Finding what is normal - Consistency in individual behaviour and movement patterns in laboratory-housed Macaques and how this relates to individual personality traits

Thesis submitted for the degree of Doctor of Philosophy by

Tim-Joshua Yannick Andres

Biosciences Institute
Faculty of Medical Sciences
Newcastle University, Newcastle-upon-Tyne
United Kingdom

March 2024

ABSTRACT

This thesis investigates individual differences in behaviour of laboratory-housed Rhesus macaques (*Macaca mulatta*) through the utilization of automated methods designed for in-cage observation. Animals consistently exhibit differences in movement and behaviour over time, this is often described as personality or temperament. Previous research found associations between temperament and physiological indicators, such as well-being and welfare of non-human primates. Rhesus macaques serve as a pivotal model species in biomedical research, and considerable efforts are made in optimizing their welfare for both ethical and scientific purposes. This emphasizes the significance of comprehending these individual differences and to specifically identifying behavioural needs.

However, evaluating welfare and temperament typically entails labour-intensive processes reliant on human observers. Recent advancements in computational techniques to automate video analysis have transformed animal behaviour studies, reducing costs and labour intensity. Furthermore, such methods are non-invasive and less susceptible to observer bias.

This thesis presents a self-built in-cage camera system (based on a Raspberry Pi), as well as a pipeline (based on YOLOv8, a state-of-the-art image-segmentation model) to automate video analysis and predict the location of pair-housed Rhesus macaques in their home cage. Utilising the camera module and pipeline, over 1500 hours of video material was collected and analysed. The findings reveal distinct individuality in movement trajectories as well as variations in space utilization across different contexts related to welfare and personality traits.

In conclusion, the findings suggest the potential of automated methods to assess behavioural needs, temperament, and welfare on an individual level, thereby contributing to more precise and comprehensive care strategies for laboratory-housed macaques. Furthermore, I present an effective solution of monitoring home cage space use on a long-term basis during light periods (5:30AM-7PM).

ACKNOWLEDGEMENTS

The journey of my PhD was an absolute privilege, and I am grateful for the experience and knowledge I was able to acquire during my time.

I want to thank the Leverhulme Doctoral Scholarship in Behaviour Informatics (DS-2017-015) for funding me and my research throughout my PhD.

The support and trust that Dr Demetris Soteropoulos provided throughout my PhD was an invaluable source of motivation for me – thank you. Your interest in new technologies and your joy in engineering were sparks of inspiration. Even though my project was only adjacent to the focus of your work, I was always one of your priorities and I am so thankful that you accepted me as your PhD student.

Dr Claire Witham, who inspired me early on to follow my interest in the application of computer vision in behavioural research. Thank you for sharing your knowledge and appreciation of Macaques with me throughout the years. Your motivation and dedication to the impactful work you do are truly inspiring for me.

My gratitude also goes to Dr Sasha Kraskov and Professor Stuart Baker for taking me on as a PhD student on a short notice for the end of my PhD. The discussions with you were always very fruitful and I feel my work has gained significance through these.

I am thankful to Dr Matthew Leach for being one of my supervisors, for the help in planning the project and putting me in touch with Dr Lauren Robinson. I really appreciated the input from both of you.

I want to thank my mentors Professor Andy Jackson and Dr Tom Smulders who have provided invaluable feedback and input. Your advice made the running of my PhD smoother than it would have been otherwise.

Norman Charlton, thank you for your help in constructing all the camera casings. Even with short notice you were always extremely helpful.

Terri Jackson, Andrew Atkinson and Dr Rachel Jones, thank you for your advice, all the help with the animals and filling out the questionnaires. The gossip about animals were at times more inspirational than one would expect.

I also want to express my gratitude to the animal technicians at the CBC - Emma, Annie and Sophie. Thank you for always lending me a helping hand and for your time in filling out the questionnaires.

I am thankful to Dr David Massey for sharing the neophobia test data collected for his PhD with me. Thank you for being a friend over these past years and I hope to meet you again soon.

Sensorimotor-group, Natalie, Maria, Mark, Rachel, Sandra, Keval, Viviane, Sandra, Anna, and Annie, thank you for taking me in and being my work family for the last four years. I cherish all the experiences we had together and am eternally grateful that you made me feel at home in Newcastle.

My appreciation also goes to all Behaviour Informatics students that have been a great group of support. I am proud to be part of this program and the amazing work that is happening. I am especially thankful to Dr Giulia Ciminelli who has shared many of the frustrations and delights of the PhD with me.

I want to thank my family and friends for their never-ending support and standing by my side even in the most adventurous journeys. A special thanks goes to my big brother, Dominik, for facilitating my PhD during the more difficult times.

My wife – Fabienne Andres, thank you for being by my side for almost half my life. I am deeply grateful for all the adventures you are enduring with me and the continuous encouragement, help and motivation you provide.

Finally, I want to express my appreciation to all the monkeys at CBC and CFM. It is an absolute privilege to be working with animals of such intelligence and personality.

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

UK – United Kingdom

EU – European Union

NHP – Non-human primate

EMG - Electromyography

CNN – Convolutional Neural Network

m – meters

INBAS – infant neurobehavioral assessment scale

BBA – biobehavioural assessment program

vs – versus

CBC – Comparative Biology Centre

CFM – Centre for Macaques

UKHSA – UK Health and Safety Agency

CCTV – close-circuit television

HDMI – high definition multimedia interface

Ah – Ampere hour

mAP – mean average precision

P - precision

R – recall

IoU – Intersection over union

LSTM – Long-Short Term Memory neural network

RNN – Recurrent neural network

GPS – global positioning system

RFID – Radio-frequency identification

KL – Kullback-Leiber (divergence)

FF – familiar food

NF – novel food

NO – novel object

CHAPTER 1. GENERAL INTRODUCTION

1.1 CHAPTER SUMMARY

Non-human primates serve as invaluable models in neuroscience research, yet ethical considerations persist regarding their use. To address these concerns, refinement of procedures involving these models remains paramount. Central to refinement efforts is the monitoring of animal welfare. However, traditional welfare assessment methods are labour-intensive and complex. Recent technological advancements offer promise in the development of automated complementary systems, presenting opportunities for enhanced refinement practices.

Automated approaches hold the potential for continuous, long-term observation of non-human primate behaviour focused on the individual, thereby advancing refinement efforts. While automated welfare monitoring systems have been established for various species (e.g. in agricultural settings), research specific to automated monitoring of primate welfare remains scarce. Thus, there exists a critical gap in knowledge regarding the application of such methods within non-human primate research.

The primary objective of this thesis is to develop a cost-effective and versatile automated observation method tailored for laboratory-housed Rhesus macaques. This method aims to complement welfare assessment practices by providing longitudinal insights into behaviour and well-being of each individual. By leveraging current state-of-the-art technology, this approach seeks to showcase the usability of the current existing methodologies. Ultimately, this thesis provides a basis for the application of automated methods to assess non-human primate welfare by utilising measures focused on the individual to further refine husbandry and research.

1.2 NON-HUMAN PRIMATES IN RESEARCH

While non-human primates (NHPs) contributed to significant advancements in biomedical research (Lear *et al.*, 2022; Mitchell *et al.*, 2018; Treue and Lemon, 2023), they only make up a small proportion of the animals used in science. 2197 procedures used NHPs in 2022 in the United Kingdom of which 101 were in Rhesus macaques, in comparison to around 890’000 procedures that used mice (UK Home Office, 2022). Compared to non-human primates, rodents are phylogenetically more distant from humans. In some situations, the greater evolutionary relatedness between NHPs and humans makes them more viable models.

In current scientific research the utilisation of animal models remains essential for the translational benefit for human well-being, as currently complete replacement is not yet possible. For a species to be a valid model species three criteria need to be met: face validity, etiological/construct validity, and predictive validity (Willner and Mitchell, 2002).

Face validity indicates a model's capability to exhibit symptoms and behavioural characteristics of the modelled disease as expected in humans. Baboons, for instance, are used as models for osteoporosis because symptoms of the disease naturally emerge in older females, illustrating face validity (VandeBerg and Williams-Blangero, 1997). Baboons and other primate species have also been used to model endometriosis due to physiological similarity in their reproductive organs, including the menstrual cycle (Braundmeier and Fazleabas, 2009; Story and Kennedy, 2004).

Construct or Etiological validity is met when the underlying cause for a disease in a model species shows high similarity with that observed in humans. Further, Rhesus and cynomolgus macaques are frequently used in behavioural manipulations and therapies and show etiological validity (Carlsson *et al.*, 2004).

Lastly, effective interventions or treatments in humans need to be mirrored in the model species to be **predictively valid**. NHPs have demonstrated predictive validity by aiding to answer fundamental scientific questions in the past including addiction, depression, drug therapies, adoption, aging, and abnormal development (Hau and Schapiro, 2006).

Because of their closer phylogenetic distance to humans NHPs match these validity criteria in a large variety of diseases. Neuroscience research involving NHPs has led to the development of deep brain stimulation methods which has since been used to treat symptoms like essential tremor, dystonia and depression (Delaloye and Holtzheimer, 2014; Gardner, 2013; Schlaepfer *et al.*, 2013; Wichmann, Bergman and DeLong, 2018). In immunology, macaques have played a crucial role in comprehending and creating vaccines for SARS-CoV-2, since they display similar symptoms of COVID-19 to humans. This makes them an invaluable model for examining the pathophysiology of the virus and creating preventive and therapeutic measures (Salguero *et al.*, 2021; Urano *et al.*, 2021). Similarly, major advances in vaccines to treat human immunodeficiency virus (HIV) have been made due to NHP models and the presence of similar diseases in primates such as the simian immunodeficiency virus (SIV). This research has progressed successful treatment of other diseases (Veazey and Lackner, 2017).

NHPs do not only serve as a physiological model but are also viable behavioural models which can be of major importance in the study of neurological and neuropsychiatric illnesses and neurodevelopmental diseases such as autism spectrum disorder (Parker, 2022; Scott and Bourne, 2022; Silverman *et al.*, 2022). Rodent models for depression and common behavioural tests such as the forced swim test have been criticised and the naturally occurring complex behavioural repertoire of NHPs shows promising patterns (Molendijk and de Kloet, 2015; Silverman *et al.*, 2022). However, the proper assessment of these behaviours remains challenging.

NHPs encompass a diverse array of species, including apes (such as chimpanzees, the closest relatives to humans), New World and Old World monkeys, and prosimians (such as lemurs, lorises, and tarsiers). Among these, macaques and marmosets are the most commonly used NHPs in the UK (UK Home Office, 2022). Notably, since 1997, research involving great apes has been prohibited in the UK. This prohibition stems from the recognition that great apes exhibit heightened susceptibility to suffering compared to other NHPs, leading to ethical concerns that override potential scientific benefits from using species even more closely related to humans.

However, concerns about the use of NHPs in scientific research have also grown more widespread, particularly among members of the public. These worries result from a growing awareness and understanding in the scientific community and the general public on the welfare and psychological wellbeing of NHPs involved in research (Bailey and Taylor, 2016; Webster, 1995). In light of this concern, recent legislative initiatives have prioritized resolving welfare issues relating to non-human primates in scientific research (Ares, 2022; Hollands, 1986; Olsson *et al.*, 2017). Discussions facilitated by professionals like Lear *et al.* (2022), Treue and Lemon (2023) and Bailey and Taylor (2016) further highlight the current importance of this topic.

Although the current utilization of NHPs in the UK is comparatively low in regards to other species, their welfare, encompassing both physical and psychological health, garners heightened attention due to their closer resemblance to humans (Bateson, Johansen-Berg and Jones, 2011; Workman *et al.*, 2019). As previously stated, this proximity to humans is pivotal, as it underpins crucial advancements in research facilitated by similarities in physiology, neuroanatomy, reproduction, development, cognition, and social abilities, but it also underscores ethical considerations, given the elevated levels of sentience, self-awareness, and human-like suffering (Hau and Schapiro, 2007; Pickard *et al.*, 2013). Despite these ethical concerns, NHPs retain exceptional value in neuroscience, where their unique brain structures and biological systems offer insights distinct from other animal models (Careaga, Murai and Bauman, 2017; Lear *et al.*, 2022; Thiele and Bellgrove, 2018). Because many mechanisms are not yet understood fully, NHPs continue to be required for important scientific and medical advancements (Roelfsema and Treue, 2014).

The 3Rs framework—Replace, Reduce, Refine—is central for humane animal research. In the UK, the National Centre for the Replacement, Refinement, and Reduction of Animals in Research (https://www.nc3rs.org.uk/) serves as the leading national organization dedicated to applying and advancing these principles. Despite extensive efforts to search for alternatives, animal models remain indispensable tools for understanding disease mechanisms and developmental processes. While ongoing exploration of substitutes persists, animal models continue to play a pivotal role in advancing scientific knowledge (Bonini, 2019). Moreover, efforts to reduce the number of animals used in research are already constrained in most studies conducted in both the UK and EU (Vermeire *et al.*, 2017).

