Research Article

Opportunities for Refinement in Neuroscience: Indicators of Wellness and Post-Operative Pain in Laboratory Macaques

Kris A. Descovich[1,2,3], Susan E. Richmond[1,4], Matthew C. Leach[5], Hannah M. Buchanan-Smith[1], Paul Flecknell[6], David A. H. Farningham[7], Claire Witham[6,7], M. Carolyn Gates[8] and Sarah-Jane Vick[1]

[1]Psychology, Faculty of Natural Sciences, University of Stirling, Stirling, United Kingdom; [2]Environmental and Animal Sciences, Unitec Institute of Technology, Auckland, New Zealand; [3]Centre for Animal Welfare and Ethics, School of Veterinary Science, University of Queensland, Gatton, Australia; [4]Humane Slaughter Association, The Old School, Brewhouse Hill, Wheathampsted, United Kingdom; [5]School of Natural and Environmental Science, Newcastle University, Newcastle upon Tyne, United Kingdom; [6]Comparative Biology Centre, Newcastle University, Newcastle upon Tyne, United Kingdom; [7]Centre for Macaques, Medical Research Council, Salisbury, United Kingdom; [8]School of Veterinary Science, Massey University, Palmerston North, New Zealand

Abstract

Being able to assess pain in nonhuman primates undergoing biomedical procedures is important for preventing and alleviating pain, and for developing better guidelines to minimize the impacts of research on welfare in line with the 3Rs principle of Refinement. Nonhuman primates are routinely used biomedical models; however, it remains challenging to recognize negative states, including pain, in these animals. This study aimed to identify behavioral and facial changes that could be used as pain or general wellness indicators in the rhesus macaque (*Macaca mulatta*). Thirty-six macaques scheduled for planned neuroscience procedures were opportunistically monitored at four times: Pre-Operative (PreOp), Post-Operative (PostOp) once the effects of anesthesia had dissipated, Pre-Analgesia (PreAn) on the subsequent morning prior to repeating routine analgesic treatment, and Post-Analgesia (PostAn) following administration of analgesia. Pain states were expected to be absent in PreOp, moderate in PreAn, and mild or absent in PostOp and PostAn when analgesia had been administered. Three potential pain indicators were identified: lip tightening and chewing, which were most likely to occur in PreAn, and running, which was least likely in PreAn. Arboreal behavior indicated general wellness, while half-closed eyes, leaning of the head, or body shaking indicated the opposite. Despite considerable individual variation, behavior and facial expressions could offer important indicators of pain and wellness. They should be routinely quantified and appropriate interventions applied to prevent or alleviate pain and promote positive welfare.

1 Introduction

Although primate use is a small proportion of total animal use in bioscience research, these animals are routinely subjected to procedures or conditions that directly and indirectly affect their welfare (e.g.,, Capitanio et al., 1996; Balcombe et al., 2004; Carlsson et al., 2004; Rennie and Buchanan-Smith, 2006a,b,c; Olsson and Westlund, 2007; Wolfensohn and Lloyd, 2013). It is a societal expectation that animals used in bioscience experience good welfare, which is characterized by an absence of unnecessary suffering (Lund et al., 2012; Leaman et al., 2014). Such considerations are also important for scientific validity as poor animal welfare may confound experimental results and affect the translation of scientific findings to human health benefits (Poole, 1997; Würbel, 2001; Tasker, 2012; Everds et al., 2013; Hall et al., 2015; Sneddon, 2017). To address these issues, animal use is guided in legislation and policy by the “3Rs” principles of Replacement, Reduction and Refinement, which aim to replace sentient animals with non-sentient alternatives, reduce the number of affected animals, and minimize the impact of experimental procedures and promote welfare when non-animal alternatives are not available, respectively (Russell et al., 1992; Osborne et al., 2009; Prescott et al., 2010).

In the 3Rs framework, Refinement is defined as “any approach which avoids or minimizes the actual or potential pain, distress and other adverse effects experienced at any time during the life of the animals involved and which enhances their well-being” (Buchanan-Smith et al., 2005, pp. 379-380). The relative impact of some experimental protocols is species-dependent, for example, social isolation (Dawkins, 2006a; Rennie and Buchanan-Smith,

Received November 6, 2018; Accepted January 24, 2019;
Epub March 29, 2019; © The Authors, 2019.

ALTEX 36(4), 535-554. doi:10.14573/altex.1811061

Correspondence: Kris Descovich, PhD,
Centre for Animal Welfare and Ethics,
University of Queensland, Gatton, Australia
(k.descovich1@uq.edu.au)

2006b,c), however others, such as the experience of pain, are widely relevant to all sentient species (Bateson, 1991; Carstens and Moberg, 2000; Sneddon et al., 2014; Mellor and Beausoleil, 2015). Pain can be defined as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" (Loeser and Treede, 2008, p. 475), and is a particularly pertinent issue in biomedical science (Hawkins, 2002; Stokes et al., 2009). Research in both humans and animals suggests that identification of pain states can be challenging as individuals vary in pain sensitivity, tolerance or expression due to life history, social connectedness, sex, health, genotype, and temperament (Mogil and Kest, 1999; Mogil et al., 2000; DeWall and Baumeister, 2006; Defrin et al., 2017; Lush and Ijichi, 2018). The alleviation of pain should be based on the needs of the individual in terms of frequency or dosage (Roughan and Flecknell, 2002; Pham et al., 2010), and Refinement protocols should consider the potential for minimizing undesirable side effects from analgesia administration (Fleming and Coombs, 1992; Cooper et al., 2009; Schaap et al., 2012). In humans, pain assessment generally incorporates eliciting verbal feedback from the patient about their experience of pain (e.g., Jensen et al., 1986), unless the patient is not able to communicate effectively (e.g., pre-verbal infants: Taddio et al., 2009). However, animal pain levels are typically assessed by an observer, resulting in two sources of variation, within-patient and within-observer (Morton, 2000), and therefore creates additional challenges in pain assessment.

Behavioral observation is a useful, non-invasive technique for the identification of welfare states in animals (Dawkins, 2006b; Fraser, 2009). A normal behavioral baseline for a species or individual can be assumed to reflect wellness (e.g., Lambeth et al., 2013) and deviation may indicate compromised welfare. In this context we use "wellness" to indicate a state of being in good health, and as an antonym of illness. Ideally, specific negative states (e.g., pain, nausea, fear) should be identifiable so that appropriate treatment (e.g., analgesia/anti-emetic/environmental adjustments) can be implemented (Morton and Griffiths, 1985; Carstens and Moberg, 2000; Roughan and Flecknell, 2002; Mellor and Beausoleil, 2015; Sneddon, 2017). Unfortunately, validated behavioral indicators of specific negative states are lacking for many species, and assessment is reliant on subjective criteria (Carstens and Moberg, 2000; Honess and Wolfensohn, 2010; Wolfensohn and Lloyd, 2013). Despite the widespread use of nonhuman primate (NHP) species in translational bioscience, evidence-based guidelines for distinguishing pain and illness from other negative states are lacking. Wolfensohn and Honess (2005, p. 60) and the National Research Council Committee (NRCC) Guidelines (2009, p. 50 and 57) suggest that pain behavior in NHPs is expressed through a "miserable appearance", huddling, "sad" or contorted facial expressions, moaning or grunting, teeth clenching, restlessness, eye rolling and shaking. Other general pain behaviors may be wincing, vocalization, difficulty in movement (Lambeth et al., 2013), hunching or arm crossing (Morton and Griffiths, 1985). Grooming, social interaction, eating and drinking behaviors may also decrease (Wolfensohn and Honess, 2005; NRCC, 2009), however any deviation from an individual's normal behavioral repertoire could indicate pain or distress (Hawkins, 2002; Wolfensohn and Lloyd, 2013). In female olive baboons (*Papio anubis*), general activity decreases after abdominal surgery, as does standing and vigilance, while locomotion and foraging seem unaffected (Allison et al., 2007), although this may not be a pain-specific response as all animals were provided with pain relief. These indicators provide a useful starting point for pain assessment, but more objective experimental data on NHP pain expression is required and across a wider range of commonly used laboratory species. Ideally, specific indicators should be identified and developed for each protocol that experimental animals are likely to experience (Morton, 2000), and humane endpoints employed and re-evaluated regularly (Hawkins, 2002).

Some pain reactions are likely to be automatic protective responses (Sneddon et al., 2014), but the expression of pain may benefit individuals by recruiting assistance from conspecifics (Langford et al., 2006; de Waal, 2008) although empirical evidence that it serves this function is limited. However, many animal species including primates are thought to "hide" their pain as a survival-enhancing strategy, making pain identification challenging (Plesker and Mayer, 2008; Murdoch et al., 2013; Fenwick et al., 2014; Gaither et al., 2014). Furthermore, distinguishing chronic pain presents an even greater challenge than for acute pain because it can be difficult to obtain a pain-free baseline for comparison (Brearley and Brearley, 2000), and there is the potential that other associated negative affective states (e.g., depression) will trigger or amplify the pain experience and diminish pain recognition and treatment (Bair et al., 2003). Assessment of animal welfare would therefore be improved by the development of sufficiently specific and sensitive indicators of pain and other negative affective states (Wolfensohn and Lloyd, 2013).

Recently, several studies have identified facial changes that indicate pain in a variety of mammalian species including mice (Langford et al., 2010), rats (Sotocinal et al., 2011), horses (Dalla Costa et al., 2014), and sheep (McLennan et al., 2016). These "grimace scales" assess the presence and intensity of pain, and are practical to use because they rely on scoring position changes of only a few facial features, typically the ears, eyes, cheeks and nose (see Descovich et al., 2017 for a review), and exploit our strong attention bias towards animal faces relative to other body regions (Leach et al., 2011). This approach has not been applied to NHP species, although primate facial expressions have been well-studied in the context of social communication (Chevalier-Skolnikoff, 1973; Partan, 2002; Ghazanfar and Logothetis, 2003; Waller and Micheletta, 2013) and facial expressions may provide insight into internal states in animals (Descovich et al., 2017).

Pain-related facial changes are widely evident in humans, present even in premature neonates although they are modified during development (Craig et al., 1993, 2001; Johnston et al., 1993). The human pain face is typically characterized by lowered brows, tightening of the eyes, raising of the cheeks, nose wrinkling, upper lip raising and horizontal stretching of the mouth (LeResche, 1982; Craig and Patrick, 1985; Prkachin, 2009) although pain type and individual variation affects expression of some elements (Prkachin, 1992; Prkachin and Solomon, 2009). Human facial expressions are commonly measured using the Facial Action Coding System (FACS), which records observable movements of the underlying facial musculature as "Action Units" (AUs) (Ekman and Friesen, 1978). Macaques share a similar facial muscle structure to humans (Burrows et al., 2009) and the FACS method has been adapted for rhesus macaques (*Macaca mulatta*), providing a comparable tool to examine facial movement (MaqFACS; Parr et al., 2010).

One key challenge in the study of negative animal welfare states is that experimental conditions must, by definition, induce or invoke these states. However, opportunistic sampling of animals undergoing planned experimental protocols, or "animal sharing", is consistent with Reduction within the 3Rs framework (e.g., Walker and Srinivas, 2013) and is desirable from an ethical standpoint because it offers benefits without causing additional harms. While this approach limits control over the experimental design and allows the influence of potential confounders, there is an important advantage in that data can be collected in contexts that reflect the actual severity or specificity of existing biomedical protocols. Pain elicited during standard analgesiometric tests is not comparable to post-operative pain as distinct nociceptive pathways are implicated (Roughan and Flecknell, 2002), which suggests that pain responses resulting from the first pain type may not be relevant to detection in the second. NHPs are used as models in neuroscience research where they undergo acute procedures such as cranial implants (Niekrasz and Wardrip, 2012; Azimi et al., 2016). These protocols are likely to cause significant pain and discomfort and it would be ethically unacceptable to undertake similar procedures solely to examine their effects on welfare, or to withhold analgesia in such cases to isolate pain responses. Therefore, opportunistic observation of animals before and after planned experimental protocols allows behavior to be measured under applied conditions.

The aim of this study was to identify general and facial behavioral changes that occur with pain states, and those that may indicate general wellness (or a deviation from wellness) in rhesus macaques, one of the most commonly used primate species in bioscience (Carlsson et al., 2004). Sampling was opportunistic, therefore, in addition to time period, the potential impact of several other variables was also examined, including sex, age, the severity classification of the procedure, and the interval since analgesia administration (to the start of the post-analgesia observation period). Finally, while this study monitors the impact of a single event, an individual's experience of previous procedures (e.g., Lutz et al., 2003) and the presence of any indicators of illness prior to the procedure were also considered as factors that may indicate longer term effects related to cumulative severity (Honess and Wolfensohn, 2010).

