Howarth, E, Witham, C and Bethell, E

Does Measuring Social Attention Lead to Changes in Behavior? A Preliminary Investigation into the Implications of Attention Bias Trials on Behavior in Rhesus Macaques

Routledge
Taylor & Francis Group

Journal of Applied Animal Welfare Science

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/haaw20

Does Measuring Social Attention Lead to Changes in Behavior? A Preliminary Investigation into the Implications of Attention Bias Trials on Behavior in Rhesus Macaques

Emmeline Howarth, Claire Witham & Emily Bethell

To cite this article: Emmeline Howarth, Claire Witham & Emily Bethell (01 Apr 2025): Does Measuring Social Attention Lead to Changes in Behavior? A Preliminary Investigation into the Implications of Attention Bias Trials on Behavior in Rhesus Macaques, Journal of Applied Animal Welfare Science, DOI: 10.1080/10888705.2025.2486351

To link to this article: https://doi.org/10.1080/10888705.2025.2486351

Published online: 01 Apr 2025.

JOURNAL OF APPLIED ANIMAL WELFARE SCIENCE
https://doi.org/10.1080/10888705.2025.2486351

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Does Measuring Social Attention Lead to Changes in Behavior? A Preliminary Investigation into the Implications of Attention Bias Trials on Behavior in Rhesus Macaques

Emmeline Howarth[a,b], Claire Witham[c], and Emily Bethell[b]

[a]Research Centre in Brain and Behaviour, School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, UK; [b]School of Natural Sciences, University of Chester, Chester, UK; [c]Centre for Macaques, Harwell Institute, Medical Research Council, Salisbury, UK

ABSTRACT

A welfare assessment tool in development must satisfy several criteria before it is considered ready for general use. Some tools that meet many of these criteria have been criticized for their negative effect on welfare. We conducted a preliminary assessment of the impact of attention bias (AB) trials using threat-neutral conspecific face pairs followed by presumed neutral-positive filler stimuli on the behavior of 21 rhesus macaques (*Macaca mulatta*; 15 female). Behavioral observations were conducted following AB trials and repeated two weeks later when no AB trials had occurred (no trial: NT). The association between observation period and behavior was assessed using linear mixed-effects models in R. Trials did not impact any observed behavior except for fear, which was displayed by five monkeys over six trials (four NT). For this sample, there was a significant reduction in fear behavior following AB trials. We, therefore, found no evidence suggesting that AB trials negatively affect behaviour. AB protocols may be suitable for continued development for primate welfare assessment and we encourage researchers to include assessing test impact on welfare in their AB protocols.

ARTICLE HISTORY
Received 21 December 2024
Accepted 24 March 2025

KEYWORDS
Affective state; animal welfare; attention bias; *Macaca mulatta*; primates

Introduction

Welfare assessment tools that are validated, standardized, easy to use and cost-effective while also being sensitive to the signal of interest and having high reliability and repeatability are challenging to develop (Blokhuis et al., 2010; Kilkenny et al., 2010). To further add to the challenge, there is no single measure of animal welfare (Wolfensohn & Honess, 2005), and the term has been defined in several ways. For example, Broom (1986) defined the welfare of an animal as “its state as regards its attempts to cope with its environment,” and Webster (2005) stated that “the welfare of a sentient animal is determined by its capacity to avoid suffering and sustain fitness.”

Due to these challenges, historic welfare assessment protocols focused on objective, prescribed (resource-based) measures that relate to the animals’ environment, such as the provision of shelter or bedding material. These measures are easy to define and record, are quantitative and have high inter-rater repeatability (Jones et al., 2022); however, they fail to consider the impact of the environment on the animal (Sherwen et al., 2018). As a result, attention moved to outcome (animal-based) measures, including health, physiology, behavior

CONTACT Emmeline Howarth e.howarth@chester.ac.uk School of Natural Sciences, University of Chester, Thomas Building, Exton Park, Chester CH1 4BJ, UK

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10888705.2025.2486351

and cognition (Richmond et al., 2017; Whay et al., 2003). These measures are now frequently used in welfare assessment (e.g., Welfare Quality, 2018). They are often considered superior and more indicative of animal experience as they measure the impact of the intervention rather than assessing the quality of the resources provided (Richmond et al., 2017). Outcome based measures can also be used to assess positive welfare (Yeates & Main, 2008). Behavior is particularly popular as it is noninvasive, can be non-intrusive and has been described as the "ultimate phenotype," as it is the outcome of the animal's individual decision-making process (Dawkins, 2004).

Despite this shift, outcome measures may be limited due to issues with variability and sensitivity (Lensink et al., 2003; Temple et al., 2013), as behavior and physiology can be affected by factors other than changes in welfare state. For example, behavior can be influenced by personality (Konečná et al., 2012) and physiology by aging (Taffet, 2024). Unvalidated methodologies and errors in data collection and interpretation call into question the informative power of some animal-based welfare proxies due to the potential for misdiagnosis of welfare (Tiemann et al., 2023; Watters et al., 2021).

Initially developed to assess exploration, sociability and memory in rodent species, open field (OF; Glickman & Hartz, 1964), novel object (NO; Ennaceur & Meliani, 1992), preference (Vincent, 1915) and forced swim tests (Bayroff, 1921) have been used to measure affective state (Cryan & Slattery, 2007; Watters & Krebs, 2019). OF and NO tests use unattached equivocal cues to assess approach and avoidance as measures of fearfulness (Forkman et al., 2007). Preference or choice tests involve the animal making decisions between different resources or stimuli (Kirkden & Pajor, 2006). Animal preference can also be used to assess motivation to obtain or avoid a resource or stimulus using the cost an animal is willing to pay as a measure of motivation (Fraser & Matthews, 1997). The forced swim test, which was commonly used for rodent models of depression, assessed time spent immobile as a measure of depression and active coping (Armario, 2021).

These tests are generally easy to conduct, relatively quick and have high inter-rater reliability as they utilize quantitative measures such as latency, duration, and frequency (Kirkden & Pajor, 2006). However, there is a lack of standardization of methods (Blaser & Heyser, 2015), issues with interpretation, especially in different contexts (Kirkden & Pajor, 2006), and, perhaps most importantly, the tests themselves may negatively impact welfare (Aliphon et al., 2023; Dalmau et al., 2009). For example, the OF and NO tests expose animals to unpredictable and unfamiliar stimuli to mimic the key features of predatory attack and induce a fear response (Forkman et al., 2007).