Given the challenges associated with replacing NHPs and the low numbers of animals generally involved in a single study, particularly in neuroscience research, refinement and hence improvement of well-being and quality of life emerges as the primary concern within the 3Rs framework. Refinement strategies aim to mitigate the severity of adverse procedures experienced by animal models, including initiatives to minimize pain and distress, improve housing and handling conditions, enhance experimental techniques, refine training and care protocols, and implement comprehensive monitoring and assessment procedures (Vitale, Manciocco and Alleva, 2009). Such refinement efforts are essential for upholding the welfare and ethical standards in animal research while ensuring the reliability and relevance of scientific findings.

1.3 PRIMATE WELFARE

Ensuring animal welfare in research is ethically paramount, especially in NHPs, who are widely acknowledged to possess the capacity to experience positive welfare and suffering (Truelove *et al.*, 2020). Ethical responsibility and respect for the animals involved dictates a prioritisation of welfare while conducting research (Schapiro and Hau, 2023).

A maintained state of positive welfare is beneficial for quality and reliability of scientific data collected, leading to higher robustness and representativeness of their findings. Ultimately, incorporating principles of animal welfare into research practices not only upholds ethical standards, but also contributes to the integrity, credibility, and applicability of scientific investigations. Ensuring higher animal welfare in research settings fosters more natural behaviours, which often provide a more accurate representation of the processes under investigation. When animals experience optimal welfare conditions, they are more likely to exhibit behaviours that align with their natural responses, enhancing the validity of research outcomes (Bayne, 2005; Bayne and Würbel, 2014; Benton *et al.*, 2013; Lambeth *et al.*, 2006;

Laule, Bloomsmith and Schapiro, 2003; Magden *et al.*, 2013; Schapiro *et al.*, 2000; Schapiro, Perlman and Boudreau, 2001; Schapiro *et al.*, 2005). Conversely, inadequate welfare conditions can introduce additional confounding circumstances, such as high stress levels, which may obscure or mask the intended results (Bayne and Würbel, 2014; Hopkins and Latzman, 2017; Poole, 1997; Prescott and Lidster, 2017; Schapiro and Hau, 2023; Schapiro *et al.*, 2000; Shively, 2017). Furthermore, poor welfare has the potential to prolong a study and affects the refinement of experimental procedures. Occultation might also impact replacement efforts, as larger sample sizes could be needed. Moreover, compromised welfare may lead to increased costs associated with the management of physical health issues and potential veterinary interventions, further underscoring the importance of prioritizing animal welfare in research practices (Pickard *et al.*, 2013).

Laboratory training, restriction protocols, chair and head constraints, head implants and subdermal electromyography (EMG) are common procedures in neuroscientific research and raise significant welfare issues due to the stress associated with the surgical procedure, the risk of post-operative infections, and the potentially painful maintenance requirements particularly in NHPs (Johnston *et al.*, 2016). Additionally, the single or pair-housing of Rhesus macaques in neuroscience laboratories are deemed stressful for a species naturally accustomed to living in large multi-male/multi-female groups in the wild (Chatfield and Morton, 2018). Hence, it is important to measure levels of welfare especially for macaques used for neuroscientific research.

Various indicators are used to assess animal well-being. In the context of non-human primates, key indicators for welfare assessment include body weight, alopecia, physiological markers, behavioural patterns, hippocampal indicators, and telomere biomarkers (Castellano Bueno, 2022). Nevertheless, most of these estimates require invasive procedures like blood sampling or brain scans. For that reason, behaviour is commonly considered as a less invasive measure for general assessment of an animal's state. This can entail measurement of all possible behaviours and monitor changes in any occurrences or by focusing on more specific behaviours known to be indicators of welfare. Most common examples of the latter are abnormal, anxiety-like and affiliative behaviours (see Lutz and Baker, 2023).

Abnormal behaviours deviate qualitatively and quantitatively from the typically expressed behaviours (Camus *et al.*, 2013; Erwin, 1979). For example, stereotypies describe invariant and repetitive behaviours that can be indicative of negative welfare (Mason and Latham, 2004). These can involve whole body movement (Poirier and Bateson, 2017) or self-directed behaviour that lack apparent ecological relevancy or context such as hair pulling (Bayne and

Novak, 1998). In rare cases such self-directed behaviours can even lead to injuries (Anderson and Chamove, 1980; Lutz, Well and Novak, 2003).

Anxiety-like behaviours can also be self-directed such as self-grooming, yawning or scratching and are sometimes called displacement behaviours (Maestripieri *et al.*, 1992) and frequency of these behaviours decreases with the use of anxiolytics (Schino *et al.*, 1996; Schino *et al.*, 1991). Inactivity is also used as an indicator for anxiety and is expressed as a hunched posture (Harlow and Zimmermann, 1959) or in more extreme cases in the form of freezing (Kalin and Shelton, 1989). Inactive behaviours are sometimes described as depressive-like behaviours and have also been experimentally validated using social isolation (Hennessy, Chun and Capitanio, 2017; Hennessy *et al.*, 2014), shortening light exposure (Li *et al.*, 2013) and pharmacological interventions (Perera *et al.*, 2011; Qin *et al.*, 2015).

Affiliative behaviours serve as a positive measure of welfare as most non-human primates are a highly social species and potentiating the expression of social behaviours can improve their conditions (Baker *et al.*, 2012).

While behavioural measures are very useful in assessing welfare states behavioural expression and reactivity may vary between individuals (Koolhaas, 2008). Further, even when exposed to identical stressors and environments, individual NHPs may exhibit diverse affective states due to differences in genetics, personality, and coping mechanisms. These variations hinder the reliable assessment of welfare states solely based on environmental factors (Dawkins, 2012; McFarland, 1989; Robinson *et al.*, 2018b).

1.4 TEMPERAMENT AND PERSONALITY

A central framework of behavioural ecology is individual optimisation. Following this, an individual will adapt its behaviour to the environment to maximise its individual fitness (Brommer and Class, 2017). While this represents behavioural traits in the evolutionary context, optimisation theory is limited in explaining variation in behaviour. Variation in optimisation theory is either viewed as noise around a single adaptive optimum (Mather and Anderson, 1993) or as being one of multiple co-existing evolutionary stable strategies (see Weiss and Adams, 2013). To provide a broader concept for between individual variation in behaviour, even in sub-optimal tendencies, the concept of animal personality was introduced (see Carter *et al.*, 2013; Whitham and Washburn, 2017).

The concept of animal personality is defined as corresponding to “individual behavioural differences that are consistent over time and across situations” (Réale *et al.*, 2007, p. 294).

Among the most common personality traits are boldness, exploration-avoidance, activity, aggressiveness, and sociability (Réale *et al.*, 2007; Weiss, 2017). This concept explains variation in behaviour based on cost-benefit trade-offs within an ecological framework (Dingemanse and Wolf, 2010; Nettle, 2006; Wolf and Weissing, 2010). As an example, a ‘bold’ individual will be more likely to gain access to new territories but might also be at a higher risk of predation.

The terminology of personality and temperament is still debated (see Réale *et al.*, 2007). Temperament has been described as a component of personality, which defines individual differences in percept-based skills, habits and being more related to learning and the procedural memory (Cloninger, 1994). On the other hand, they can be distinguished on a measurement level; with personality being a measure of an individual’s variation in reference to a dimension existing in a population, and temperament being a response of an animal measured on a biological scale within a single context (MacKay and Haskell, 2015). In this thesis, I treat these terms as interchangeably.

1.5 THE LINK BETWEEN TEMPERAMENT AND WELFARE

Understanding inter-individual variation improves the understanding of behavioural needs for each animal and aids in finding efficient adaption of procedures accordingly (Blaszczyk, 2020; Loyer and Ha, 2017). Measures of personality have shown to affect social interactions between animals and can hence be used for appropriate selection for cage mate selection (Capitanio *et al.*, 2017b; Weinstein and Capitanio, 2012). Since temperament has also been related to measures of performance in specific experimental tasks, animal selection for experiments can be refined (Coleman, Tully and McMillan, 2005). Furthermore, personality scores can aid in identifying animals at higher risk of injuries, illness, and higher social instability, which could help the allocation of potential observation and handling needed from caretakers (Gottlieb, Capitanio and McCowan, 2013; Gottlieb *et al.*, 2018a; Robinson *et al.*, 2018b). Finally, focusing on measuring personality in laboratory-housed animals can help to identify constraints in behaviour, observed for example when the expression of two behaviours is not completely independent. This is a concept known as behaviour syndrome that describes the correlation of the average expression of two or more behaviours (Sih et al., 2004). While a behavioural syndrome can occur naturally, it might be useful to identify the origin of abnormal behaviours from potential constraints that have been induced by captive housing conditions, for example some animals might have a higher inclination to movement which could lead to repetitive movement (Poirier and Bateson, 2017).

As stated in Section 1.3, the welfare of an animal can impact the quality and potential cost of research. The assessment of an individual's ability to cope or act in a research context could positively impact animal welfare as well as lead to further refinement potentially improving control of confounds and strengthen research outcomes (Bethell *et al.*, 2019b; Buchanan-Smith, 2006; Jennings, Prescott and Refinement, 2009; Prescott and Lidster, 2017; Würbel, 2002).

1.6 AUTOMATED APPROACHES TO MEASURE PRIMATE BEHAVIOUR

Behavioural observations have traditionally been employed for monitoring and assessment of animal welfare as a means of refinement (see section 1.3). However, such methods are susceptible to observer bias, labour-intensive, and potentially disruptive due to the presence of human observers (Anderson and Perona, 2014; Ardoin and Sueur, 2023; Gaither et al., 2014; Mathis et al., 2018; Porto et al., 2013; Weinstein, 2018). Hence, there is a growing interest towards automating behavioural analysis (Knaebe et al., 2022). This necessitates the identification of individuals and the classification of their actions. In the past some of the automated approaches involved invasive procedures such as collaring animals (Ballesta et al., 2014; Morita et al., 2020), coloured jackets (Rose et al., 2012) or RFID tacks (Maddali et al., 2014). Recent advances in computer vision technology and increased accessibility of these approaches facilitated the development of non-invasive video tracking methods. These are mainly divided into pixel-based analysis or action classification using pose estimation. For a collection of available methods suitable for a variety of species see Figure 1-1 and Table 1-1 adapted from (Wiltshire et al., 2023).

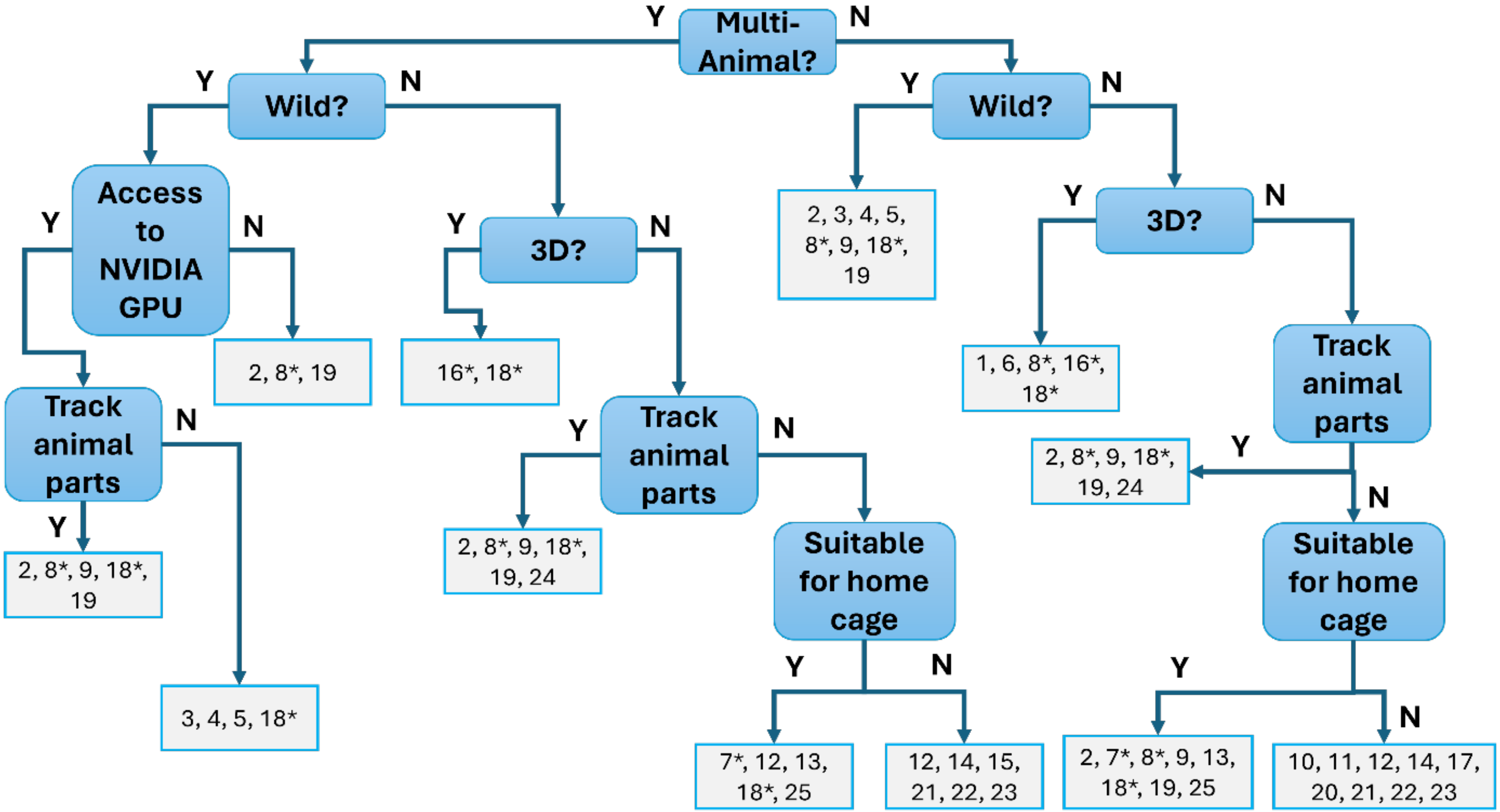

Figure 1-1, Decision tree of available methods for animal tracking in videos, adapted from Wiltshire et al. (2023). Includes only methods described to be suitable for a variety of species or primates. Asterixis indicate methods tested on primates.