2 Methods

Subjects and housing

This study was conducted between 2010 and 2014 and included 36 rhesus macaques (*Macaca mulatta*) (22 male, 14 female) undergoing planned neuroscience procedures. They were aged between 4 and 13 years (Tab. 1 and Tab. S1[1]). This study was conducted opportunistically, and the number of animals used was therefore directly dependent on other unrelated studies within the collaborating facility, which determined both the number of animals that could be included and procedures that could be studied. Animals were originally obtained from a national (UK) breeding facility at a mean of 2.3 years prior to the study. Animal housing was in accordance with UK Home Office requirements (2014) with all being group housed with 1-9 other individuals, except for two adult males who were housed singly because they were incompatible with others. The facility underwent a major refurbishment during data collection and the majority of subjects (n = 22) were observed in the new accommodation. Space allowance per macaque was similar, but in the older facilities, animals were housed in 3 separate holding areas, (up to 10-12 animals per area), whereas the new accommodation housed up to 50-60 animals in a single larger area. Enclosures allowed a minimum floor area of 4.40 m^2 for each individual or pair of animals. All enclosures incorporated vertical space using raised platforms, and environmental enrichment, including a wood shavings substrate (Eco-pure, Datesand, Manchester, UK) for bedding and foraging, was available. The macaques were provided with appropriate nutrition and daily forage including commercially prepared food (Mazuri Primates Extended, Banana Chunks and Trio Munch Grains from Special Diets Services, Witham, U.K. LBS Biotechnology, U.K.) supplemented with forage mix and fruit. Water was available *ad libitum* unless restricted for other studies. The light-dark cycle at this facility was 12 h:12 h and the temperature was maintained at 22°C, with relative humidity at 24%.

Tab. 1: Rhesus macaque demographic information

Animal	Age (years)	Ill (Y/N)	Year of procedure	Time at facility (days)	Procedure	UK Home Office severity classification	Analgesia	Post-analgesia interval (min)	First procedure (Y/N)	Time since last procedure (days)	Face coded (Y/N)
F1	4	N	2010	261	CRI	Moderate	MET	60	N	22	N
F2	13	N	2011	266	EMG	Moderate	MTD	60	N	147	N
F3	5	N	2011	577	EMG	Moderate	BUP	60	N	59	N
F4	7	N	2010	468	EMG	Moderate	MET	60	N	12	Y
F5	9	N	2014	371	CRI	Moderate	MET	60	N	18	Y

---

[1] doi:10.14573/altex.1811061s

Animal	Age (years)	Ill (Y/N)	Year of procedure	Time at facility (days)	Procedure	UK Home Office severity classification	Analgesia	Post-analgesia interval (min)	First procedure (Y/N)	Time since last procedure (days)	Face coded (Y/N)
F6	5	Y	2010	957	EMG	Moderate	BUP	60	N	42	Y
F7	10	N	2010	496	CRI	Moderate	MET	60	N	56	Y
F8	8	N	2015	79	OPT	Moderate	MET	60	Y		Y
F9	4	N	2011	307	CRI	Moderate	MTD	60	N	22	N
F10	4	N	2011	272	EMGCRI	Moderate	MTD	60	Y		N
F11	6	Y	2014	707	CRI	Moderate	MTC	60	N	167	Y
F12	6	N	2014	734	EMG	Moderate	MET	60	N	36	Y
F13	4	N	2014	655	EMG	Moderate	MET	30	Y		Y
F14	4	Y	2014	472	EMG	Moderate	MET	30	N	13	Y
M1	12	N	2010	3016	SCP	Low	MET	60	N	137	N
M2	13	N	2010	3442	SCP	Low	MET	60	N	97	N
M3	10	Y	2014	2599	SCP	Low	MTD	30	N	86	Y
M4	8	N	2010	1684	SCP	Low	MET	60	N	71	Y
M5	4	N	2014	35	OPT	Moderate	MMD	60	Y		Y
M6	8	N	2010	1707	SCP	Low	MET	60	N	26	Y
M7	4	N	2014	36	OPT	Moderate	BMD	60	Y		Y
M8	7	N	2010	1085	CRI	Moderate	MET	60	N	157	N
M9	4	N	2014	46	OPT	Moderate	BMD	60	Y		Y
M10	9	N	2014	2098	CRI	Moderate	BMD	60	N	189	Y
M11	4	N	2014	38	OPT	Moderate	MMD	60	Y		Y
M12	6	N	2014	973	MAR	Low	MET	30	N	318	Y
M13	4	N	2014	44	OPT	Moderate	MMD	60	Y		Y
M14	4	N	2014	43	OPT	Moderate	BMD	60	Y		Y
M15	4	N	2014	9	OPT	Moderate	MMD	60	Y		Y
M16	4	N	2014	18	OPT	Moderate	BMD	60	Y		Y
M17	9	Y	2011	2030	SCP	Low	MET	60	N	73	N
M18	8	N	2014	1783	SCP	Low	MET	30	N	19	Y
M19	4	N	2014	65	OPT	Moderate	BMD	60	Y		Y
M20	4	N	2014	7	OPT	Moderate	BMD	60	Y		Y
M21	5	N	2014	2231	CRI	Moderate	BUPM	60	N	91	Y
M22	9	N	2014	2161	SCP	Low	MET	60	N	154	Y

Animal: F, female; M, male. *Age:* Macaque age at the time of the procedure. *Ill:* Indicates whether the macaque was considered potentially ill at the time of the procedure. *Year of procedure:* The year the procedure was conducted. *Time at facility:* Number of days from arrival at the facility until the procedure. *Procedure:* The procedure type undertaken. CRI, cranial implant; EMG, electromyography wire implant; OPT, optogenetic surgery; SCP, cleaning of an existing cranial implant; MAR, repair of the surgical margins of a cranial implant; EMGCRI, electromyography wire implant + cranial implant. *UK Home Office severity classification:* Designated severity classification by the UK Home Office. *Analgesia:* Pain relief medication given in the post-procedure period. MET, meloxicam; MTD, meloxicam + dexamethasone; BUP, buprenorphine; MTC, meloxicam + dexamethasone + carbamazepine; MMD, meloxicam + dexamethasone + methadone; BMD, meloxicam + dexamethasone + buprenorphine; BUPM, meloxicam + buprenorphine. *Postanalgesia interval:* Indicates whether filming of the postanalgesia period occurred 30 or 60 minutes after receiving analgesia. *First procedure:* Indicates whether this was the macaque's first procedure (Yes/No). *Time since last procedure:* Number of days since the macaque's previous procedure. If this was the macaque's first procedure this is blank. *Face coded:* Indicates whether facial movements were coded using MaqFACS (Yes/No).

Ethical considerations

No regulated procedures were performed for the purposes of this study, and no veterinary procedures were delayed, omitted or modified from standard protocols. All procedures were performed under UK Home Office license and peri-operative protocols conformed with their Guidance on the Operation of the Animals (Scientific Procedures) Act, 1986 incorporating European Directive 2010/63/EU. All surgical protocols were approved by the Newcastle University Animal Welfare and Ethical Review Committee; animals received appropriate anesthesia/sedation and analgesia according to the approved experimental protocols for the relevant project, with analgesia drug and dosage determined by an experienced laboratory veterinarian (Tab. S1[1]). Behavioral data collection methods were approved by the Psychology Ethics Committee, University of Stirling and adhered to the guidelines for the treatment of animals in behavioral research and teaching (ASAB, 2018).

Experimental protocol

The behavior of rhesus macaques undergoing planned surgical neuroscience research procedures was recorded. Procedures were prospectively categorized by the UK Home Office as moderate (n = 27) or mild in severity (n = 9). Moderate procedures included cranial implants (CRI), electromyography wire implants (EMG) and optogenetic surgery (OPT). Mild procedures included implant maintenance and repair (SCP and MAR). All animals received appropriate anesthesia/sedation and analgesia according to the approved experimental protocols for the relevant project, with analgesia drug and dosage determined by an experienced laboratory veterinarian. The influence of anesthesia and analgesia were not included in the statistical analysis due to the complexity and variation in drug protocols used.

Video footage of each macaque was collected during four periods: i) PreOp: on a day preceding the procedure when the animal was expected to experience no pain and to be performing normal behaviors for that individual, ii) PostOp: on the day of the procedure when the animal had recovered sufficiently from the anesthesia to move normally around the cage. All animals in this period received analgesia during the procedure, iii) PreAn: on the morning following the procedure directly before the animal received routine administration of analgesia, and iv) PostAn: on the same morning as the PreAn period, and 60 minutes following a routine administration of analgesia, except for a small subset (n = 5) filmed 30 minutes following a routine administration of analgesia. The PreOp period represented the baseline for each individual, and wellness indicators were defined as behaviors that were different in the pre-operative period when compared to all others. The PreAn period was considered the most informative for isolating pain responses in terms of both likelihood and intensity; PreOp was assumed to be a pain-free baseline, while analgesia was administered for both PostOp and PostAn and pain was assumed to be reduced or absent. Anesthesia and surgery, however well they are conducted, will have a range of non-specific effects potentially including but not limited to nausea, fatigue, disorientation, and metabolic or endocrine effects due to the surgical stress response. Therefore, behaviors that either peak or trough during PreAn and are significantly different from other periods were considered to have the most potential as pain indicators.

Filming was undertaken in the same location for all periods. Most macaques were filmed in a recovery cage directly adjacent to their home enclosure although a few (n = 3) were filmed in a restricted section of similar size in their home enclosure. Animals were alone in their enclosure during all filming sessions but could maintain vocal contact and limited visual contact with cage-mates or other conspecifics. Two camcorders recorded simultaneously; one captured the entire enclosure and the second was mounted on a motorized pan/tilt tripod head (CamRanger MP-360) and was maneuvered by remote control to zoom in on the macaque's face. Various camera models were used (e.g., Canon Legria HF M52; Sony Handycam) with any analogue footage converted into a digital format using Mediacruise (Canopus Co. Ltd.).

Twenty-five minutes of macaque footage was recorded for each period but the first five minutes were excluded from analysis to allow the camera operator to move out of view to minimize observer effects (Iredale et al., 2010; Peterson et al., 2017). Although the camera operator remained out of sight of the macaque, familiar facility staff were in the general vicinity and intermittently in direct view. Experimenter bias (Kilkenny et al., 2013) was minimized by relabeling videos with blind labels and by coding in a randomized video order, however fully blinded coding was not feasible as the effects of some surgical procedures (e.g., shaved patches of hair) were visible.

Behavioral recording

Behavior was coded according to a catalogue of defined behaviors (Tab. 2) with long duration behaviors (e.g., walking) recorded continuously and short, discrete behaviors (e.g., vocalizations) counted on each occurrence. One experienced observer coded all macaque behavior; however, inter-observer reliability (IOR) was checked against a second trained observer to ensure consistency. Both had more than one year's experience in the observation of animal behavior including primates. Observers watched four 20-min videos of different macaques and coded the behavior at 30 s intervals. The target percentage of agreement between observers was 80% and IOR test was 73%. Behavior definitions were discussed and clarified prior to coding different videos for which agreement reached 89%. Facial expressions were recorded by two MaqFACS accredited coders (Parr et al., 2010). Seven macaques were excluded from MaqFACS coding as facial movements were not clear enough in at least one of their videos. MaqFACS facial "action units" were recorded continuously from the video with the exception of blinking and vocalizing, which were counted as events (Tab. 2). Due to the complexity of measuring facial movements, each face video was coded twice – once for the upper face and once for the lower face – and playback speed was reduced to between 0.1x and 0.2x to capture the activation and deactivation of muscles as they occurred. Due to the location of surgeries and head movements, ear positions/movements were excluded from coding as they could not be reliably differentiated. All coding was completed using Cowlog (Hanninen and Pastell, 2009) with output files processed in Excel (Microsoft® Excel for Mac 2017 version) prior to analyses.

