



HAL
open science

Evaluating indicators of wellbeing in captive grey mouse lemurs, *Microcebus murinus*, in the context of enrichment

Juan Pablo Perea-Rodriguez, Julie Ferreira, Aude Anzeraey, Fabienne Aujard, Caroline Gilbert, Fabien Pifferi, Emmanuelle Pouydebat, J. Terrien, Audrey Maille

► To cite this version:

Juan Pablo Perea-Rodriguez, Julie Ferreira, Aude Anzeraey, Fabienne Aujard, Caroline Gilbert, et al.. Evaluating indicators of wellbeing in captive grey mouse lemurs, *Microcebus murinus*, in the context of enrichment. *Animal Behaviour*, 2024, 218, pp.121-134. 10.1016/j.anbehav.2024.09.013 . hal-04764563

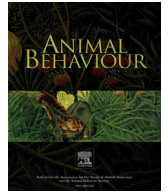
HAL Id: hal-04764563

<https://hal.science/hal-04764563v1>

Submitted on 7 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Evaluating indicators of wellbeing in captive grey mouse lemurs, *Microcebus murinus*, in the context of enrichment

Juan Pablo Perea-Rodriguez ^{a,*} , Julie Ferreira ^{a,b}, Aude Anzeraey ^c, Fabienne Aujard ^c ,
Caroline Gilbert ^{c,d}, Fabien Pifferi ^c, Emmanuelle Pouydebat ^c, Jérémy Terrien ^c,
Audrey Maille ^{a,e} 

^a Laboratoire d'Eco-anthropologie (UMR 7206 EA), Muséum National d'Histoire Naturelle, CNRS, Université de Paris Cité, Paris, France

^b Master « Ecophysiologie, Ecologie et Ethologie », Université de Strasbourg, Strasbourg, France

^c Laboratoire Mécanismes Adaptatifs et Evolution (UMR 7179 MECADEV), Muséum National d'Histoire Naturelle, CNRS, Brunoy, France

^d Ecole Nationale Vétérinaire d'Alfort, Maisons-Alfort, France

^e Pôle Parcs zoologiques, DGD Musées, Jardins et Zoos, Muséum National d'Histoire Naturelle, Paris, France

ARTICLE INFO

Article history:

Received 2 November 2023

Initial acceptance 6 March 2024

Final acceptance 20 August 2024

MS. number: ANBEH-D-23-00570

Keywords:

behavioural motivation
body composition
circadian activity
grey mouse lemur
machine learning
Microcebus murinus
principal component analysis
Strepsirrhini
welfare

Developing and implementing effective approaches to improve the welfare of animals participating in research requires reliable monitoring of their wellbeing. Since wellbeing is a multidimensional state that relates to an individual's physiological, behavioural and psychological/mental health, assessing it requires the use of measures that are both ecologically relevant and species-specific. Grey mouse lemurs are small, nocturnal, arboreal and omnivorous primates and a powerful model for fundamental and translational research. As such, it is important to study the factors that may help predict their wellbeing in captivity. Therefore, we monitored 12 adult mouse lemurs (six females and six males) housed with a same-sex weight-matched cage mate for 5 weeks under standard (Control) levels of environmental enrichment and 12 adults (six females and six males) under relatively High enrichment. We hypothesized that higher structural enrichment would stimulate metabolic energy expenditure, influencing how individuals would respond to a battery of behavioural and physiological measures, especially since their feeding rations remained unchanged throughout the study. We characterized circadian activity, judgement bias, predatory motivation, thermal reactivity to stress and changes in body composition. Generally, mouse lemurs under High enrichment showed more activity and increased predatory motivation and optimism than those in the Control group during daylight. Lemurs under High enrichment also showed increased responses during a stress reactivity test and higher losses in body and fat mass. Principal component analysis identified seven (out of nine) components that explained 96% of the variability in the data. Using a Naïve Bayes classifier algorithm trained with these new principal components, we classified the animals as either from High or Control enrichment levels with 100% accuracy. In summary, by measuring several relevant behavioural and physiological responses to environmental enrichment, we identified a multidimensional effect related to differences in welfare.

© 2024 The Authors. Published by Elsevier Ltd on behalf of The Association for the Study of Animal Behaviour. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The practical use of animals by humans has been an important tool for the improvement of our health and quality of life. Notably, the development of animal models for translational research has been essential for increasing our understanding of the ultimate and proximate mechanisms of complex biological processes. Although small rodents are the most commonly used taxa for this type of research, other vertebrates, including nonhuman primates, can be

more suitable models because of phylogenetic, morphological, physiological and psychological similarities to humans (Harding, 2017; Shively & Clarkson, 2009). Certain species provide unique opportunities to undertake research questions related to human biology owing to similarities in relevant traits, such as reproductive ecology (e.g. parental care: Saltzman et al., 2017) or ageing processes (e.g. changes in cognition: Heuer et al., 2012).

* Corresponding author.

E-mail address: pereajp@gmail.com (J. P. Perea-Rodriguez).

Animal research, however, often involves the maintenance of captive animals in appropriate conditions, including providing a social environment that is relevant to their behavioural ecology, which may ultimately promote reproduction. Indeed, housing captive animals in unsuitable environments introduces additional confounding factors that hinder how any data collected from them can be used and interpreted (Baker & Dettmer, 2017; Buchanan-Smith et al., 2005; Nelson & Mandrell, 2005; Russell & Burch, 1959). Regardless of what species is used, it is also necessary for researchers to evaluate, monitor and improve the physical and psychological health of the animals used in science (Clark, 2017; Coleman, 2011). For example, in situations in which breeding colonies are maintained, managing stress becomes indispensable for the health of the population, as high stress levels can interfere with regular metabolic and reproductive functions (Tardif et al., 2012). Unfortunately, ignoring the wellbeing of captive animals used in research leads to poor science owing to inappropriate welfare practices (Poole, 1997).

To ensure that captive animals are treated ethically, ongoing efforts and policies are being increasingly set forward. This has culminated in the conception of animal welfare as a formal scientific field (Buchanan-Smith et al., 2005). Animal welfare is a multidisciplinary science that includes sociology, biology, agriculture and veterinary medicine, among other fields, and has led to the advancement of theoretical frameworks and protocols that improve the overall health and wellbeing of animals under human care (in farms, zoos, aquariums, laboratories, etc.; Barber, 2009). Wellbeing can be generally defined as the satisfaction of an individual's physical, physiological and psychological needs (Broom, 1991). Importantly, wellbeing is a subjective state as it depends on how individuals evaluate their own situation, and wellbeing may thus change with time for a given individual (Dodge et al., 2012). As such, attempts to evaluate wellbeing effectively require not only focusing on reducing any negative effects of the captive experience and fulfilling an animal's basic needs but also enriching their lives to allow them to live a 'life worth living' (Mellor, 2016). Another important aspect of positive welfare practices is developing reliable physiological and behavioural indices that can characterize different dimensions of individual wellbeing, including physical, physiological and psychological dimensions, at a given time and in any given environment.

Some of the most common practices used to promote the wellbeing of captive primates include environmental enrichment, which may consist of providing them with comfortable enclosures and temperature, humidity and lighting, as well as cagemates and specialized diets (Honeß & Marin, 2006). Environmental enrichment can take different forms related to gustatory, olfactory, tactile, auditory and visual stimuli, as well as opportunities for interactions with objects, conspecifics, caretakers and other humans (Hoy et al., 2010; Maple, 2013). Providing captive animals with variable environments enhances their psychological wellbeing, as highly predictable environments may promote abnormal behaviours due to a lack of stimulation, whereas variable ones can provide novel behavioural opportunities (Young, 2003). Managing resource-based measures (e.g. enclosure size, luminosity and temperature) has been essential for the development of animal welfare practices, but one main criticism is that such an approach ignores any measure of how individuals respond to their environment (i.e. individual-based measures). Thus, developing measures of how individuals respond to such resource-based measures play a crucial role in the evaluation of wellbeing, as they describe how (or whether) individuals respond to changes in the (presumed/perceived) quality of their environment.

Indeed, one of the initial methods used to evaluate the wellbeing of captive animals was through resource-based measures,

which describe an individual's environment (Buchanan-Smith, 2011). For example, welfare approaches may regulate how much space, food and substrate an individual should have in its enclosure or other physical characteristics of their environment. Individual-based measures may include physiological and/or behavioural variables that quantify emotional state in a given environment at a given moment. For example, biomarkers for immune, adrenal, thyroid and/or gonadal function as well as sexual behaviours can inform us about reproductive function (Stevenson & Ball, 2011).

There are several ways in which individual-based measures are being used to assess the wellbeing of nonhuman primates in captivity. Behavioural tests measuring cognitive biases assess how cognitive functions and processing, particularly attention and memory (Bethell, et al., 2012), may be affected by emotional changes. Monitoring facial temperature and expression reveals distinct patterns that are linked to emotional responses in various mammals and birds (Chotard et al., 2018). The presence or duration of specific behaviours, such as an overall reduction in locomotor activity, food and water intake (i.e. anorexia), motivation to interact with novel stimuli (i.e. neophobia) or increased aggression, self-grooming and scratching, are related to stress (sometimes referred to as fear, anxiety, depression or anhedonia, i.e. lack of motivation to perform pleasurable activities) in various species of vertebrates, including primates (Coleman & Pierre, 2014). Monitoring long- and short-term changes in adrenal function in mammals has been commonly used to model fluctuations of how individuals may be using metabolic energy (Joseph & Whirledge, 2017). Similarly, following any major fluctuations in body composition (e.g. % fat and lean mass) is a useful way to quantify major changes in metabolic energy expenditure (Helfer et al., 2019). Finally, changes in circadian behavioural patterns can be an efficient energy-saving strategy used by metabolically challenged mammals and other endotherms (van der Vinne et al., 2014, 2015, 2019).

Currently, the state-of-the-art approaches to monitor and evaluate the wellbeing of captive animals have automated the way behavioural data are collected and processed for interventions. For example, Computer Vision is now a common tool in many production farms that continuously process video data captured from carefully selected places to determine the presence or absence of a specific behavioural posture (Rao et al., 2020) or vocalization in the case of computer hearing (Cuan et al., 2022). The benefits of these methods, aside from their simplicity to perform and maintain, are that they provide the opportunity to provide both individual and population level monitoring, they are noninvasive and they reduce the responsibility of workers. Since these machine learning algorithms are trained with the data provided, it is imperative that these data are ecologically relevant to the study species. Similarly, if the goal is to evaluate and predict individual level wellbeing, then predictive models will be most accurate when they have been trained with an individual's data.

All the aforementioned resource- and individual-based indices can aid in the evaluation of wellbeing of captive nonhuman primates. However, they have to be used synergistically to fit within the conceptual framework related to animal welfare, which aims to achieve a comprehensive understanding of the wellbeing of an animal. Thus, there is a need to develop multidimensional and reliable indicators for the emotional states of primates (and other animals) using both resource- and individual-based measures and to integrate them to assess the wellbeing of animals in a comprehensive state. This approach can be useful to understand the mechanisms by which wellbeing may be affected by environmental changes.