Number	**Software**	**Multi-Animal**	**Used on 'Wild' Data?**	**Species**	**Tracking Style**	**Notes**
1	Anipose			Any (Authors tested on flies, mice, humans)	3D Pose estimation	Uses 2D DeepLabCut predictions.
2	APT	✓	✓	Any (Authors tested on flies, mice, larva)	Pose estimation	
3	Biosense	✓	✓	Any (Authors tested on insects, fish tadpoles)	Centroid	Region preference, speed, distance travelled, nearest neighbour distance.
4	BioTrack	✓	✓	Any (Authors tested on humans, ants, termites)	Bounding box & centroid	Only for linux and Mac. Uses ladar (laser hits) to detect objects of interest so requires multiple lasers set up to cover the experimental area. Real-time tracking. Has been used in the 'field' for humans (football game) and suggests the same could be done for animal tracking.
5	BioTracker	✓	✓	Any (Authors tested on fish)	Centroid	
6	DANNCE			Any (focus on rats)	3D pose estimation	7 million pre-marked rodent frames. Software can be further trained for other animals.

7*	DeepBhvTracking	✓		Any (Authors tested on mice and marmosets)	Centroid	Can only force multi-animal with physical markers.
8*	DeepLabCut	✓	✓	Any (Authors tested on mice, flies, cheetah, humans)	3D & 2D Pose estimation	Available on Google Collaboratory.
9	DeepPoseKit	✓	✓	Any (Authors tested on insects, flies, zebra)	Pose estimation	Can only force multi-animal and will not work on novel videos.
10	EthoVision			Any	Centroid & 3-point tracking	Charge to use. Marks behaviours, heatmaps, zones of interest, gait analysis. Live tracking. Multiple arena tracking.
11	Ethowatcher			Any (Authors tested on rats)	Centroid	Length and orientation tracking. Frequency and duration of user-defined behaviours.
12	Idtracker.ai	✓		Any (Authors tested on fish, mice, ants, flies)	Centroid	Limit 100 individuals. Numbered individuals.
13	IDTracker	✓		Any (Authors tested on mice, ants, fish, flies)	Centroid	Uses MATLAB.
14	MARGO	✓		Any (Authors tested on flies, bees, worms and zebrafish)	Centroid	Multiple arenas. MATLAB based. Real-time tracking.
15	Multi-Animal Tracker	✓		Any (Authors tested on worms, flies, fish)	Centroid	Mark animals' location on a few frames and the software completes the rest for that video. Uses MATLAB. Tracks speed and direction as well as

						branching into behavioural recognition.
16*	OpenMonkeyStudio			Macaques	3D pose estimation	Requires specific cage apparatus.
17	Pathtrackr			Any (Authors tested on spiders)	'blob', centroid & bounding boxes	Does not require a homogenous background that is always lighter or always darker than the animal Uses R. Tracks direction, distanced moved and velocity.
18*	SIPEC	✓		Any (Authors tested on mice, macaques)	Pose estimation & bounding boxes	Can use for segmentation, pose estimation, individual identification, behavioural classification, 3D tracking.
19	SLEAP	✓	✓	Any (Authors tested on flies, mice, bees, gerbils)	Pose estimation	Available on Google Collaboratory.
20	Spectral time-lapse			Any (Authors tested on pigeons, mice, ants)	Centroid	Uses MATLAB. colour maps with animal's outline which show position at user defined intervals.
21	SwisTrack	✓		Any (Authors tested on cockroaches and robots)	Centroid	Mixed species tracking in the same frame. Specialised for swarm dynamics. Does not require python to download.
22	ToxTrack	✓		Any (Authors tested on insects, fish, rodents)	Centroid	Does not require python to download. Multiple arenas. Arena coverage, arena heat maps, speed

						and distance tracking. Only for Windows.
23	Traktor	✓		Any (Authors tested on fish, spiders, termites)	Centroid	Tracks regions of interest, distance, speed, activity, distance to conspecifics. Can track a single individual in a ‘noisy’ laboratory environment.
24	Trex	✓		Any (Authors tested on fish, insects, flies)	Pose estimation	Specialised in large number of individuals (eg- swarms and schools). Up to 100 individuals whilst retaining ID, 256 without ID.
25	UMATracker	✓		Any (Authors tested on mice, ants, termites, fish, flies)	Centroid & bounding boxes	Number of individuals in the video must be constant. Regions of interest, Interaction graphs.

Table 1-1, Available methods for animal tracking in videos, adapted from Wiltshire et al. (2023). Includes only methods described to be suitable for a variety of species or primates. * Indicate methods tested on primates

Pixel based approaches for identification or action detection in primates make use of the information held in the image. For example, face detection algorithms have been used to detect regions of interest in an image (i.e. the animals face) to extract local binary patterns to classify individuals using common machine learning approaches (Witham, 2018). Local binary pattern is an image processing technique that measures information of texture in an image. Each pixel is compared to the brightness of neighbouring pixels and assigned a binary value. Alternatively, the pixel values can be used as a direct input for a deep convolutional neural network from the face (Schofield et al., 2019; Schofield et al., 2023; Ueno et al., 2019; Ueno et al., 2022) or the full body of an animal (Marks et al., 2022). Convolutional neural networks (CNN) are deep learning models inspired by the structure and function of the human visual system. CNNs primarily focus on images as inputs, hence can deal with three dimensions, and generally consist of three types of layers. The convolutional layer filters the input image over the multiple dimensions with the use of kernel filters, which generate feature maps that identify the presence of specific patterns. The pooling layer then uses varying operations to reduce the spatial dimensions of the feature maps. Finally, the fully connected layer has neurons connected to all activations in the previous layer to map feature to the output classes (O'Shea and Nash, 2015). Similar approaches have been used to classify animal behaviour from video data alone (Ardoin and Sueur, 2023; Marks et al., 2022) or with additional use of audio data (Bain et al., 2021).

Other approaches have made use of CNNs to estimate the pose of an animal, hence tracking individual body parts within each frame (see Blanco Negrete et al., 2021; Labuguen et al., 2021). The resulting pose data need further analysis to be classified into behaviours, this has been done using 2D pose estimates (Li et al., 2023) or with the use of stereovision approaches (using more than one camera) in 3D (Bala et al., 2020).

1.7 THESIS AIMS

This thesis aims to adapt current available methodologies to assess individual measures of behaviour in the context of welfare and personality. No study has applied automated methods to assess changes in behavioural patterns of laboratory-housed Rhesus macaques in the scope of welfare, this thesis addresses this gap by:

- Summarise the current knowledge of personality in Rhesus macaques in a scoping review to identify existing methods and associated measures (Chapter 1)
- Developing an in-cage camera module that is capable of capturing video footage over a long period of time (Chapter 3)
- Describing and validating a pipeline for automated tracking and identification of laboratory Rhesus macaques (Chapter 3)
- Confirming the ability of the developed method to obtain data holding information on individuality and consistency (Chapter 4)
- Utilising the pipeline to associate divergence of behavioural patterns in context of events related to welfare (Chapter 5)
- Showcasing associations between personality trait ratings and behavioural measures obtained with the pipeline (Chapter 6)

All the code for analysis and the camera module and pipeline is available in the following repository: https://github.com/jreuka/CBC_pipeline/

CHAPTER 2. TEMPERAMENT AND PERSONALITY, A MEASURE OF CONSISTENCY IN RHESUS MACAQUES – SCOPING LITERATURE REVIEW

2.1 CHAPTER SUMMARY

This Chapter presents a scoping review of the existing literature concerning personality and temperament of Rhesus macaques. The review identifes prevalent methodologies utilized for measurement including the timescales of observations and age of observed individuals as well as factors and behaviours associated with temperament and personality metrics.

2.2 INTRODUCTION

Animal personality or temperament describe consistent individual differences over context and time (see Réale *et al.*, 2007). Personality in nonhuman primates has been studied throughout the last decades and has proven to have construct validity and reliability (Freeman and Gosling, 2010). The study of nonhuman primate personality has been of interest to a range of disciplines such as psychology, anthropology, endocrinology, ecology and animal housing and welfare (see Freeman and Gosling, 2010; Coleman, 2012; Norman, Rowden and Cowlishaw, 2021). This presents a variety of perspectives and can lead to generalised conclusions. Nonhuman primate personality was associated with individual measures and ratings of animal wellbeing, health, and welfare (Robinson *et al.*, 2021a; Robinson *et al.*, 2018a). Further, quantitative measures of personality hold the potential to identify periods where an individual might be at higher risk of stress (Gottlieb *et al.*, 2018b), improving the success of social introductions (Ruhde *et al.*, 2020; Capitanio *et al.*, 2017a), select animals suitable for laboratory work (Zijlmans *et al.*, 2021; Coleman, Tully and McMillan, 2005) and reduce stress through tailored enrichment programs (Franks *et al.*, 2013; Gartner and Weiss, 2018; Hopper *et al.*, 2014). While the literature on nonhuman primate personality has been synthesized before (see Freeman and Gosling, 2010; Coleman, 2012; Norman, Rowden and Cowlishaw, 2021), these reviews included multiple species, thus having a broader picture of results rather than a focus on Rhesus macaques. Freeman and Gosling (2010) presented an overview of validity in studies involving 28 species of primates and identified gaps in the knowledge, such as missing practical applications in welfare and wildlife management. Similarly, the review of Norman, Rowden and Cowlishaw (2021) covered 28 species and discusses how personality can be applied in the fields of welfare and animal management. Their review further discusses the standardisation of tools for easier implementation and synthesis of results.

Here I present a literature review summarising the methods used to measure personality and temperament in Rhesus macaques, both in the wild and in captivity. Further, behavioural expression and physiology associated to personality measures are extracted and summarised to give an overview of the current knowledge. Based on the findings in the existing literature, I drew conclusions on which measures to include in my thesis.

2.3 LITERATURE SEARCH PROCEDURES

This review synthesized current knowledge on temperament and personality in Rhesus macaques. To review the literature a search was conducted on the 17th of February 2022 on the databases Ovid, Scopus and WebofScience. The keywords used for the search were based on the wording of Norman, Rowden and Cowlishaw (2021) and (Freeman and Gosling, 2010) and included the words “personality”, “temperament”, “Rhesus macaque”, “Rhesus monkey” and “Macaca mulatta” (i.e. ((“personality” OR “temperament”) AND (“Rhesus macaque” OR “Rhesus monkey” OR “Macaca mulatta”)).

The initial search yielded 1698 articles and were collated using Endnote (https://endnote.com/). All duplicates and articles ineligible by the automation tool were removed leading to 1002 records for screening. For an article to be included, it needed to be a primary peer-reviewed publication that measured temperament or personality and was written in English. Further, the measures needed to be included in the results to exclude studies that just mentioned the use of scanning for temperament without presenting their results. Articles that mentioned general concepts such as anxiety but did not mention these in context of temperament or personality were also excluded. This led to an inclusion of 148 articles ranging from November 1972 to October 2021 (see Flow Diagram in Figure 2-1).

2.4 OVERVIEW OF RESEARCH APPROACHES AND CORRELATES

2.4.1 SUBJECT INFORMATION

Overall articles, single component scores for more than 8000 individuals were presented. Articles measuring multiple personality components included in total more than 15’000 individuals. For both the number of males and females comes to roughly the same, with slightly less males being rated. It is difficult to report the exact numbers as it is not always indicated which animals derive from the same study. The median number of animals scored on temperament in a study is 36.5 and 44 for the single and multiple component components respectively.