Tab. 2: Catalogue of macaque behavior recorded
(including behaviors that had no significant effects when analyzed or were too infrequent for analysis)

Locomotion	**Type**	**Descriptions**
Quadrupedal walk[a]	State	Walk on 4 limbs
Bipedal walk[a]	State	Walk on 2 legs
Quadrupedal run[b]	State	Run on 4 limbs
Bipedal run[b]	State	Run on 2 legs
Climb[c]	State	Vertical climbing
Descend[c]	State	Vertical descending
Hang[c]	State	Hanging underneath an elevated object or off the bars, including when standing on bars
Pace	State	Repetitive pacing or circling; must complete 2 lots of the pattern, e.g., 2 circles or 2 back and forth to be coded as pacing
Posture/Action	**Type**	**Descriptions**
Crouch	State	Crouching in a low position
Lean head	State	Head is resting on bars, wall, floor of cage or arm.
Lying	State	Animal is lying on their back, side or stomach.
Sit up	State	Sitting upright
Sit hunched	State	Sitting with shoulders slumped; back may be curved and head is often lower than the shoulders or leaning on something, face may be oriented towards the floor, and chest may rest on the knees.
Stand four limbs[d]	State	Standing with weight resting on 3-4 limbs
Stand two limbs[d]	State	Standing on legs only
Body rock	State	Animal is sitting with arms on cage wall or bar and rocks torso back and forth.
Body jerk	Event	Clear body jerk or spasm
Body shake	Event	A shaking movement of the head and body
Head jerk	Event	A jerking or circling like movement of the head; appears unrelated to context and can be repetitive
Pull head piece	Event	Pulling at the cranial implant
Hand shake	Event	Snapping or shaking of the hand
Hair tug	Event	Pulling or plucking at hair
Jump	Event	Jumping movement
Rub face	Event	Animal uses the hand to rub the face in a non-scratching movement (without nails), animal may also rub face against the bars.
Scratch	Event	Use of the hand or foot to scratch the surface of the skin; attention of the animal may or may not be on the area of scratching.
Stretch	Event	Stretching or arching of the back, often by holding on to the top bars
Shiver	Event	Shivering occurs in rhythmic movements where the shoulders are contracted in towards the neck.
Touch wound	Event	Touching an area of the body where the integument has been damaged (associated state behavior is grooming).
Vigilance	**Type**	**Descriptions**
Cage monitor	State	Animal visually scans its cage environment.
General vigilance	State	Animal is engaged in watchful behavior that surveys the general environment.
Focused vigilance	State	Animal is engaged in watchful behavior of a particular object, location or event.

Activity	Type	Descriptions
Affiliation	State	Friendly interaction with conspecific within visual contact
Aggression	State	Attack, threaten or chase a conspecific within visual contact
Present rear	State	Standing on all fours presenting rear to conspecifics
Groom	State	Self-grooming using hands or by licking; grooming is also coded when they are touching/examining their skin/coat/wound.
Drink	State	Animal is consuming water.
Eating and foraging	State	Animal is searching for and consuming food.
Manipulate object	State	Animal is using hands or mouth to investigate and move an inanimate, moveable object in the environment.
Manipulate cage	State	Animal is using hands or mouth to pull or grab parts of the enclosure such as padlocks, adjustable panels and cage dividers.
Cage shake	Event	Vigorous shaking of cage or aggressive bouncing off the cage bars or wall
Masturbate	State	Stimulating own genitals, usually with the hands or mouth
Oral	State	Licking or chewing on a non-food object
Out of sight	State	The animal is not in view or behavior is not clear enough to reliably score.
Face coding	**Type**	**MaqFACS description, Action Unit (AU) label and musculature included (Parr et al., 2010)**
Eyebrow raise	State	The brow line is raised (AU1+2: *Frontalis muscle*).
Lower glabella	State	Brow lowering evident as medial bulging in the glabellar region (AU41: *Procerus*).
Cheek raise	State	The cheeks are raised so that the area around the eyes cinches inwards, including the upper and lower lids, and producing movement around the brows (AU6: *Orbicularis oculi*).
Close eyes	State	Eyelids are completely closed (AU43: *Relaxation of Levator palpebrae superioris; Orbicularis oculi, pars palpebralis).*
Half-close eyes	State	Eyelids are partially closed (H43).
Blink	Event	Rapid closing and opening of the eyelids (AU45)
Lips towards each other	State	The lips move towards each other and may appear flattened against the gums (AU8: *Orbicularis oris*).
Nose wrinkle + upper lip raise	State	The nose is pulled upwards, causing wrinkling and raising the nostril wings, in combination with an upper lip raise (AU9+AU10: *Levator labii superioris and Levator labii superioris alaeque nasi*).
Upper lip raise	State	Upper lip is pulled upwards in smooth arc to reveal the teeth, stronger actions may reveal the upper gums (AU10: *Levator labii superioris*).
Lip corner pull	State	The mouth corners are pulled obliquely upwards and backwards towards the ears (AU12: *Zygomatic major*).
Lower lip depress	State	The lower lip is pulled downwards in a smooth curve exposing the teeth, stronger actions may reveal the lower gums (AU16: *Depressor labii inferioris*).
Chin raise	State	The chin is pulled upwards causing the skin to flatten beneath the lower lip (AU17: *Mentalis muscle*).
True pucker	State	Purses the lips medially forward towards each other, narrowing the mouth corners medially and protruding the lips (AU18i: *Orbicularis oris incisivii labii superioris and inferioris*).
Outer pucker	State	The lips protrude and cinch together at a point distal to the midline, causing them to part and appear inflated (AU18ii: *Orbicularis oris, Incisivii labii inferioris and superioris*).

Face coding	Type	MaqFACS description, Action Unit (AU) label and musculature included (Parr et al., 2010)
Lip smack	State	Rapid and repeated smacking of the lips together, the teeth are covered, the tongue may protrude, and actions may be accompanied by vocalization (AD181: *Orbicularis oris*).
Lip tighten	State	Tightening and narrowing of the lips (AU23: *Orbicularis oris*)
Lips part	State	The lips part so some space is observable between them (AU25: *Several AUs may cause the lips to part*).
Jaw drop	State	The bottom jaw is relaxed and lowered (AU26).
Mouth stretch	State	The lower jaw is actively stretched to open the mouth, often occurs during yawning (AU27: *Mylohyoid, Depressor anguloris, Levator labii inferioris*).
Vocalization	Event	Macaque makes any type of vocalization.
Tongue out	State	The tongue protrudes in front of the teeth and is visible (AD19).
Chew	State	Macaque performs a chewing motion of the mouth but is not eating.
Lip tuck	State	Lower lip appears to tuck behind the upper lip.

[a] Behaviors combined in category “All walk” are “Quadrupedal walk” and “Bipedal walk”. [b] Behaviors combined in category “All run” are “Quadrupedal run” and “Bipedal run”. [c] Behaviors combined in category “All arboreal” are “Climb”, “Descend” and “Hang”. [d] Behaviors combined in category “All stand” are “Stand 4 limbs” and “Stand 2 limbs”

Statistical analysis

Statistical analysis was undertaken using R (R Core Team, 2017) in RStudio (Version 1.0.153, RStudio, Inc). For all analyses, the aim was to identify how behavior and facial movements differed between experimental periods (PreOp, PostOp, PreAn, and PostAn) in order to detect pain and wellness indicators. As outlined above, PreOp and PreAn were considered to be the most salient periods for identifying wellness and pain indicators, respectively.

Presence-absence analysis

To determine if behavior and facial movements were more likely to occur in particular periods, dependent variables were converted to a 1/0 data structure (0 = absence, 1 = presence). A logistic regression was conducted (glmmADMB function, glmmADMB package, Skaug et al., 2016) with individual allocated as a random effect to account for the repeated measures design, and a binomial distribution specified with the canonical logit link function. The main fixed effect of interest was “Period”. Secondary fixed effects included in the full models were “Sex” (male/female), “Age” (integer from 4 to 13), “Severity” (Mild/Moderate), “Pre-operative illness” (Yes/No – whether the macaque was unwell at the time of the procedure as some procedures were undertaken for medical reasons), “PreviousOps” (Yes/No – whether the macaque had undergone procedures previously), and “PostAnTime” (30/60 minutes). Although all data collection periods were 20 minutes in length, on some occasions the animal’s behavior was not visible or was unclear. To account for deviations in observation time, sampling effort (log transformed time in sight in seconds) was included as an offset in the model.

Models were simplified by stepwise selection using Wald Chi tests from the ANOVA function (Car package, Fox and Weisberg, 2011) to identify the least important variables, with confirmation using Akaike Information Criterion (AIC) and likelihood ratio tests with the ANOVA function (R base package. R Core Team, 2017) to ensure removal did not result in a weakened model. This was undertaken until either all fixed effects were significant or the variable to be removed was Period, the main variable of interest. Models including and excluding Period were then compared using AIC and likelihood ratio tests. When models indicated that Period was a significant variable, pairwise contrasts between different periods were undertaken using the pairs function (Lenth, 2016) with Tukey corrections for multiple comparisons. The likelihood of some behaviors could not be analyzed because these were displayed by either almost all or too few animals: walking (all walking and quadrupedal walking), eyebrow raise, true pucker, lip part and jaw drop.

Duration and frequency analysis

Analysis of behavioral durations and frequencies were conducted similarly to the presence/absence model, except with generalized linear mixed models (GLMMs) using the glmmADMB function. The residuals were not normally distributed as the data were over-dispersed; therefore, a negative binomial distribution was specified with a log link. Some behavior or facial movements could not be analyzed because these were either too rare or displayed by too few animals: bipedal running, affiliation, head jerk, pull on the head piece, hand shake, stretch, nose wrinkle with upper lip raise, upper lip raise, chin raise, true pucker, lip tuck. Plots indicated that model fitting was not optimal, and therefore confirmatory non-parametric tests were also undertaken using Friedman’s tests with Bonferroni-corrected Wilcox signed-rank post-hoc contrasts. As both methods resulted in similar outcomes and interpretations, only the results from the GLMMs are reported because these allow for more complexity in the model and were slightly more conservative overall.

Multivariate analysis

Behaviors were analyzed with a multi-level sparse partial least squares discriminant analysis (sPLSDA) with the sPLSDA function (mixOmics package, Le Cao et al., 2017), which allows for repeated measures on the same individuals. Data were log transformed, and variables were scaled and centered during the analysis. Firstly, a sPLSDA was undertaken on all of the data to determine which behaviors influenced classification of cases to period. Secondly, to determine if the model would allow predictions, the dataset was randomly divided into training (67%) and test sets (33%). The training set was tuned on 20 components, using 10-fold, leave-one-out cross validation, and maximum distance specifications, and this model was then used to predict classification of Period in the test data.

3 Results

3.1 Presence-absence analysis

Monkeys had clearly different behavior in the pre-operative period (PreOp), when they were assumed to be well, than at other times. Arboreal behavior such as climbing, descending, hanging, and the combined "all-arboreal" category were all most likely to occur in PreOp, as was cage manipulation (Tab. 3). There was an effect of sex on arboreal behavior with males less likely to be off the ground [probability with 95% confidence intervals: 0.23 (0.01-0.87)] than females [0.74 (0.25-0.96)], (χ^2 (1, n = 36) = 5.95, p = 0.015). Conversely, body shake had the opposite pattern and was least likely to occur in the PreOp period (Tab. 3). Body shake was also more likely to occur when the procedure severity was moderate [0.73 (0.18-0.97] compared to mild [0.45 (0.09-0.87)], (χ^2 (1, n = 36) = 4.31, p = 0.038), and when it was not the animal's first procedure [0.75 (0.19-0.97) compared to 0.43 (0.09-0.85)], (χ^2 (1, n = 36) = 7.09, p = 0.008).

During PostOp, the period when the residual effect of anesthesia was likely to be most evident, several behaviors were more likely to occur: half-closed eyes, leaning head and face rubbing, although PostOp was only significantly different from PreOp (Tab. 3). Face rubbing was more likely to occur in males [0.78 (0.46-0.94)] than females [0.58 (0.33-0.80)], (χ^2 (1, n = 36) = 5.75, p = 0.016), while the pattern of half-closed eyes was influenced by several variables (sex, age, and post-analgesia time) (Tab. 3). Conversely, standing and cage shaking decreased in the PostOp period (Tab. 3). Moreover, monkeys undergoing moderate severity procedures had a lower probability of cage shaking overall [0.15 (0.01-0.76)] than monkeys undergoing mild procedures [0.58 (0.17-0.90)], (χ^2 (1, n = 36) = 6.07, p = 0.014).

The probability of occurrence of pain related behaviors was expected to be most evident in the PreAn period. The likelihood of two behaviors, i.e., lip tightening and chewing, peaked in PreAn but these were only significantly different from the PreOp period, with PostOp and PostAn falling more or less between the two (Tab. 3). Similarly, the likelihood of running behavior (quadrupedal running and all running) troughed in PreAn but was not significantly lower than in either the PostOp or PostAn periods (Tab. 3).