Attention bias (AB) describes a tendency to differentially allocate attention toward one of two or more stimuli that vary in emotional content. AB tasks were first developed to assess emotion in humans and have revealed that the allocation of attention to emotional stimuli changes with affective state (Bar-Haim et al., 2007; MacLeod et al., 1986). Here, affective state is defined as an animal's mood or emotional state and can encompass high and low arousal (intensity) and positive and negative valance (Mendl et al., 2010). Affective state is considered to be a key component of welfare (Mendl et al., 2010; Monk et al., 2023). AB has since been adapted as a noninvasive method of determining affective state for welfare assessment in a range of animal species (review Crump et al., 2018).

In primates, AB tasks can detect shifts in affective state following the provision of enrichment (Bethell et al., 2012) and reveal stable individual differences in baseline social attention, suggesting sensitivity to trait affect and significant reproducibility between tasks conducted several years apart (Howarth et al., 2021). Further, polymorphisms in genes within the serotonin, dopamine, and oxytocin pathways have been associated with AB, suggesting a genetic basis for the biological underpinning of individual variation in emotion (Howarth et al., 2023).

AB tasks measure innate preferential attention allocation and require little or no training (e.g., Howarth et al., 2021). They may be more appropriate and practical in real-world settings compared to judgment bias tasks as these require extensive training. Therefore, judgment bias tasks can be disruptive to management and husbandry routines, costly in terms of both money and time, and participant number attrition may impact study statistical power (Bethell, 2015; Harding et al., 2004).

In Howarth et al. (2021), we argued that a tool in development must satisfy several criteria before it is ready for general use. Here, we add that a tool must also have minimal impact on welfare. To that end, this study aimed to provide the research community with a better understanding of the impact that AB trials may have on animal behavior.

Materials and methods

Ethics

The AB protocol was developed following discussion with the facility Home Office Inspector (Nov 2011) and carried out in accordance with ethical guidelines for work with non-human primates (NC3Rs, 2006; 2015). Approval for the study was granted by the Medical Research Council (MRC) AWERB in November 2017 and the Liverpool John Moores University ethics panel (approval #EB_EH/2017–5). Animal health was monitored daily by the care staff and annually with a full veterinary examination. Animals were not restrained at any time during training or testing. Participation in this study was voluntary insofar as only animals who approached their individual target and the apparatus for rewards took part. Methods and results are reported according to the ARRIVE guidelines (Kilkenny et al., 2010).

Animals and housing

Data were collected from 21 adult rhesus macaques (*Macaca mulatta;* 15 female) housed at the Centre for Macaques, MRC Harwell Institute, UK (mean age on the first day of testing = 8.33 years, range = 3.5–13.2 years). The monkeys were housed in six breeding groups comprising one adult male and between three and 11 related females, plus infants and juveniles. One monkey was housed in a single-sex group with only adult females and their offspring from previous years. Any adult (at least 3 years old) in the six social groups was eligible for inclusion in the study if they reached the station training criterion of sitting by their individual targets (Kemp et al., 2017) for the time (around one minute) required for AB trials to be conducted. Animal information is provided in Supplementary Material Table S1, and full details about the animal housing, social groupings, training, enrichment and diet can be found in Howarth et al. (2021). Further details of the facility can be found in Witham (2015), and images and video of the facility, including macaque housing, can be viewed at https://cfm.har.mrc.ac.uk/ and https://macaques.nc3rs.org.uk/.

Attention bias

The AB protocol assessed in this study included the presentation of unfamiliar threat-neutral male conspecific face pair stimuli (Figure 1) to macaques for three seconds. Stimuli were presented as digital images using an automated apparatus on two 8-inch computer screens (Figure 2). Each stimulus pair contained one frontal view of an unfamiliar male macaque face with an aggressive expression, direct gaze and bared teeth in an open mouth (threat face) and one frontal view of the same male with a relaxed or neutral expression, eyes and mouth closed (neutral face). Construction of the stimuli from images collected at the Caribbean Primate Research Centre is detailed in Bethell et al. (2012), and images can be downloaded from the Macaque Faces Stimulus Set (Witham & Bethell, 2019). When a monkey was sitting by their individual target and looking centrally at the apparatus, an AB trial was triggered remotely via a MATLAB display using a laptop that was out of the monkey's field of view. Looking time at the stimuli was recorded using a video camera and later coded using Behavioral Observation Research Interactive Software (BORIS; Friard & Gamba, 2016). Looking time data used for determining affective state from the AB trials has been previously published by Howarth et al. (2021). All data were collected on quiet days when no stressors were known to have occurred (i.e., no scheduled cleaning or vet checks).

Figure 1. Four example unfamiliar threat-neutral male conspecific face pair stimuli from the Macaque Faces Stimulus Set (Witham & Bethell, 2019), with one picture pair grouped within each rectangle. Here, the threat face is presented on the right and the neutral on the left, while in the attention bias trials the side of threat face presentation was randomized.

Figure 2. The automated apparatus used for displaying the stimuli for the attention bias trials. (a): Face (threat-neutral conspecific face pair) stimuli; (b): inter-trial interval; (c): example presumed neutral-positive filler (fruit or vegetable) stimuli.

A set of presumed neutral-positive filler stimuli was created and used in addition to the threat-neutral face pairs. The filler stimuli were images of vegetables and fruit that were included in the macaques' diet and, therefore, would presumably find pleasant and interesting to look at (Waitt & Buchanan-Smith, 2006). Filler stimuli were shown to each monkey for three seconds after they had completed a trial so that the threat-neutral face pair was not the last image they saw, thereby reducing the chance of a negative association developing between the macaques and the apparatus. The filler stimuli were used as an end-of-session cue to indicate the end of a trial and "release" the monkey from the training or testing session (Topoleski et al., 2018). The responses were not filmed as the filler stimuli were included only to ensure monkeys viewed neutral-positive stimuli as well as threat-neutral stimuli on the apparatus.

Behaviour

An ethogram of behavioral indices (Table 1) was constructed to include key behaviors associated with aggression, anxiety, depression, distraction, fear, foraging, prosocial and inactive behavior. These categories were chosen as they have previously been identified as key indices for welfare in rhesus macaques (Camus et al., 2013; Coleman & Pierre, 2014; Maestripieri & Wallen, 1997).