The age of animals rated ranged from 7-day old infants up to 31-year-old adults. As the two assessment programs BBA and INBAS were used in 62 of the 148 studies and these score temperament during infancy there is a bias towards younger ages in the articles presented here (see Table 2-1 and Table 2-2).

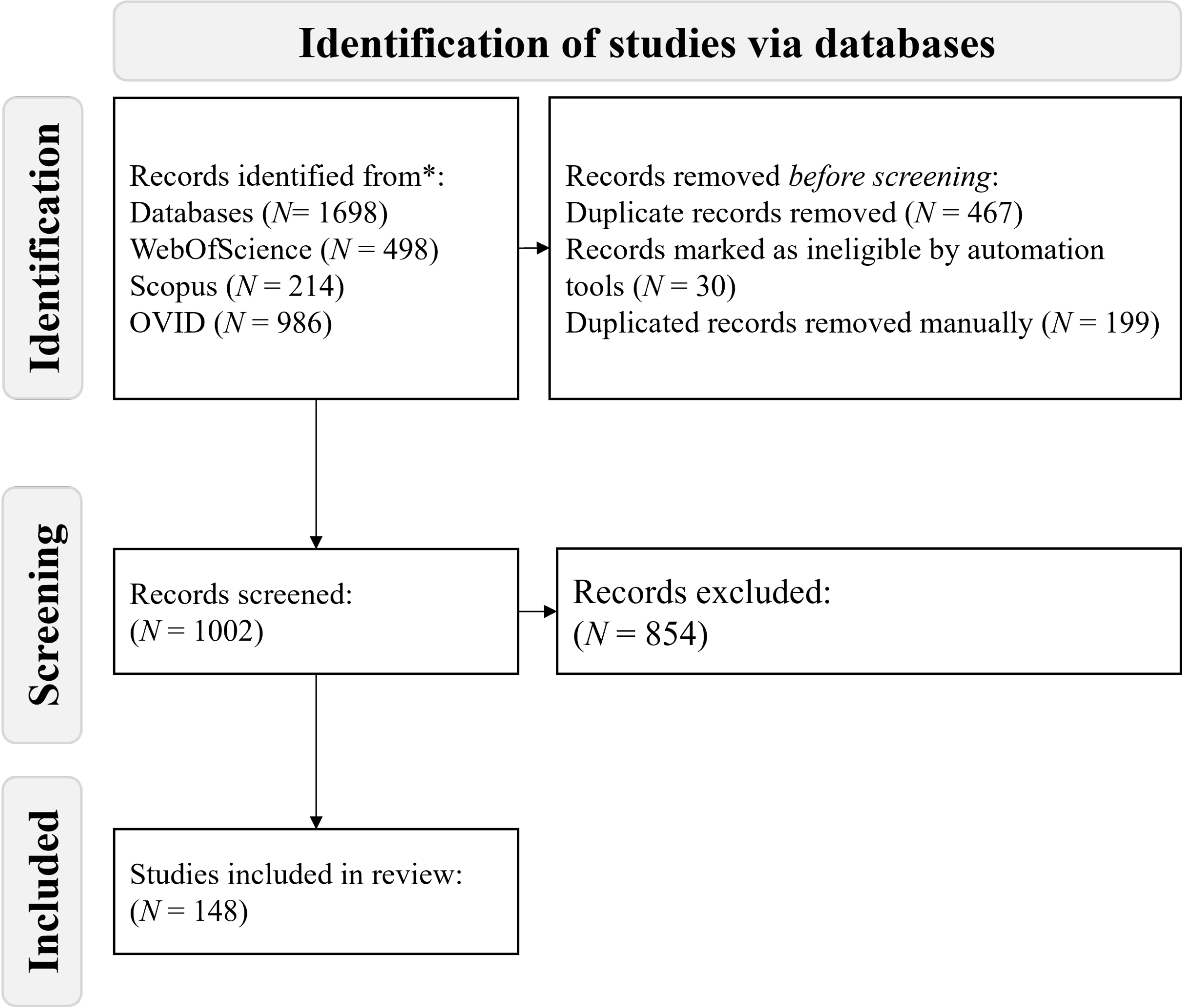

Figure 2-1, Flow Diagram of the selection process for literature search.

The timescale of observation to score an individual’s temperament ranged from 20 minutes in the INBAS assessment (Schneider *et al.*, 1991) to multiple hundred hours per individual (Bolig *et al.*, 1992). Overall, most common time periods were between 1-8 hours of observation per individual.

2.4.2 ASSESSMENT OF PERSONALITY

Methods used in included articles can be grouped into three categories: behavioural coding, behavioural testing and subjective ratings of personality and behavioural expression (Freeman and Gosling, 2010). Out of the 148 identified articles 50 used the above-described methods to score Rhesus macaques on a single component (see Table 2-1) and 98 studies measured multiple components (see Table 2-2). The most common method to assess single components of personality were behavioural tests with 38 out of the 50 studies using these methods and an additional three using tests in combination with subjective ratings. For multiple components subjective rating was the most common method with 79 of the 98 studies applying this method.

Behavioural coding records behaviour of an individual according to a variety of discrete behaviours observed in its species (Gosling, 2001; Vazire et al., 2007). This is generally done in animals in the wild or their home enclosure, however, can be part of testing procedures as well. However, it is often not feasible to measure every behaviour that occurs and hence it is common practice to create an ethogram describing behaviours of interest with definitions of behaviours and if these are recorded by presence, frequency, duration, intensity or latency (Bateson and Martin, 2021). Two examples for behavioural coding are one-zero sampling, where in a given interval the presence or absence of each behaviour is noted, as well as focal recordings, in which an individual is followed for a certain timeframe and its behaviour recorded in regular intervals and/or ad libitum (Altmann, 1974).

Behavioural tests are experimental tasks that aim to assess variability in personality traits by controlling experimental stimuli. These tasks include tests like the human intruder test (Kalin and Shelton, 1989), the playroom test (Haley *et al.*, 2009), the anticipatory anxiety test (Gibboni, Zimmerman and Gothard, 2009), the snake test (Kalin *et al.*, 2001) and the novel object test (Coleman, Tully and McMillan, 2005) (described in detail below). With these methods, some issues arise due to the use of “many tests for one trait”, which do not always correlate (e.g. anxiety has been tested using both novel object and human intruder) and “one test for many traits”, which can lead to confusion of what trait is actually measured (e.g. novel object test have been used to measure both the anxious and exploratory component) (Carter et al., 2013; Brommer and Class, 2017; Ruhde et al., 2020; Zijlmans et al., 2021).

Kalin and Shelton (1989) tested infant Rhesus monkeys with the **human intruder test** in two different situations. The general procedure of the test included the separation of the infant from its mother and placing it in a cage in a different room. The animal was left alone for 10 minutes at which point a human intruder entered the room. The first stimuli involved no eye contact from the human intruder. The experimenter entered the room and remained motionless with a defined distance (2.5 m) from the cage presenting the profile. The second stimuli involved the intruder to continuously stare at the infant. Only one of the stimuli was given during a single test but tests were repeated a week later with the other stimuli. The human intruder exited the room after 10 minutes and the infant was left alone for another 10 minutes before returning to its mother. The whole testing procedure was videotaped and a variety of behaviours, such as vocalisation and activity were scored. Kalin and Shelton (1989) found that in the first alone phase infants were very active and frequently emitted "coo" vocalisations, a known distress call of Rhesus macaques. The behaviour of infants differed during the different human intruder conditions. During the 'continuous stare' condition, infants increased their frequency of vocalisation both for "coo" vocalisations and "barking". "Barking" is an aggressive vocalisation that rarely occurred outside of this experimental condition. During the 'no eye contact' stimuli animals showed a significant increase in freezing behaviour. This is defined as the body remaining motionless with only slight head movements for at least three seconds.

Haley et al. (2009) tested adult female Rhesus macaques (age range 19-25 years) using a **playroom**. The playroom was 2.44 x 3.45 m in dimension and at the far end from the door, ten ports were mounted. Each port was 5 cm apart and had a door with a sensor to detect opening events. Animals could be observed through two one-way windows and were videorecorded. The test consisted of four phases. First, during the acclimation phase (phase 1) the animal was placed in the room with all ports baited with food rewards and the doors to the ports stood open. If animals received food from eight ports in two consecutive trails, they moved to the second phase. The training phase (phase 2) consisted of all ports being baited but the doors being shut. Again, when animals retrieved eight rewards during two consecutive trials they moved to the third phase. In phase 3, only one pre-assigned port was baited. An animal met the criterion when it retrieved the reward opening a maximum of two doors and did so in two consecutive trials. In the shift phase (phase 4), two different pre-assigned port were baited. Animals completed this phase when they retrieved the reward on two consecutive trials within the first two searches in each trial. The sequence and number of ports opened were recorded and coded. Animals that did not show preference for one area in the room were scored as low in anxiety (Haley *et al.*, 2012a).

Gibboni, Zimmerman and Gothard (2009) used the **anticipatory anxiety test** to highlight differences in behaviour and temperament. For this test a familiar caretaker approached the cage with a bucket of food and remained standing in front of the cage for five minutes without delivering the food. During this time the caretaker did not make eye contact with the animals.

Pacing, yawning, sit still, vocalisations, head bobs and turn in place behaviours were recorded during these five minutes of test. This test was only conducted once. Gibboni, Zimmerman and Gothard (2009) describe in detail differences between animals during the test but also derived a general rating of motivation from this test.

Kalin *et al.* (2001) examined animals for fear response using a **snake test**. The test involved a clear plastic box with a stimulus inside and the most preferred foods that were placed on top of the box. The stimuli included an empty box, a black masking tape roll, a fake snake, and a live snake. The latter two stimuli were never presented during the first four trials of the day nor in three consecutive trials. Otherwise, the order was pseudo randomly assigned and each stimulus was presented six times for each test day. There was a waiting period of 45 seconds between each 60 second trial. The latency to reach for each food item as well as withdrawing behaviour was recorded. Withdrawing behaviour mainly occurred during the live snake stimulus. These measures were used to score animals' anxious temperament.

Coleman, Tully and McMillan (2005) this study conducted a **novel object test** to measure temperament. This involved the experimenter to stand approximately 1 m from the cage without making eye contact for 5 minutes. After this a new food item was placed through an opening in the cage and left for 3 minutes. To avoid hunger as a factor this test was conducted within two hours of feeding. During the approach phase, self-directed aggressive behaviour, overgrooming and stereotypic behaviours were recorded. During the food phase the latency to inspect the food item was recorded. The latency measure was used to categorise animals into exploratory, moderate or inhibited categories.

Finally, **subjective ratings** make use of the familiarity of the observers with individuals to measure their personality using rating scales of multiple behavioural descriptors (Carter et al., 2013; McCrae and Weiss, 2007). This method is derived from research in human personality traits and has been adopted for understanding animal personality (e.g. King and Figueredo, 1997). This approach is often based on a questionnaire to rate individuals and has demonstrated interrater and test-retest reliability (Freeman and Gosling, 2010) and behavioural validity (Morton et al., 2013). A common theory is the five-factor model, which is based on the reduction of multiple ratings to personality traits, clustering them into the big five-personality traits: boldness, exploration-avoidance, activity, aggressiveness and sociability (Réale et al., 2007; Weiss, 2017). The reduction is often done using factor analysis or component analysis which provides explanation of each factor's observed pattern of variation (see Carter et al., 2013; Weiss, 2017).

In the scanned articles two major assessment programmes used for obtaining a variety of measures in infant Rhesus macaques, including temperament. The **Infant Neurobehavioral assessment scale (INBAS)** consisted of a 20-minute battery of developmental tests. The infants (age range around 7-30 days) were tested at multiple points during development. The INBAS included tests for motor function and orientation that involved the experimenter to handle the infant. During the testing procedures the experimenter observed temperamental characteristics such as irritability, consolability, fearfulness, response intensity, and persistence and these were rated at the end of the assessment session (Schneider *et al.*, 1991; Schneider and Suomi, 1992b). For this reason, I grouped studies using the INBAS into the subjective rating rather than the behavioural testing category. The second program is the **biobehavioural assessment program (BBA)**. Again, this program focused on assessing behavioural and physiological responsiveness in slightly older infants (age range around 90-120 days). The duration of the assessment is 25 hours in which the animals are separated from their mothers and relocated into individual indoor cages. While infants were in individual cages five to eight animals were tested at the same time. The holding cage contained enrichment, food and water. During the 25-hour period animals go through a variety of tests for which they are removed from their holding cage. The tests include novel object tests, blood sampling, video playback and human intruder tests (see Golub *et al.*, 2009). After the 25-hour period infants were reunited with their mothers. The BBA included infants being rated on a list with 16 adjectives at the end of the testing period by the experimenter (Golub *et al.*, 2009). Hence, articles using the BBA are categorised into the subjective rating category except for five studies (Baxter *et al.*, 2021; Linden *et al.*, 2018; Golub and Hogrefe, 2015; Capitanio *et al.*, 2012; Kinnally *et al.*, 2010) that specifically mention the use of behavioural coding data rather than ratings. Overall, 31 articles used the INBAS and another 31 used the BBA and as part of this assessed infant's temperament.