Tab. 3: Minimum adequate model summaries of statistical analysis using presence/absence behavioral data of rhesus macaques over a 20-min observation time

Significant differences between experimental Periods (PreOp, PostOp, PreAn, PostAn) in the probability of behavior occurrence are indicated with letters in italics (different letters = significant differences). Akaike information criterion (AIC) given for models with and without Period as the main variable of interest

Behavior	Minimum adequate model	AIC with Period	AIC w/out Period	ΔAIC	P value	Probability of occurrence (95% confidence limits)			
						PreOp	PostOp	PreAn	PostAn
						A	*B*	*B*	*B*
Climb	~ Period + Sex	166.7	191.8	-25.1	<0.0001	0.89 (0.49 – 0.98)	0.18 (0.01 – 0.79)	0.25 (0.02 – 0.85)	0.48 (0.06 – 0.93)
Descend	~ Period + Sex	165.9	190.7	-24.8	<0.0001	0.83 (0.40 – 0.97)	0.09 (0.00 – 0.66)	0.29 (0.03 – 0.85)	0.38 (0.04 – 0.89)
Hang	~ Period + Sex + Ill + PreviousOps + PostAnTime	117.6	144.0	-26.4	<0.0001	0.76 (0.00 – 1.00)	0.03 (0.00 – 1.00)	0.04 (0.00 – 1.00)	0.06 (0.00 – 1.00)
All arboreal	~ Period + Sex	166.9	191.7	-24.8	<0.0001	0.89 (0.49 – 0.99)	0.17 (0.01 – 0.80)	0.28 (0.02 – 0.87)	0.53 (0.07 – 0.94)
Manipulate cage	~ Period	171.3	191.6	-20.2	<0.0001	0.96 (0.84 – 0.99)	0.44 (0.10 – 0.85)	0.74 (0.28 – 0.95)	0.63 (0.22 – 0.91)
Body shake	~ Period + Severity + PreviousOps	163.9	190.1	-26.2	<0.0001	0.19 (0.03 – 0.64)	0.66 (0.16 – 0.95)	0.83 (0.24 – 0.99)	0.69 (0.15 – 0.97)
						A	*B*	*AB*	*AB*
Half close eyes	~ Period + Sex + Age + PostAnTime	107.7	111.7	-4.0	0.02	0.24 (0.01 – 0.89)	0.79 (0.06 – 1.00)	0.51 (0.03 – 0.97)	0.54 (0.03 – 0.98)

Behavior	Minimum adequate model	AIC with Period	AIC w/out Period	ΔAIC	P value	Probability of occurrence (95% confidence limits)			
						PreOp	PostOp	PreAn	PostAn
						A	*B*	*AB*	*AB*
Lean head	~ Period	157.2	160.1	-2.9	0.03	0.66 (0.45 – 0.81)	0.93 (0.68 – 0.99)	0.83 (0.52 – 0.96)	0.8 (0.48 – 0.95)
Rub face	~ Period + Sex	178.4	183.7	-5.3	0.01	0.47 (0.25 – 0.70)	0.84 (0.50 – 0.97)	0.70 (0.35 – 0.91)	0.7 (0.34 – 0.91)
All stand	~ Period	152.7	157.5	-4.8	0.01	0.96 (0.83 – 0.99)	0.70 (0.27 – 0.94)	0.85 (0.44 – 0.98)	0.83 (0.43 – 0.97)
Cage shake	~ Period + Severity	162.9	175.6	-12.7	<0.001	0.66 (0.21 – 0.93)	0.09 (0.01 – 0.65)	0.34 (0.04 – 0.86)	0.39 (0.05 – 0.88)
						A	*AB*	*B*	*B*
Touch wound	~ Period + PreviousOps	155.6	167.5	-11.9	<0.001	0.36 (0.08 – 0.79)	0.68 (0.19 – 0.95)	0.89 (0.35 – 0.99)	0.83 (0.25 – 0.99)
						A	*AB*	*B*	*AB*
Tighten lips	~ Period + PostAnTime	134.7	139.6	-4.8	0.01	0.34 (0.06 – 0.80)	0.43 (0.06 – 0.91)	0.80 (0.27 – 0.98)	0.69 (0.19 – 0.96)
Chew	~ Period + Sex	135.6	138.3	-2.7	0.03	0.63 (0.33 – 0.85)	0.76 (0.37 – 0.94)	0.93 (0.60 – 0.99)	0.75 (0.34 – 0.95)
Quadrupedal run	~ Period + Age	154.0	161.6	-7.6	<0.01	0.43 (0.02 – 0.96)	0.13 (0.00 – 0.87)	0.06 (0.00 – 0.81)	0.13 (0.00 – 0.89)
All run	~ Period + Age	156.6	164.6	-8.0	<0.01	0.47 (0.02 – 0.97)	0.13 (0.00 – 0.88)	0.08 (0.00 – 0.85)	0.15 (0.00 – 0.92)
						A	*B*	*AB*	*B*
Close eyes	~ Period + Sex	121.5	133.9	-12.4	<0.001	0.10 (0.02 – 0.45)	0.65 (0.12 – 0.96)	0.42 (0.06 – 0.89)	0.67 (0.13 – 0.96)
Quadrupedal stand	~ Period	162.5	177.0	-14.4	<0.001	0.95 (0.82 – 0.99)	0.52 (0.16 – 0.86)	0.84 (0.42 – 0.97)	0.72 (0.32 – 0.93)
Bipedal stand	~ Period + Sex + Age + Ill	155.1	172.7	-17.6	<0.0001	0.89 (0.09 – 1.00)	0.18 (0.00 – 0.96)	0.59 (0.01 – 0.99)	0.35 (0.01 – 0.98)

3.2 Duration and frequency analysis

Several behaviors occurred for significantly more time in the PreOp period compared to other periods: This general pattern applied to walking (quadrupedal and all walking), arboreal behaviors (climbing, descending, hanging, and all arboreal), standing (quadrupedal, bi- and all stand), crouching, cage manipulation, and cage monitoring (Tab. 4). Arboreal behavior was also influenced by sex and pre-operative illness; females were off the ground more often than males (χ^2 (1, n = 36) = 11.06, p < 0.001), as were those individuals not considered to be ill prior to the procedure (χ^2 (1, n = 36) = 8.3, p = 0.004; Fig. 1a,b). Similarly, those individuals with pre-operative signs of illness spent less time standing following the procedure (χ^2 (1, n = 36) = 7.18, p = 0.007; Fig. 1c).

In contrast, the duration of half-closed eyes and leaning head, and the frequency of body shake was lower during the PreOp period (Tab. 4). Half-closed eyes was affected by several other explanatory variables including sex, age, severity, wellness and whether the animal had undergone previous procedures (Tab. 4). On average, males half-closed their eyes significantly more [39.1 (2.05-743.0)] than females [4.59 (0.24-88.08)] (χ^2 (1, n = 29) = 14.5, p = 0.0001). Half-closed eyes were more prevalent in individuals who were undergoing their first procedure [35.69 (2.14-594.46)] compared to a subsequent one [5.02 (0.23-111.89)], [χ^2 (1, n = 29) = 14.19, p = 0.0008), when the procedure severity was moderate compared to mild [χ^2 (1, n = 29) = 10.61, p = 0.001), (Fig. 1d), and in those who were potentially ill prior to the procedure [44.11 (2.61-746.34)], compared to those who were not ill prior to the procedure [4.07 (0.17-95.31)] (χ^2 (1, n = 29) = 11.6, p = 0.001). Body shake was more frequently observed in females [2.65 (1.48-4.75)] than males [1.36 (0.63-2.94)] (χ^2 (1, n = 36) = 6.61, p = 0.01).

There were no behaviors that clearly peaked or troughed during PreAn relative to all other periods.

3.3 Multivariate analysis

sPLSDA analysis of the training dataset classified the periods using three components and was plotted to visualize how behaviors related to each other (Fig. 2a-c). Only component 1 of the three could be sensibly labelled and appeared to indicate an axis of ac-

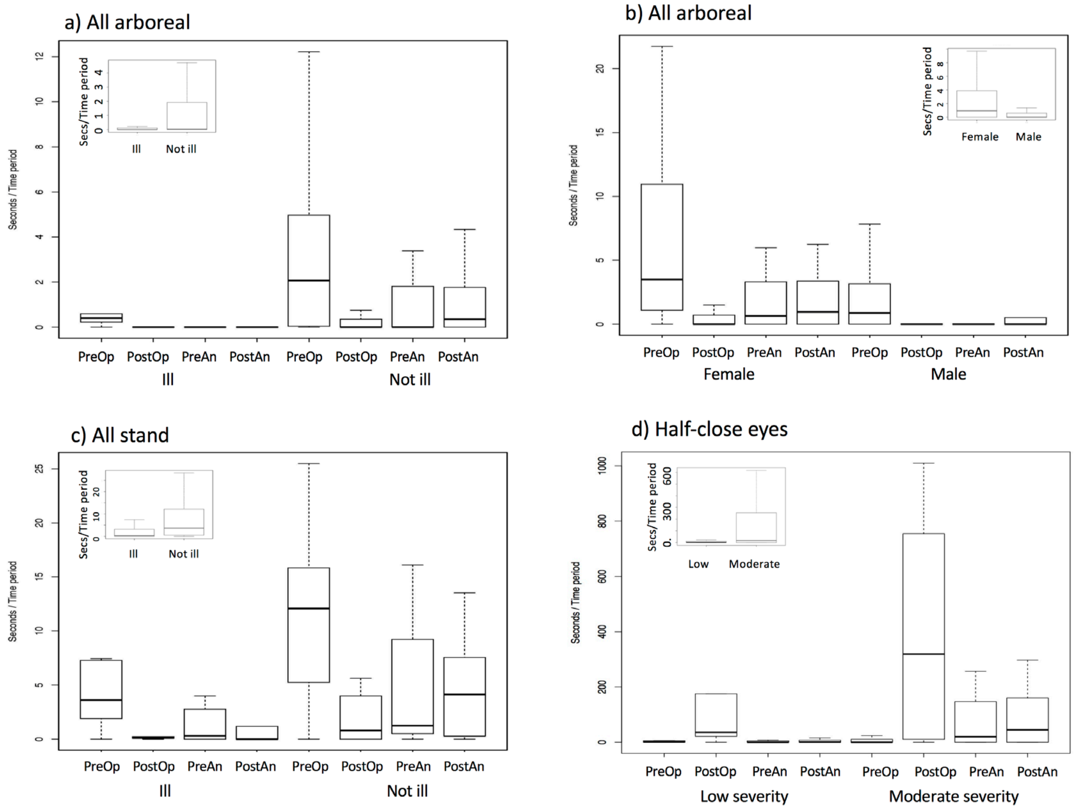

Fig. 1: Boxplots of behavior prevalence for rhesus macaques experiencing neuroscience procedures in relation to explanatory variables

a) & c) Pre-operative illness (ill/not ill); b) sex (female/male); and d) procedure severity (mild/moderate). Large figures depict behavioral prevalence for each 20-min observation period for the explanatory variable. Subset figures depict behavioral prevalence for the explanatory variable only.

Tab. 4: Minimum adequate model summaries of statistical analysis using duration (timed in seconds) and frequency (counted) behavioral data of rhesus macaques over a 20-min observation time

Significant differences between experimental Periods (PreOp, PostOp, PreAn, PostAn) in the prevalence of behavior are indicated with letters in italics (different letters = significant differences). Akaike information criterion (AIC) given for models with and without Period as the main variable of interest.

Behavior	Minimum adequate model	AIC with Period	AIC w/out Period	ΔAIC	P value	Unit	Mean (95% confidence intervals)			
							PreOp	PostOp	PreAn	PostAn
							A	*B*	*B*	*B*
Quadrupedal walk	~ Period	1456.7	1482.5	-25.8	<0.0001	Sec	129.55 (83.18 – 201.79)	22.85 (10.82 – 48.26)	37.71 (18.05 – 78.78)	34.69 (15.98 – 75.32)
All walk	~ Period	1470.9	1498.5	-27.5	<0.0001	Sec	129.1 (83.53 – 199.52)	23.18 (11.29 – 47.59)	38.59 (19.04 – 78.22)	36.67 (17.55 – 76.62)
Climb	~ Period + Sex + Ill	590.2	610.9	-20.6	<0.0001	Sec	2.94 (0.52 – 16.55)	0.29 (0.04 – 2.25)	0.51 (0.07 – 3.72)	0.91 (0.13 – 6.38)
Descend	~ Period + Sex + Ill	517.9	536.8	-18.9	<0.0001	Sec	1.96 (0.33 – 11.76)	0.17 (0.02 – 1.43)	0.51 (0.07 – 3.91)	0.5 (0.06 – 3.85)