In this study, we aimed to identify physiological and behavioural indicators of wellbeing in captive grey mouse lemurs. Mouse

lemurs are naturally found in Madagascar and are generally solitary foragers. Females sleep and often forage in the company of related kin, whereas males prefer to sleep alone (Radespiel, 2000). They are seasonal breeders and use several energy-saving strategies to counteract the high metabolic costs resulting from their small size (Naya et al., 2018). They are known for their daily torpor, which is an energy-saving strategy, and they have been an important model for research relating to their metabolic (e.g. Pifferi et al., 2018, 2019; Terrien et al., 2018; Vuarin & Henry, 2014) and reproductive (e.g. Huchard et al., 2017; Perret, 2021; Perret & Aujard, 2001) adaptations. Because of their nocturnal and arboreal habits and their shared ecology with early euprimates (Ho et al., 2021), they are an ideal model for studying the evolution of communicative abilities (e.g. Scheumann et al., 2007), grasping (e.g. Reghem et al., 2012), foraging cognition (e.g. Joly et al., 2008) and circadian biology (e.g. Hozer & Pifferi, 2022). They also show several interesting characteristics, such as fluctuating sexual dimorphism (Schmid & Kappeler, 1998), hygienic behaviours (e.g. Poirotte & Kappeler, 2019) and sleeping site sharing (Radespiel et al., 2003).

Here, we characterized the behavioural and physiological changes in adult male and female mouse lemurs, resulting from a 5-week-long exposure to two different levels of environmental enrichment in captivity to identify reliable measures for their wellbeing. More specifically, we used different acute and chronic indicators to model how exposure to structural enrichment may explain differences in wellbeing. We hypothesized that higher levels of environmental enrichment would stimulate metabolic energy expenditure, influencing how individuals would respond to a battery of behavioural and physiological measures. As such, we measured their circadian activity, predatory motivation, willingness to approach ambiguous stimuli (i.e. judgement bias), response to a 2 min manual restraint (i.e. stress reactivity) and body mass composition throughout the study. These approaches are currently used for monitoring the wellbeing of other captive mammals, including in the colony of mouse lemurs studied here (i.e. body composition measures). We used principal component analysis (PCA) as well as the latest machine learning algorithms to identify how well they could differentiate between the two levels of enrichment based on the physiological and behavioural data collected.

METHODS

Study Population and Subjects

Twelve adult males and 12 adult females (aged 5–7 years), born and raised from 10 lineages at the MICROCEBES platform hosted by the Muséum National d'Histoire Naturelle, participated in the study. The colony was founded in Brunoy, France in the 1970s and is currently considered one of the largest colonies of nonhuman primates in the world. Mouse lemurs are housed indoors in climate-controlled rooms, which are generally maintained at a constant temperature and humidity of 25 °C and 55%, respectively. The animals are exposed to fluctuating patterns of artificial photoperiod throughout the year with day length changes from winter (short days: 10:14 h light:dark) to summer (long days: 14:10 h light:dark). This protocol simulates their natural environment and stimulates reproductive function; they generally mate during the early summer and give birth during the summer (Génin & Perret, 2003; Perret & Aujard, 2001).

In captivity, grey mouse lemurs can live to 14 years of age and can show similar age-related cognitive changes as humans (Languille et al., 2012). They breed seasonally in captivity (and in the wild: Radespiel et al., 2022), producing one to four infants per birth, which are weaned at 40 days of age (Landes et al., 2019). The

infants in the studied population were cared for by their mother and other females in the group (i.e. alloparental care); juveniles reach puberty before they are moved from their natal group to another same-sex group (Perret, 2021). Adult males and females are normally housed separately, unless they are being bred. Individuals are sometimes housed alone under special circumstances (e.g. being elderly) but are more commonly housed in cages containing up to eight individuals, which is similar to what has been reported in the wild as there is high overlap between and within the sexes in the home ranges (Schliehe-Diecks et al., 2012). All animals in the colony are routinely weighed manually once a month to follow seasonal metabolic changes and to monitor their overall health.

Before being added to our study, the animals were housed in same-sex groups with approximately eight cagemates. They were housed in 0.72 m³ (height: 170 cm; width: 69.5 cm; depth: 61 cm) enclosures and provided with branches, fresh laurel, *Laurus nobilis*, clippings, ropes and other types of structural enrichment. There were no changes in the portions provided to the animals, but competition was potentially reduced by moving the animals to smaller groups.

Experimental Design

General study design

The mouse lemurs were randomly assigned to either a High level of enrichment or to a standard (Control) level and housed with a single, sex-, age- and weight-matched cagemate for 5 weeks under a summer-like photoperiod. We thus formed one experimental group with a relatively High level of enrichment and one Control group with a standard level of enrichment, each consisting of 12 mouse lemurs each (six females and six males; Table A1). Body mass measures taken before the initiation of the study showed that females were generally heavier (mean \pm SE: 116.08 \pm 6.19 g) than males (90.00 \pm 5.76 g), which is expected in this species (Thomas et al., 2015, 2016). When deciding how to allocate pairs of mouse lemurs to either enrichment level, we controlled their average initial body masses to ensure they were similar (Control: 100.67 \pm 7.65 g; High: 105.42 \pm 6.55 g). This was done to account for any differences in behaviour and physiology relating to an individual's body mass, which is an important predictor for these traits (Thomas et al., 2015, 2016).

To determine how mouse lemurs responded to environmental enrichment, we first monitored activity in daylight and at night using a series of battery-powered motion sensors. Second, the animals underwent a series of training sessions aimed at familiarizing them with the testing apparatus and paradigms. One week after the training, we used the apparatus to evaluate their responses to a judgement bias test and a predatory motivation test (none of the animal subjects had been exposed to these tests before). Third, we measured changes in facial temperature during a 2 min manual restraint and a wellness check performed before each training and testing session. These wellness checks were also done to ensure that the animals could participate in the study; if any signs of major injuries, lethargy, sickness or abnormal behaviours were seen, the animals were immediately returned to their enclosures and the caretaking staff were notified. Finally, we quantified major metabolic fluctuations by recording changes in body mass and composition (i.e. % fat and free fluid) between the first and last week of the study period and after all the mouse lemurs had completed their respective tests. A summary of the different measures performed is provided in Table A1.

All training and testing sessions were done between 1500 and 1900, as the lights-off period commenced at 1700. Owing to the inhibitory effect of light on activity of mouse lemurs,

training and testing were done in a dark room with the aid of red light emitted by a series of lamps, including a headlamp worn by the experimenter. We completed the study between April and October 2021 with three cohorts consisting of eight animals each (four females and four males); all cohorts were studied separately (cohort 1: April–May; cohort 2: June–July; cohort 3: September–October).

Training and testing were done on Mondays, Wednesdays and Fridays. The animals were fed a standardized wheat-based porridge and a mix of fruits after they completed each of their respective training or testing sessions on these days. From each cohort, we allocated half of the mouse lemurs to subcohort A (i.e. subcohorts 1A, 2A and 3A) and the remaining animals to subcohort B (i.e. subcohorts 1B, 2B and 3B). Each subcohort consisted of one male and one female from the Control group and one male and one female from the High enrichment group (i.e. one animal per cage allocated to each cohort).

Animals from subcohort A were trained on week 2, allowed to rest on week 3, tested on week 4 and allowed to rest again on week 5, whereas animals from subcohort B were trained on week 3, allowed to rest on week 4 and tested on week 5 of the study (Fig. 1a). All animals belonging to the same subcohort received one session of training/testing on the same day, and they were given a resting day after each training/testing session, as well as a resting week after completing the three training sessions. Unfortunately, one of our male (M293CA) study subjects from cohort 1 in the High enrichment group was unable to participate in the training and testing owing to a predisposition to seizures, but circadian activity, stress reactivity and body composition data were collected from this individual.

All animals remained undisturbed for an additional week after the completion of the study, after which they were returned to their original same-sex group. All cages and animals were given a brief visual wellness check during the weekdays (ca. 1100) throughout the study, in which their food and drinking water levels were monitored, as well as the integrity of the substrates given. During these brief ‘morning’ checks, the animals generally peeked from (or jumped out of) their nestboxes, which gave an overall idea of their alertness.

Environmental enrichment

Animals were housed in a cage fitted with several substrates for them to use (Fig. 2). The substrates for the Control enrichment consisted of one wooden nestbox (10 cm³), fresh wild laurel clippings (replaced as often as needed), a tray with a cotton pad at the bottom of the cage to collect faeces (replaced as needed) and natural climbing branches. High enrichment consisted of the same substrates given to the Control enrichment cages with additional hiding tubes (two cardboard or PVC tubes, each measuring 30 cm in length and 15 cm in diameter) hung from the top of the cage and approximately 4 m of rugged and flexible PVC pipes (2 cm in



Figure 2. Example of two cages the animals were housed in for studying (a) High and (b) Control levels of enrichment. Substrates for the Control level of enrichment consisted of two wooden nestboxes (10 cm³; one is shown, the second nestbox was placed on the door), fresh laurel clippings (replaced as often as needed), a tray with a cotton pad at the bottom of the cage to collect faeces (replaced as needed) and natural climbing branches. High enrichment consisted of the same substrates given to the Control enrichment cages with additional hiding tubes (two cardboard or PVC tubes, each measuring 30 cm in length and 15 cm in diameter) hung from the top of the cage and approximately 4 m of rugged and flexible PVC pipes (2 cm in diameter) extending to all corners of the cage. All animals were exposed to an artificial summer-like (i.e. 14:10 h light:dark) photoperiod.

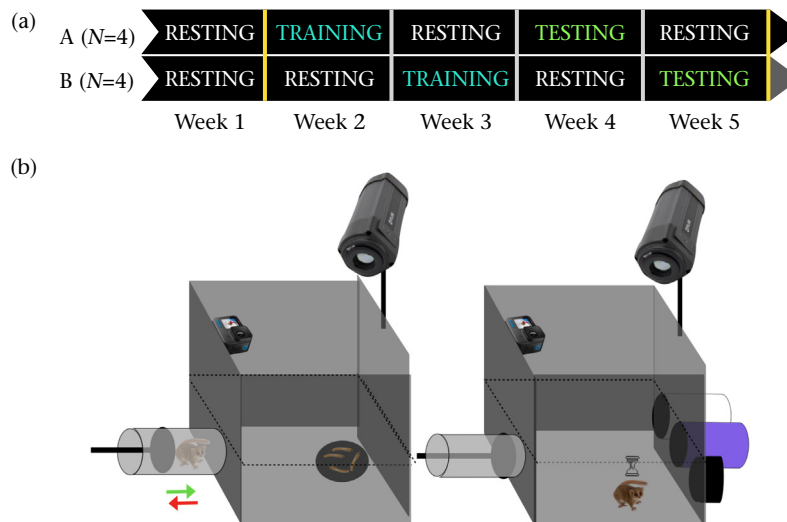


Figure 1. (a) Diagram showing the protocol scheduling for each subcohort. Animals from subcohort A (i.e. one male and one female from each enrichment level; top arrow) were trained on week 2 and tested on week 4, while animals from subcohort B (bottom arrow) were trained on week 3 and tested on week 5. Body composition measurements were performed at the end of week 1 and week 5 (yellow line) (b) Representation of the testing apparatus used for judgement bias (left) and predatory motivation (right) tests. The apparatus (0.90 × 0.50 m and 0.90 m high) was made of PVC and consisted of a removable animal transfer tube (30 cm long and 35 cm in circumference) and a series of inter-changeable arenas.

diameter) extending to all corners of the cage. These additional substrates provided the small arboreal primates with the opportunity to move around the available space within their cage and the flexibility to sleep alone or with their cagemate. We selected these types of environmental enrichment conditions based on decades of experience with the biology of captive mouse lemurs and other nocturnal primates and rodents, and their anatomy, physiology, biomechanics and social behaviour.