Live continuous focal animal behavioral observations were completed using BORIS (Friard & Gamba, 2016; Lehner, 1992). Observations were conducted on four consecutive weekdays (Tuesday – Friday) following completion of the AB trials for that social group on that day. Following completion of the AB trial, where the monkeys had been stationed (Kemp et al., 2017), groups were given 10 minutes to move away from the training and testing area and return to performing their normal behavioral repertoire before each monkey was observed for five minutes. The monkeys were familiar with the observer as it was the same individual who conducted the training and AB trials. They were, therefore, considered habituated to the observer's presence (Prescott & Buchanan-Smith, 2007). Behavioral observations were started after no monkey within the group displayed any vigilance behavior toward the observer's location for two minutes. Five minutes was used, as short behavioral observation times are standard in AB protocols (e.g., 30 seconds: Campbell et al., 2019; 3 mins: Baqueiro-Espinosa et al., 2023; Lee et al., 2018; 5 mins: Bethell et al., 2012) and these data were initially recorded to validate the AB method for rhesus macaques (AB data have been previously published in Howarth et al., 2021). All individuals were observed by the same researcher (EH) within 60 minutes of their AB trial. The total duration of each behavior per five-minute trial was summed for each observation for each monkey.

Behavioral observations were repeated when no AB trials had been conducted for that group for two weeks, but all other conditions were the same (no trial: NT). Observations were matched for day of the week to account for the impact of changes in staff activity, diet and training schedules on behavior (Gottlieb et al., 2013). Time of day for observations was also matched to remove the impact of circadian changes in behavior and attention (e.g., Kappeler & Erkert, 2003; Novak et al., 2013; Plant, 1981). The Centre has a two-week rolling schedule for cleaning; this was accounted for in the time between observations to ensure enclosures were not being cleaned for either observation period. Daily husbandry schedules, such as cleaning, are thought to be disruptive and would have delayed the time for data collection (Bethell et al., 2012; Howarth et al., 2021).

NT data could not be collected prior to the AB trial period for practical and animal management reasons; however, it was decided that the two-week period was sufficient for macaque behavior to return to a representative baseline. Trial number has a significant effect on measures of social attention, with repeated testing associated with reduced response with a possible habituation effect (Bethell et al., 2019; Howarth et al., 2021; King et al., 2012). Previous work has shown that an inter-trial interval of seven days is sufficient to eradicate possible confounds such as carry-over effects (Howarth et al., 2021); therefore, two weeks is enough time for these carry-over effects to subside.

Table 1. Ethogram used for behavioral observation of captive rhesus macaques (*Macaca mulatta*).

Category	Behaviour	Description
Aggression	Aggressive	The animal chases, attacks, threatens, stares at, displaces or lunges toward a conspecific.
Anxiety (active)	Body shake	Like a dog shake – the animal rapidly moves whole body, usually starting with the head followed by rest of body.
	Stereotypic	The behavior has no obvious function. Includes pacing, bar biting or head tossing.
	Vigilance	The animal has alert posture scanning their environment or looking at a particular thing (may be out of view).
Anxiety (low activity) & depression	Self-directed	The animal uses their hands or mouth to clean, scratch or manipulate their skin or fur.
	Sit hunched	The animal is sitting with their back and head curved round so that the head is below their slumped shoulder.
	Yawn	Animal opens its mouth wide exposing their teeth (not directed at conspecific). The head may tilt upwards, and they may close their eyes.
Distraction	Object	The animal uses hands or mouth to investigate and move an inanimate, moveable item in the environment and/or pull or grab parts of the enclosure such as padlocks, sliding adjustable panels and cage dividers.
	Locomotion	Any behavior (except those otherwise defined) that involves the animal moving from one location to another, for example, quadrupedal and bipedal walking and running, climbing, descending, and jumping. The animal must not be engaged in any other activity.
Fear	Grimace	The animal's lips are pulled back to expose the teeth.
	Submissive	Animal moves away, flees, is displaced by or ducks away from a conspecific. They may give out high-pitched screams. Animal may also present its hindquarters in a non-sexual context (not followed by mating).
Foraging	Foraging	The animal is searching for and/or consuming food or water.
Inactive	Resting	The animal is lying horizontally with the stomach, back or side touching the floor or the animal is sitting upright.
	Stationary	Weight-bearing on two or four limbs.
Prosocial	Affiliative	Friendly interaction between the animal and a conspecific includes huddling, being in physical contact with a conspecific and hugging but not grooming behavior.
	Allogrooming	Reciprocal grooming. The animal's skin or fur is cleaned, scratched or manipulated by the hands or mouth of a conspecific. The animal uses their hands or mouth to clean, scratch or manipulate the skin or fur of a conspecific.
	Interaction with infant	The animal interacts with a baby e.g., grooming, playing, carrying, or feeding.
	Lip smack	Animal opens and closes its mouth repeatedly without showing its teeth, occasionally making a smacking sound.
	Sexual behavior	A female animal presents, or a male is presented with the hindquarters for mounting for copulation. The animal is mounted or mounts.
Other	Other	Any behavior not otherwise defined.
	Out of sight	The animal is not visible to the observer.

For five monkeys, observations could not be conducted four times during the AB trial period due to husbandry or veterinary activities on the fourth day. As trial number has been shown to significantly influence attention to threat (Howarth et al., 2021), for these animals, to balance the data, only three days of NT behavioral data were included in the analysis.

Statistical analysis

Statistical analyses were conducted in R Studio with R version 4.3.2 (R Core Team, 2023). Data were analyzed using linear mixed models (LMMs) using the R package "lme4" version 1.1–15 (Bates et al., 2015). A maximal model was built per response variable separately. Response variables were the eight behavior groups: aggression, high activity anxiety, low activity anxiety and depression, distraction, fear, foraging, inactive and prosocial. Observation period (after AB trials or NT) was entered as a key predictor variable in each model. In addition, we controlled for factors known to influence socioemotional behavior and attention, including sex, trial number (Howarth et al., 2021) and early weaning (Prescott et al., 2012). All predictor variables were initially assessed to ensure none were correlated above 0.4, which could result in collinearity (Crawley, 2007). Response variables were visually inspected for their distribution and transformed to obtain more normal distributions when necessary. Appropriate transformations were identified using Tukey's Ladder of Transformation (Tukey, 1977) to extract an appropriate lambda for transformation. Participant monkey identity was entered as a random factor in all models.

Model stability was assessed by visually inspecting qq-plots and histograms of residuals and running influence diagnostics to identify any influential cases. Where the model was stable, predictor variables with the greatest p value were removed in a stepwise manner (Crawley, 2007) to attain a final model comprising only the key predictor variable (after AB trials or NT) and predictor variables with $p \leq 0.05$. The initial model included the interaction between observation period and sex; however, this was non-significant for all response variables, so the interaction was removed from subsequent models. The "anova" function was used to compare the fit of the final model against the null model (a model retaining the random effect but with all fixed effects removed and an intercept of 1 specified). The final model was accepted only if it provided a significantly better fit than the null at $p \leq 0.05$ (which it did in all cases where significant results are reported). To account for multiple testing, confidence intervals (CI) were calculated, and significance was accepted where the CIs did not cross zero (Tan & Tan, 2010).