Behavior	Minimum adequate model	AIC with Period	AIC w /out Period	ΔAIC	P value	Unit	Mean (95% confidence intervals)			
							PreOp	PostOp	PreAn	PostAn
							A	***B***	***B***	***B***
Hang	~ Period + PreviousOps	484.2	496.6	-12.4	<0.001	Sec	16.2 (1.16 – 226.28)	0.04 (0.00 – 2.71)	0.29 (0.01 – 15.34)	0.39 (0.01 – 25.49)
All arboreal	~ Period + Sex + Ill	833.9	849.1	-15.3	<0.0001	Sec	15.64 (1.97 – 124.30)	0.6 (0.05 – 7.83)	2.02 (0.13 – 32.35)	2.92 (0.22 – 39.16)
Quadrupedal stand	~ Period	1212.7	1222.0	-9.2	<0.01	Sec	78.68 (40.35 – 153.41)	10.51 (2.93 – 37.72)	18.4 (5.96 – 56.79)	17.02 (5.22 – 55.52)
Bipedal stand	~ Period + Severity + Ill + PostAnTime	1062.7	1075.8	-13.1	<0.001	Sec	10.5 (0.72 – 154.10)	1.25 (0.07 – 22.24)	2.98 (0.18 – 50.31)	2.53 (0.15 – 43.62)
All stand	~ Period + Ill	1423.7	1432.5	-8.8	<0.01	Sec	67.55 (19.26 – 236.91)	11.88 (2.43 – 58.09)	22.21 (4.77 – 103.48)	21.57 (4.51 – 103.22)
Crouch	~ Period + PostAnTime	552.4	584.6	-32.1	<0.0001	Sec	8.3 (1.48 – 46.47)	0.13 (0.01 – 1.43)	0.44 (0.04 – 4.36)	0.38 (0.04 – 3.91)
Manipulate cage	~ Period + Age + PostAnTime	997.8	1016.9	-19.1	<0.0001	Sec	67.9 (8.91 – 517.75)	6.83 (0.62 – 75.58)	10.65 (1.13 – 99.87)	15.32 (1.57 – 149.69)
Monitor cage	~ Period + PreviousOps	1149.4	1158.8	-9.4	<0.01	Sec	29.29 (15.14 – 56.70)	11.68 (5.08 – 26.82)	14.13 (6.12 – 32.63)	13.16 (5.68 – 30.53)
Half close eyes	~ Period + Sex + Age + Severity + Ill + PreviousOps	785.2	820.4	-35.2	<0.0001	Sec	1.50 (0.08 – 27.78)	71.37 (3.60 – 1414.49)	15.18 (0.69 – 336.19)	19.79 (0.97 – 402.10)
Lean head	~ Period	1635.1	1650.9	-15.8	<0.0001	Sec	20.59 (9.14 – 46.38)	291.17 (92.69 – 914.66)	145.51 (43.54 – 486.36)	122.73 (34.98 – 430.56)
Body shake	~ Period + Sex	589.5	619.8	-30.3	<0.0001	Count	0.61 (0.32 – 1.15)	3.03 (1.35 – 6.80)	3.18 (1.47 – 6.89)	2.22 (1.02 – 4.81)
							A	***B***	***AB***	***AB***
Lower glabella	~ Period + Age	354.8	360.0	-5.2	0.01	Sec	2.52 (0.25 – 25.06)	0.35 (0.02 – 5.00)	1.35 (0.12 – 15.72)	0.76 (0.06 – 9.39)
Cage shake	~ Period + PostAnTime	461.1	472.0	-10.9	<0.0001	Count	7.15 (0.88 – 58.06)	0.44 (0.03 – 6.34)	1.87 (0.15 – 22.61)	1.83 (0.15 – 22.55)
Focused vigilance	~ Period + PostAnTime	1430.3	1448.2	-17.9	<0.001	Sec	102.37 (42.89 – 244.34)	23.73 (8.21 – 68.53)	58.25 (21.15 – 160.39)	60.17 (21.92 – 165.17)
Not vigilant	~ Period	1917.4	1928.2	-10.8	<0.001	Sec	222.91 (180.49 – 275.29)	364.08 (267.13 – 496.23)	291.1 (213.32 – 397.25)	288.23 (211.31 – 393.14)
							A	***AB***	***B***	***B***
Groom	~ Period + PreviousOps	1590.0	1596.3	-6.3	<0.01	Sec	37.68 (18.51 – 76.72)	51.36 (20.67 – 127.63)	103.53 (38.89 – 275.57)	81.3 (32.04 – 206.29)
Touch wound	~ Period + Sex + Severity	871.2	878.5	-7.3	<0.01	Count	4.64 (0.97 – 22.16)	5.97 (1.05 – 33.98)	15.35 (2.86 – 82.22)	12.55 (2.41 – 65.39)

Behavior	Minimum adequate model	AIC with Period	AIC w/out Period	ΔAIC	P value	Unit	Mean (95% confidence intervals)			
							PreOp	PostOp	PreAn	PostAn
							A	*B*	*AB*	*B*
Close eyes	~ Period + Ill + PreviousOps	486.2	498.5	-15.1	<0.001	Sec	0.73 (0.04 – 13.10)	16.67 (0.78 – 357.97)	5.30 (0.30 – 92.09)	24.17 (1.63 – 358.89)
Blink	~ Period + PreviousOps + Severity	1034.7	1049.8	-13.4	<0.0001	Count	101.49 (69.48 – 148.24)	72.29 (48.41 – 107.97)	88.71 (59.28 – 132.75)	82.69 (55.29 – 123.67)
Lip smack	~ Period	173.1	181.9	-8.8	<0.01	Sec	0.33 (0.03 – 3.31)	0.01 (0.00 – 0.32)	0.09 (0.01 – 1.38)	0.03 (0.00 – 0.59)
							A	*B*	*B*	*AB*
Rub face	~ Period	683.3	688.5	-5.2	0.01	Count	1.36 (0.78 – 2.35)	4.02 (1.78 – 9.08)	3.17 (1.37 – 7.34)	2.5 (1.09 – 5.72)
							A	*A*	*B*	*B*
Shiver	~ Period + Age	508.3	531.9	-23.6	<0.001	Count	0.3 (0.00 – 19.47)	0.44 (0.01 – 35.63)	3.74 (0.06 – 249.42)	2.86 (0.04 – 182.29)
							A	*B*	*A*	*AB*
Lower lip depress	~ Period + Severity	244.3	249.7	-5.4	<0.01	Sec	1.67 (0.33 – 8.31)	0.13 (0.01 – 2.38)	2.20 (0.27 – 17.66)	1.49 (0.18 – 12.56)

Tab. 5a: Classification success rate (% correct) for each component identified using a sparse partial least squares discriminant analysis (sPLSDA) from training data (n = 18) of behavior collected from rhesus macaques over four 20-min observation periods

Period	Component 1	Component 2	Component 3
PreOp	99.4	88.9	88.9
PostOp	94.4	88.9	94.4
PreAn	0	61.1	44.4
PostAn	0	0	0

Tab. 5b: Classification success rate (% correct) for individual components 1-3 and the mean of components 1-3, identified using a sparse partial least squares discriminant analysis (sPLSDA) applied from training data (n = 18) to a test dataset (n = 9) of behavior collected from rhesus macaques over four 20-min observation periods

	Classification rate (%)															
	Component 1				Component 2				Component 3				Mean Components 1-3			
Actual case	PreOp	PostOp	PreAn	PostAn	PreOp	PostOp	PreAn	PostAn	PreOp	PostOp	PreAn	PostAn	PreOp	PostOp	PreAn	PostAn
PreOp	**88.9**	0.0	0.0	11.1	**44.4**	0.0	55.6	0.0	**11.1**	88.9	0.0	0.0	**48.1**	29.6	18.5	3.7
PostOp	44.4	**11.1**	22.2	22.2	22.2	**0.0**	77.8	0.0	11.1	**88.9**	0.0	0.0	25.9	**33.3**	33.3	7.4
PreAn	77.8	11.1	**11.1**	0.0	33.3	0.0	**66.7**	0.0	0.0	100.0	**0.0**	0.0	37.0	37.0	**25.9**	0.0
PostAn	100.0	0.0	0.0	**0.0**	88.9	0.0	11.1	**0.0**	55.6	44.4	0.0	**0.0**	81.5	14.8	3.7	**0.0**

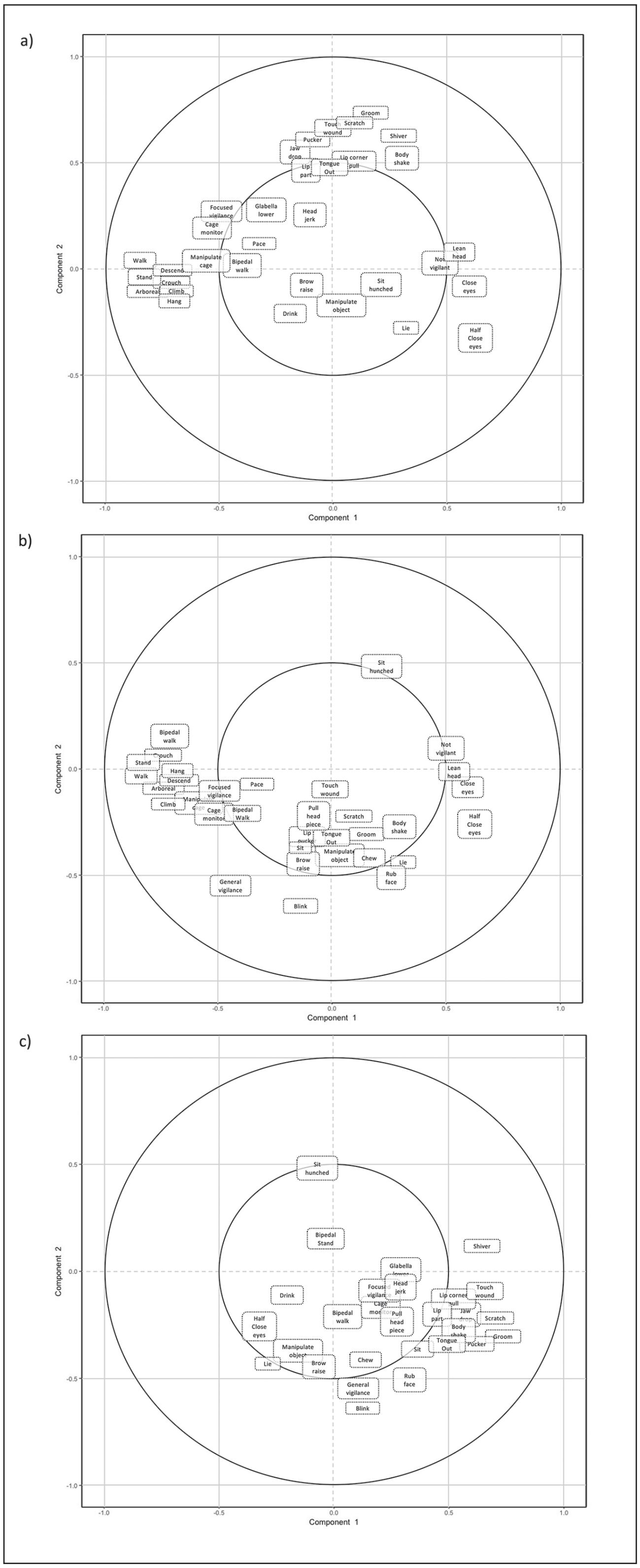

Fig. 2: Relationship between behavioral variables compiled from a sparse partial least squares discriminant analysis (sPLSDA) using training data (n = 18) collected from rhesus macaques over four 20-min observation periods.
For a) components 1 and 2; b) 1 and 3; c) 2 and 3.

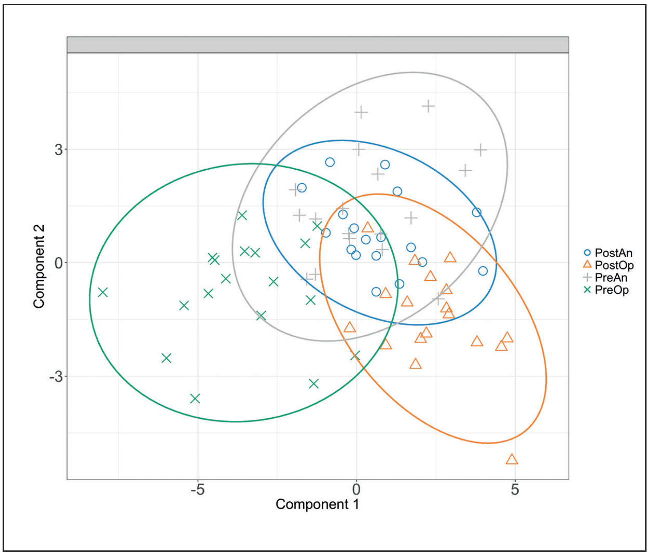

Fig. 3: Relationship between experimental periods (PreOp, PostOp, PreAn, PostAn) compiled from a sparse partial least squares discriminant analysis (sPLSDA) using training data (n = 18) of rhesus macaque behavior, plotted on components 1 and 2

tivity (-1.0) to inactivity (+1.0). Using the training data only, the PreOp and PostOp periods were clearly different from the other periods (Fig. 3a), and the classification rate was high across all three components (Tab. 5a). In the training data, the PreAn period was best identified by component 2 but the classification rate was moderate (61%). PostAn was not correctly classified for any components (Tab. 5a) and had large areas of overlap with other periods (Fig. 3). When applied to the test dataset, the mean successful classification rates of the test dataset across all three components were moderate for PreOp (48.1%), PostOp (33.3) and PreAn (25.9%), and unsuccessful for PostAn (0%) (Tab. 5b).