Method	**Reference**	**Trait/Component**	**Age**
behavioural coding	Kemp *et al.* (2017)	Affiliative/cooperative	29-220 months
	Wood *et al.* (2020b)	Anxious	6 months
	Maestripieri (2000)	Emotionality	Adult
	Erickson *et al.* (2005)	Withdrawal	6-7 months
behavioural testing	Davidson, Kalin and Shelton (1993)	Anxiety	1.08-1.24 years
	Kalin *et al.* (1998)		8 months
	Kalin *et al.* (2001)		~22.5-27.9 months
	Kalin, Shelton and Davidson (2007)		~35 months
	Fox *et al.* (2008)		-
	Rogers *et al.* (2008)		Young
	McCormack *et al.* (2009)		3 months

	Oler *et al.* (2009)		~3 years
	Oler *et al.* (2010)		0.74-4.2 years
	Corcoran *et al.* (2012)		6.6-14.6 years
	Fox *et al.* (2012)		
	Haley *et al.* (2012a)		21-27 years
	Haley *et al.* (2012b)		18-27
	Rogers *et al.* (2013)		0.74-4.2 years
	Shackman *et al.* (2013)		
	Birn *et al.* (2014)		1.7-3.5 years
	Roseboom *et al.* (2014)		
	Fox *et al.* (2015)		Young
	Kalin *et al.* (2016)		1.76-2.63 years
	Coleman *et al.* (2017)		3-30 years
	Hamel *et al.* (2017)		3-30 years
	Fox *et al.* (2018)		
	Fox *et al.* (2019)		~3 years
	Tromp *et al.* (2019)		0.84-4.42 years
	Kovner *et al.* (2020)		1.8-2.7 years
	Neudecker *et al.* (2021)		2 years
	Coleman, Tully and McMillan (2005)	Behavioural inhibition	5-7 years
	Shackman *et al.* (2017)		1.45-3.42 years
	Bethell *et al.* (2019a)		4-12 years
	Ruhde *et al.* (2020)		4.7-12.8 years
	Coleman *et al.* (2003)		3-6 months
	Gottlieb, Maier and Coleman (2015)	Boldness	4.8-15.6 years
	Wooddell *et al.* (2019)	Emotional reactivity	14-30 days
	Zijlmans *et al.* (2021)	Exploratory tendency	2.5-21.4 years
	Gibboni, Zimmerman and Gothard (2009)	Motivation	6-10 years
	Maness *et al.* (2019)	State control	14 days
	Paukner, Capitanio and Blozis (2020)		7-30 days
	Wood *et al.* (2020a)		14 days
combined testing & rating	Wood *et al.* (2021a)	Anxious	3-4 months
	Chun *et al.* (2013)	Behavioural inhibition	93-124 days
	Parker *et al.* (2018)	Sociability	3-4 months
subjective rating	Doudet *et al.* (1995)	Aggression	Adult
	Heinz *et al.* (1998)		5 years
	Boyce *et al.* (1998)	Behavioural inhibition	1-19 years
	Boyce (2004)		1-19 years
	Isakova *et al.* (2017)	State control	3-6 months

Table 2-1, Category of scoring method, citation, component, and age for articles that measured a single component.

Method	**Reference**	**Age**
behavioural coding	Chamove, Eysenck and Harlow (1972)	9-12 months
	Suomi, Novak and Well (1996)	>2 years
	Bardi and Huffman (2006)	0-9 months
	Brent *et al.* (2014)	
	Kohn *et al.* (2016)	3-20 years
	von Borell, Kulik and Widdig (2016)	0-7 years
behavioural test	Barr *et al.* (2008)	24 weeks & 3-9 years
	Kinnally (2008)	5-10 years
	Kinnally *et al.* (2010)	90-120 days
	Capitanio *et al.* (2012)	90-126 days
	Golub and Hogrefe (2015)	
	McClintick and Grant (2016)	3.8-6.3 years
	Linden (2018)	90-120 days
	Baxter *et al.* (2021)	3-4 months
subjective rating	Schneider, Kraemer and Suomi (1991)	4 weeks
	Schneider *et al.* (1991)	4 weeks
	Schneider, Coe and Lubach (1992)	2&4 weeks
	Schneider and Suomi (1992a)	4 weeks
	Schneider and Suomi (1992b)	4 weeks
	Bolig *et al.* (1992)	1-14 years
	Stevenson-Hinde and Zunz (1978)	>1 year
	Stevenson-Hinde, Stillwell-Barnes and Zunz (1980)	>1 year
	Champoux, Higley and Suomi (1997)	>6 months
	Champoux *et al.* (1997)	7 days & 14 days
	Clarke and Snipes (1998)	3-12 months
	Lidow (1998)	2,4 weeks & 2 months
	Capitanio (1999b)	6.2 years
	Laudenslager *et al.* (1999)	~3 years
	Schneider, Moore and Becker (2001)	4-23 days
	Champoux *et al.* (2002a)	7-30 days
	Champoux *et al.* (2002b)	7-30 days
	Maninger *et al.* (2003)	~7.3 years
	Capitanio and Widaman (2005)	5-10 years
	Golub *et al.* (2005)	5-13 months
	Schneider *et al.* (2005)	4-5 years
	Machado and Bachevalier (2006)	2.4-3.2 years
	Golub, Hogrefe and Germann (2007)	8-12 months
	Isakova *et al.* (2007)	7 days - ~ 6months
	Kraemer *et al.* (2008)	1month
	Machado *et al.* (2008)	Adult
	Weinstein and Capitanio (2008)	90-120 days
	Golub *et al.* (2009)	Infants
	Bliss-Moreau *et al.* (2010)	9-18 months
	Coe *et al.* (2010)	2 weeks
	Hinde and Captanio (2010)	3-4months
	Kay *et al.* (2010)	7 days
	Bachevalier, Machado and Kazama (2011)	Adult

	Capitanio, Mendoza and Cole (2011)	89-133 days
	Capitanio *et al.* (2011)	92-122 days
	McCowan *et al.* (2011)	Adults
	Rommeck *et al.* (2011)	90-115 days
	Sullivan *et al.* (2011)	91-124 days
	Sullivan, Mendoza and Capitanio (2011)	90-125 days
	Vandeleest and Capitanio (2012)	0.25-0.36 years
	Weiss, Adams and Johnson (2011)	1.6-24.4 years
	Weiss *et al.* (2011)	1.5-13 years
	Willette *et al.* (2011)	2 weeks
	Golub, Hogrefe and Unger (2012)	3-4 months
	Howell *et al.* (2012)	0.7-19 years
	Vandeleest and Capitanio (2012)	3.2-4 months
	Weinstein and Capitanio (2012)	3-4 months
	Willette *et al.* (2012)	19-31 years
	Zhang *et al.* (2012)	7days - 6 months
	Bauman *et al.* (2013)	1-6 months
	Gottlieb, Capitanio and McCowan (2013)	90-120 days
	Jiang, Kanthaswamy and Capitanio (2013)	90-130 days
	Suarez-Jimenez *et al.* (2013)	7,14,21,30 days
	Sussman *et al.* (2013)	>1 year
	Vandeleest, Mendoza and Capitanio (2013)	4.5-5 months
	Beisner and McCowan (2014)	3 months
	Capitanio *et al.* (2014)	4.2-18.7 years
	Isakova *et al.* (2014)	7-14 days
	Hinde *et al.* (2015)	3-4 months
	Raper *et al.* (2015)	6-28 days
	Chun and Capitanio (2016)	90-124 days
	Elfenbein *et al.* (2016)	90-120 days
	Herrington, Del Rosso and Capitanio (2016)	90-120 days
	Rendina, Lubach and Coe (2016)	2 weeks
	Capitanio *et al.* (2017a)	3-4 months
	Gottlieb *et al.* (2018b)	90-120 days
	Kinnally, Gonzalez and Capitanio (2018)	90-120 days
	Robinson *et al.* (2018a)	0.84-20.94 years
	Walker *et al.* (2018)	90-120 days
	Altschul *et al.* (2019)	1-21.1 years
	Simpson, Robinson and Paukner (2019)	6.5-7.5 months
	Blomquist, Hinde and Capitanio (2021)	3-4 months
	Ceniceros, Capitanio and Kinnally (2021)	90-120 days
	Fox *et al.* (2021)	90-120 days
	Pomerantz and Capitanio (2021)	90-120 days
	Robinson *et al.* (2021a)	0.92-20.97 years
	Talbot *et al.* (2021)	1.25-6.27 years
	Wood *et al.* (2021b)	1-4 weeks
	Wood *et al.* (2021c)	1-4 weeks
shared dataset (Capitanio and Widaman, 2005)	Capitanio *et al.* (2008)	Adult

shared dataset (Capitanio, 1999b)	Capitanio, Mendoza and Baroncelli (1999)	6.2 years
shared dataset (Capitanio, 1999b)	Capitanio (2002)	6.2 years
shared dataset (Capitanio, 1999b)	Capitanio, Mendoza and Bentson (2004)	6.2 years
shared dataset with (Capitanio *et al.*, 2008)	Sloan *et al.* (2008)	5-10 years

Table 2-2, Category of scoring method, citation, and age for articles that measure multiple components.

2.4.3 CORRELATES OF PERSONALITY

From the included articles five general categories of associations were found, physiological, welfare, behaviour, life history, and the general concept of personality.

Physiological measures that were associated with temperament are listed in Table 2-3. This included neural metabolic activity such as mRNA and hormone expression in specific brain regions (Kovner *et al.*, 2020; Shackman *et al.*, 2017; Kalin *et al.*, 2016; Fox *et al.*, 2015; Roseboom *et al.*, 2014), and other neuroscientific measures such as brain activity, morphology and connectivity (Tromp *et al.*, 2019; Fox *et al.*, 2018; Birn *et al.*, 2014; Haley *et al.*, 2012b; Fox *et al.*, 2012; Fox *et al.*, 2008; Kalin *et al.*, 1998; Davidson, Kalin and Shelton, 1993) and their association to anxiety. The included literature additionally covered experimental lesions in different regions of the brain and how they impacted affiliative, aggressive, anxious, excitable, exploratory, and fearful scores. Also, general metabolic activity related to serotonin seemed to be decreased in animals rated higher in activity, aggression and withdrawal, but increased in anxious individuals (Kinnally *et al.*, 2010; Oler *et al.*, 2009; Erickson *et al.*, 2005; Heinz *et al.*, 1998; Doudet *et al.*, 1995). Animals rated higher in activity, emotionality, excitability, and withdrawal, but lower in anxiety, nervousness and confidence showed decreased measures related to the cortisol metabolism (Linden *et al.*, 2018; Hamel *et al.*, 2017; Capitanio, Mendoza and Cole, 2011; Erickson *et al.*, 2005; Capitanio, Mendoza and Bentson, 2004). Similarly, the immune response was weaker in animals rated higher in aggression but stronger in animals rated higher in sociability (Sloan *et al.*, 2008; Capitanio *et al.*, 2008; Maninger *et al.*, 2003). Temperament scores of anxiety, state control and aggression have additionally been associated with an increase in self-administered alcohol intake and decrease in sensitivity to alcohol with increased aggression ratings (Wood *et al.*, 2020b; McClintick and Grant, 2016; Heinz *et al.*, 1998). Many of the temperament scores were also found to be heritable (Wood *et al.*, 2021c; Fox *et al.*, 2021; Blomquist, Hinde and Capitanio, 2021; Brent *et al.*, 2014; Sullivan, Mendoza and Capitanio, 2011; Oler *et al.*, 2010; Rogers *et al.*, 2008; Champoux *et al.*, 1997; Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980).