Results

Behavior data were collected from 21 adult monkeys (15 female). For 16 monkeys (11 female), behavioral observations were conducted on four consecutive AB trial days (total after AB trials duration = 20 minutes per monkey) and four consecutive NT days (total NT duration = 20 minutes per monkey). For five monkeys (four female), behavioral observations were conducted on three consecutive AB trial days (total after AB trials duration = 15 minutes per monkey) and three consecutive NT days (total NT duration = 15 minutes per monkey). This resulted in 158 trials for analysis.

Model output for each behavior is shown in Table 2. For inactive behavior, observation period and sex were retained in the final model. Male monkeys displayed a significantly higher duration of inactive behaviors (mean = 120.634 ± 99.168 seconds) than female monkeys (mean = 49.816 ± 71.243 seconds; $t = -2.632$, $p = 0.012$, df = 157). For all other behaviors, observation period was the only factor retained in the final model.

There was no significant relationship between observation period and any of the recorded behaviors except for fear. Only five monkeys (four female) over six trials (four NT) displayed any fear behavior. For this small sample, monkeys displayed a greater duration of fear behavior during the NT observation period (mean = 0.345 ± 1.700 seconds) compared to the after AB trials observation period (mean = 0.156 ± 1.036; $t = -2.523$, LRT = 6.224, $p = 0.013$, df = 157).

Table 2. Model output for the association between observation period (after attention bias (AB) trials) and without attention bias trial (no trial: NT)) and behavior in rhesus macaques (n = 21).

Behaviour	After AB trial mean (s)	NT mean (s)	Estimate	SE	t	95% CIL	95% CIU	LTR X^2	P	npar
Aggression	0.643 ± 3.033	0.798 ± 2.348	0.035	0.081	0.433	−0.125	0.195	0.187	0.665	1
High activity anxiety	2.416 ± 5.263	10.947 ± 25.968	−0.177	0.145	−1.219	−0.464	0.110	1.478	0.224	1
Low activity anxiety and depression	11.674 ± 33.054	18.260 ± 36.862	0.045	0.206	0.217	−0.362	0.451	0.047	0.828	1
Distraction	17.339 ± 23.146	29.716 ± 32.805	−0.064	0.429	−0.149	−0.911	0.783	0.022	0.881	1
Fear	0.156 ± 1.036	0.345 ± 1.700	−0.144	0.057	−2.523	−0.256	−0.031	6.224	0.013	1
Foraging	123.197 ± 125.270	73.310 ± 106.176	−1.076	1.148	−0.937	−3.340	1.188	0.876	0.349	1
Inactive	55.935 ± 88.449	84.933 ± 82.123	−0.448	0.582	−0.769	−1.596	0.701	0.590	0.442	1
(Sex)	120.634 ± 99.168	49.816 ± 71.243	−1.687	0.641	−2.632	−2.956	−0.409	6.303	0.012	1
Prosocial	84.134 ± 126.400	77.894 ± 111.626	0.464	0.374	1.242	−0.273	1.202	1.536	0.215	1

SE: standard error; CIL: lower confidence interval; CIU: upper confidence interval; LRT: likelihood ratio test; npar: number of parameters.

Observation period was retained in the final model for all behaviors. Retained factors, other the observation period, are denoted in brackets under the behavior. Significant *p* values are highlighted in bold.

Discussion

The present study aimed to determine whether AB trials have a negative impact on the behavior of rhesus macaques. To do this, we assessed whether there was any difference in the behavior of 21 adult rhesus macaques following AB trials, compared to a time-matched observation period when no AB trials had been conducted for at least two weeks prior (NT), but all other conditions were the same. There was no significant change in any behavior, except for fear, which was shown to reduce following AB trials compared to NT in the five monkeys that displayed any fear behavior.

When developing a new method of welfare assessment, it is crucial that the tool is repeatable, reliable, sensitive, cost-effective, standardized, and validated before it is ready for general use (Blokhuis et al., 2010; Howarth et al., 2021; Kilkenny et al., 2010). Although some current welfare assessment tools meet several of these criteria, they have been criticized for their negative impact on animal welfare during testing (Aliphon et al., 2023; Dalmau et al., 2009; Forkman et al., 2007).

Attention bias

The present study has shown that AB trials using an automated apparatus with threat-neutral conspecific face pairs followed by presumed neutral-positive fruit and vegetable filler stimuli do not have a significant negative impact on the behavior of rhesus macaques. These threat-neutral stimuli have previously been shown to influence social attention to threat (Howarth et al., 2021). Howarth et al. (2021) revealed that stimulus ID had a significant influence on attention bias difference score (calculated as the total duration of time looking at the threat face minus the total duration of time looking at the neutral face). The authors suggested that this may be the result of using photographs of real animals, resulting in variation in brightness, color, contrast energy and luminance of the images as well as the emotional intensity, age, attractiveness, dominance, and orientation of the faces. However, the impact on behavior of the filler stimuli and AB protocol as a whole had never been assessed.

The filler stimuli used here were images of food items that the macaques are familiar with and may find interesting to look at (Waitt & Buchanan-Smith, 2006). The images were downloaded from Google and chosen based on having a realistic appearance and a white background. A more formal approach to the selection and validation of filler stimuli is needed. A closer investigation into the impact of AB trials with and without the inclusion of these stimuli is a key next step for developing

AB tasks for welfare assessment, as they may be essential for ensuring AB protocols do not impact animal welfare.

AB trials have utilized a range of species-relevant stimuli, for example, alarm calls (Brilot & Bateson, 2012), novel objects (Verbeek et al., 2014), predators (Lee et al., 2016) and aggressive conspecifics (Vögeli et al., 2014). Many use unpredictability or suddenness like the NO or OF tests (e.g., sudden object presentation: Romeyer & Bouissou, 1992; loud, unpredictable noise: Desiré et al., 2004) to induce a threat response. For example, here we used the sudden presentation of unfamiliar conspecific faces with a threatening expression. Other studies, such as Baqueiro-Espinosa et al. (2023), used the rapid and sudden movements of an opening and closing umbrella as the negative stimuli. Unpredictable and unfamiliar stimuli are associated with negative emotional states (Dalmau et al., 2009) and trigger a defense cascade, which is a well-conserved and adaptive response that evolved to promote detection, evaluation, and response to threat (Statham et al., 2020). Although exposure to threat is usually shorter for AB trials (between 3 seconds (e.g., Howarth et al., 2021) and 10 seconds (e.g., Baqueiro-Espinosa et al., 2023; Lee et al., 2018), the results of the present study are important as they demonstrate that AB trials in primates do not negatively affect behavior and therefore, do not have the same negative impact on welfare as NO and OF tests.