Cages were 0.5×0.7 m and 0.65 m high) for the first two cohorts and slightly smaller for cohort 3 ($0.5 \times 0.5 \times 0.5$ m) owing to limited space availability at the facilities. The cages were placed on a two-level shelf and were either 40 cm (two bottom cages) or 150 cm (two top cages) from the floor. Pairs of mouse lemurs were randomly assigned to the top and bottom cages, two pairs being in the top cages and two pairs being in the bottom cages for each cohort, except for the last cohort when only single column cages were available and the cages with High enrichment were placed on the bottom shelves to balance the design.

Training and testing apparatus

We used a custom training and testing apparatus designed to minimize any stress induced by our different tests and to focus the subject's attention on the paradigm presented. The apparatus consisted of an arena and a transfer tube (Fig. 1b). The arena was made of three grey PVC boxes, each measuring 0.90×0.50 m and 0.30 m high. One box formed the bottom section of the arena (0.90×0.50 m and 0.30 m high), where subjects were exposed to different stimuli. The bottom section of the arena had hay/straw on the floor to provide comfortable support and traction to facilitate movement. The two remaining boxes were glued together and their floor was removed to make the top section of the arena (0.90×0.50 m and 0.60 m high), which was stacked onto the replaceable bottom section (i.e. testing arena). The top section was used to prevent the mouse lemurs from jumping out and escaping, and it held a small camera (GoPro Hero 4) and a thermographic camera (FLIR A320) pointed towards the testing arena. We made five interchangeable arenas; the first was used for the training sessions and provided access to two tubes, one on the left side of the arena (for the positive stimulus) and another on the right side of the arena (for the negative stimulus). The second, third and fourth arenas were used for the judgement bias test and provided access to a single tube on either the left (for the positive stimulus), middle (for the ambiguous stimulus) or right side (for the negative stimulus). The fifth arena was used for predatory tests and did not have any tubes attached to it or any holes other than the entrance to the arena from which an animal transfer tube was attached to it.

We designed an animal transfer tube that minimized the disturbances resulting from the handling and transport of the animals, as well as from the presence of an experimenter/observer. It also intended to provide the subject with the opportunity to voluntarily participate in the session, although it was very uncommon that they spontaneously entered the arena. The removable animal transfer tube was made of 3 mm thick PVC and was 40 cm in length and 15 cm in diameter. It also had two feet, one on each extreme of the tube, to stabilize the animal if it decided to move inside. One end of the transfer tube was sealed with a PVC lid that had a small hole in the centre where a pushing tool (i.e. a clear hot glue stick) could be inserted. The opposite side of the tube remained open and had a 3D-printed removable lid that secured an animal inside while being transferred to and from the testing room. The design of this lid was based on a camera lens cap, which fitted securely into the grooves in the tube; this abolished any possibility of the animal being accidentally crushed. As another safety precaution, the transfer

tube could not fit into the arena if the removable lid was still inserted in the transfer tube.

The transfer tube was not completely sealed to facilitate ventilation and had a handful (ca. 5 g) of hay/straw. To give animals access to the testing arena, the lid was removed from the transfer tube before attaching it to the open side of the testing arena (Fig. 1b). The inside of the transfer tube had a moving part next to the closed side that if pushed from the outside with a pushing tool acted as a plunger. The pushing tool was only inserted in the tube after the open transfer tube was attached to the arena and 5 min had passed. This system was used to motivate animals to enter the arena after a 5 min acclimatization period. Most of the animals responded to the vibrations of the plunger and learned to quickly exit the transfer tube and enter the arena.

Training procedure

During the training period (occurring during week 2 for sub-cohort A or week 3 for subcohort B), the mouse lemurs had three 30 min training sessions (done on separate days) to familiarize themselves with the testing procedures, including being handled by the experimenter, travelling in the transfer tube and entering and exiting the testing arena. During all three training sessions, the animals were given a chance to learn the location and characteristics of two stimuli that they faced again in the testing session for judgement bias. The positive stimulus consisted of a hole on the left side of the arena, which gave access to a second 40 cm long transfer tube. The negative stimulus was a hole on the right side of the arena giving access to a short 5 cm long tube.

As mouse lemurs tend to search for hiding places and avoid open areas (Thomas et al., 2016), presumably because of the high predation risk owing to their small body size (Goodman et al., 1993), the positive stimulus provided them with the opportunity to seek shelter inside the long transfer tube, whereas the negative stimulus did not allow the animal to hide their body completely (Fig. 1b). Animals were given 35 min to explore the arena and both stimuli during each training session. To prevent animals from identifying the properties of the tubes (i.e. the length) using visual cues, we concealed the holes using a series of curtains of increasing size during the second and third training sessions. Thus, the holes were not covered with any curtains during the first session, then they were partly covered with small curtains (6×7 cm) during the second one and finally they were covered with large curtains (12×14 cm) during the third training session.

Each session consisted of extracting each animal from its enriched cage and conducting a stress reactivity test (i.e. wellness check). The animal was placed inside the transfer tube secured inside the cage and was moved to the testing room. Animals were given a 5–10 min acclimatization period in the testing room before the lid of the transfer tube was removed and the transfer tube was attached to the arena and remained undisturbed for 5 min (animals did not dart out before the tube was attached). This allowed the animal to voluntarily enter the arena and participate in the session. However, as the majority of the animals (ca. 80%) did not enter the testing arena voluntarily during these 5 min, they were gently pushed out of the transfer tube with the plunger into the training arena where they remained for 30 min. Once the session ended, the animals were placed in a transfer tube (i.e. the positive stimulus or the original transfer tube) and returned to their enriched cage and fed. Training sessions were performed on Mondays, Wednesdays and Fridays. As one session lasted for approximately 1 h per animal and owing to mobilization restrictions during the COVID-19 pandemic in France at the time of the study, we could only perform four sessions per afternoon, one animal after the other (1500–1900).

Testing procedure

During the testing period (week 4 or 5), each mouse lemur first underwent a testing session for judgement bias and then a testing session for predatory motivation, each on different days and separated by a resting day. The animals were handled, transferred and tested using the same procedure as done during the training sessions, except that they were exposed to different arenas.

For the judgement bias test, the animals were given two consecutive 15 min trials before being exposed to an ambiguous stimulus. In the first trial, the mouse lemurs were exposed to an arena that only gave access to the negative stimulus (i.e. hole on the right side, concealed by a curtain that gave access to a short hiding tube). In the second trial, they were exposed to another arena that only gave access to a positive stimulus (i.e. hole on the left, concealed by a curtain that gave access to a long hiding tube). For the third trial, the animals were exposed to an ambiguous stimulus that they had never encountered before (i.e. hole in the middle, concealed by a curtain that gave access to a long hiding tube).

For the predatory motivation test, we placed live mealworms, *Tenebrio molitor*, in the middle of the arena in a PVC dish. The test commenced when the transfer tube containing the subject was attached to the testing arena and it lasted for 15 min.

Data Collection

Behavioural data

Circadian activity. For each of the three cohorts, spontaneous activity during the first week of the experiment was measured for all four cages (two mouse lemurs per cage): two cages with Control levels of enrichment and two cages with high levels of enrichment (three cages hosting females and three hosting males in each enrichment group). We did this to determine whether the animals would alter their activity patterns owing to the level of enrichment. We monitored spontaneous home cage activity continuously using microcontroller-based passive infrared sensors, which we have used successfully to study the circadian biology of captive mouse lemurs in our colony (e.g. Hozer et al., 2019, 2020). These passive infrared sensors were programmed to be activated and record movement at a set rate (Matikainen-Ankney et al., 2019). We measured the total number of activity bouts to determine how the animals were using their cages. We also used hourly summaries to determine how animals were using the daytime and night-time, as shifts in circadian behaviour can be a useful energy-saving strategy in mammals and other endotherms (van der Vinne et al., 2014, 2015).

We focused on monitoring spontaneous activity during the first week of the study because the animals remained undisturbed during this period compared with during the training and testing periods when they were disturbed thrice weekly for training/testing. We used the total count of activity bouts recorded for each cage for the analyses, as well as the ratio between the total sum of the activity bouts recorded at daylight to the total sum recorded during the night (i.e. circadian behaviour index). A circadian behaviour index with a value larger than 1 indicates that individuals show higher activity at daylight than at night (i.e. diurnality), a value of 1 indicates that individuals were equally active during daylight and night (i.e. cathemerality) and a value below 1 indicates that individuals were more active during the night than daylight (i.e. nocturnality).

Judgement bias. This test generally involves training an individual to differentiate between a positive and a negative stimulus and then characterizing its willingness to approach an ambiguous stimulus (Bateson & Nettle, 2015; Bethell, 2015). This test is a useful tool to understand how an individual's emotional state can

influence pessimism (unwillingness to approach ambiguous stimuli) or optimism (willingness to approach ambiguous stimuli; Brydges et al., 2011).

We considered that mouse lemurs would learn to differentiate between the positive and negative stimuli only if they showed shorter latencies to approach the positive stimulus than the negative one during the last training session (i.e. the third training session during which the positive and negative stimuli were fully covered with curtains). Only the data of animals that met this criterion were used in the analyses of the judgement bias test. During the test, we recorded the latencies for animals to approach each stimulus (within 3 cm) and calculated an index that indicates whether the latency to approach the ambiguous stimulus was either close to the positive or negative stimulus (Verjat et al., 2021).

Judgement bias index = $1 - ([\text{latency ambiguous} - \text{latency positive}] / [\text{latency negative} - \text{latency positive}])$

The value of the judgement bias index can range from 0 to 1. On one hand, a value of 0 indicates that latencies to approach negative and ambiguous stimuli were similar, which is a sign of unwillingness to approach ambiguous stimuli (i.e. pessimism). On the other hand, a value of 1 indicates that the latencies to approach ambiguous and positive stimuli were similar, which is a sign of willingness to approach ambiguous stimuli (i.e. optimism). Thus, the higher the value of the index, the more optimistic the response (Verjat et al., 2021). In cases where the latency to approach the ambiguous stimulus was shorter than the latency to approach the positive stimulus, the calculated index was rounded to 1 (i.e. the index value would otherwise be higher than 1). Similarly, when the latency to approach the ambiguous stimulus was longer than the latency to approach the negative stimulus, the calculated index was rounded to 0.

Predatory motivation. To determine the motivation of individuals to catch live prey, we characterized predatory responses to four 2 cm long (store-bought) mealworms for each mouse lemur. The mealworms were placed at the centre of the arena in a grey PVC dish 12 cm in diameter. The total number of prey ingested during the duration of the test was recorded for each 15 min test and used for analyses.

Physiological data

Stress reactivity. Stress reactivity is commonly determined by measuring the responses of individuals to an acute stressor (Novak et al., 2013), such as limiting an individual's mobility and space utilization (Smith & French, 1997). We measured facial thermal responses to a 2 min wellness check, in which a manual restraint was applied to secure the animal. Each stress reactivity test was recorded with a thermographic camera (FLIR A320), which allowed us to monitor temperature changes in both eyes and ears and the nose throughout the restraint (Boileau et al., 2019). Facial temperature changes were assessed for each training and testing session, but we only included measures collected during the first training session when the procedure was unfamiliar to our study subjects.