Component	**In/De-crease**	**(Physiological) Correlate**
Activity	↓	iron deficient anaemic males (Golub *et al.*, 2009)
	↓	5-HTT (serotonin transporter) expression (Kinnally *et al.*, 2010)
	↑	prenatal ketamine exposure (Capitanio *et al.*, 2012)
	-	heritability (Brent *et al.*, 2014)
	↓	hair cortisol (Linden *et al.*, 2018)
	↑	prenatal stress exposure (Ceniceros, Capitanio and Kinnally, 2021)
Affiliative	↓	hippocampal lesion (Machado and Bachevalier, 2006)
	↓	orbitofrontal lesion (Bachevalier, Machado and Kazama, 2011)
Aggression	↓	serotonin metabolite (CSF 5-HIAA) concentrations (Doudet *et al.*, 1995)
	↑	β-CIT-binding (estimate of serotonin transporter availability) (Heinz *et al.*, 1998)
	↓	serotonin metabolite (CSF 5-HIAA) concentrations (Heinz *et al.*, 1998)
	↓	sensitivity to alcohol intoxication (Heinz *et al.*, 1998)
	↑	ISG expression (immune response) (Capitanio *et al.*, 2008)
	↓	CXR3 expression (immune response) (Capitanio *et al.*, 2008)
	↑	sham-operated (used as control) (Bachevalier, Machado and Kazama, 2011)
	-	heritability (Brent *et al.*, 2014)
	↑	self-administered ethanol and blood ethanol (McClintick and Grant, 2016)
Anxious	↑	frontal asymmetry score (Kalin *et al.*, 1998)
	↓	amygdala lesion (Kalin *et al.*, 2001)
	↓	prenatal stress exposure (Ceniceros, Capitanio and Kinnally, 2021)
	↑	frontal asymmetry score (Davidson, Kalin and Shelton, 1993)
	↑	growth hormone responsiveness (Coleman *et al.*, 2003)
	↓	prepartum maternal cortisol (Bardi and Huffman, 2006)
	↑	postpartum maternal cortisol (Bardi and Huffman, 2006)
	↓	orbitofrontal lesion (Kalin, Shelton and Davidson, 2007)
	↑	brain activity in amygdala, bed nucleus of stria terminalis, hippocampus and periaqueductal grey (Fox *et al.*, 2008)
	-	heritability (Rogers *et al.*, 2008)
	-	scan paths of facial expressions (Gibboni, Zimmerman and Gothard, 2009)
	↑	l/s genotype in rh5-HTTLPR (McCormack *et al.*, 2009)
	↑	receptor density and binding for 5-HTT (serotonin transporter) (Oler *et al.*, 2009)
	-	heritability (Oler *et al.*, 2010)
	↑	sham-operated (used as control) (Bachevalier, Machado and Kazama, 2011)
	↓	amygdala neuroplasticity (Fox *et al.*, 2012)
	↓	MAP2 and SYN mRNA levels (proteins with integral roles in cognition) (Haley *et al.*, 2012a)
	↓	amygdala size (Haley *et al.*, 2012b)
	↓	degree of Chinese ancestry (Jiang, Kanthaswamy and Capitanio, 2013)

	-	single nucleotide polymorphisms in CRHR1 (gene for the corticotropin-releasing hormone receptor) (Rogers *et al.*, 2013)
	↓	reduced connectivity between dorsolateral prefrontal cortex and the central nucleus of the amygdala (Birn *et al.*, 2014)
	↓	mRNA levels of neuropeptide Y receptor 1 and neuropeptide Y receptor 5 in the central nucleus of the amygdala (Roseboom *et al.*, 2014)
	-	metabolism in prefrontal-limbic-midbrain circuit (Fox *et al.*, 2015)
	↑	corticotropin-releasing hormone (CRH) in the dorsal amygdala (Kalin *et al.*, 2016)
	↑	high cortisol phenotype (Hamel *et al.*, 2017)
	↑	increased metabolism in the bed nucleus of the stria terminalis (Shackman *et al.*, 2017)
	-	co-heritable with functional connectivity of the central nucleus of the amygdala and the bed nucleus stria terminalis (Fox *et al.*, 2018)
	-	neurotrophin receptor kinase 3 (NTRK3) and other gene expression (Fox *et al.*, 2019)
	↓	uncinate fasciculus alterations in males (Tromp *et al.*, 2019)
	-	lateral central nucleus RNA expression (Kovner *et al.*, 2020)
	↑	alcohol intake as adolescents (Wood *et al.*, 2020b)
	-	methylation in 13 genes (Baxter *et al.*, 2021)
Assertiveness	↑	facial width-to-height ratio (Altschul *et al.*, 2019)
	↑	lower-height/full-height ratio (Altschul *et al.*, 2019)
behavioural inhibited (anxious)	↑	airway hyper-tension (Chun *et al.*, 2013)
Confidence	-	heritability (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980)
	↑	morning cortisol concentrations (Capitanio, Mendoza and Bentson, 2004)
	↑	available milk energy (Hinde and Captanio, 2010)
	↑	higher levels of cortisol in mother's milk for sons but not daughters (Sullivan *et al.*, 2011)
	-	mother's temperament scores during infancy (heritability) (Sullivan, Mendoza and Capitanio, 2011)
	↑	high monoamine oxidase-A genotype and prenatal iron deficiency (Golub, Hogrefe and Unger, 2012)
	↓	low monoamine oxidase-A genotype and prenatal iron deficiency (Golub, Hogrefe and Unger, 2012)
	-	heritability (Brent *et al.*, 2014)
	↓	higher levels of cortisol in mother's milk (Hinde *et al.*, 2015)
	↓	facial width-to-height ratio (Altschul *et al.*, 2019)
	↓	lower-height/full-height ratio (Altschul *et al.*, 2019)
	-	narrow-sense heritability (Blomquist, Hinde and Capitanio, 2021)
	↑	prenatal stress exposure (Ceniceros, Capitanio and Kinnally, 2021)
Dominance	-	heritability (Brent *et al.*, 2014)
Emotionality	↓	airway hyper-tension (Capitanio *et al.*, 2011)

	↑	prenatal endotoxemia (through lipopolysaccharide administration) (Willette *et al.*, 2011)
	↓	prenatal ketamine exposure if animals had the low activity genotype in the monoamine oxidase-A gene (gene involved in dopamine, norepinephrine, and serotonin production) (Capitanio *et al.*, 2012)
	↓	hair cortisol (Linden *et al.*, 2018)
Excitability	↓	afternoon cortisol concentrations (Capitanio, Mendoza and Bentson, 2004)
	↑	hippocampal lesion (Machado and Bachevalier, 2006)
Exploratory	↑	hippocampal lesion (Machado and Bachevalier, 2006)
	↑	C/G genotype in rhCRH locus (Barr *et al.*, 2008)
Fearfulness	↓	post rotary nystagmus (Schneider and Suomi, 1992a)
	↓	amygdala and hippocampus lesion (Bliss-Moreau *et al.*, 2010)
	↑	sham-operated (used as control) (Bachevalier, Machado and Kazama, 2011)
	↓	high monoamine oxidase-A genotype and prenatal iron deficiency (Golub, Hogrefe and Unger, 2012)
	↑	low monoamine oxidase-A genotype and prenatal iron deficiency (Golub, Hogrefe and Unger, 2012)
	↓	high monoamine oxidase-A genotype and prenatal iron deficiency (Golub and Hogrefe, 2015)
Gentleness	-	mother's temperament scores during infancy (heritability) (Sullivan, Mendoza and Capitanio, 2011)
	-	narrow-sense heritability (Blomquist, Hinde and Capitanio, 2021)
Gregarious	↓	chronic idiopathic colitis (Howell *et al.*, 2012)
Inconsolable	↑	adrenocorticotropic hormone exposure (Schneider, Coe and Lubach, 1992)
Inhibited temperament	-	heritability (Fox *et al.*, 2021)
Interaction-inhibiting	↓	amygdala lesion (Machado *et al.*, 2008)
Irritability	↑	adrenocorticotropic hormone exposure (Schneider, Coe and Lubach, 1992)
	↑	s allel in serotonin gene (5-HT) and prenatal alcohol exposure (Kraemer *et al.*, 2008)
Negative affectivity	-	paternal half-siblings (heritability) when nursery reared (Wood *et al.*, 2021c)
Nervousness	↑	cortisol levels (Capitanio, Mendoza and Cole, 2011)
	↓	relationship of cortisol levels and neutrophil numbers (Capitanio, Mendoza and Cole, 2011)
	-	heritability (Brent *et al.*, 2014)
	↑	higher levels of cortisol in mother's milk (Hinde *et al.*, 2015)
	-	narrow-sense heritability (Blomquist, Hinde and Capitanio, 2021)
Orienting/regulation	-	paternal half-siblings nursery reared (Wood *et al.*, 2021c)
	-	maternal half-siblings mother reared (Wood *et al.*, 2021c)
Slow	↑	high monoamine oxidase-A genotype and prenatal iron deficiency (Golub, Hogrefe and Unger, 2012)
	↓	low monoamine oxidase-A genotype and prenatal iron deficiency (Golub, Hogrefe and Unger, 2012)

Sociability	-	heritability (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980)
	↑	antibody response to immunisation (Maninger *et al.*, 2003)
	↑	IgG response to tetanus vaccination (Sloan *et al.*, 2008)
	-	heritability (Brent *et al.*, 2014)
	-	specific neuropeptide and kinase signalling pathways (Parker *et al.*, 2018)
State control	↑	hybrid vs (Indian origin) (Champoux *et al.*, 1997)
	↑	prenatal cocaine exposure (Lidow, 1998)
	↓	prenatal alcohol exposure (Coe *et al.*, 2010)
	↑	l/s genotype in rh5-HTTLPR (Champoux *et al.*, 2002a)
	↓	postnatal iron deficiency (Golub, Hogrefe and Germann, 2007)
	↓	prenatal alcohol exposure (Coe *et al.*, 2010)
	↑	mothers infected with the Zika virus (Maness *et al.*, 2019)
	↑	alcohol intake as adolescents (Wood *et al.*, 2020a)
Surgency/extraversion	-	paternal half-siblings nursery reared (Wood *et al.*, 2021c)
	-	paternal half-siblings mother reared (Wood *et al.*, 2021c)
Vigilance	-	mother's temperament scores during infancy (heritability) (Sullivan, Mendoza and Capitanio, 2011)
Withdrawal	↓	reduction in cortisol concentrations (Erickson *et al.*, 2005)
	↓	reduction in serotonin metabolites (Erickson *et al.*, 2005)

Table 2-3, Physiological correlates to personality, the direction of association, and the component it has been associated with. References for each are shown in brackets.

Table 2-4 summarises the **welfare** correlates identified in the included literature. This revealed that wellbeing of animals measured with subjective rating scales or through observation of rates of injuries, stereotypies and incidents of diarrhoea is lower in animals high in anxiety, nervousness, and vigilant as well as low in confidence, dominance, friendliness, and openness scores (Robinson *et al.*, 2021a; Neudecker *et al.*, 2021; Simpson, Robinson and Paukner, 2019; Robinson *et al.*, 2018a; Gottlieb *et al.*, 2018b; Elfenbein *et al.*, 2016; Gottlieb, Maier and Coleman, 2015; Gottlieb, Capitanio and McCowan, 2013; Weiss *et al.*, 2011; Boyce, 2004; Boyce *et al.*, 1998). Further, two studies revealed that pairing success in males was higher with more similarity in exploratory temperament and if they scored lower in gentleness and nervousness (Ruhde *et al.*, 2020; Capitanio *et al.*, 2017a). Capitanio *et al.* (2017a) additionally revealed that similarity in nervousness ratings let to more successful female pairings. Ratings of the exploratory component were associated with training success in a touch screen task and implementation of a voluntary weighing procedure (Zijlmans et al., 2021; Coleman, Tully and McMillan, 2005).

Component/Trait	direction of correlation	(Welfare) Correlate
Activity	↑	in early life when reared without vestibular stimulation (Schneider, Kraemer and Suomi, 1991)
	↓	vestibular stimulation during rearing (Schneider *et al.*, 1991)
	↓	maternal responsiveness (Bardi and Huffman, 2006)
Anxious	↓	subjective wellbeing (Simpson, Robinson and Paukner, 2019)
	↑	rates of injury during group confinement (Boyce *et al.*, 1998; Boyce, 2004)
	↓	maternal responsiveness (Bardi and Huffman, 2006)
	↓	subjective wellbeing (Weiss *et al.*, 2011)
	↑	nursery reared compared to mother reared (Corcoran *et al.*, 2012)
	↑	motor stereotypic behaviour (Gottlieb, Capitanio and McCowan, 2013)
	↓	rates of injury (Robinson *et al.*, 2018a)
	↑	one time exposure to isoflurane anaesthesia (Neudecker *et al.*, 2021)
Assertiveness	↑	subjective wellbeing (Simpson, Robinson and Paukner, 2019)
Confidence	↓	social contact with peers during rearing (Rommeck *et al.*, 2011)
	↑	subjective wellbeing (Weiss *et al.*, 2011)
	↓	motor stereotypic behaviour (Gottlieb, Maier and Coleman, 2015)
	↓	diarrhoea incidents (Gottlieb *et al.*, 2018b)
	↓	rates of injury (Robinson *et al.*, 2018a)
	↓	higher body condition score during pregnancy in mothers (Walker *et al.*, 2018)
	↑	subjective wellbeing & welfare (Robinson *et al.*, 2021a)
Dominance	↑	subjective wellbeing & welfare (Robinson *et al.*, 2021a)
Emotional reactivity	↑	prenatal exposure of acoustical startle (Rendina, Lubach and Coe, 2016)
Emotionality	↑	higher body condition score during pregnancy in mothers (Walker *et al.*, 2018)
Enterprise	↑	number of maternal rejections (Bardi and Huffman, 2006)
Excitability	↑	adverse experience in first eight months of life for males (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980)
Exploratory	↑	training success using a positive reinforcement technique for a target touch task (Coleman, Tully and McMillan, 2005)
	-	similarity led to more successful pairings in males (Ruhde *et al.*, 2020)
	↑	voluntary weighing success (Zijlmans *et al.*, 2021)
Fearfulness	↑	in early life when reared without vestibular stimulation (Schneider, Kraemer and Suomi, 1991)
	↓	vestibular stimulation during rearing (Schneider *et al.*, 1991)
	-	rearing condition (Schneider and Suomi, 1992b)
Friendliness	↑	subjective wellbeing (Weiss *et al.*, 2011)
	↑	subjective wellbeing & welfare (Robinson *et al.*, 2021a)
Gentle	↓	motor stereotypic behaviour (Gottlieb, Capitanio and McCowan, 2013)