Further research must be conducted to confirm this in other species and using other stimuli. AB has been used to assess emotion in numerous other species, including birds (e.g., Campbell et al., 2019), sheep (e.g., Raoult & Gygax, 2019), cattle (e.g., Lee et al., 2018), pigs (e.g., Luo et al., 2019), rats (e.g., Parker et al., 2014) and dogs (Baqueiro-Espinosa et al., 2023). The result of the present study cannot be extrapolated to other species using other stimuli, as there is variation in animal response to threat depending on, for example, whether they are a predator or prey species (Fardell et al., 2020). For some species, exposure to alarm vocalizations does not impact their welfare (e.g., pigs: Düpjan et al., 2011), while for others, alarm vocalizations trigger a significant fear response. In wild black-capped chickadees (*Poecile atricapillus*), predator-induced fear responses are associated with prolonged elevated neuronal activity and heightened sensitivity to threat (Zanette et al., 2019). Chickadees were exposed to 15 minutes of recorded predator alarm call vocalizations; birds exposed to this stimulus 7 days prior had a significantly heightened fear response compared to individuals exposed to neutral conspecific vocalization. Exposed birds had a longer duration of freeze response and heightened activity in the amygdala and hippocampus compared to naïve birds. The authors reported that this suggests a post-traumatic stress disorder response in animals where predator cues are used to elicit a fear response (Zanette et al., 2019). It is therefore important to assess whether fear responses elicited by welfare assessment tools have a significant long-term impact on welfare. To progress the field, welfare assessment tools must focus on protocols that have minimal impact on animal welfare. We, therefore, encourage researchers conducting AB trials to evaluate the impact of their stimuli on the species involved.

Fear behaviour

For the five monkeys that displayed fear behavior, the duration of this behavior was longer during the NT period compared to after AB trials. This behavioral response should not be used to infer that AB trials improve welfare or reduce the occurrence of fearful behavior. Due to the small sample of monkeys that displayed fear behavior, the significant finding may be a Type II error (Columb & Atkinson, 2015). However, the inclusion of fear behavior in this study is significant to demonstrate that AB trials do not induce a fear behavior response. This is a key consideration for welfare, especially for this subset of individuals that may have more fearful temperaments, i.e., those that showed any fear behavior.

Due to strict dominance hierarchies (Cooper & Bernstein, 2008; de Waal & Luttrell, 1985) and variation in genetic predisposition to fear and submissive behaviors (Howarth et al., 2023), some individuals are especially at risk of impaired welfare. Social rank is known to impact captive welfare and reliably predict behavioral and physiological responses to stress (Zhang et al., 2024), including the frequency of fear behavior. Subordinate macaques are more likely to display fear behavior than dominant individuals (Shively & Day, 2015). Despite the socioemotional challenges associated with rank, social housing is essential for welfare as individual housing is known to be significantly stressful (review DiVincenti & Wyatt, 2011). However, inappropriate grouping or changes to the social group are also associated with behavioral and physiological responses indicative of compromised welfare (review Hannibal et al., 2017). As changes to the social environment are not uncommon in research facilities (Hannibal et al., 2017), it is key that we develop sensitive welfare assessment tests that can detect shifts in social attention and cognitive welfare in these more vulnerable individuals. We need to understand individual profiles and not just population-level patterns. As a result, initial studies will have smaller sample sizes, but it is important that authors report these findings and make their data available for meta-analysis in the future.

Conclusions

AB trials using an automated apparatus with threat-neutral conspecific face pairs followed by the presentation of presumed neutral-positive fruit and vegetable filler stimuli were not associated with a negative change in behavior in rhesus macaques. Further research is needed to confirm this finding in other species and with different stimuli. Therefore, we encourage researchers to repeat this investigation with their AB protocols to ensure that welfare is never compromised during assessment using AB.

Acknowledgments

We thank the staff at MRC-CFM, particularly Faye Peters and Sebastian Merritt, who assisted with data collection.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

ERI Howarth was supported by an LJMU PhD studentship. CL Witham and the Centre for Macaques are funded by the Medical Research Council.

Author contributions

Conceptualization, E.H., E.B. and C.W.; methodology, E.H., E.B. and C.W.; software, C.W.; validation, E.H. and E.B.; formal analysis, E.H.; investigation, E.H.; resources, C.W.; data curation, E.H.; writing – original draft preparation, E. H.; writing – review and editing, E.H., E.B. and C.W.; visualization, E.H.; supervision, E.B.; project administration, E. H.; funding acquisition, E.H., E.B. and C.W. All authors have read and agreed to the published version of the manuscript.

References

Aliphon, B., Dai, T., Moretti, J., Penrose-Menz, M., Mulders, W. H. A. M., Blache, D., & Rodger, J. (2023). A repeated measures cognitive affective bias test in rats: Comparison with forced swim test. *Psychopharmacology*, *240*(11), 2257–2270. https://doi.org/10.1007/s00213-022-06281-8

Armario, A. (2021). The forced swim test: Historical, conceptual and methodological considerations and its relationship with individual behavioral traits. *Neuroscience & Biobehavioral Reviews*, *128*, 74–86. https://doi.org/10.1016/j.neubiorev.2021.06.014

Baqueiro-Espinosa, U., Lo, T. H., Hunter, R., Donnelly, P., McEvoy, V., Crump, A., & Arnott, G. (2023). Positive human interaction improves welfare in commercial breeding dogs: Evidence from attention bias and human sociability tests. *Applied Animal Behaviour Science*, *262*, 105904. https://doi.org/10.1016/j.applanim.2023.105904

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.48550/arXiv.1406.5823

Bayroff, A. G. (1921). The experimental social behavior of animals. II. The effect of early isolation of white rats on their competition in swimming. *Journal of Comparative Psychology*, *29*(3), 293–306. https://doi.org/10.1037/h0057256

Bethell, E. J. (2015). A "how-to" guide for designing judgment bias studies to assess captive animal welfare. *Journal of Applied Animal Welfare Science: JAAWS*, *18*(Sup. 1), S18–S42. https://doi.org/10.1080/10888705.2015.1075833

Bethell, E. J., Cassidy, L. C., Brockhausen, R. R., & Pfefferle, D. (2019). Toward a standardized test of fearful temperament in primates: A sensitive alternative to the human intruder task for laboratory-housed Rhesus macaques (*Macaca mulatta*). *Frontiers in Psychology*, *10*, 1–17. https://doi.org/10.3389/fpsyg.2019.01051