4 Discussion

4.1 Indicators of pain

The primary aim of this study was to identify potential behavioral and facial changes that indicate acute post-operative pain states in rhesus macaques. The period where pain was expected to be at its highest was PreAn, on the morning following the surgical procedure prior to routine administration of analgesia. At this point the effects of anesthesia and peri-surgical analgesia were expected to have dissipated significantly or entirely, as the next dose of pain relief was due to be given. Behavior that changes when pain is likely to be present and returns towards baseline levels after administration of analgesia can be used as evidence of pain (Roughan and Flecknell, 2002; Sotocinal et al., 2011; Wolfensohn and Lloyd, 2013; Sneddon et al., 2014). Three behaviors appeared to be potential pain indicators either in their presence (lip tightening and chewing, which peaked in probability at PreAn) or absence (running, which troughed in probability at PreAn), but these behaviors did not differ in amount or frequency during the period.

Lip tightening (AU23, Tab. 2) is not an action typically recruited in the human pain face (Prkachin and Solomon, 2009), however mouth tension, e.g., horizontal mouth stretching (AU20) and upper lip raising (AU10), is and comparable actions are observed in animal pain behavior, such as a "strained mouth" in horses (Dalla Costa et al., 2014). Chewing can also be a pain behavior in both humans and other animals. For example, donkeys chew more during mechanical nociceptive tests in comparison to sham tests (Grint et al., 2017), sheep grind their teeth when experiencing duodenal distension (Kania et al., 2009), and chewing can have an analgesic effect in humans (Weijenberg and Lobbezoo, 2015). In terms of locomotion, only running decreased during PreAn, however difficulty in movement is considered a negative correlate of quality of life in NHPs (Lambeth et al., 2013). In a previous study, female baboons decreased overall activity after abdominal surgery, although no effect on locomotion was found (Allison et al., 2007). Previous procedures (PreviousOps) contributed to some models (Tab. 3, 4), but there was not strong evidence of a link with a continued underlying pain state (i.e., no cumulative severity) for putative pain indicators.

However, it is important to note that changes in these behaviors did not map perfectly onto anticipated pain states because the probability of occurrence in PreAn did not differ significantly from either PostOp or PostAn. This may be for several possible reasons. Firstly, it may be that behavior is a relatively poor indicator of pain in NHPs (Allison et al., 2007), or secondly, that the animals in the study are masking their pain (Fenwick et al., 2014; Gaither et al., 2014) due to the presence of care staff and other macaques in the vicinity. A third possibility is that actual pain states do not align with the predicted pain states within periods. The efficacy and effects of anesthesia and analgesia regimes in NHPs undergoing surgery are not yet well understood (Bertrand et al., 2018) and the absence of significant behavioral change following analgesia could indicate that the drug or dosage was not optimal in managing the levels of pain experienced, or at least not for all individuals. However, there was also no effect of procedure severity on these behaviors, which would be expected for reliable pain indicators.

While it is possible that the monkeys were not experiencing pain, this is very unlikely, as similar protocols generate significant levels of pain when conducted on humans (Dunn et al., 2016). The opportunistic experimental design may have resulted in behavioral variation that masked some of the pain-specific responses; e.g., grimace scale studies have typically assessed pain using analgesiometric tests or following a standardized procedure (Langford et al., 2010; Sotocinal et al., 2011; Dalla Costa et al., 2014). Lastly, the prevalence and frequency of behaviors may also be confounded by individual variation in drug response, or in pain reaction or tolerance. Personality in primates is recognized to have a significant effect on behavior (Coleman, 2012) as well as general health and welfare (Robinson et al., 2016, 2018). For example, in female baboons individuals significantly varied in their response to the same standardized surgical procedure (Allison et al., 2007). Studies in dogs, horses and humans have also suggested that pain expression, if not the actual pain experience, is affected by personality (Williams, 2002; Ijichi et al., 2014; Lush and Ijichi, 2018), and this is a key area for future study on pain behavior in NHPs.

Although our results suggest that behavior and facial expressions alone are insufficient to assess pain states in NHPs, these may nonetheless make an important contribution to perioperative welfare, for example, in monitoring wellness or medication effects (Flecknell, 2018), and triangulation with physiological measures could enhance our understanding of pain responses (Allison et al., 2007). Not all pain indicators that have been previously proposed (e.g., Morton and Griffiths, 1985; Wolfensohn and Honess, 2005; NRCC, 2009; Lambeth et al., 2013, see introduction) were observed. This is possibly because proposed indicators were constructed from subjective impressions rather than empirical research with experimental controls such as blinding, minimization of observer effects and randomized analysis. Alternatively, it may also be that some indicators are specific to particular types of procedures or pain.

4.2 Indicators of wellness

Wellness indicators were considered to be those that were significantly different in the baseline PreOp period compared to all other periods. There were clear changes in behavior from PreOp levels, indicating that the procedures carried out considerably impacted the behavioral repertoire of macaques, with effects remaining for at least 12 hours post-procedure regardless of the administration of analgesia. This was particularly evident in the multivariate analysis as PreOp was clearly different from the other groups and had the most successful classification rate. This was also supported by the univariate analyses where many behaviors either peaked or troughed in PreOp. Several behaviors, primarily those indicative of activity level and alertness, were more likely to be present and/or performed more in the baseline period, and while behaviors of this nature can indicate an absence of pain (Roughan and Flecknell, 2002; Sneddon, 2017) in this case they appear to reflect general wellness because they were insensitive to analgesia administration. These behaviors include arboreal behaviors (such as climbing and hanging), standing, crouching and two cage-related behaviors (cage manipulation and cage monitoring) that may indicate motivation to return to the home cage from the recovery cage. Arboreal behavior and standing were also lower in monkeys that had indicators of illness prior to their procedures, supporting the interpretation that they may be good general indicators of wellness. Rhesus macaques are primarily terrestrial in the wild (Wells and Turnquist, 2001), however in experimental facilities much of their time is spent in elevated positions (Clarence et al., 2006), which is likely to be an anti-threat behavior. The recovery cage lacked an elevated perch and therefore the reduction in arboreality after the PreOp period may indicate a need to reduce energy expenditure, or reflect discomfort in movement (e.g., Allison et al., 2007). However, there was also an interaction with sex; females were more likely to spend time off the ground than males, which may reflect sex differences in threat or stress responses, consistent with previous findings of rhesus macaque reactions to an unfamiliar observer (Iredale et al., 2010).

The post-operative reduction in standing and cage monitoring are consistent with previous research on female baboons undergoing abdominal surgery (Allison et al., 2007). Posture may reflect reduced alertness or general reduction in activity, congruent with telemetry measures taken in female baboons (Allison et al., 2007). Cage monitoring, which is diminished following anesthesia, can be interpreted as indicating alertness (similar to the "checking" behaviour in Allison et al., 2007), and potentially associated with motivation to return to the home enclosure from the temporary cage. The exhibition of arboreal behavior and environmental manipulation seem to be good indicators of general well-being in macaques, at least in the context of temporary separation in a holding cage, although it is unclear whether post-operative reductions are attributable to pain or other factors because the frequency of these behaviors remained low across all postoperative periods, even after analgesia.

The converse pattern was identified for three behaviors that were either less likely to occur, or had lower durations, in the baseline pre-operative period; half-closed eyes, body shake and head leaning, which may indicate reduced wellness, again without specificity to pain. Half-closed eyes and head leaning occurred less before a procedure than during all subsequent periods, while body shake was both less frequent and less common. Monkeys that were potentially unwell prior to the procedure half-closed their eyes more frequently, supporting the interpretation that this behavior is influenced by reduced wellness. In previous research, reduced eye aperture is related to sedation effects (Bertrand et al., 2016), may function as a protective mechanism (Defensor et al., 2012) and is a common hallmark of the pain face in mammals including mice (Langford et al., 2010), horses (Dalla Costa et al., 2014), rats (Sotocinal et al., 2011), sheep (McLennan et al., 2016), and seals (MacRae et al., 2018). Similarly, body shaking may indicate reduced wellness as the likelihood of occurrence was higher for procedures of moderate severity, and this behavior has been linked to anxious states in clinically ill macaques (Gaither et al., 2014). Head leaning occurred for significantly lower durations in PreOp than in the post-operative periods, and could serve a similar function to the behavior of pressing hand to head described in clinically ill rhesus macaques, and may alleviate pain via manual pressure (Gaither et al., 2014), however, there was no influence of procedure severity that could specifically indicate pain. Head leaning shares postural similarities with huddling/hunching which has been suggested as a potential pain or distress behavior (Morton and Griffiths, 1985; Wolfensohn and Honess, 2005) and this supports the interpretation that it may be a potential indicator of malady. In terms of potential signs of cumulative severity, individuals who had undergone previous procedures had an increased likelihood of body shaking and reduced duration of half-closed eyes, however wellness indicators were mostly insensitive to previous procedures.

4.3 Indicators of sedation

Although the macaques were only filmed once cage-side observation suggested they had recovered from sedation, some behaviors peaked in the PostOp period when any residual effects of anesthesia would be expected to be most evident. These "sedation-related" behaviors were leaning head, half-closed eyes and face rubbing (the likelihood of occurrence but not the duration of time/frequency), and a non-vigilant state (duration but not likelihood). Behaviors that were most suppressed by sedation effects were standing (likelihood), lowered glabella (duration), focused vigilance (duration), and cage shaking (both frequency and likelihood). These behaviors suggest a general and expected decrease in activity and environmental engagement rather than being specific to pain or feeling unwell, but are highly relevant in applied contexts; it is important to distinguish sedation effects from other negative affective states to avoid inflated pain scores due to similarities in behavioral response (Langford et al., 2010; Miller et al., 2015).

4.4 Practicality of potential indicators

Clinical monitoring of animals is largely reliant on cage-side observation, and behavioral or facial patterns identified in this study provide insight into how macaque welfare may be monitored. Generalization of the results through replication and refinement of observable indicators is needed, with consideration given to which indicators are likely to be most effective and most practical.

One challenge in the interpretation of these findings is the influence of secondary, explanatory variables, which are likely to reflect the complexity of the pain experience and variation in response. For example, arboreal behavior and body shake were (positively and negatively, respectively) associated with wellness, while face rubbing was associated with sedation, however all of these behaviors were also influenced by the sex of the animal. Ideal indicators of pain or wellness would be effectively generalized; however, it is possible that if such indicators existed then robust guidelines on identifying pain in macaques would already exist, and as our results suggest, in practice some variation should be expected.

A second practical challenge is the prevalence of behavior evident during cage-side monitoring; behaviors which are rare or change too subtly in relation to wellness states are unlikely to be sufficiently robust to use in assessments. For example, sedation effects appear to decrease lowering of the glabella, but given the mean frequency was less than 3 seconds in any period, this would be difficult to practically detect at the cage-side. Similarly, the likelihood of chewing behavior increased with assumed pain state, however even in the baseline period when pain was presumed to be absent, the probability of occurrence within the observation period was high. This limitation could potentially be overcome by developments in automated monitoring; however, this may also be impractical in cage-side contexts. Based on our results, the most promising indicator of wellness appears to be the presence of arboreal behavior when macaques were not in their home enclosure, while the following would warrant close monitoring if occurring within a 20-minute period: two or more body shakes; more than 1 minute of head leaning or 10 seconds of half-closed eyes; and less than 1 minute of standing or cage manipulation. It is recommended that these indicators are included in facility welfare assessments, and imperative that housing offers macaques vertical space as per guidelines for housing research primates (Jennings and Prescott, 2009; NHMRC, 2016)

The third practical challenge is the potential impact of observer effects. In this study, the camera operator was not in visual range of the animal, however it was not possible to fully control potential observer effects because intermittent staff activity in the vicinity could evoke or suppress some behavioral responses (Iredale et al., 2010; Peterson et al., 2017). In applied contexts, animal monitoring is likely to be carried out using cage-side observation, however this may be insufficient for clinical assessments of macaque pain severity, at least without amelioration of potential observer effects, for example, through remote monitoring (Gaither et al., 2014).