Body composition. We monitored changes in body mass and composition during our study by performing nuclear magnetic resonance (NMR) scans after 1 week of enrichment and 4 weeks later (i.e. after the testing period had ended). We used an NMR spectroscopy system (Minispec LF90 TD-NMR analyzer, BioSpin, Bruker Corp., Billerica, MA, U.S.A.) to perform the body composition evaluations, which gives fat, lean and free fluid masses. The animals were weighed before being scanned, and we determined total mass change between week 1 and week 5. We also calculated the percentage of dry fat mass from the measurements in each scan, and we determined the change in percentage dry fat mass between week 1 and week 5 for each mouse lemur. The animals had not been fed before being scanned, and they were stimulated to urinate

and defecate before being weighed and scanned by very gently massaging their abdomen and tickling the dorsal area of their rump.

Data Analyses

We used common supervised machine learning methods for animal behaviour research (Valletta et al., 2017; Wang et al., 2020) to build a trained model that most accurately predicted an individual's level of enrichment (i.e. as a proxy for wellbeing). Thus, we used the explanatory parameters calculated from the different measures collected (Table 1) to train the classification algorithm that most accurately classified individuals as being under High or Control levels of enrichment.

We divided our analyses into three steps: (1) exploratory data analyses, (2) feature/variable/dimension reduction using PCA (Budaev, 2010) and (3) classification and cross-validation of predictive algorithms. We used PCA as a feature reduction method because of its simplicity of use and implementation, but other methods, such as Akaike information criterion (i.e. AIC-) based model selection (Symonds & Moussalli, 2011) can be useful when there is prior knowledge on the subject being studied. All data used in the PCA are presented in Table A2. Data cleaning, exploration, visualizations and analyses were done using Python with the 'pandas' (McKinney, 2010), 'numpy' (Harris et al., 2020), 'matplotlib' (Hunter, 2007), 'seaborn' (Waskom, 2021), and 'scikit-learn' (Pedregosa et al., 2011) libraries.

Exploratory data analyses

We first explored the data to determine how treatment and any confounding factors (subcohort, location of the cage and lineage) affected the different behavioural and physiological measures.

Principal component analysis

PCA is a method to reduce the number of dimensions in a data set and it avoids multicollinearity of independent variables in a regression analysis by using newly calculated composite indices instead of several available measures (Budaev, 2010). The nine features used in the PCA were the sex, lineage, subcohort, cage location, circadian behaviour index, number of mealworms ingested, changes in facial temperature during a wellness check and changes in body mass and body fat composition (during 5 weeks; Table 1). The goal was to use our raw data to calculate new components and to reduce the number of explanatory variables (i.e. features) for our target/outcome variable (i.e. enrichment level). We selected the number of components that explained 95% of the variance in the data, i.e. $\text{sklearn.decomposition.PCA}(n_{\text{components}} = 0.95)$.

Classification and cross-validation

We used six different classifier algorithms (i.e. binary logistic regression, support-vector machines, Naïve Bayes, K-nearest neighbour, random forest and XGBoost) to predict the two levels of enrichment, using the new components identified by the PCA to train them. Binary logistic regression is a means of quantifying the unique contribution of each individual member of a group of independent variables to a binary outcome, in this case, the two levels of enrichment, to identify the strongest linear combination of variables with the greatest probability of detecting the observed outcome (Brewster et al., 2018; Bortsel & Stevenson, 2021). The support-vector machine method tries to maximize a linear decision boundary between classifications and is computationally relatively cheap to use (Hamilton et al., 2019; Schneider et al., 2019). Naïve Bayes methods are a set of supervised learning algorithms based on applying Bayes' theorem with the 'naïve' assumption of conditional independence between every pair of features given the value of the class variable and have successfully been used in ethological studies (Fang et al., 2021). The idea underlying the K-nearest-neighbour method is to assign new unclassified examples to the class to which the majority of its K-nearest neighbours belong (Budaev, 2010). Finally, we used random forest and gradient tree-boosted (XGBoost) classifiers (Brewster et al., 2018).

For cross-validation, the data were split into two portions: (1) the training set (80%) for developing all base learner models and (2) the test set (20%) for model selection to estimate overall performance of the selected final predictive algorithm. To identify the best-performing algorithm, we used *k*-fold cross-validation, which splits the training set into *k* smaller sets for this final model to explore. This is an efficient way to identify the most accurate predictive model, but it can be computationally expensive (Pons et al., 2017).

Ethical Note

The procedures used were selected because of their noninvasive nature. In many cases, the procedures themselves could serve as enrichment, as they promote natural behaviours. The various procedures applied were organized in such a way as to leave sufficient recovery time for the animals between two procedures. Weight, prostration and lack of feeding were carefully monitored in each animal subject to ensure they were in good health. All our procedures were approved by the ethical committee at the Muséum National d'Histoire Naturelle (MNHN; approval number: Comité Cuvier 68–112) and followed the EU regulations. In addition, the animal procedures in the study followed the guidelines for the treatment of animals for research recommended by the ASAB/ABS.

Table 1

Explanatory variables (features) used to build predictive models for our target; that is, the two levels of environmental enrichment (High or Control)

Features	Description
Lineage	Subject's family of origin (a, b, c, d, e, f, g, h, i or j)
Sex	Subject's sex (male or female)
Subcohort	Subject's subcohort (1A, 1B, 2A, 2B, 3A or 3B)
Cage location	Cage location on shelf (top or bottom)
Circadian activity index	Duration of daytime activity divided by duration of night-time activity
Judgement bias index*	Calculated index based on latencies to approach negative, positive and ambiguous stimuli
Mealworms ingested	Count of experimentally presented mealworms ingested during test (0–4)
Nose temperature change	Change in nose temperature during 2 min manual restraint
Body mass change	Change in body mass throughout the 5-week study
Body fat mass change	Change in body fat mass composition throughout the 5-week study

* Judgement bias tests were not used for the PCA or to train predictive models because only half of the animals studied met our learning criteria (see Methods for details).

RESULTS

Exploratory Data Analysis

Circadian activity

The animals were mostly active at night (number of activity bouts: $18\ 019.41 \pm 3006.84$; mean \pm SE), but also showed periods of activity at daylight ($10\ 487.25 \pm 5149.642$; Fig. 3a). The average level of activity in each cage remained stable during the first week of the study, independent of the enrichment level. Males were more active ($42\ 167.83 \pm 15\ 654.51$) than females ($19\ 338.50 \pm 4307.89$). The animals under High enrichment generally showed higher activity ($44\ 672.33 \pm 15\ 600.78$) than those under Control enrichment ($16\ 834.00 \pm 2949.81$). The animals in cages on the bottom shelf were generally more active ($38\ 022.0 \pm 16\ 013.61$) than those in cages on the top shelf ($23\ 484.33 \pm 4611.87$). Circadian activity indices were relatively higher in the animals in the High enrichment group (0.47 ± 0.17 ; mean \pm SE) than in the animals in the Control enrichment group (0.09 ± 0.03). Additionally, males showed slightly higher circadian activity indices (0.30 ± 0.16) than females (0.25 ± 0.09 ; Fig. 3a).

Judgement bias

Thirteen (57%) of the 23 trained mouse lemurs met our learning criterion by showing shorter latency to approach the positive stimulus compared with the negative stimulus on the third training session. Of those, seven were under High enrichment (four males and three females) and six were under Control enrichment (four males and two females). One female (i.e. F277ABA) under Control enrichment was excluded from the analyses because she did not approach the positive or negative stimulus in the test session. Females, on average, had higher (0.74 ± 0.11) indices than males (0.42 ± 0.13). Judgement bias index values for mouse lemurs under High

enrichment were slightly higher (0.62 ± 0.14 ; mean \pm SE) than for those under Control enrichment (0.54 ± 0.12 ; Fig. 3b). Animals housed in cages on the top (0.61 \pm 0.12) or bottom (0.56 \pm 0.14) shelves showed similar judgement bias indices.

Predatory motivation

Twelve (52%) of the 23 tested animals ate some or all of the mealworms offered; six of them were the same individuals that met our learning criterion in the judgement bias test. Of the 12 animals that consumed the mealworms, seven belonged to the High enrichment group. Males and females showed, on average, similar patterns of mealworm ingestion (males: 1.91 ± 0.61 mealworms; females: 1.83 ± 0.56 mealworms; mean \pm SE). In addition, animals under High enrichment ate, on average, almost twice as many experimentally presented mealworms than animals under Control enrichment (High: 2.45 ± 0.59 mealworms; Control: 1.33 ± 0.53 mealworms; Fig. 3c). Mouse lemurs in the top and bottom cages showed very similar predatory motivation (top: 1.83 ± 0.56 mealworms; bottom: 1.91 ± 0.61 mealworms).

Stress reactivity

Changes in facial temperature during a 2 min manual restraint were recorded from both eyes, both ears and the nose. We focused our analyses on the difference in nose temperature during the second week of the study at between 0 and 120 s of the test because most of the variation in facial temperature was seen on the nose after 2 min of manual restraint (Fig. A1) and because of the ease of monitoring one area rather than two. Females showed similar nose temperature change after the 2 min manual restraint compared with males (males: 0.18 ± 0.43 °C; females: 0.25 ± 0.30 °C; mean \pm SE). The animals in the High enrichment group generally showed a larger change in nose temperature (0.48 ± 0.28 °C) than those in the Control enrichment group (-0.06 ± 0.46 °C; Fig. 3d). The animals

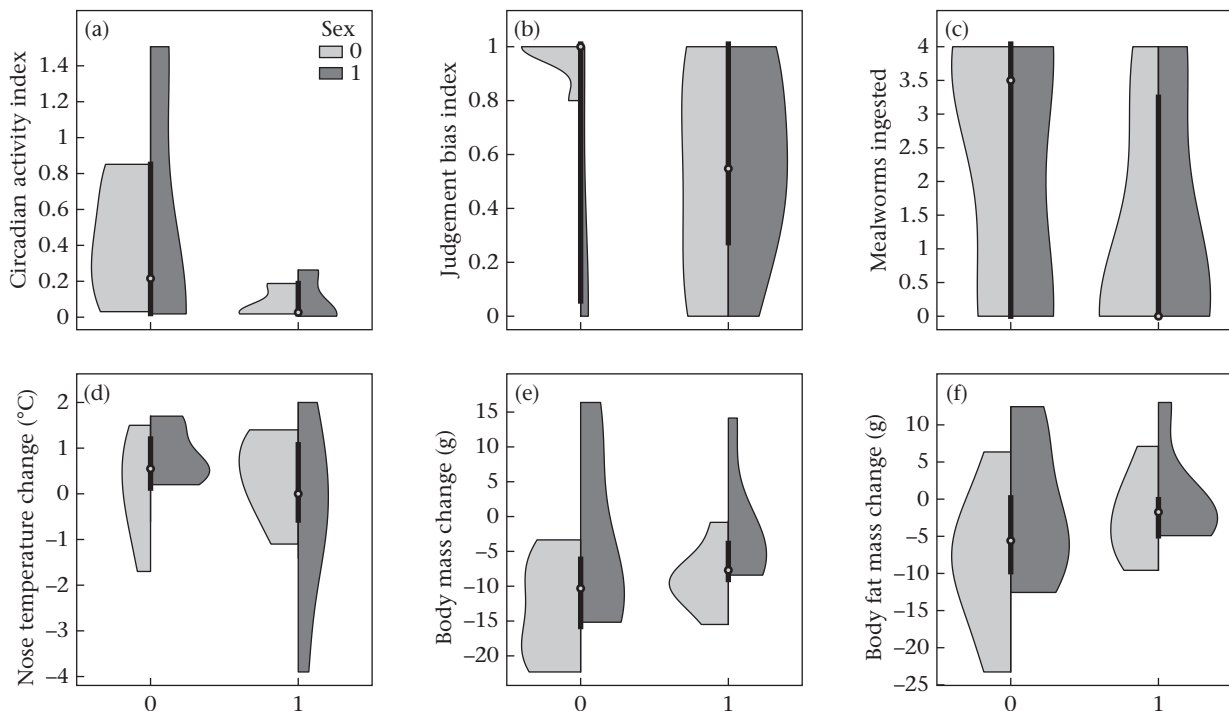


Figure 3. Visual summaries of data obtained for (a) the circadian activity index, (b) the judgement bias index, (c) number of mealworms ingested, (d) nose temperature change, (e) body mass change and (f) body fat mass change, depending on their respective level of enrichment (High: $N = 12$; Control: $N = 12$) and sex (females = 12; males = 12). The violin plots show the distribution of the data with kernel density estimation. Inside each violin, the white dot represents the median, the thick bar indicates the interquartile range (IQR), and the thin lines (whiskers) extend to 1.5 times the IQR from the 25th and 75th percentiles.

housed in top cages generally showed a larger thermal response ($0.38\text{ }^{\circ}\text{C} \pm 0.32\text{ }^{\circ}\text{C}$) than those located in bottom cages ($0.04\text{ }^{\circ}\text{C} \pm 0.44\text{ }^{\circ}\text{C}$).

