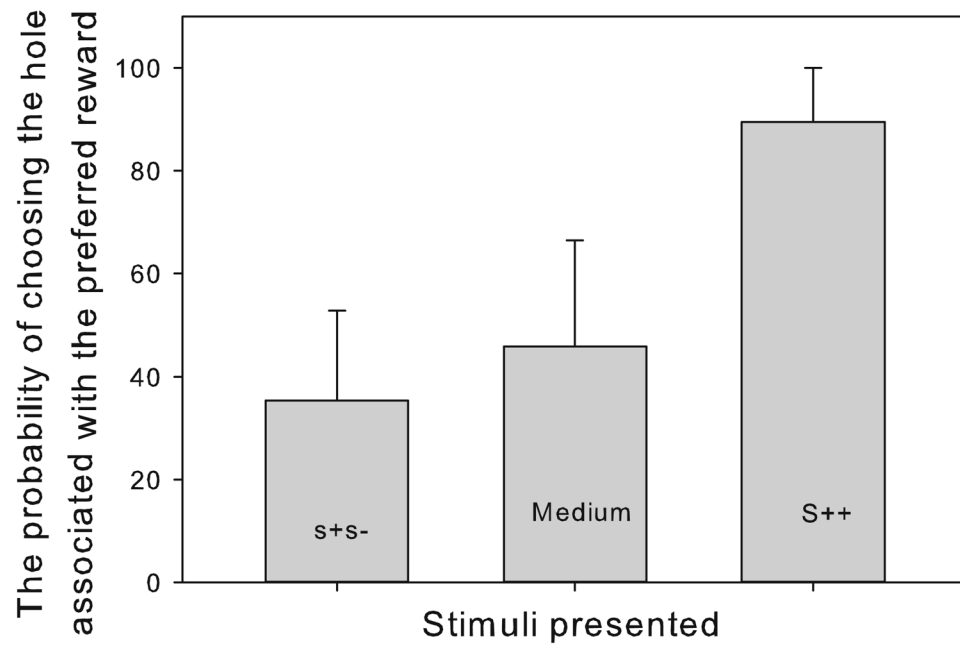


Fig. 3. The average probability of all subjects that reached the learning criterion ($n = 12$) of choosing the hole associated with the preferred reward during cognitive-bias trials when presented with S++, medium probe, and s+s-

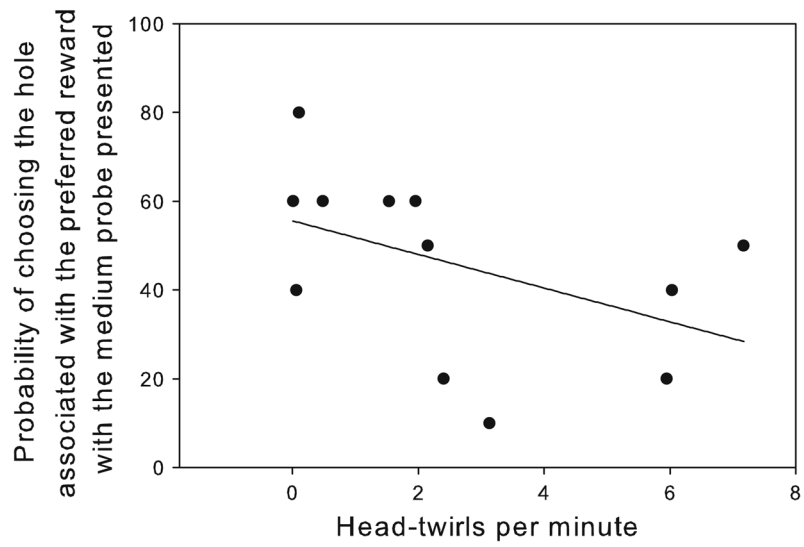


Fig. 4. Correlation between head twirls per minute and the probability of choosing the hole associated with the preferred reward when the medium probe was presented, $n = 12$