5 Conclusion

Surgical research protocols have undesirable welfare implications for animals used as experimental models, such as the rhesus macaque. The 3Rs principle of Refinement guides scientists to minimize the effects on such animals; however, this is reliant on accurate assessment of negative affective states including, but not limited to, pain. This project has identified several potential behavioral indicators of pain and general wellness, however practical implementation to applied contexts is likely to have challenges. Although macaques are thought to hide their responses to pain, they were clearly negatively impacted by the procedures, as evidenced by changes in their behavior. It is difficult to disentangle common indicators of pain, malaise and sedation in macaques due to the complexity of interactions with other factors and given the confounds of the opportunistic sampling and applied context of the current study. Directions for future research should aim to extend this work to different experimental interventions and to examine the influence of individual behavioral variation on pain response. The precautionary principle should be applied to pain relief until sensitive and robust measures of pain are identified and validated (Flecknell, 1984; Roughan and Flecknell, 2002; Sneddon et al., 2014). Evaluating improved anesthetic and intraoperative care regimens using "wellness" indicators could lead to significant refinements of research procedures.

References

Allison, S. O., Halliday, L. C., French, J. A. et al. (2007). Assessment of buprenorphine, carprofen, and their combination for postoperative analgesia in olive baboons (Papio anubis). *J Am Assoc Lab Anim Sci 46*, 24-31.

ASAB (2018). Guidelines for the treatment of animals in behavioural research and teaching. *Anim Behav 135*, I-X. doi:10.1016/j.anbehav.2017.10.001

Azimi, K., Prescott, I. A., Marino, R. A. et al. (2016). Low profile halo head fixation in non-human primates. *J Neurosci Methods 268*, 23-30. doi:10.1016/j.jneumeth.2016.04.018

Bair, M. J., Robinson, R. L., Katon, W. and Kroenke, K. (2003). Depression and pain comorbidity: A literature review. *Arch Intern Med 163*, 2433-2445. doi:10.1001/archinte.163.20.2433

Balcombe, J. P., Barnard, N. D. and Sandusky, C. (2004). Laboratory routines cause animal stress. *Contemp Top Lab Anim Sci 43*, 42-51.

Bateson, P. (1991). Assesment of pain in animals. *Anim Behav 42*, 827-839. doi:10.1016/S0003-3472(05)80127-7

Bertrand, H. G. M. J., Ellen, Y. C., O'Keefe, S. and Flecknell, P. A. (2016). Comparison of the effects of ketamine and fentanyl-midazolam-medetomidine for sedation of rhesus macaques (Macaca mulatta). *BMC Vet Res 12*, 93. doi:10.1186/s12917-016-0721-9

Bertrand, H. G. M. J., Sandersen, C. and Flecknell, P. A. (2018). Reported analgesic and anaesthetic administration to human primates undergoing experimental surgical procedure: 2010-2015. *J Med Primatol 47*, 217-225. doi:10.1111/jmp.12346

Brearley, J. C. and Brearley, M. J. (2000). Chronic pain in animals. In P. A. Flecknell and A. E. Waterman-Pearson (eds.), *Pain Management in Animals* (147-160). Harcourt Publishers Ltd. doi:10.1016/B978-0-7020-1767-4.50009-6

Buchanan-Smith, H. M., Rennie, A. E., Vitale, A. et al. (2005). Harmonising the definition of refinement. *Anim Welf 14*, 379-384.

Burrows, A. M., Waller, B. M. and Parr, L. A. (2009). Facial musculature in the rhesus macaque (Macaca mulatta): Evolutionary and functional contexts with comparisons to chimpanzees and humans. *J Anat 215*, 320-334. doi:10.1111/j.1469-7580.2009.01113.x

Capitanio, J. R., Mendoza, S. P. and McChesney, M. (1996). Influences of blood sampling procedures on basal hypothalamic-pituitary-adrenal hormone levels and leukocyte values in rhesus macaques (Macaca mulatta). *J Med Primatol 25*, 26-33. doi:10.1111/j.1600-0684.1996.tb00189.x

Carlsson, H. E., Schapiro, S. J., Farah, I. and Hau, J. (2004). Use of primates in research: A global overview. *Am J Primatol 63*, 225-237. doi:10.1002/ajp.20054

Carstens, E. and Moberg, G. P. (2000). Recognising pain and distress in laboratory animals. *ILAR J 41*, 62-71. doi:10.1093/ilar.41.2.62

Chevalier-Skolnikoff (1973). Facial expression of emotion in nonhuman primates. In P. Ekman (ed.), *Darwin and Facial Expression: A Century of Research in Review* (11-89). New York, USA: Academic Press.

Clarence, W. M., Scott, J. P., Dorris, M. C. and Paré, M. (2006). Use of enclosures with functional vertical space by captive rhesus monkeys (Macaca mulatta) involved in biomedical research. *J Am Assoc Lab Anim Sci 45*, 31-34.

Coleman, K. (2012). Individual differences in temperament and behavioral management practices for nonhuman primates. *Appl Anim Behav Sci 137*, 106-113. doi:10.1016/j.applanim.2011.08.002

Cooper, C. S., Metcalf-Pate, K. A., Barat, C. E. et al. (2009). Comparison of side effects between buprenorphine and meloxicam used postoperatively in Dutch belted rabbits (Oryctolagus cuniculus). *J Am Assoc Lab Anim Sci 48*, 279-285.

Craig, K. D. and Patrick, C. J. (1985). Facial expression during induced pain. *J Pers Soc Psychol 48*, 1080-1091. doi:10.1037//0022-3514.48.4.1089

Craig, K. D., Whitfield, M. F., Grunau, R. V. E. et al. (1993). Pain in the preterm neonate: Behavioural and physiological indices. *Pain 52*, 287-299. doi:10.1016/0304-3959(93)90162-I

Craig, K. D., Prkachin, K. M. and Grunau, R. V. E. (2001). The facial expression of pain. In D. C. Turk (ed.), *Handbook of Pain Assessment* (153-169). New York, USA: Guilford Publications.

Dalla Costa, E., Minero, M., Lebelt, D. et al. (2014). Development of the horse grimace scale (HGS) as a pain assessment tool in horses undergoing routine castration. *PLoS One 9*, 1-10. doi:10.1371/journal.pone.0092281

Dawkins, M. S. (2006a). A user's guide to animal welfare science. *Trends Ecol Evol 21*, 77-82. doi:10.1016/j.tree.2005.10.017

Dawkins, M. S. (2006b). Through animal eyes: What behaviour tells us. *Appl Anim Behav Sci 100*, 4-10. doi:10.1016/j.applanim.2006.04.010

Defensor, E. B., Corley, M. J., Blanchard, R. J. and Blanchard, D. C. (2012). Facial expressions of mice in aggressive and fearful contexts. *Physiol Behav 107*, 680-685. doi:10.1016/j.physbeh.2012.03.024

Defrin, R., Lahav, Y. and Solomon, Z. (2017). Dysfunctional pain modulation in torture survivors: The mediating effect of PTSD. *J Pain 18*, 1-10. doi:10.1016/j.jpain.2016.09.005

Descovich, K. A., Wathan, J., Leach, M. C. et al. (2017). Facial expression: An under-utilized tool for the assessment of welfare in mammals. *ALTEX 34*, 409-429. doi:10.14573/altex.1607161

de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annu Rev Psychol 59*, 279-300. doi:10.1146/annurev.psych.59.103006.093625

DeWall, C. N. and Baumeister, R. F. (2006). Alone but feeling no pain: Effects of social exclusion on physical pain tolerance and pain threshold, affective forecasting, and interpersonal empathy. *J Pers Soc Psychol 91*, 1-15. doi:10.1037/0022-3514.91.1.1

Dunn, L. K., Naik, B. I., Nemergut, E. C. and Durieux, M. E. (2016). Post-craniotomy pain management: Beyond opioids. *Curr Neurol Neurosci Rep 16*, 93. doi:10.1007/s11910-016-0693-y

Ekman, P. and Friesen, W. V (1978). *Facial Action Coding System: Investigator's Guide*. Palo Alto, CA, USA: Consulting Psychologists Press.

Everds, N. E., Snyder, P. W., Bailey, K. L. et al. (2013). Interpreting stress responses during routine toxicity studies: A review of the biology, impact, and assessment. *Toxicol Pathol 41*, 560-614. doi:10.1177/0192623312466452

Fenwick, N., Duffus, S. E. G. and Griffin, G. (2014). Pain management for animals used in science: Views of scientists and veterinarians in Canada. *Animals 4*, 494-514. doi:10.3390/ani4030494

Flecknell, P. A. (1984). The relief of pain in laboratory animals. *Lab Anim 18*, 147-160. doi:10.1258/002367784780891226

Flecknell, P. (2018). Analgesics in small mammals. *Vet Clin North Am Exot Anim Pract 21*, 83-103. doi:10.1016/j.cvex.2017.08.003

Fleming, B. M. and Coombs, D. W. (1992). A survey of complications documented in a quality-control analysis of patient-controlled analgesia in the postoperative patient. *J Pain Symptom Manage 7*, 463-469. doi:10.1016/0885-3924(92)90132-2

Fox, J. and Weisberg, S. (2011). An R Companion to Applied Regression, Second Edition. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Fraser, D. (2009). Animal behaviour, animal welfare and the scientific study of affect. *Appl Anim Behav Sci 118*, 108-117. doi:10.1016/j.applanim.2009.02.020

Gaither, A., Baker, K. C., Gilbert, M. et al. (2014). Videotaped behavior as a predictor of clinical outcome in rhesus macaques (Macaca mulatta). *Comp Med 64*, 193-199.

Ghazanfar, A. A. and Logothetis, N. K. (2003). Facial expressions linked to monkey calls. *Nature 423*, 937-938. doi:10.1038/423937a

Grint, N. J., Beths, T., Yvorchuk-St Jean, K. et al. (2017). Analysis of behaviors observed during mechanical nociceptive threshold testing in donkeys and horses. *J Equine Vet Sci 50*, 102-109. doi:10.1016/j.jevs.2016.11.001

Hall, L. E., Robinson, S. and Buchanan-Smith, H. M. (2015). Refining dosing by oral gavage in the dog: A protocol to harmonise welfare. *J Pharmacol Toxicol Methods 72*, 35-46. doi:10.1016/j.vascn.2014.12.007

Hanninen, L. and Pastell, M. (2009). CowLog: Open-source software for coding behaviors from digital video. *Behav Res Methods 41*, 472-476. doi:10.3758/BRM.41.2.472

Hawkins, P. (2002). Recognizing and assessing pain, suffering and distress in laboratory animals: A survey of current practice in the UK with recommendations. *Lab Anim 36*, 378-395. doi:10.1258/002367702320389044

Honess, P. and Wolfensohn, S. (2010). The extended welfare assessment grid: A matrix for the assessment of welfare and cumulative suffering in experimental animals. *Altern Lab Anim 38*, 205-212. doi:10.1007/978-1-4020-9219-0_5

Ijichi, C., Collins, L. M. and Elwood, R. W. (2014). Pain expression is linked to personality in horses. *Appl Anim Behav Sci 152*, 38-43. doi:10.1016/j.applanim.2013.12.007

Iredale, S. K., Nevill, C. H. and Lutz, C. K. (2010). The influence of observer presence on baboon (Papio spp.) and rhesus macaque (Macaca mulatta) behavior. *Appl Anim Behav Sci 122*, 53-57. doi:10.1016/j.applanim.2009.11.002

Jennings, M. and Prescott, M. J. (eds.) (2009). Refinements in husbandry, care and common procedures for non-human primates: Ninth report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement. *Lab Anim 43, Suppl 1*, 1-47. doi:10.1258/la.2008.007143

Jensen, M. P., Karoly, P. and Braver, S. (1986). The measurement of clinical pain intensity: A comparison of six methods. *Pain 27*, 117-126. doi:10.1016/0304-3959(86)90228-9

Johnston, C. C., Stevens, B., Craig, K. D. and Grunau, R. V. (1993). Developmental changes in pain expression in premature, full-term, two- and four-month-old infants. *Pain 52*, 201-208. doi:10.1016/0304-3959(93)90132-9

Kania, B. F., Brytan, M. and Tomaszewska, D. (2009). Centrally administered verapamil prevents the autonomic reaction to visceral pain in sheep. *Res Vet Sci 86*, 121-128. doi:10.1016/j.rvsc.2008.04.009

Kilkenny, C., Browne, W. J., Cuthill, I. C. et al. (2013). Improving bioscience research reporting: The arrive guidelines for reporting animal research. *Animals 4*, 35-44. doi:10.3390/ani4010035

Lambeth, S. P., Schapiro, S. J., Bernacky, B. J. and Wilkerson, G. (2013). Establishing 'quality of life' parameters using behavioural guidelines for humane euthanasia of captive non-human primates. *Anim Welf 22*, 429-435. doi:10.7120/09627286.22.4.429

Langford, D. J., Crager, S. E., Shehzad, Z. et al. (2006). Social modulation of pain as evidence for empathy in mice. *Science 312*, 1967-1970. doi:10.1126/science.1115233

Langford, D. J., Bailey, A. L., Chanda, M. L. et al. (2010). Coding of facial expressions of pain in the laboratory mouse. *Nat Methods 7*, 447-449. doi:10.1038/nmeth.1455