Body composition change

Of the 24 animals in the study, 21 (87.5%) lost body mass throughout the study (average mass change: $-6.99 \pm 0.86\text{ g}$; mean \pm SEM), which is characteristic of the species (Perret & Aujard, 2001). Two of the three individuals that gained weight throughout the study were housed under High enrichment levels. Females showed, on average, larger overall body mass loss ($-11.18 \pm 1.90\text{ g}$) than males ($-2.81 \pm 3.03\text{ g}$). Average body mass loss during the study was $-8.50 \pm 3.31\text{ g}$ for mouse lemurs under High enrichment and $-5.49 \pm 2.14\text{ g}$ for those under Control enrichment (Fig. 3e). Individuals housed in top cages lost more body mass than those in bottom cages (top: $-10.0 \pm 1.75\text{ g}$; bottom: $-3.99 \pm 3.36\text{ g}$).

Six of 24 subjects, three under High enrichment and three under Control enrichment, gained between 0.55% and 14.70% body fat mass throughout the study, while the remaining subjects lost between 0.27% and 23.27% of their body fat mass throughout the 5 weeks of the study. Average measures of body composition indicate that females lost a higher percentage of dry fat mass ($-5.37 \pm 2.45\%$) than males ($-0.64 \pm 2.30\%$). Loss in dry fat was generally more pronounced for mouse lemurs under High enrichment ($-4.84 \pm 2.93\%$) than Control enrichment ($-1.17 \pm 1.77\%$; Fig. 3f). Dry fat mass change during the study was slightly higher for animals in bottom cages ($-2.53 \pm 2.99\%$) than those housed in top cages ($-3.48 \pm 1.84\%$). As such, the position of the cage was included as a random factor in our models.

Principal Component Analysis

To perform the classification exercise using PCA, we excluded the judgement bias indices to maximize our sample (i.e. only 13 of 24 subjects went through the learning criterion for the judgement bias test) and to fit all models to the same data set (Table A2). In addition, we imputed missing data from male M293CA, which did not participate in all the tests owing to his predisposition to seizures, with the most frequent value for all animals. The results from the PCA reduced the dimensionality of our data from nine features (sex, lineage, subcohort, cage location, circadian activity index, number of mealworms ingested during predatory motivation test, change in nose temperature during restraint, change in body mass and change in body dry fat) to seven newly calculated components, which together explained 96% of the variability in our data. The first three principal components together explained 65% of the variability in the data (Fig. 4). The PCA loadings are provided in Table A3. The data from the seven newly calculated components used to train the models are shown in Table A4.

Classification and Cross-Validation

From the six different algorithms trained with the seven new components identified by PCA, the Naïve Bayes classifier produced the most accurate predictions (100%) followed by the XGBoost classifier (80%); the remaining four did not exceed 60% accuracy.

DISCUSSION

The evolution of efficient approaches to the welfare of animals participating in research is dependent on the identification of reliable indices for evaluating the multidimensionality of their wellbeing. Such approaches should be commonplace, especially when addressing research questions regarding complex biological processes because ignoring the wellbeing of research subjects

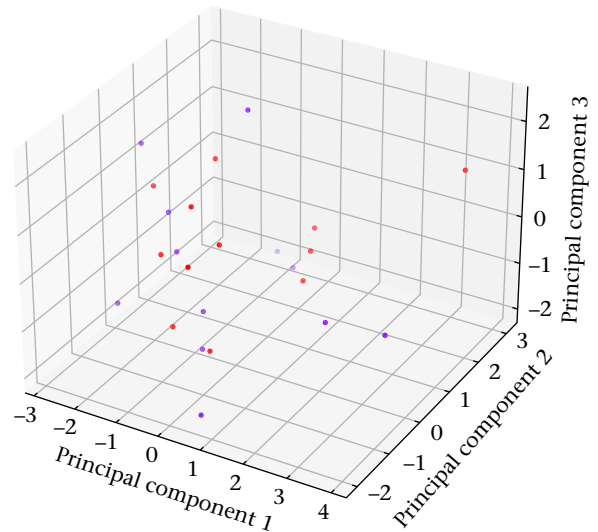


Figure 4. Three best PCA-derived components, which represent 65% of the variance, for animals under High (purple) or Control (red) levels of environmental enrichment.

devalues the data collected from them (Buchanan-Smith et al., 2005; Hemsworth et al., 2015; Mellor, 2016; Poole, 1997). Therefore, in this study, we sought to identify reliable indicators for the evaluation of the wellbeing of captive grey mouse lemurs a powerful animal model in fundamental and translational research. To do so, we characterized ecologically relevant behavioural and physiological changes resulting from a 5-week-long exposure to contrasting levels of environmental complexity (i.e. High versus Control) in adult males and females. To isolate the possible effects that structural enrichment had on the behaviour and physiology of the animals studied, we did not include any changes to their diets.

We hypothesized that higher levels of structural environmental enrichment would stimulate metabolic energy expenditure, influencing how individuals would respond to our array of behavioural and physiological measures. We evaluated how individuals responded to enrichment depending on their sex by measuring (1) spontaneous circadian activity during the first week of the study (Hozer et al., 2019, 2020; Hozer & Pifferi, 2020, 2022; van der Vinne et al., 2014, 2015, 2019; Perea-Rodriguez et al., 2022), (2) changes in nose temperature during a stress reactivity test consisting of a 2 min manual restraint occurring in the first week of the study (Trevain & Valsecchi, 2021), (3) judgement bias (Gygax, 2014; Mendl et al., 2009) and (4) predatory motivation during the fourth or fifth week of the study. Additionally, to determine whether exposure to structural enrichment influenced how animals invested metabolic energy throughout the 5-week study, we recorded changes in body mass and composition (i.e. % fat) between the first and last week of the study period. Finally, we used current machine learning tools to train a model to accurately predict (i.e. classify) each individual's level of enrichment based on these changes. Thus, we explored the data collected, performed PCA to combine and reduce the number of explanatory variables (i.e. features) and trained several classifier algorithms to find the ones(s) that most accurately predicted the two levels of enrichment based on these newly calculated primary components/features.

Our exploratory data analyses showed that housing adult male and female mouse lemurs under relatively high or standard levels of environmental complexity can result in changes in their behaviour and physiology. More specifically, individuals housed under High enrichment were more active during daylight, ingested more experimentally presented live prey in a predatory motivation

test and showed more dramatic losses in body mass and body fat composition. Responses to the judgement bias test were not considered because almost half of the animal subjects showed no evidence that they learned the locations of the positive and negative stimuli during the training phase or failed to remember them in the testing phase. Finally, in an attempt to integrate resource- and individual-based measures to provide a single index to evaluate the wellbeing of the animals in our study, we performed PCA with nine of the 10 explanatory variables (excluding the judgement bias indices) and trained six different classifier algorithms to find the most accurate one. We identified the single best-trained algorithm (i.e. naïve Bayes classifier) that most accurately predicted the wellbeing of the animals, which in this case was related to their level of environmental enrichment.

Our results suggest, based on the different behavioural and physiological measures, that the mouse lemurs in this study were stimulated by their assigned level of environmental enrichment in a predictable manner. We used our knowledge of the species' ecology to identify a series of behavioural and physiological measures that would cohesively describe changes in the wellbeing of animals exposed to enrichment. These tests related to how mouse lemurs distributed their activity throughout the day, how they foraged for prey and dealt with acute stressors as well as how their metabolism (body mass and composition) responded to the two distinct levels of environmental complexity. Aided by current machine learning approaches; that is, PCA-based feature reduction and cross-validation of classifier algorithms, we were able to identify a model that accurately predicted the level of enrichment after being trained on the data collected for seven of the nine measures we designed for the species.

Although we did not show any interaction between the sex and the level of enrichment, we found that males and females responded differently to the experiment. Females were the heaviest at the beginning of the study and more likely to lose body mass throughout, while males were more likely to gain body mass. Females also showed higher estimates of body mass and body fat mass, suggesting they incurred a higher metabolic cost as a result of our environmental manipulation.

Several caveats have to be considered when interpreting the results from our study. First, this study was performed on a seasonally breeding species and was conducted during the summer-like months, when individuals are investing heavily in reproduction. As such, their wellbeing (and the way we evaluate it) may be completely different during the winter-like months when their thermo-energetics are high and they use torpor to mediate it. Second, we did not include animals that reproduced that season and the physiological changes that come from mating, fertilization, implantation, gestation, parturition and lactation are not included in our models. Third, we used very specific structural environmental enrichment (i.e. hiding and climbing structures) and other types (e.g. sensorial, nutritional) that may have similar, contrasting or synergistic effects on the same individuals sampled, especially if altering their nutritional intake was involved. Fourth, the order in which the behavioural tests were administered may have influenced how individuals responded; as such, balancing the order of the negative and positive stimuli during the judgement bias test (to address any lateralization effects), as well as alternating between giving the judgement bias or predatory motivation test first, could help tease out these effects. Furthermore, we studied a small number of individuals, and increasing our sample may improve the predictability of our models. Finally, here we decided to use PCA as a feature-reducing method, and although it only reduced the number of features from nine to seven, it allowed us to quickly train models using all the data available. We recommend that analysts, if possible, use other approaches that use prior knowledge of the

subject studied for feature reduction (e.g. information-theoretic approaches: [Symonds & Moussalli, 2011](#)).

Conclusion

In summary, our results suggest that the captive mouse lemurs studied showed distinct behavioural and physiological responses to environmental enrichment. This was evident by the accuracy of the trained machine learning model in differentiating between animals using the data collected. We believe that the approach used can easily be adapted to monitor the wellbeing of other captive animals, that is, to collect relevant data in a standardized way from various individuals from a colony, reduce the number of features or explanatory variables using PCA (or other methods) and cross-validate several algorithms to identify the best predictive model.