	↓	lead to more successful pairings in males (Capitanio *et al.*, 2017a)
	↑	diarrhoea incidents (Gottlieb *et al.*, 2018b)
Irritability	↑	in early life when reared without vestibular stimulation (Schneider, Kraemer and Suomi, 1991)
	↓	vestibular stimulation during rearing (Schneider *et al.*, 1991)
	-	rearing condition (Schneider and Suomi, 1992b)
Nervousness	↑	higher diarrhoea incidences when nursery reared (Elfenbein *et al.*, 2016)
	-	similarity led to more successful pairings in females (Capitanio *et al.*, 2017a)
	↓	led to more successful pairings in males (Capitanio *et al.*, 2017a)
	↑	diarrhoea incidents (Gottlieb *et al.*, 2018b)
	↑	paternal rearing experience in F1-F3 generations (Kinnally, Gonzalez and Capitanio, 2018)
Openness	↑	subjective wellbeing (Simpson, Robinson and Paukner, 2019)
	↑	subjective wellbeing & welfare (Robinson *et al.*, 2021a)
State control	↑	mother reared compared to peer reared (Champoux *et al.*, 2002a)
	↑	nursery reared compared to mother reared (Paukner, Capitanio and Blozis, 2020)
Vigilant	↑	diarrhoea incidents (Gottlieb *et al.*, 2018b)

Table 2-4, Welfare correlates to personality, the direction of association, and the component it has been associated with. References for each are shown in brackets.

Social behaviours made up the majority of the **behavioural measures** that showed correlation with personality ratings (see Table 2-5). They included for example the initiation of proximity, play or both with another conspecific and reciprocated relationships. General social behaviours were more frequent and positive in animals scoring higher in sociable, curious and lower in adaptable, anxious, excitable and equable personality measures (Talbot *et al.*, 2021; Chun and Capitanio, 2016; Capitanio *et al.*, 2014; Weinstein and Capitanio, 2012; Weinstein and Capitanio, 2008; Kinnally *et al.*, 2008; Capitanio, 2002; Capitanio, 1999b). Further, Pomerantz and Capitanio (2021) showed that higher similarity in emotionality lead to more positive social interactions. Also, aggression seemed to decrease in animals lower in confidence and higher in cautiousness (Kinnally *et al.*, 2008; Capitanio, 1999b). However, other behaviours such as vocalisations or displays were also associated with personality scores. Vandeleest and Capitanio (2012) showed that infants emitted scream and gecker vocalisations more frequently if they scored higher in distress. Beisner and McCowan (2014) reported increased frequencies of peaceful silent bared teeth displays in animals rated higher in gentleness and lower in confidence.

Component/Trait	direction of correlation	(Behaviour) Correlate
Adaptability	↓	proximity initiated (Weinstein and Capitanio, 2008; Weinstein and Capitanio, 2012)
Anxious	↑	attention to unfamiliar monkey faces (Pomerantz and Capitanio, 2021)
	↑	anxiety and emotionality during novelty and human intruder tests (Chun and Capitanio, 2016)
	↑	time spent alone (Chun and Capitanio, 2016)
Apprehensiveness	↑	reactivity (Bolig *et al.*, 1992)
Cautiousness	↓	aggression in familiar and novel groups (Kinnally *et al.*, 2008)
Confidence	↓	reactivity (Bolig *et al.*, 1992)
	↑	aggressive behaviours (Capitanio, 1999b)
	↓	peaceful silent bared teeth display (Beisner and McCowan, 2014)
	↑	inappropriate behaviour (Talbot *et al.*, 2021)
Curiosity	↓	reactivity (Bolig *et al.*, 1992)
	↑	social engagement in familiar group (Kinnally *et al.*, 2008)
Distress	↑	vocalisations (Vandeleest and Capitanio, 2012)
	↑	mother initiated leaves (Vandeleest and Capitanio, 2012)
Emotionality	-	similar scores showed more positive social interactions (Pomerantz and Capitanio, 2021)
Equability	↓	reactivity (Bolig *et al.*, 1992)
	↓	less aggressive (Capitanio, 1999b)
	↑	passive interactions (Capitanio, 1999b)
	↓	proximity initiated (Weinstein and Capitanio, 2008)
	↓	play initiated (Weinstein and Capitanio, 2008)
	↓	multiplex preferences (proximity and play) (Weinstein and Capitanio, 2008)
	↓	reciprocated relationships (Weinstein and Capitanio, 2008)
	↓	proximity initiated (Weinstein and Capitanio, 2012)
	↓	play initiated (Weinstein and Capitanio, 2012)
	↓	multiplex preferences (proximity and play) (Weinstein and Capitanio, 2012)
	↓	reciprocated relationships (Weinstein and Capitanio, 2012)
Excitability	↑	reactivity (Bolig *et al.*, 1992)
	↑	inconsistent social behaviour (Capitanio, 1999b)
Fearfulness	↑	reactivity (Bolig *et al.*, 1992)
Gentle	↑	peaceful silent bared teeth display (Beisner and McCowan, 2014)
Inhibited temperament	-	refused treats offered by potentially threatening human experimenters (Fox *et al.*, 2021)
Insecurity	↑	reactivity (Bolig *et al.*, 1992)
Irritability	↑	reactivity (Bolig *et al.*, 1992)
	↑	inappropriate behaviour (Talbot *et al.*, 2021)
Nervousness	-	similar scores showed fewer dominance signalling (Pomerantz and Capitanio, 2021)
Positive engagement	↑	time away from mothers (Vandeleest and Capitanio, 2012)
	↓	nursing (Vandeleest and Capitanio, 2012)

	↓	contact with mother (Vandeleest and Capitanio, 2012)
	↓	mother initiated leaves (Vandeleest and Capitanio, 2012)
Sociability	↑	affiliative interactions (Capitanio, 1999b)
	↓	latency to gaze avert (Capitanio, 2002)
	↓	tendency to sit and stare (Capitanio, 2002)
	↑	interactions with social partners (Capitanio *et al.*, 2014)
	↓	poor social motivation (Talbot *et al.*, 2021)
	↑	poor social attractiveness (Talbot *et al.*, 2021)
Tense	↑	reactivity (Bolig *et al.*, 1992)
Understanding	↓	reactivity (Bolig *et al.*, 1992)

Table 2-5, Behavioural correlates to personality, the direction of association, and the component it has been associated with. References for each are shown in brackets.

Other studies associated **life history** factors such as differences in age, sex, and rank between individuals and their measures of temperament (see Table 2-6). Older animals generally scored higher on the components confidence, motherly, protectiveness and eccentricity but lower on activity and playfulness. Females generally scored higher on excitability and sociability but lower in activity, confidence, and equability (McCowan et al., 2011; Bolig et al., 1992; Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980; Stevenson-Hinde and Zunz, 1978). Moreover, animals with higher social rank and power rated higher in confidence, dominance, effectiveness, excitability (Wooddell et al., 2019; McCowan et al., 2011; Bolig et al., 1992). Finally, there was also a positive interactive effect of sex and age on confidence, excitability and a negative effect on anxiety, boldness, and sociability. Sex and rank had positive interactive effect on sociability, excitability, and equability (Kohn *et al.*, 2016; McCowan *et al.*, 2011).

Component	**direction of correlation**	**(life-history) Correlate**
Activity	↑	adult males compared to adult females (Stevenson-Hinde and Zunz, 1978)
	↓	age (Bolig *et al.*, 1992)
Anxious	↑	less relaxed mother infant relationships (Vandeleest, Mendoza and Capitanio, 2013)
	↓	older females (Kohn *et al.*, 2016)
	↑	less initiative in mother-infant behaviours (Wood *et al.*, 2021c)
Bold	↓	older females (Kohn *et al.*, 2016)
	↑	higher-ranking females (Kohn *et al.*, 2016)
Confidence	↑	adult males compared to adult females and animals aged 1-5 (Stevenson-Hinde and Zunz, 1978)
	↑	adult males compared to adult females (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980)
	↑	social rank (Bolig *et al.*, 1992)
	↑	older animals (Capitanio, Mendoza and Bentson, 2004)
	↑	social power (McCowan *et al.*, 2011)
Dominance	↑	social rank (Bolig *et al.*, 1992)
Eccentricity	↑	age (Bolig *et al.*, 1992)

Effectiveness	↑	social rank (Bolig *et al.*, 1992)
Emotionality	↑	birth season compared to mating season (Maestripieri, 2000)
	↑	higher social status as juveniles (Wooddell *et al.*, 2019)
Equability	↑	social power in males (McCowan *et al.*, 2011)
	↓	female animals (McCowan *et al.*, 2011)
Excitability	↑	females (Bolig *et al.*, 1992)
	↑	social power (McCowan *et al.*, 2011)
	↑	two- and three year old females compared to males (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980)
Fearfulness	↑	social rank (Bolig *et al.*, 1992)
Motherly	↑	age (Bolig *et al.*, 1992)
Playfulness	↓	age (Bolig *et al.*, 1992)
positive engagement	↑	less relaxed mother infant relationships (Vandeleest, Mendoza and Capitanio, 2013)
Protectiveness	↑	age (Bolig *et al.*, 1992)
Sociability	↓	adult males compared to adult females and animals aged 1-5 (Stevenson-Hinde and Zunz, 1978)
	↑	sex (Bolig *et al.*, 1992)
	↓	older females (Kohn *et al.*, 2016)
	↑	higher-ranking females (Kohn *et al.*, 2016)
State control	↓	female animals (Coe *et al.*, 2010)
	↑	dominance rank of mother (Suarez-Jimenez *et al.*, 2013)

Table 2-6, Life-history correlates to personality, the direction of association, and the component it has been associated with. References for each are shown in brackets.

Lastly, some of the literature identified in the review focused on the **concept of personality** (see Table 2-7). Capitanio and Widaman (2005) for example conducted a confirmatory factor analysis which led to strong and moderate loadings in confidence, equability, irritability, and sociability components. Further, the components activity, confidence, and fearfulness were found to be stable over multiple years (von Borell, Kulik and Widdig, 2016; Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980; Stevenson-Hinde and Zunz, 1978). Chamove, Eysenck and Harlow (1972) also received comparable scores in components of affiliative, aggression, and fearful from subjective ratings and behavioural tests.

Component	**Comparison**
Activity	stable over three years (Stevenson-Hinde and Zunz, 1978)
	changes in ratings during early life within rearing condition and sex (Clarke and Snipes, 1998)
Affiliative	same component derived from stimulus test (Chamove, Eysenck and Harlow, 1972)
Aggression	same component derived from stimulus test (Chamove, Eysenck and Harlow, 1972)
	stabilised with age and was inconsistent during infancy (von Borell, Kulik and Widdig, 2016)
Anxious	presented less invasive touch screen task as an alternative for the human intruder paradigm (Bethell *et al.*, 2019a)
Attentive	changes in ratings during early life within rearing condition and sex (Clarke and Snipes, 1998)

Confidence	stable over three years (Stevenson-Hinde and Zunz, 1978) stable over four years (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980) not stable from ante-natal to post-natal year (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980) strong loadings in confirmatory factor analysis (Capitanio and Widaman, 2005) decreased with complexity of test object (Bliss-Moreau *et al.*, 2010)
Equability	moderate loadings in confirmatory factor analysis (Capitanio and Widaman, 2005)
Excitability	not stable until adulthood (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980) stable from ante-natal to post-natal year (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980) changes in ratings during early life within rearing condition and sex (Clarke and Snipes, 1998)
Fearfulness	same component derived from stimulus test (Chamove, Eysenck and Harlow, 1972) consistent over seven-year period (von Borell, Kulik and Widdig, 2016)
Independent	changes in ratings during early life within rearing condition and sex (Clarke and Snipes, 1998)
Irritability	moderate loadings in confirmatory factor analysis (Capitanio and Widaman, 2005)
Negative affectivity	three factor structure in Rhesus macaque infants (Kay *et al.*, 2010) three factor structure in Rhesus macaque infants (Wood *et al.*, 2021b)
Orienting/regulation	three factor structure in Rhesus macaque infants (Kay *et al.*, 2010) three factor structure in Rhesus macaque infants (Wood *et al.*, 2021b)
Outwardly directed	changes in ratings during early life within rearing condition and sex (Clarke and Snipes, 1998)
Sociability	not stable until adulthood (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980) not stable from ante-natal to post-natal year (Stevenson-Hinde, Stillwell-Barnes and Zunz, 1980) strong loadings in confirmatory factor analysis (Capitanio and Widaman, 2005) diffused during maturation (even though mother-infant interactions were not taken into account) (von Borell, Kulik and Widdig, 2016)
Surgency/ extraversion	three factor structure in Rhesus macaque infants (Kay *et al.*, 2010) three factor structure in Rhesus macaque infants (Wood *et al.*, 2021b)

Table 2-7, Conceptual comparisons drawn to personality components with references for each are shown in brackets.