Le Cao, K.-A., Rohart, F., Gonzalez, I. et al. (2017). mixOmics: Omics Data Integration Project. R package version 6.2.0. https://cran.r-project.org/package=mixOmics

Leach, M. C., Coulter, C. A., Richardson, C. A. and Flecknell, P. A. (2011). Are we looking in the wrong place? Implications for behavioural-based pain assessment in rabbits (Oryctolagus cuniculi) and beyond? *PLoS One 6*, e13347. doi:10.1371/journal.pone.0013347

Leaman, J., Latter, J. and Clemence, M. (2014). Attitudes to animal research in 2014: A report by Ipsos MORI for the Department for Business, Innovation & Skills. https://www.ipsos.com/sites/default/files/migrations/en-uk/files/Assets/Docs/Polls/sri_BI-Sanimalresearch_TRENDreport.pdf

Lenth, R. (2016). Least-squares means: The R package lsmeans. *J Stat Softw 69*, 1-33. doi:10.18637/jss.v069.i01

LeResche, L. (1982). Facial expression in pain: A study of candid photographs. *J Nonverbal Behav 7*, 46-56. doi:10.1007/BF01001777

Loeser, J. D. and Treede, R. D. (2008). The Kyoto protocol of IASP basic pain terminology. *Pain 137*, 473-477. doi:10.1016/j.pain.2008.04.025

Lund, T. B., Lassen, J. and Sandøe, P. (2012). Public attitude formation regarding animal research. *Anthrozoös 25*, 475-490. doi:10.2752/175303712X13479798785896

Lush, J. and Ijichi, C. (2018). A preliminary investigation into personality and pain in dogs. *J Vet Behav 24*, 62-68. doi:10.1016/j.jveb.2018.01.005

Lutz, C., Well, A. and Novak, M. (2003). Stereotypic and self-injurious behavior in rhesus macaques: A survey and retrospective analysis of environment and early experience. *Am J Primatol 60*, 1-15. doi:10.1002/ajp.10075

MacRae, A. M., Makowska, I. J. and Fraser, D. (2018). Initial evaluation of facial expressions and behaviours of harbour seal pups (Phoca vitulina) in response to tagging and microchipping. *Appl Anim Behav Sci 205*, 167-174. doi:10.1016/j.applanim.2018.05.001

McLennan, K. M., Rebelo, C. J. B., Corke, M. J. et al. (2016). Development of a facial expression scale using footrot and mastitis as models of pain in sheep. *Appl Anim Behav Sci 176*, 19-26. doi:10.1016/j.applanim.2016.01.007

Mellor, D. J. and Beausoleil, N. J. (2015). Extending the "five domains" model for animal welfare assessment to incorporate positive welfare states. *Anim Welf 24*, 241-253. doi:10.7120/09627286.24.3.241

Miller, A., Kitson, G., Skalkoyannis, B. and Leach, M. (2015). The effect of isoflurane anaesthesia and buprenorphine on the mouse grimace scale and behaviour in CBA and DBA/2 mice. *Appl Anim Behav Sci 172*, 58-62. doi:10.1016/j.applanim.2015.08.038

Mogil, J. S. and Kest, B. (1999). Sex differences in opioid analgesia. Of mice and women. *Pain Forum 8*, 48-50. doi:10.1016/S1082-3174(99)70045-0

Mogil, J. S., Chesler, E. J., Wilson, S. G. et al. (2000). Sex differences in thermal nociception and morphine antinociception in rodents depend on genotype. *Neurosci Biobehav Rev 24*, 375-389. doi:10.1016/S0149-7634(00)00015-4

Morton, D. B. and Griffiths, P. H. (1985). Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *Vet Rec 116*, 431-436. doi:10.1136/vr.116.16.431

Morton, D. B. (2000). A systematic approach for establishing humane endpoints. *ILAR J 41*, 80-86. doi:10.1093/ilar.41.2.80

Murdoch, F. R., Maker, G. L., Nitsos, I. et al. (2013). Intraperitoneal medetomidine: A novel analgesic strategy for postoperative pain management in pregnant sheep. *Lab Anim 47*, 66-70. doi:10.1177/0023677212473712

NHMRC – National Health and Medical Research Council (2016). Principles and Guidelines for the Care and Use of Non-Human Primates for Scientific Purposes. https://nhmrc.gov.au/about-us/publications/principles-and-guidelines-care-and-use-non-human-primates-scientific-purposes

NRCC – National Research Council Committee (2009). *Recognition and Alleviation of Pain in Laboratory Animals* (47-69). Washington DC, USA: National Academies Press.

Niekrasz, M. A. and Wardrip, C. L. (2012). Surgery in nonhuman primates. In C. R. Abee, K. Mansfield, S. Tardif et al. (eds.), *Nonhuman Primates in Biomedical Research* (339-358). Elsevier Inc. doi:10.1016/B978-0-12-381365-7.00014-5

Olsson, I. A. S. and Westlund, K. (2007). More than numbers matter: The effect of social factors on behaviour and welfare of laboratory rodents and non-human primates. *Appl Anim Behav Sci 103*, 229-254. doi:10.1016/j.applanim.2006.05.022

Osborne, N. J., Payne, D. and Newman, M. L. (2009). Journal editorial policies, animal welfare, and the 3Rs. *Am J Bioeth 9*, 55-59. doi:10.1080/15265160903318343

Parr, L. A., Waller, B. M., Burrows, A. M. et al. (2010). Brief communication: MaqFACS: A muscle-based facial movement coding system for the rhesus macaque. *Am J Phys Anthropol 143*, 625-630. doi:10.1002/ajpa.21401

Partan, S. R. (2002). Single and multichannel signal composition: Facial expressions and vocalizations of Rhesus macaques (Macaca mulatta). *Behaviour 139*, 993-1027. doi:10.1163/15685390260337877

Peterson, E. J., Worlein, J. M., Lee, G. H. et al. (2017). Rhesus macaques (Macaca mulatta) with self-injurious behavior show less behavioral anxiety during the human intruder test. *Am J Primatol 79*, 1-8. doi:10.1002/ajp.22569

Pham, T. M., Hagman, B., Codita, A. et al. (2010). Housing environment influences the need for pain relief during post-operative recovery in mice. *Physiol Behav 99*, 663-668. doi:10.1016/j.physbeh.2010.01.038

Plesker, R. and Mayer, V. (2008). Nonhuman primates mask signs of pain. *Lab Primate Newsl 47*, 1-3.

Poole, T. (1997). Happy animals make good science. *Lab Anim 31*, 116-124. doi:10.1258/002367797780600198

Prescott, M. J., Brown, V. J., Flecknell, P. A. et al. (2010). Refinement of the use of food and fluid control as motivational tools for macaques used in behavioural neuroscience research: Report of a working group of the NC3Rs. *J Neurosci Methods 193*, 167-188. doi:10.1016/j.jneumeth.2010.09.003

Prkachin, K. M. (1992). The consistency of facial expressions of pain: A comparison across modalities. *Pain 51*, 297-306. doi:10.1016/0304-3959(92)90213-U

Prkachin, K. M. (2009). Assessing pain by facial expression: Facial expression as nexus. *Pain Res Manag 14*, 53-58.

doi:10.1155/2009/542964

Prkachin, K. M. and Solomon, P. E. (2009). The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain 139*, 267-274. doi:10.1016/j.pain.2008.04.010

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Rennie, A. E. and Buchanan-Smith, H. M. (2006a). Refinement of the use of non-human primates in scientific research. Part II: Housing, husbandry and acquisition. *Anim Welf 15*, 215-238.

Rennie, A. E. and Buchanan-Smith, H. M. (2006b). Refinement of the use of non-human primates in scientific research. Part I: The influence of humans. *Anim Welf 15*, 203-213.

Rennie, A. E. and Buchanan-Smith, H. M. (2006c). Refinement of the use of non-human primates in scientific research. Part III: Refinement of procedures. *Anim Welf 15*, 239-261.

Robinson, L. M., Waran, N. K., Leach, M. C. et al. (2016). Happiness is positive welfare in brown capuchins (Sapajus apella). *Appl Anim Behav Sci 181*, 145-151. doi:10.1016/j.applanim.2016.05.029

Robinson, L. M., Coleman, K., Capitanio, J. P. et al. (2018). Rhesus macaque personality, dominance, behavior, and health. *Am J Primatol 80*, e22739. doi:10.1002/ajp.22739

Roughan, J. V. and Flecknell, P. A. (2002). Buprenorphine: A reappraisal of its antinociceptive effects and therapeutic use in alleviating post-operative pain in animals. *Lab Anim 36*, 322-343. doi:10.1258/002367702320162423

Russell, W., Burch, R., Hume, C. et al. (1992). *The Principles of Humane Experimental Technique (Special ed.)*. Potters Bar, Herts, UK: Universities Federation for Animal Welfare.

Schaap, M. W. H., Uilenreef, J. J., Mitsogiannis, M. D. et al. (2012). Optimizing the dosing interval of buprenorphine in a multimodal postoperative analgesic strategy in the rat: Minimizing side-effects without affecting weight gain and food intake. *Lab Anim 46*, 287-292. doi:10.1258/la.2012.012058

Skaug, H., Fournier, D., Bolker, B. et al. (2016). glmmADMB: Generalized Linear Mixed Models using ‘AD Model Builder’.

Sneddon, L. U., Elwood, R. W., Adamo, S. A. and Leach, M. C. (2014). Defining and assessing animal pain. *Anim Behav 97*, 201-212. doi:10.1016/j.anbehav.2014.09.007

Sneddon, L. U. (2017). Pain in laboratory animals: A possible confounding factor? *Altern Lab Anim 45*, 161-164. doi:10.1177/026119291704500309

Sotocinal, S. G., Sorge, R. E., Zaloum, A. et al. (2011). The rat grimace scale: A partially automated method for quantifying pain in the laboratory rat via facial expressions. *Mol Pain 7*, 55. doi:10.1186/1744-8069-7-55

Stokes, E. L., Flecknell, P. A. and Richardson, C. A. (2009). Reported analgesic and anaesthetic administration to rodents undergoing experimental surgical procedures. *Lab Anim 43*, 149-154. doi:10.1258/la.2008.008020

Taddio, A., O’Brien, L., Ipp, M. et al. (2009). Reliability and validity of observer ratings of pain using the visual analog scale (VAS) in infants undergoing immunization injections. *Pain 147*, 141-146. doi:10.1016/j.pain.2009.08.027

Tasker, L. (2012). Linking welfare and quality of scientific output in cynomolgus macaques (Macaca fascicularis) used for regulatory toxicology. Psychology eThese, University of Stirling. http://hdl.handle.net/1893/9801

UK Home Office (2014). Code of practice for the housing and care of animals used in scientific procedures 2014. https://www.gov.uk/government/publications/code-of-practice-for-the-housing-and-care-of-animals-bred-supplied-or-used-for-scientific-purposes

Walker, A. and Srinivas, G. B. (2013). Opportunities and strategies to further reduce animal use for Leptospira vaccine potency testing. *Biologicals 41*, 332-337. doi:10.1016/j.biologicals.2013.06.006

Waller, B. M. and Micheletta, J. (2013). Facial expression in nonhuman animals. *Emot Rev 5*, 54-59. doi:10.1177/1754073912451503

Weijenberg, R. A. F. and Lobbezoo, F. (2015). Chew the pain away: Oral habits to cope with pain and stress and to stimulate cognition. *Biomed Res Int 2015*, 149431. doi:10.1155/2015/149431

Wells, J. P. and Turnquist, J. E. (2001). Ontogeny of locomotion in rhesus macaques (Macaca mulatta): II. Postural and locomotor behavior and habitat use in a free-ranging colony. *Am J Phys Anthropol 115*, 80-94. doi:10.1002/ajpa.1059

Williams, A. C. D. C. (2002). Williams, A. C. D. C. (2002). Facial expression of pain: An evolutionary account. *Behav Brain Sci 25*, 439-488.

Wolfensohn, S. and Honess, P. (2005). Physical well-being. In *Handbook of Primate Husbandry and Welfare* (59-98). Blackwell Publishing Ltd. doi:10.1002/9780470752951

Wolfensohn, S. and Lloyd, M. (2013). Balancing animal welfare with science. In *Handbook of Laboratory Animal Management and Welfare* (37-50). John Wiley & Sons.

Würbel, H. (2001). Ideal homes? Housing effects on rodent brain and behaviour. *Trends Neurosci 24*, 207-211. doi:10.1016/S0166-2236(00)01718-5

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

The authors would like to thank Emmeline Howarth, Dr Henri Bertrand and the technical and care staff who assisted in the conduct of these studies. This research was funded by the National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs: G1100563/1; NC/K001159/1).