Author Contributions

Aude Anzeraey: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Conceptualization. **Audrey Maille:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Caroline Gilbert:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition. **Emmanuelle Pouydebat:** Writing – review & editing, Writing – original draft, Validation, Resources, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Fabien Pifferi:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Fabienne Aujard:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Jérémy Terrien:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Juan Pablo Perea-Rodríguez:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Julie Ferreira:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation.

Data Availability

All the data used in our analyses are presented in [Table A2](#).

Declaration of Interest

The authors declare that they have no conflicts of interest regarding the research, authorship and publication of this article.

Acknowledgments

We thank our colleagues in the Eco-Anthropology lab (EA-UMR 7206) and Adaptive Mechanisms and Evolution lab (MECADEV-UMR 7179) for their support during the development and completion of the study. Special thanks to Isabelle Hiron, Sandrine Gandor and Lauriane Dezairé for providing animal care support. We also thank the members of the Primate-Environment Interactions (IPE) and Biological Adaptation (BioAdapt) research teams for their feedback. We thank Sylvie Le Bomin and Jose Utge for sharing their camera equipment, and our colleagues and friends from the Society

for French-Speaking Primatologists for their comments on our work during the 22nd Conference at St. Etienne, France. This work, including JPPR's postdoctoral fellowship, was funded by a grant from Idex Sorbonne Université in the context of the French programme 'Investissement d'Avenir' (n° SU-19-3-EMRG-04) awarded to A.M. Finally, we thank the editors and referees for their constructive advice and comments on our work.

References

- Baker, K. C., & Dettmer, A. M. (2017). The well-being of laboratory non-human primates. *American Journal of Primatology*, 79(1), 1–5. <https://doi.org/10.1002/ajp.22520>
- Barber, J. C. E. (2009). Programmatic approaches to assessing and improving animal welfare in zoos and aquariums. *Zoo Biology*, 28(6), 519–530. <https://doi.org/10.1002/zoo.20260>
- Bateson, M., & Nettle, D. (2015). Development of a cognitive bias methodology for measuring low mood in chimpanzees. *PeerJ*, 3, e998. <https://doi.org/10.7717/peerj.998>
- Bethell, E. J. (2015). A “how-to” guide for designing judgment bias studies to assess captive animal welfare. *Journal of Applied Animal Welfare Science*, 18(Suppl. 1), S18–S42. <https://doi.org/10.1080/10888705.2015.1075833>
- Bethell, E. J., Holmes, A., MacLarnon, A., & Semple, S. (2012). Cognitive bias in a non-human primate: Husbandry procedures influence cognitive indicators of psychological well-being in captive rhesus macaques. *Animal Welfare*, 21(2), 185–195. <https://doi.org/10.7120/09627286.21.2.185>
- Boileau, A., Farish, M., Turner, S. P., & Camerlink, I. (2019). Infrared thermography of agonistic behaviour in pigs. *Physiology and Behavior*, 210, Article 112637. <https://doi.org/10.1016/j.physbeh.2019.112637>
- Brewster, L. R., Dale, J. J., Guttridge, T. L., Gruber, S. H., Hansell, A. C., Elliott, M., Cowx, I. G., Whitney, N. M., & Gleiss, A. C. (2018). Development and application of a machine learning algorithm for classification of elasmobranch behaviour from accelerometer data. *Marine Biology*, 165(4), 62. <https://doi.org/10.1007/s00227-018-3318-y>
- Broom, D. M. (1991). Animal welfare: Concepts and measurement. *Journal of Animal Science*, 69(10), 4167–4175. <https://doi.org/10.2527/1991.69104167x>
- Brydges, N. M., Leach, M., Nicol, K., Wright, R., & Bateson, M. (2011). Environmental enrichment induces optimistic cognitive bias in rats. *Animal Behaviour*, 81(1), 169–175. <https://doi.org/10.1016/j.anbehav.2010.09.030>
- Buchanan-Smith, H. M. (2011). Environmental enrichment for primates in laboratories. *Advances in Science and Research*, 5(1), 41–56. <https://doi.org/10.5194/asr-5-41-2010>
- Buchanan-Smith, H. M., Rennie, A., Vitale, A., Pollo, S., Prescott, M. J., & Morton, D. B. (2005). Harmonising the definition of refinement. *Animal Welfare*, 14(4), 379–384. <https://doi.org/10.1017/S0962728600029717>
- Budaev, S. V. (2010). Using principal components and factor analysis in animal behaviour research: Caveats and guidelines. *Ethology*, 116(5), 472–480. <https://doi.org/10.1111/j.1439-0310.2010.01758.x>
- Chotard, H., Ioannou, S., & Davila-Ross, M. (2018). Infrared thermal imaging: Positive and negative emotions modify the skin temperatures of monkey and ape faces. *American Journal of Primatology*, 80(5), Article e22863. <https://doi.org/10.1002/ajp.22863>
- Clark, F. E. (2017). Cognitive enrichment and welfare: Current approaches and future directions. *Animal Behavior and Cognition*, 4(1), 52–71. <https://doi.org/10.12966/abc.05.02.2017>
- Coleman, K. (2011). Caring for nonhuman primates in biomedical research facilities: Scientific, moral and emotional considerations. *American Journal of Primatology*, 73(3), 220–225. <https://doi.org/10.1002/ajp.20855>
- Coleman, K., & Pierre, P. J. (2014). Assessing anxiety in nonhuman primates. *ILAR Journal*, 55(2), 333–346. <https://doi.org/10.1093/ilar/ilu019>
- Cuan, K., Zhang, T., Li, Z., Huang, J., Ding, Y., & Fang, C. (2022). Automatic Newcastle disease detection using sound technology and deep learning method. *Computers and Electronics in Agriculture*, 194, Article 106740. <https://doi.org/10.1016/j.compag.2022.106740>
- Dodge, R., Daly, A., Huyton, J., & Sanders, L. (2012). The challenge of defining wellbeing. *International Journal of Wellbeing*, 2(3), 222–235. <https://doi.org/10.5502/ijw.v2i3.4>
- Génin, F., & Perret, E. M. (2003). Daily hypothermia in captive grey mouse lemurs (*Microcebus murinus*): Effects of photoperiod and food restriction. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 136(1), 71–81. [https://doi.org/10.1016/S1096-4959\(03\)00172-6](https://doi.org/10.1016/S1096-4959(03)00172-6)
- Goodman, S. M., O'Connor, S., & Langrand, O. (1993). A review of predation on lemurs: Implications for the evolution of social behavior in small, nocturnal primates. In P. M. Kappeler, & J. U. Ganzhorn (Eds.), *Lemur social systems and their ecological basis* (pp. 51–66). Springer. https://doi.org/10.1007/978-1-4899-2412-4_5
- Gygax, L. (2014). The A to Z of statistics for testing cognitive judgement bias. *Animal Behaviour*. <https://doi.org/10.1016/j.anbehav.2014.06.013>
- Harding, J. D. (2017). Nonhuman primates and translational research: Progress, opportunities, and challenges. *ILAR Journal*, 58(2), 141–150. <https://doi.org/10.1093/ilar/ilx033>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Helfer, G., Barrett, P., & Morgan, P. J. (2019). A unifying hypothesis for control of body weight and reproduction in seasonally breeding mammals. *Journal of Neuroendocrinology*, 31(3), Article e12680. <https://doi.org/10.1111/jne.12680>
- Hemsworth, P. H., Mellor, D. J., Cronin, G. M., & Tilbrook, A. J. (2015). Scientific assessment of animal welfare. *New Zealand Veterinary Journal*, 63(1), 24–30. <https://doi.org/10.1080/00480169.2014.966167>
- Heuer, E., Rosen, R. F., Cintron, A., & Walker, L. C. (2012). Nonhuman primate models of Alzheimer-like cerebral proteopathy. *Current Pharmaceutical Design*, 18(8), 1159–1169. <https://doi.org/10.2174/138161212799315885>
- Ho, C. L. A., Fichtel, C., & Huber, D. (2021). The gray mouse lemur (*Microcebus murinus*) as a model for early primate brain evolution. *Current Opinion in Neurobiology*, 71, 92–99. <https://doi.org/10.1016/j.conb.2021.09.012>
- Honess, P. E., & Marin, C. M. (2006). Enrichment and aggression in primates. *Neuroscience & Biobehavioral Reviews*, 30(3), 413–436. <https://doi.org/10.1016/j.neubiorev.2005.05.002>
- Hoy, J. M., Murray, P. J., & Tribe, A. (2010). Thirty years later: Enrichment practices for captive mammals. *Zoo Biology*, 29(3), 303–316. <https://doi.org/10.1002/zoo.20254>
- Hozer, C., Perret, M., Pavard, S., & Pifferi, F. (2020). Survival is reduced when endogenous period deviates from 24 h in a non-human primate, supporting the circadian resonance theory. *Scientific Reports*, 10(1), Article 18002. <https://doi.org/10.1038/s41598-020-75068-8>
- Hozer, C., & Pifferi, F. (2020). Physiological and cognitive consequences of a daily 26 h photoperiod in a primate: Exploring the underlying mechanisms of the circadian resonance theory. *Proceedings. Biological Sciences*, 287(1931), Article 20201079. <https://doi.org/10.1098/rspb.2020.1079>
- Hozer, C., & Pifferi, F. (2022). Relationships between endogenous circadian period, physiological and cognitive parameters and sex in aged gray mouse lemurs (*Microcebus murinus*). *Chronobiology International*, 39(3), 363–373. <https://doi.org/10.1080/07420528.2021.2001478>
- Hozer, C., Pifferi, F., Aujard, F., & Perret, M. (2019). The biological clock in gray mouse Lemur: Adaptive, evolutionary and aging considerations in an emerging non-human primate model. *Frontiers in Physiology*, 10, 1033. <https://doi.org/10.3389/fphys.2019.01033>
- Huchard, E., Schliehe-Diecks, S., Kappeler, P. M., & Kraus, C. (2017). The inbreeding strategy of a solitary primate, *Microcebus murinus*. *Journal of Evolutionary Biology*, 30(1), 128–140. <https://doi.org/10.1111/jeb.12992>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Joly, M., Scheumann, M., & Zimmermann, E. (2008). Wild mouse lemurs revisit artificial feeding platforms: Implications for field experiments on sensory and cognitive abilities in small primates. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, 70(9), 892–896. <https://doi.org/10.1002/ajp.20560>
- Joseph, D. N., & Whirlledge, S. (2017). Stress and the HPA axis: Balancing homeostasis and fertility. *International Journal of Molecular Sciences*, 18(10), 2224. <https://doi.org/10.3390/ijms18102224>
- Landes, J., Henry, P.-Y., Hardy, I., Perret, M., & Pavard, S. (2019). Female reproduction bears no survival cost in captivity for gray mouse lemurs. *Ecology and Evolution*, 9(11), 6189–6198. <https://doi.org/10.1002/ece3.5124>
- Languille, S., Blanc, S., Blin, O., Canale, C. I., Dal-Pan, A., Devau, G., Dhenain, M., Dorieux, O., Epelbaum, J., Gomez, D., Hardy, I., Henry, P. Y., Irving, E. A., Marchal, J., Mestre-François, N., Perret, M., Picq, J. L., Pifferi, F., ... Aujard, F. (2012). The grey mouse lemur: A non-human primate model for ageing studies. *Ageing Research Reviews*, 11(1), 150–162. <https://doi.org/10.1016/j.arr.2011.07.001>
- Matikainen-Ankney, B. A., Garmendia-Cedillos, M., Ali, M., Krynskiy, J., Salem, G., Miyazaki, N. L., Phida, T., & Kravitz, A. V. (2019). Rodent activity detector (RAD), an open source device for measuring activity in rodent home cages. *eNeuro*, 6(4). <https://doi.org/10.1523/ENEURO.0160-19.2019>
- McKinney, W. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference* (Vol. 445). <https://doi.org/10.25080/Majora-92bf1922-00a>
- Mellor, D. J. (2016). Updating animal welfare thinking: Moving beyond the “five freedoms” towards “a life worth living”. *Animals: An Open Access Journal from MDPI*, 6(3), 21. <https://doi.org/10.3390/ani6030021>
- Mendl, M., Burman, O. H. P., Parker, R. M. A., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, 118(3–4), 161–181. <https://doi.org/10.1016/j.applanim.2009.02.023>