Bethell, E. J., Holmes, A., MacLarnon, A., & Semple, S. (2012). Evidence that emotion mediates social attention in rhesus macaques. *PLOS ONE*, *7*, 1–9. https://doi.org/10.1371/journal.pone.0044387

Blaser, R., & Heyser, C. (2015). Spontaneous object recognition: A promising approach to the comparative study of memory. *Frontiers in Behavioral Neuroscience*, *9*, 183. https://doi.org/10.3389/fnbeh.2015.00183

Blokhuis, H. J., Veissier, I., Miele, M., & Jones, B. (2010). The welfare Quality® project and beyond: Safeguarding farm animal well-being. *Acta Agriculturae Scandinavica A*, *60*(3), 129–140. https://doi.org/10.1080/09064702.2010.523480

Brilot, B. O., & Bateson, M. (2012). Water bathing alters threat perception in starlings. *Biology Letters*, *8*(3), 379–381. https://doi.org/10.1098/rsbl.2011.1200

Broom, D. M. (1986). Indicators of poor welfare. *The British Veterinary Journal*, *142*(6), 524–526. https://doi.org/10.1016/0007-1935(86)90109-0

Campbell, D. L. M., Taylor, P. S., Hernandez, C. E., Stewart, M., Belson, S., & Lee, C. (2019). An attention bias test to assess anxiety states in laying hens. *PeerJ*, *7*, e7303. https://doi.org/10.7717/peerj.7303

Camus, S. M. J., Blois-Heulin, C., Li, Q., Hausberger, M., Bezard, E., & Sgambato-Faure, V. (2013). Behavioural profiles in captive-bred cynomolgus macaques: Towards monkey models of mental disorders? *PLOS ONE*, *8*(4), e62141. https://doi.org/10.1371/journal.pone.0062141

Coleman, K., & Pierre, P. J. (2014). Assessing anxiety in nonhuman primates. *ILAR journal*, *55*(2), 333–346. https://doi.org/10.1093/ilar/ilu019

Columb, M. O., & Atkinson, M. S. (2015). Statistical analysis: Sample size and power estimations. *BJA Education*, *16*(5), 159–161. https://doi.org/10.1093/bjaed/mkv034

Cooper, M. A., & Bernstein, I. S. (2008). Evaluating dominance styles in Assamese and rhesus macaques. *International Journal of Primatology*, *29*(1), 225–243. https://doi.org/10.1007/s10764-008-9236-y

Crawley, M. J. (2007). *The R book*. John Wiley & Sons Ltd; Chichester.

Crump, A., Arnott, G., & Bethell, E. J. (2018). Affect-driven attention biases as animal welfare indicators: Review and methods. *Animals*, *8*, 136. https://doi.org/10.3390/ani8080136

Cryan, J. F., & Slattery, D. A. (2007). Animal models of mood disorders: Recent developments. *Current Opinion in Psychiatry*, *20*, 1–7. https://doi.org/10.1097/YCO.0b013e3280117733

Dalmau, A., Fabrega, E., & Velarde, A. (2009). Fear assessment in pigs exposed to a novel object test. *Applied Animal Behaviour Science*, *117*(3–4), 173–180. https://doi.org/10.1016/j.applanim.2008.12.014

Dawkins, M. S. (2004). Using behaviour to assess welfare. *Animal Welfare*, *13*, S3–7. https://doi.org/10.1017/S0962728600014317

Desiré, L., Veissier, I., Despres, G., & Boissy, A. (2004). On the way to assess emotions in animals: Do lambs (*Ovis aries*) evaluate an event through its suddenness, novelty, or unpredictability? *Journal of Comparative Psychology*, *118*, 363–374. https://doi.org/10.1037/0735-7036.118.4.363

de Waal, F. B., & Luttrell, L. M. (1985). The formal hierarchy of rhesus macaques: An investigation of the bared-teeth display. *American Journal of Primatology*, *9*(2), 73–85. https://doi.org/10.1002/ajp.1350090202

DiVincenti, L., & Wyatt, J. D. (2011). Pair housing of macaques in research facilities: A science-based review of benefits and risks. *Journal of the American Association for Laboratory Animal Science: JAALAS*, *50*(6), 856–863.

Düpjan, S., Tuchscherer, A., Langbein, J., Schön, P. C., Manteuffel, G., & Puppe, B. (2011). Behavioural and cardiac responses towards conspecific distress calls in domestic pigs (*Sus scrofa*). *Physiology Behaviour*, *103*(5), 445–452. https://doi.org/10.1016/j.physbeh.2011.03.017

Ennaceur, A., & Meliani, K. (1992). A new one-trial test for neurobiological studies of memory in rats III. Spatial vs. non-spatial working memory. *Behaviour Brain Research*, *51*(1), 83–92. https://doi.org/10.1016/s0166-4328(05)80315-8

Fardell, L. L., Pavey, C. R., & Dickman, C. R. (2020). Fear and stressing in predator-prey ecology: Considering the twin stressors of predators and people on mammals. *PeerJ*, *8*, e9104. https://doi.org/10.7717/peerj.9104

Forkman, B., Boissy, A., Meunier-Salaün, M.-C., Canali, E., & Jones, R. B. (2007). A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiology Behaviour*, *92*(3), 340–374. https://doi.org/10.1016/j.physbeh.2007.03.016

Fraser, D., & Matthews, L. R. (1997). Preference and motivation testing. In M. C. Appleby & B. O. Hughes (Eds.), *Animal Welfare* (pp. 159–173). CAB International.

Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Glickman, S. E., & Hartz, K. E. (1964). Exploratory behavior in several species of rodents. *Journal of Professional Nursing: Official Journal of the American Association of Colleges of Nursing*, *51*(1), 83–92. https://doi.org/10.1037/h0047075

Gottlieb, D. H., Coleman, K., & McCowan, B. (2013). The effects of predictability in daily husbandry routines on captive rhesus macaques (*Macaca mulatta*). *Applied Animal Behaviour Science*, *143*(2–4), 117–127. https://doi.org/10.1016/j.applanim.2012.10.010

Hannibal, D. L., Bliss-Moreau, E., Vandeleest, J., McCowan, B., & Capitanio, J. (2017). Laboratory rhesus macaque social housing and social changes: Implications for research. *American Journal of Primatology*, *79*, 1–14. https://doi.org/10.1002/ajp.22528

Harding, E. J., Paul, E. S., & Mendl, M. (2004). Animal behaviour: Cognitive bias and affective state. *Nature*, *427*(6972), 312. https://doi.org/10.1038/427312a