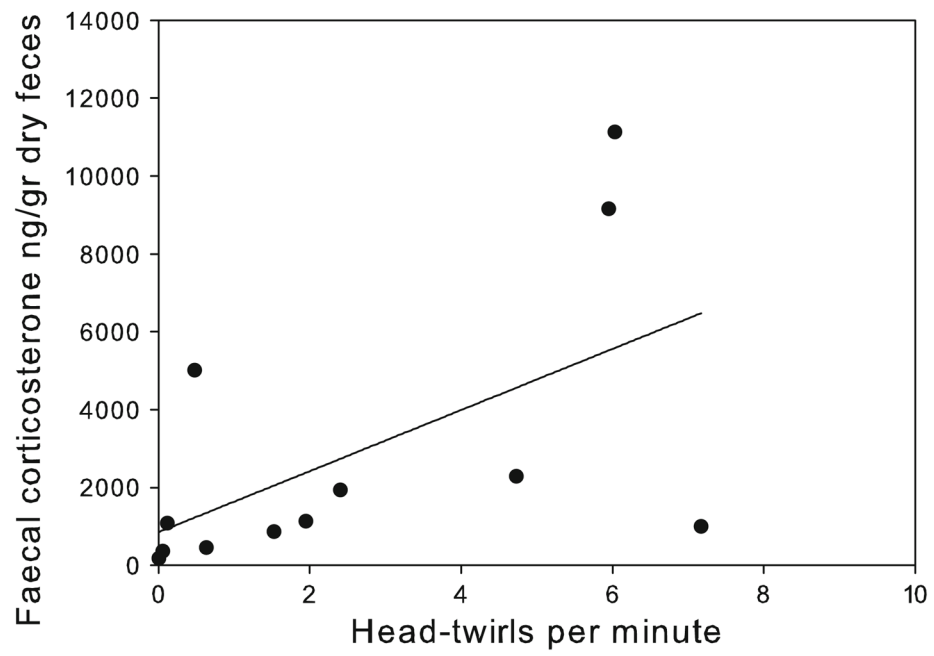


Fig. 5. Correlation between rates of head twirls and faecal corticosterone levels, $n = 12$ stereotypic pacing and level of faecal corticosterone

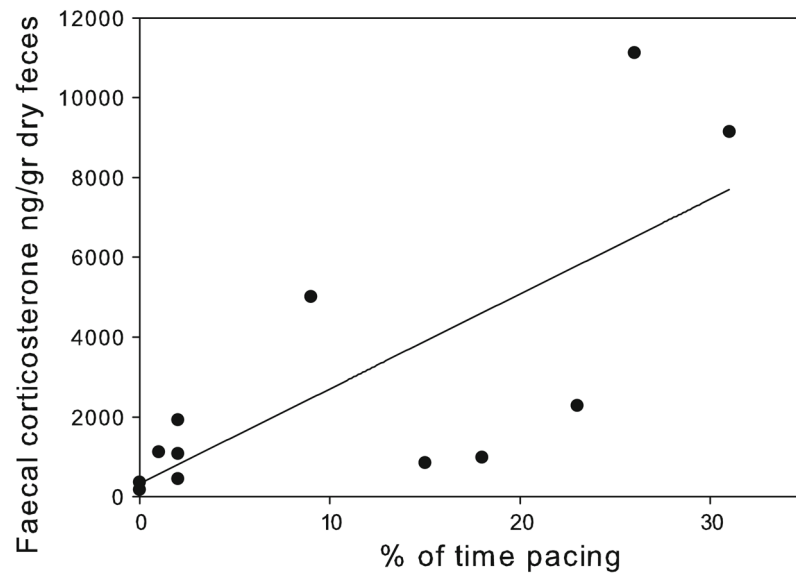


Fig. 6. Correlation between proportion of stereotypic pacing and faecal corticosterone levels, $n = 12$

Table 1

Age, sex, housing situation, levels of stereotypic behaviours, activity levels, proportion of choosing the preferred reward on 'probe' trials, and corticosterone levels

Subject's name	Age (months)	Sex	Housing (cage/run)	HT/min	% of time pacing	% of time active	Proportion of choosing the preferred reward on 'probe' trials	Mean corticosterone (nanogram per gram dry faeces) \pm SD
Applesauce	73	F	Run	3.13	0.0	93.2	10	<i>b</i>
Garth	216	M	Run	0.48	8.6	71.1	60	5,016.47 \pm 2,136.85
Hotrod	101	M	Cage	1.53	15.0	95.6	60	860.91 \pm 763.99
Icarus	56	M	Cage	4.73	22.9	94.9	<i>a</i>	2,290.67 \pm 1,249.55
Irene	106	F	Run	0.01	1.3	86.8	60	<i>b</i>
Isabelle	360	F	Run	0.63	2.2	93.2	<i>a</i>	456.79 \pm 200.79
Ivory	146	F	Run	0.01	0.0	80.1	<i>a</i>	185.81 \pm 20.92
Junior, Mint	79	M	Cage	6.03	26.3	99.9	40	11,126.10 \pm 1,683.21
Lenny	45	M	Cage	7.17	17.8	99.4	50	995.05 \pm 349.36
Lilliput	55	F	Run	2.40	1.6	97.1	20	1,937.24 \pm 681.67
Liv	138	F	Run	0.1	0.2	92.5	80	<i>b</i>
Lucy	300	F	Run	0.12	2.0	65.1	<i>a</i>	1,083.91 \pm 565.95
Lychee	137	F	Run	2.15	13.4	87.1	50	<i>b</i>
Manuel	119	M	Cage	0.06	0.0	95.0	40	368.80 \pm 239.11
Mr. Goodbar	74	M	Cage	5.95	30.8	99.3	20	9,150.95 \pm 1,0081.37
Shane	93	M	Cage	1.95	1.2	88.4	60	1,131.11 \pm 522.42

^aData are missing for four monkeys that did not reach the testing stage

^bData are missing for four pregnant females