- Naya, D. E., Naya, H., & White, C. R. (2018). On the interplay among ambient temperature, basal metabolic rate, and body mass. *American Naturalist*, 192(4), 518–524. <https://doi.org/10.1086/698372>
- Nelson, R. J., & Mandrell, T. D. (2005). Enrichment and nonhuman primates: "First, do no harm". *ILAR Journal*, 46(2), 171–177. <https://doi.org/10.1093/ilar.46.2.171>
- Novak, M. A., Hamel, A. F., Kelly, B. J., Dettmer, A. M., & Meyer, J. S. (2013). Stress, the HPA axis, and nonhuman primate well-being: A review. *Applied Animal Behaviour Science*, 143(2–4), 135–149. <https://doi.org/10.1016/j.applanim.2012.10.012>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Perea-Rodríguez, J. P., Corley, M. K., de la Iglesia, H., & Fernandez-Duque, E. (2022). Thermoenergetic challenges and daytime behavioural patterns of a wild cathemeral mammal. *Animal Behaviour*, 185, 163–173. <https://doi.org/10.1016/j.anbehav.2021.12.008>
- Perret, M. (2021). Litter sex composition influences competitive performance during first reproduction in male mouse lemurs. *Physiology and Behavior*, 228, Article 113196. <https://doi.org/10.1016/j.physbeh.2020.113196>
- Perret, M., & Aujard, F. (2001). Regulation by photoperiod of seasonal changes in body mass and reproductive function in gray mouse lemurs (*Microcebus murinus*): Differential responses by sex. *International Journal of Primatology*, 22(1), 5–24. <https://doi.org/10.1023/A:1026457813626>
- Pifferi, F., Epelbaum, J., & Aujard, F. (2019). Strengths and weaknesses of the gray mouse lemur (*Microcebus murinus*) as a model for the behavioral and psychological symptoms and neuropsychiatric symptoms of dementia. *Frontiers in Pharmacology*, 10, 1291. <https://doi.org/10.3389/fphar.2019.01291>
- Pifferi, F., Terrien, J., Marchal, J., Dal-Pan, A., Djelti, F., Hardy, L., Chahory, S., Cordonnier, N., Desquilbet, L., Hurion, M., Zahariev, A., Chery, I., Zizzari, P., Perret, M., Epelbaum, J., Blanc, S., Picq, J.-L., Dhenain, M., & Aujard, F. (2018). Caloric restriction increases lifespan but affects brain integrity in grey mouse lemur primates. *Communications Biology*, 1, 30. <https://doi.org/10.1038/s42003-018-0024-8>
- Poirotte, C., & Kappeler, P. M. (2019). Hygienic personalities in wild grey mouse lemurs vary adaptively with sex. *Proceedings. Biological Sciences*, 286(1908), Article 20190863. <https://doi.org/10.1098/rspb.2019.0863>
- Pons, P., Jaen, J., & Catala, A. (2017). Assessing machine learning classifiers for the detection of animals' behavior using depth-based tracking. *Expert Systems with Applications*, 86, 235–246. <https://doi.org/10.1016/j.eswa.2017.05.063>
- Poole, T. (1997). Happy animals make good science. *Laboratory Animals*, 31(2), 116–124. <https://doi.org/10.1258/002367797780600198>
- Radespiel, U. (2000). Sociality in the gray mouse lemur (*Microcebus murinus*) in northwestern Madagascar. *American Journal of Primatology*, 51(1), 21–40. [https://doi.org/10.1002/\(SICI\)1098-2345\(200005\)51:1<21::AID-AJP3>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1098-2345(200005)51:1<21::AID-AJP3>3.0.CO;2-C)
- Radespiel, U., Ehresmann, P., & Zimmermann, E. (2003). Species-specific usage of sleeping sites in two sympatric mouse lemur species (*Microcebus murinus* and *M. ravelobensis*) in northwestern Madagascar. *American Journal of Primatology*, 59(4), 139–151. <https://doi.org/10.1002/ajp.10071>
- Radespiel, U., Rakotoniravony, R., Rasoloharijaona, S., & Randrianambinina, B. (2022). A 24-year record of female reproductive dynamics in two sympatric mouse lemur species in Northwestern Madagascar. *International Journal of Primatology*, 43(4), 559–583. <https://doi.org/10.1007/s10764-021-00261-z>
- Rao, Y., Jiang, M., Wang, W., Zhang, W., & Wang, R. (2020). On-farm welfare monitoring system for goats based on Internet of Things and machine learning. *International Journal of Distributed Sensor Networks*, 16(7). <https://doi.org/10.1177/1550147720944030>
- Reghem, E., Byron, C., Bels, V., & Pouydebat, E. (2012). Hand posture in the grey mouse lemur during arboreal locomotion on narrow branches. *Journal of Zoology*, 288(1), 76–81. <https://doi.org/10.1111/j.1469-7998.2012.00926.x>
- Russell, W. M. S., & Burch, R. L. (1959). *The principles of humane experimental technique*. Methuen.
- Saltzman, W., Harris, B. N., De Jong, T. R., Perea-Rodríguez, J. P., Horrell, N. D., Zhao, M., & Andrew, J. R. (2017). Paternal care in biparental rodents: Intra- and inter-individual variation. *Integrative and Comparative Biology*, 57(3), 589–602. <https://doi.org/10.1093/icb/ix047>
- Scheumann, M., Zimmermann, E., & Deichsel, G. (2007). Context-specific calls signal infants' needs in a strepsirrhine primate, the gray mouse lemur (*Microcebus murinus*). *Developmental Psychobiology*, 49(7), 708–718. <https://doi.org/10.1002/dev.20234>
- Schliehe-Diecks, S., Eberle, M., & Kappeler, P. M. (2012). Walk the line-dispersal movements of gray mouse lemurs (*Microcebus murinus*). *Behavioral Ecology and Sociobiology*, 66(8), 1175–1185. <https://doi.org/10.1007/s00265-012-1371-y>
- Schmid, J., & Kappeler, P. M. (1998). Fluctuating sexual dimorphism and differential hibernation by sex in a primate, the gray mouse lemur (*Microcebus murinus*). *Behavioral Ecology and Sociobiology*, 43(2), 125–132. <https://doi.org/10.1007/s002650050474>
- Schneider, S., Taylor, G. W., Linquist, S., & Kremer, S. C. (2019). Past, present and future approaches using computer vision for animal re-identification from camera trap data. *Methods in Ecology and Evolution*, 10(4), 461–470. <https://doi.org/10.1111/2041-210X.13133>
- Shively, C. A., & Clarkson, T. B. (2009). The unique value of primate models in translational research. Nonhuman primate models of women's health: Introduction and overview. *American Journal of Primatology*, 71(9), 715–721. <https://doi.org/10.1002/ajp.20720>
- Smith, T. E., & French, J. A. (1997). Psychosocial stress and urinary cortisol excretion in marmoset monkeys (*Callithrix kuhli*). *Physiology and Behavior*, 62(2), 225–232. [https://doi.org/10.1016/S0031-9384\(97\)00103-0](https://doi.org/10.1016/S0031-9384(97)00103-0)
- Stevenson, T. J., & Ball, G. F. (2011). Information theory and the neuropeptidergic regulation of seasonal reproduction in mammals and birds. *Proceedings. Biological Sciences*, 278(1717), 2477–2485. <https://doi.org/10.1098/rspb.2010.2181>
- Symonds, M. R. E., & Moussalli, A. (2011). A brief guide to model selection, multi-model inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, 65(1), 13–21. <https://doi.org/10.1007/s00265-010-1037-6>
- Tardif, S., Carville, A., Elmore, D., Williams, L. E., & Rice, K. (2012). Reproduction and breeding of nonhuman primates. *Nonhuman Primates in Biomedical Research*, 197–249. <https://doi.org/10.1016/B978-0-12-381365-7.00008-X>
- Terrien, J., Gaudubois, M., Champeval, D., Zaninotto, V., Roger, L., Riou, J. F., & Aujard, F. (2018). Metabolic and genomic adaptations to winter fattening in a primate species, the grey mouse lemur (*Microcebus murinus*). *International Journal of Obesity*, 42(2), 221–230. <https://doi.org/10.1038/ijo.2017.195>
- Thomas, P., Herrel, A., Hardy, L., Aujard, F., & Pouydebat, E. (2016). Exploration behavior and morphology are correlated in captive gray mouse lemurs (*Microcebus murinus*). *International Journal of Primatology*, 37(3), 405–415. <https://doi.org/10.1007/s10764-016-9908-y>
- Thomas, P., Pouydebat, E., Hardy, L., Aujard, F., Ross, C. F., & Herrel, A. (2015). Sexual dimorphism in bite force in the grey mouse lemur. *Journal of Zoology*, 296(2), 133–138. <https://doi.org/10.1111/jzo.12225>
- Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour*, 124, 203–220. <https://doi.org/10.1016/j.anbehav.2016.12.005>
- van der Vinne, V., Gorter, J. A., Riede, S. J., & Hut, R. A. (2015). Diurnality as an energy-saving strategy: Energetic consequences of temporal niche switching in small mammals. *Journal of Experimental Biology*, 218(16), 2585–2593. <https://doi.org/10.1242/jeb.119354>
- van der Vinne, V., Riede, S. J., Gorter, J. A., Eijer, W. G., Sellix, M. T., Menaker, M., Daan, S., Pilorz, V., & Hut, R. A. (2014). Cold and hunger induce diurnality in a nocturnal mammal. *Proceedings of the National Academy of Sciences of the United States of America*, 111(42), 15256–15260. <https://doi.org/10.1073/pnas.1413135111>
- van der Vinne, V., Tachinardi, P., Riede, S. J., Akkerman, J., Scheepe, J., Daan, S., & Hut, R. A. (2019). Maximising survival by shifting the daily timing of activity. *Ecology Letters*, 22(12), 2097–2102. <https://doi.org/10.1111/ele.13404>
- Verjat, A., Devienne, P., Rödel, H. G., & Féron, C. (2021). More exploratory house mice judge an ambiguous situation more negatively. *Animal Cognition*, 24(1), 53–64. <https://doi.org/10.1007/s10071-020-01414-y>
- Vuarin, P., & Henry, P.-Y. (2014). Field evidence for a proximate role of food shortage in the regulation of hibernation and daily torpor: A review. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 184(6), 683–697. <https://doi.org/10.1007/s00360-014-0833-0>
- Wang, J., Bell, M., Liu, X., & Liu, G. (2020). Machine-learning techniques can enhance dairy cow estrus detection using location and acceleration data. *Animals: An Open Access Journal from MDPI*, 10(7), 1160. <https://doi.org/10.3390/ani10071160>
- Waskom, M. L. (2021). Seaborn: Statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. <https://doi.org/10.21105/joss.03021>
- Young, R. J. (2003). Environmental enrichment: An historical perspective. In *Environmental Enrichment for Adaptive Animals* (Vols. 1–19).