Howarth, E. R. I., Kemp, C., Thatcher, H. R., Szott, I. D., Farningham, D., Witham, C. L., Holmes, A., Semple, S., & Bethell, E. J. (2021). Developing and validating attention bias tools for assessing trait and state affect in animals: A worked example with *Macaca mulatta*. *Applied Animal Behaviour Science*, *234*, 105198. https://doi.org/10.1016/j.applanim.2020.105198

Howarth, E. R. I., Szott, I. D., Witham, C. L., Wilding, C. S., Bethell, E. J., & Singh, A. S. (2023). Genetic polymorphisms in the serotonin, dopamine and opioid pathways influence social attention in rhesus macaques (*Macaca mulatta*). *PLOS ONE*, *18*(8), e0288108. https://doi.org/10.1371/journal.pone.0288108

Jones, N., Sherwen, S. L., Robbins, R., McLelland, D. J., & Whittaker, A. L. (2022). Welfare assessment tools in zoos: From theory to practice. *Veterinary Science*, *9*(4), 170. https://doi.org/10.3390/vetsci9040170

Kappeler, P. M., & Erkert, H. G. (2003). On the move around the clock: Correlates and determinants of cathemeral activity in wild red fronted lemurs (*eulemur fulvus rufus*). *Behavioral Ecology and Sociobiology*, *54*, 359–369. https://doi.org/10.1007/s00265-003-0652-x

Kemp, C., Thatcher, H., Farningham, D., Witham, C., MacLarnon, A., Holmes, A., Semple, S., & Bethell, E. J. (2017). A protocol for training group-housed rhesus macaques (*Macaca mulatta*) to cooperate with husbandry and research procedures using positive reinforcement. *Applied Animal Behaviour Science*, *197*, 90–100. https://doi.org/10.1016/j.applanim.2017.08.006

Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M., & Altman, D. G. (2010). Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLOS Biology*, *8*, e1000412. https://doi.org/10.1371/journal.pbio.1000412

King, H. M., Kurdziel, L. B., Meyer, J. S., & Lacreuse, A. (2012). Effects of testosterone on attention and memory for emotional stimuli in male rhesus monkeys. *Psychoneuroendocrinology*, *37*(3), 396–409. https://doi.org/10.1016/j.psyneuen.2011.07.010

Kirkden, R. D., & Pajor, E. A. (2006). Using preference, motivation and aversion tests to ask scientific questions about animals' feelings. *Applied Animal Behaviour Science*, *100*(1–2), 29–47. https://doi.org/10.1016/j.applanim.2006.04.009

Konečná, M., Weiss, A., Lhota, S., & Wallner, B. (2012). Personality in Barbary macaques (*Macaca sylvanus*): Temporal stability and social rank. *Journal of Research in Personality*, *46*(5), 581–590. https://doi.org/10.1016/j.jrp.2012.06.004

Lee, C., Café, L. M., Robinson, S., Doyle, R. E., Lea, J., Small, A. H., & Colditz, I. G. (2018). Anxiety influences attention bias but not flight speed and crush score in beef cattle. *Applied Animal Behaviour Science*, *205*, 210–215. https://doi.org/10.1016/j.applanim.2017.11.003

Lee, C., Verbeek, E., Doyle, R., & Bateson, M. (2016). Attention bias to threat indicates anxiety differences in sheep. *Biology Letters*, *12*(6), 20150977. https://doi.org/10.1098/rsbl.2015.0977

Lehner, P. N. (1992). Sampling methods in behavior research. *Poultry Science*, *71*(4), 643–649. https://doi.org/10.3382/ps.0710643

Lensink, B. J., Van Reenen, C. G., Engel, B., Rodenburg, T. B., & Veissier, I. (2003). Repeatability and reliability of an approach test to determine calves' responsiveness to humans: "a brief report". *Applied Animal Behaviour Science*, *83*(4), 325–330. https://doi.org/10.1016/S0168-1591(03)00138-2

Luo, L., Reimert, I., de Haas, E. N., Kemp, B., & Bolhuis, J. E. (2019). Effects of early and later life environmental enrichment and personality on attention bias in pigs (*Sus scrofa domesticus*). *Animal Cognition*, *22*, 959–972. https://doi.org/10.1007/s10071-019-01287-w

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*(1), 15–20. https://doi.org/10.1037/0021-843X.95.1.15

Maestripieri, D., & Wallen, K. (1997). Affiliative and submissive communication in rhesus macaques. *Primates*, *38*, 127–138. https://doi.org/10.1007/BF02382003

Mendl, M., Burman, O. H. P., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B*, *277*, 2895–2904. https://doi.org/10.1098/rspb.2010.0303

Monk, J. E., Campbell, D. L. M., Lee, C., & Tilbrook, A. (2023). Future application of an attention bias test to assess affective states in sheep. *Animal Production Science*, *63*(6), 523–534. https://doi.org/10.1071/AN22260

NC3Rs. (2006). *Primate accommodation, care and use*. NC3Rs: London. https://www.nc3rs.org.uk/sites/default/files/documents/Guidelines/NC3Rs%20guidelines%20-%20Non-human%20primate%20accomodation%20care%20and%20use.pdf:

NC3Rs. (2015). *The macaque website, national centre for the replacement, refinement and reduction of animals in research*. Retrieved March 5, 2023, from http://www.nc3rs.org.uk/macaques/

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

Parker, R. M., Paul, E. S., Burman, O. H., Browne, W. J., & Mendl, M. (2014). Housing conditions affect rat responses to two types of ambiguity in a reward-reward discrimination cognitive bias task. *Behaviour Brain Research*, *274*, 73–83. https://doi.org/10.1016/j.bbr.2014.07.048

Plant, T. M. (1981). Time courses of concentrations of circulating gonadotropin, prolactin, testosterone, and cortisol in adult male rhesus monkeys (*Macaca mulatta*) throughout the 24 h light-dark cycle. *Biology of Reproduction*, *25*, 244–252. https://doi.org/10.1095/biolreprod25.2.244

Prescott, M. J., & Buchanan-Smith, H. M. (2007). Training laboratory-housed non-human primates, part 1: A UK survey. *Animal Welfare*, *16*(1), 21–36. https://doi.org/10.1017/S096272860003089X

Prescott, M. J., Nixon, M. E., Farningham, D. A. H., Naiken, S., & Griffiths, M.-A. (2012). Laboratory macaques: When to wean? *Applied Animal Behaviour Science*, *137*(3–4), 194–207. https://doi.org/10.1016/j.applanim.2011.11.001

Raoult, C. M. C., & Gygax, L. (2019). Mood induction alters attention toward negative-positive stimulus pairs in sheep. *Scientific Reports*, *9*, 7759. https://doi.org/10.1038/s41598-019-44330-z

R Core Team. (2023). *R: A language and environment for statistical computing*. https://www.R-project.org/

Richmond, S. E., Wemelsfelder, F., de Heredia, I. B., Ruiz, R., Canali, E., & Dwyer, C. M. (2017). Evaluation of animal-based indicators to be used in a welfare assessment protocol for sheep. *Frontiers in Veterinary Science*, *4*, 210. https://doi.org/10.3389/fvets.2017.00210

Romeyer, A., & Bouissou, M. F. (1992). Assessment of fear reactions in domestic sheep, and influence of breed and rearing conditions. *Applied Animal Behavior Science*, *34*(1–2), 93–119. https://doi.org/10.1016/S0168-1591(05)80060-7

Sherwen, S., Hemsworth, L., Beausoleil, N., Embury, A., & Mellor, D. (2018). An animal welfare risk assessment process for zoos. *Animals*, *8*, 130. https://doi.org/10.3390/ani8080130

Shively, C. A., & Day, S. M. (2015). Social inequalities in health in nonhuman primates. *Neurobiology of Stress*, *1*, 156–163. https://doi.org/10.1016/j.ynstr.2014.11.005

Statham, P., Hannuna, S., Jones, S., Campbell, N., Robert Colborne, G., Browne, W. J., Paul, E. S., & Mendl, M. (2020). Quantifying defence cascade responses as indicators of pig affect and welfare using computer vision methods. *Scientific Reports*, *10*(1), 8933. https://doi.org/10.1038/s41598-020-65954-6

Taffet, G. E. (2024). Physiology of Aging. In M. R. Wasserman, D. Bakerjian, S. Linnebur, S. Brangman, M. Cesari & S. Rosen (Eds.), *Geriatric Medicine*. Spinger. https://doi.org/10.1007/978-3-030-74720-6_103

Tan, S. H., & Tan, C. B. (2010). The correct interpretation of confidence intervals. *Proceeding of Singapore Healthcare*, *19*(3), 276–278. https://doi.org/10.1177/201010581001900316

Temple, D., Manteca, X., Dalmau, A., & Velarde, A. (2013). Assessment of test-retest reliability of animal-based measures on growing pig farms. *Livestock Science*, *151*(1), 35–45. https://doi.org/10.1016/j.livsci.2012.10.012

Tiemann, I., Fijn, L. B., Bagaria, M., Langen, E. M. A., van der Staay, F. J., Arndt, S. S., Leenaars, C., & Goerlich, V. C. (2023). Glucocorticoids in relation to behavior, morphology, and physiology as proxy indicators for the assessment of animal welfare. A systematic mapping review. *Frontiers in Veterinary Science*, *9*, 954607. https://doi.org/10.3389/fvets.2022.954607

Topoleski, J., Schultz, C. A., & Warren, W. G. (2018). Identifying and resolving end of session cues in substance detection canine training. *Frontiers in Veterinary Science*, *6*(5), 206. https://doi.org/10.3389/fvets.2018.00206

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

Verbeek, E., Ferguson, D., & Lee, C. (2014). Are hungry sheep more pessimistic? The effects of food restriction on cognitive bias and the involvement of ghrelin in its regulation. *Physiology Behaviour*, *123*, 67–75. https://doi.org/10.1016/j.physbeh.2013.09.017

Vincent, S. B. (1915). The white rat and the maze problem: The introduction of a tactual control. *Journal of Animal Behaviour*, *5*(3), 175–184. https://doi.org/10.1037/h0075779

Vögeli, S., Lutz, J., Wolf, M., Wechsler, B., & Gygax, L. (2014). Valence of physical stimuli, not housing conditions, affects behaviour and frontal cortical brain activity in sheep. *Behavioural Brain Research*, *267*, 144–155. https://doi.org/10.1016/j.bbr.2014.03.036

Waitt, C., & Buchanan-Smith, H. M. (2006). Perceptual considerations in the use of colored photographic and video stimuli to study nonhuman primate behavior. *American Journal of Primatology*, *68*, 1054–1067. https://doi.org/10.1002/ajp.20303

Watters, J. V., & Krebs, B. L. (2019). Assessing and enhancing the welfare of animals with equivocal and reliable cues. *Animals (Basel)*, *9*(9), 680. https://doi.org/10.3390/ani9090680

Watters, J. V., Krebs, B. L., & Eschmann, C. L. (2021). Assessing animal welfare with behavior: Onward with caution. *Journal of Zoological and Botanical Garden*, *2*, 75–87. https://doi.org/10.3390/jzbg2010006

Webster, J. (2005). The assessment and implementation of animal welfare: Theory into practice. *Revue Scientifique Technology Office international Epizooties*, *24*(2), 723–734. https://doi.org/10.20506/rst.24.2.1602

Welfare Quality. (2018). *Assessment protocols*. https://www.welfarequalitynetwork.net/en-us/reports/assessment-protocols/

Whay, H. R., Main, D. C. J., Green, L. E., & Webster, A. J. F. (2003). Animal-based measures for the assessment of welfare state of dairy cattle, pigs and laying hens: Consensus of expert opinion. *Animal Welfare*, *12*(2), 205–217. https://doi.org/10.1017/S0962728600025641

Witham, C., & Bethell, E. J. (2019). Macaque Faces. *Figshare*. https://doi.org/10.6084/m9.figshare.9862586

Witham, C. L. (2015). Automated face recognition of rhesus macaques. *Journal of Neuroscience Methods*, *300*, 157–165. https://doi.org/10.1016/j.jneumeth.2017.07.020

Wolfensohn, S., & Honess, P. (Eds.). (2005). Chapter 6. In *Handbook of Primate Husbandry and Welfare* (1st ed, pp. 111). Blackwell Publishing Ltd.

Yeates, J. W., & Main, D. C. J. (2008). Assessment of positive welfare: A review. *Veterinary Journal (London, England: 1997)*, *175*(3), 292–300. https://doi.org/10.1016/j.tvjl.2007.05.009

Zanette, L. Y., Hobbs, E. C., Witterick, L. E., MacDougall-Shackleton, S. A., & Clinchy, M. (2019). Predator-induced fear causes PTSD-like changes in the brains and behaviour of wild animals. *Scientific Reports*, *9*, 11474. https://doi.org/10.1038/s41598-019-47684-6

Zhang, Z., Dong, X., Liu, Z., & Liu, N. (2024). Social status predicts physiological and behavioral responses to chronic stress in rhesus monkeys. *iScience*, *27*(6), 110073. https://doi.org/10.1016/j.isci.2024.110073