Appendix

Table A1

Summary of the different measures performed and how they were collected

Measure name	Procedure	Where was it done	When was it done	Feature used in PCA analyses
Circadian activity	Recording circadian activity with motion sensors	Experimental cage	Week 1	Total daytime and night-time bouts of activity
Judgement bias	Characterizing responses to ambiguous stimuli	Testing room with apparatus	Week 4 or 5	Not used in the analyses
Predatory motivation	Characterizing responses to prey	Testing room with apparatus	Week 4 or 5	Count of mealworms ingested
Stress reactivity	Monitoring facial temperatures during a 2 min manual restraint	Animal room	Weekly, before each training and testing session	Changes in nose temperatures on the first training session on week 4 or 5
Change in body mass	Weighing animals on weeks 1 and 5	Animal room	Weeks 1 and 5	Changes in body mass (g)
Change in body composition	Scanning animals with NMR on weeks 1 and 5	Adjacent facility	Weeks 1 and 5	Changes in body fat mass (g)

Table A2

Data sets used in the analyses

Start_date	Animal ID	Family	Sex	Location	Treatment	Cohort	Initial_ mass (g)	Mealworms_ ingested	Count_ activity_ bouts	Circadian_ activity_ index	Nose_ temperature_ change	Body_mass_ change (g)	%_Dry_fat_ mass_ change	JBI_
7 Apr 2021	M293BDA	B	Male	Bottom	High	1B	115	0	155 593	1.51	0.60	-13.35	-1.88	0.00
7 Apr 2021	M293CA	B	Male	Bottom	High	1A	141	NA	155 593	1.51	0.40	-15.17	-12.56	NA
7 Apr 2021	M294B	F	Male	Bottom	Control	1B	70	0	17 811	0.01	-3.90	14.14	13.02	0.38
7 Apr 2021	M294D	F	Male	Bottom	Control	1A	90	4	17 811	0.01	-0.40	-0.63	-4.18	1.00
7 Apr 2021	F321DB	C	Female	Top	High	1A	125	3	14 851	0.03	-1.70	-3.35	-4.01	1.00
7 Apr 2021	F388A	A	Female	Top	High	1B	120	4	14 851	0.03	-0.60	-22.31	-14.69	1.00
7 Apr 2021	F319AAA	C	Female	Top	Control	1A	140	3	91 39	0.02	1.40	-15.49	-9.58	0.00
7 Apr 2021	F319AAB	C	Female	Top	Control	1B	167	0	9139	0.02	-0.20	-7.76	-6.76	1.00
2 Jun 2021	M277AGD	J	Male	Bottom	High	2A	82	0	5400	0.02	0.70	16.38	12.45	0.07
2 Jun 2021	M285AC	G	Male	Bottom	High	2B	76	4	5400	0.02	0.20	9.08	8.24	0.03
2 Jun 2021	M338BA	A	Male	Bottom	Control	2B	93	0	8140	0.02	-1.40	-7.64	-0.46	1.00
2 Jun 2021	M294AA	F	Male	Bottom	Control	2A	89	1	8140	0.02	-0.30	-4.93	-1.82	0.67
2 Jun 2021	F293CB	B	Female	Top	High	2B	108	4	51 002	0.85	0.50	-17.65	6.35	1.00
2 Jun 2021	F285BAC	G	Female	Top	High	2A	106	0	51 002	0.85	1.10	-9.95	-8.84	0.80
2 Jun 2021	F366B	G	Female	Top	Control	2B	95	4	14 521	0.03	-1.10	-9.58	0.55	0.00
2 Jun 2021	F277ABA	J	Female	Top	Control	2A	95	0	14 521	0.03	1.30	-8.86	-0.27	1.00
8 Sep 2021	F329BB	E	Female	Bottom	High	4B	120	4	12 571	0.40	-0.10	-20.37	-23.26	1.00
8 Sep 2021	F354CC	E	Female	Bottom	High	4A	121	0	12 571	0.40	1.50	-7.05	-5.94	1.00
8 Sep 2021	F354F	E	Female	Top	Control	4B	93	0	13 947	0.19	0.20	-11.00	-5.03	0.67
8 Sep 2021	F370C	E	Female	Top	Control	4A	103	0	13 947	0.19	0.70	-0.83	7.10	0.36
8 Sep 2021	M198FDD	H	Male	Top	Control	4A	88	0	37 446	0.26	2.00	-4.87	-1.64	0.43
8 Sep 2021	M370A	E	Male	Top	Control	4B	85	4	37 446	0.26	1.00	-8.41	-4.91	0.00
8 Sep 2021	M331AA	I	Male	Bottom	High	4B	80	4	28 617	0.02	1.70	-7.67	-5.23	0.00
8 Sep 2021	M304D	D	Male	Bottom	High	4A	71	4	28 617	0.02	1.50	-10.67	-8.67	1.00

JBI: judgement bias index. NA: not available.

Table A3

Loadings from the seven principal components identified

Feature	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Lineage	0.347856	-0.484911	-0.075085	0.153308	-0.371860	-0.242184	0.492690
Sex	0.366126	0.199033	-0.479274	-0.204531	-0.100916	0.198775	-0.458703
Subcohort	-0.009072	-0.545724	-0.297842	-0.053957	0.761457	0.072552	0.045670
Cage location	-0.278780	-0.280555	0.505556	0.291300	-0.098466	0.359206	-0.252592
Circadian activity index	-0.221803	0.369863	-0.396067	0.412701	0.017192	0.379742	0.524677
Mealworms ingested	-0.154578	-0.194130	0.003352	-0.733348	-0.264756	0.480487	0.286054
Nose temperature change	-0.177156	-0.417233	-0.451045	0.300979	-0.408436	0.142120	-0.350293
Body mass change	0.577379	-0.009969	0.123323	0.055560	-0.058458	0.006562	0.028738
Body fat mass change	0.478860	-0.007282	0.211732	0.212367	0.146269	0.610106	0.017747

Table A4
Data calculated from the newly identified principal components used to train the models

PC1	PC2	PC3	PC4	PC5	PC6	PC7	Treatment
-0.610986	2.629772	-1.931313	1.295896	0.005742	0.788571	0.238072	Control
-1.311315	3.025391	-1.985512	0.994610	-0.555143	-0.068487	0.237054	Control
3.857766	2.034835	1.621681	-0.378546	0.892389	-0.252946	0.546267	High
1.146516	0.842769	-0.057054	-1.611074	-1.434670	-0.178228	0.091262	High
-0.455478	0.990935	2.397925	-0.700467	-0.427695	0.057585	0.158394	Control
-2.749820	0.616054	1.374560	-1.368657	-0.255117	-0.142678	-0.423045	Control
-1.946180	-0.004615	1.022365	-0.204382	-1.433059	-0.032675	-0.747726	High
-0.846704	0.490441	1.576362	0.671705	-0.058926	-0.713763	-0.710129	High
3.854943	-0.510807	-0.273665	0.910094	-0.714761	-0.106074	0.103569	Control
2.495373	-0.507220	-0.383655	-1.117137	-0.269828	0.859584	0.270056	Control
0.701501	1.568786	-0.103625	-0.615618	1.605491	-0.426482	-1.105249	High
1.231516	0.494548	-0.456075	-0.438706	-0.085698	-0.659190	-0.325544	High
-1.680634	0.100469	0.475984	0.223636	0.521825	2.240474	0.559351	Control
-1.170410	-0.292214	-0.000392	1.894328	-0.588273	-0.337780	0.704589	Control
-0.218645	-1.007446	1.585058	-0.724364	0.127155	0.467069	0.978320	High
0.182090	-1.614210	0.814652	1.577435	-0.967833	-0.666782	0.264805	High
-2.311823	-0.693039	-1.152681	-1.587122	0.844279	-1.323222	1.241115	Control
-0.372319	-0.510235	-0.924420	0.893385	0.675917	-0.912533	0.253378	Control
-0.839571	-1.146206	0.530242	0.941032	1.368929	-0.414425	-0.136424	High
0.428819	-1.009946	0.975507	1.462462	0.916845	0.504235	-0.241643	High
0.589333	-1.517463	-0.869950	1.354905	-0.253442	0.179442	-0.884567	High
-0.415262	-1.356964	-0.731417	-0.742428	0.343117	1.162438	-0.568732	High
0.731474	-1.973900	-1.882818	-1.157139	-0.274180	-0.084557	0.217263	Control
-0.290185	-0.649737	-1.621759	-1.573848	0.016936	0.060423	-0.720436	Control

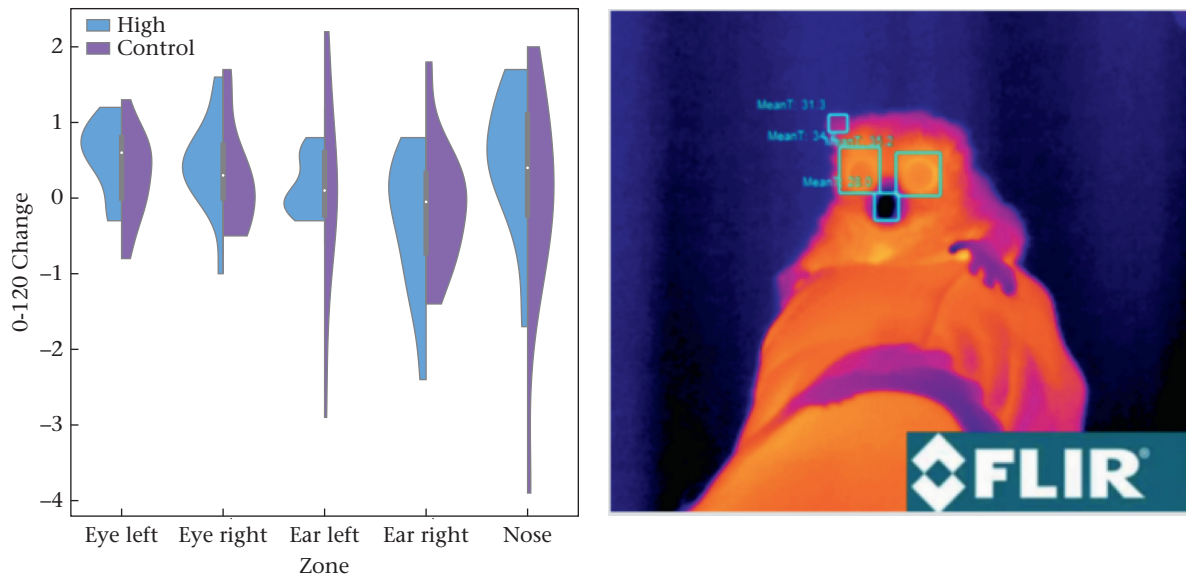


Figure A1. Data summaries and examples of thermographic data collected during a 2 min manual restraint while undergoing a routine wellness check. (a) Changes in surface temperature of the eyes, ears and nose of animals under High and Control levels of environmental enrichment. (b) Representation of how temperature measures were obtained. The violin plots show the distribution of the data with kernel density estimation. Inside each violin, the white dot represents the median, the thick bar indicates the interquartile range (IQR), and the thin lines (whiskers) extend to 1.5 times the IQR from the 25th and 75th percentiles.