2.5 DISCUSSION

This scoping review summarised the current knowledge of personality research in Rhesus macaques. It included 148 peer-reviewed articles and extracted information on the methods used, age of the animals tested and the associations that were made. This provides an overview of the current standing of Rhesus macaque personality research. The methods applied were categorised into behavioural coding, behavioural testing, and subjective ratings and articles found associations to measures in physiology, welfare, behaviour, life history and personality in general.

As expected, methodology and vocabulary used was inconsistent between studies (see Norman, Rowden and Cowlishaw, 2021; Robinson and Weiss, 2023). For example, some articles that made use of the biobehavioural assessment program used behavioural testing rather than subjective ratings as originally proposed (Baxter et al., 2021; Linden et al., 2018; Golub et al., 2009; Golub and Hogrefe, 2015; Capitanio et al., 2012; Kinnally et al., 2010). Also, human intruder tests were used to measure anxious temperament and behavioural inhibition, some studies seem to use these interchangeably, but this is not always clear (Fox et al., 2008; Coleman, Tully and McMillan, 2005). Furthermore, trait information need to be treated with caution as even when they used the same naming for variables they can be representative of components or a trait (a component derived using factor analysis vs single item on a questionnaire). Likewise, while some studies compared loadings from behavioural tests to subjective ratings, this might not always be a straightforward comparison (Chamove, Eysenck and Harlow, 1972).

Contrary to Norman, Rowden and Cowlishaw (2021) findings there was a big overlap in the approaches used between studies and only a few validated novel methods of assessment (e.g. Zijlmans et al., 2021). This is also shown by the 62 studies making use of the biobehavioural assessment program and the Infant Neurobehavioral assessment scale. Further, 24 of the articles had a clear link to welfare as might some of the literature grouped here under physiological correlates.

This review only included peer-reviewed articles published in English. I did not include conference abstracts or literature in other languages and the conclusions presented are likely to represent a biased picture.

While the Hominoid Personality Questionnaire was not the most used for Rhesus macaques it is the most used in primate research in the last decade, recommending this as a standardised tool for further comparisons (Norman, Rowden and Cowlishaw, 2021). It was also the subjective rating method used in the more recent publications of this review. Hence the Hominoid Personality Questionnaire is selected to be used in my thesis (see Chapter 6).

Overall, the findings of this review led to the identification of behaviours to be taken into account in the following chapters of this thesis. The system developed as part of this thesis (see Chapter 3) is capable of recording events of proximity to others, leading to insights on sociability. Further, the quantification of movement and space use can be of importance to capture personality components such as activity and anxiety (see Chapter 6). Identification of displays and facial expressions is out of the scope for this thesis as these behaviours are hard to classify automatically. Also, vocalisations cannot be taken into account as the developed system does not record sound.

Animals observed in this study were all females within a similar age range. The animals are pair housed and are matched according to observations of staff form the breeding centre and local animal technicians. Hence factors of life history are not expected to have a major influence in the ratings of personality for this thesis.

This chapter summarises the current knowledge of Rhesus macaque personality including the methodologies applied, age of individuals measured and associations in physiology, welfare, life history and behaviour.

CHAPTER 3. METHODS - AUTOMATED IN-CAGE MONITORING OF PAIR-HOUSED LABORATORY RHESUS MACAQUES

3.1 CHAPTER SUMMARY

Conventional methods for assessing individual behavioural expressions are often labour-intensive, invasive, and subjective to bias. Leveraging current advances in computer vision tools can help to overcome these issues and such approaches are becoming more prominent in the field of animal behaviour. In this chapter, I present a low cost, non-invasive and automated pipeline to predict identity and location of pair housed Rhesus macaques. This chapter presents the built pipeline based on a self-designed camera module (based on a Raspberry Pi) mounted within the cage and a fast instance segmentation pipeline (based on YOLOv8). The pipeline will be used in the following chapters to show that such methodologies offer a promising avenue for complementing and advancing the field of welfare assessment and animal behaviour research.

3.2 INTRODUCTION

The availability of automated methods to monitor behaviour in animals has massively increased in the last decade in both captivity and the wild (Aguilar-Moreno and Graña, 2023; Mathis and Mathis, 2020). The automation of measuring behaviour is useful for monitoring health and welfare, especially for animals in captivity (Li *et al.*, 2020; Matthews *et al.*, 2017; Roughan, Wright-Williams and Flecknell, 2009; Rushen, Chapinal and de Passile, 2012).

Rhesus macaques (*Macaca mulatta*) are an invaluable model in biomedical research and are one of the most commonly used non-human primate species, due to their genetic and physiological similarities to humans (Lear *et al.*, 2022; Mitchell *et al.*, 2018; Treue and Lemon, 2023). They are stringently protected by regulations and given their sociality and intelligence, the challenging task of measuring their welfare is important (Carstens and Moberg, 2000; Hau and Schapiro, 2007). Novel computational methods are becoming increasingly valuable tools to address this challenge (Knaebe *et al.*, 2022). Automated systems are proposed to be cost efficient, holistic, and less subjected to observer bias (Sueur, Forin-Wiart and Pelé, 2020). Furthermore, it is known that Rhesus macaques change their behaviour in the presence of human observers, which can mask signs of injury or illness (Gaither *et al.*, 2014). Additionally, the amount of data that can be feasibly analysed by human observers is a fraction of the volumes possible with the use of computational methods (Mathis *et al.*, 2018; Porto *et al.*, 2013;

Weinstein, 2018). The use of automated methods allows for longer times of observation on a higher temporal resolution for the data collected for each animal, which leads to more representative and meaningful measures on an individual level. Such datasets could assist in forming individualised estimates of normal behavioural expression for each animal as well as the identification of deviations from typical patterns (Carstens and Moberg, 2000; Laudenslager and Boccia, 1996; Lutz *et al.*, 2022; Robinson *et al.*, 2018b).

There are multiple computational approaches to detecting and tracking macaques (Bala *et al.*, 2020; Blanco Negrete *et al.*, 2021; Labuguen *et al.*, 2021; Marks *et al.*, 2022). However, due to the high complexity of primate behaviour, using computer vision techniques makes its classification challenging. Different approaches of varying complexity have been proposed using Deep-learning such as raw-pixel action recognition (Marks *et al.*, 2022), markerless pose estimation (Bala *et al.*, 2020) and integration of visual and acoustic data (Bain *et al.*, 2021). An additional difficulty is the identification of individuals due to the lack of readily identifiable features. Possible solutions include the use of electronic devices or physical markings (Ballesta *et al.*, 2014; Maddali *et al.*, 2014; Morita *et al.*, 2020; Rose *et al.*, 2012) and non-invasive approaches such as facial recognition and frame-by-frame identification (Marks *et al.*, 2022; Ueno *et al.*, 2019; Ueno *et al.*, 2022; Witham, 2018). The output of many currently existing approaches needs further processing to derive actual behavioural measures, rather than pixel-based coordinates.

In this study, I present an approach for identification and location tracking of pair-housed laboratory Rhesus macaques. My approach incorporates an in-cage camera module designed for continuous video recording throughout the day, coupled with a comprehensive pipeline for tracking, individual identification, and location prediction of laboratory pair-housed Rhesus macaques.

3.3 ETHICAL STATEMENT

The animals were pair housed in accordance with the UK Home Office regulations and the Animals (Scientific Procedures) Act (1986). This project is approved by the relevant Research Ethics Committee of Newcastle University (AWERB Project ID No: ID 865). All animals recorded were part of other neuroscience projects with the appropriate approvals and only observational data was collected for this thesis.

3.4 SUBJECTS AND HUSBANDRY

All animals were housed at the Comparative Biology Centre (CBC) at Newcastle University and were part of ongoing research projects. Both housing and research comply with the NC3Rs guidelines (NC3Rs, 2017) and UK Home Office regulations. All cage dimensions are around 2.1 x 3.0 x 2.4 m and exceed the minimum requirement under UK legislation. Within the CBC multiple animals are housed, generally in pairs, and have visual and auditory contact to conspecifics. Each cage contains environmental enrichment (fixed (changed approximately every month): swings, ropes, shelves, tubes, and boxes; moveable (changed approximately biweekly): steps, poles, balls, toys) as well as natural light and wood shavings. Daily forage-mix is provided to the animals to offer the opportunity for foraging. Depending on the involved study protocol and stage animals were in, food was provided *ad libitum* or restricted to a minimum with additional food rewards given in the lab, all in accordance with the specific protocols the animals were involved in. The non-human primate housing facility is on a 12h light/dark cycle with a humidity of approximately 24% and a temperature around 20 °C. In total, ten female Rhesus macaques were recorded in their home cage using a custom-made camera module (see Section 3.5 & Figure 3-2). All monkeys were between 4.8 and 6.9 years of age at time of recording (see Table 3-1). Recordings of animal CI and CA were part of the training data for the pipeline but are not part of analysis in further chapters. Animal CF was singly housed during the time of recording relevant for this thesis. All animals are consistently named throughout this thesis, using the first letter of their name and an additional identifier. All animals came from the Medical Research Council’s Centre for Macaques (CFM) except for animal MB who was bred at the UK Health Security Agency (UKHSA).

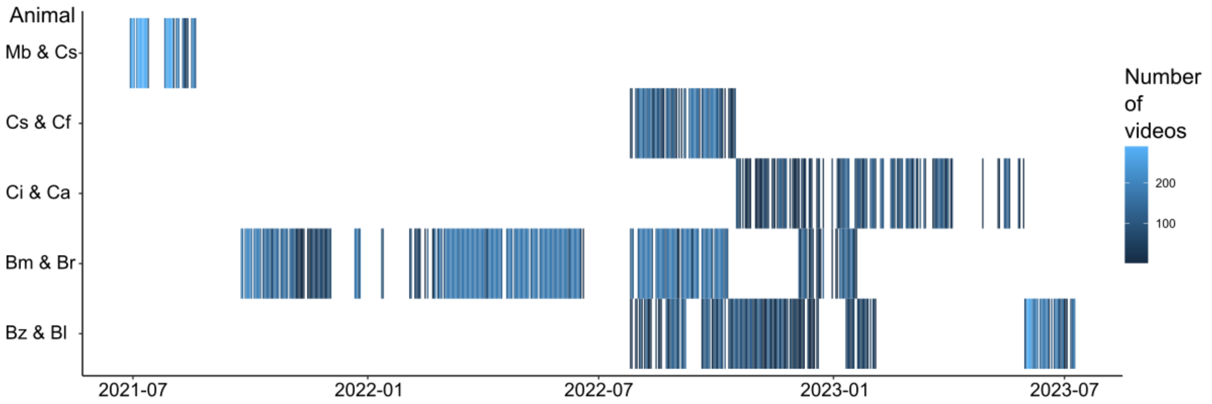

Figure 3-1, Representation of number of videos collected.

Animal	Date of birth	Date received Implant	Cage Mate
MB	19/02/2014	15/09/2020	CS
CS	21/08/2017	*25/01/2022*	MB
CF	20/05/2017	27/04/2022	
CI	*14/09/2017*	*10/10/2022*	*CA*
CA	*02/09/2017*	*07/11/2022*	*CI*
BC	15/06/2016	12/05/2022	BM
BM	23/03/2016		BC
BZ	16/05/2016	14/03/2022	BZ
BL	23/10/2016	*24/01/2023*	BL

Table 3-1, Information of animals recorded, including date of birth, the cage mate at time of recording and the date of implant surgery (entries in cursive only included for obtaining training data, dates in cursive indicate animals were not implanted during the time of recording relevant in further chapters of this thesis).

3.5 CAMERA MODULE AND VIDEO COLLECTION

Videos (1644 x 1232 size) were collected using a self-made camera module consisting of an aluminium box and a CCTV camera casing connected by two pieces of plastic with slits to attach them to the cage bars (Figure 3-2 & Figure 3-3). The module is based on a RaspberryPi 3+ or 4 (https://www.raspberrypi.com/), these single board computers are cheap components (see Table 3-2), very userfriendly and adaptable (i.e. accessible documentation and peripherals such as cameras and modules) and was hence suitable for developing this camera module. It made extensions such as multiple cameras and an internal battery easier, however in regards to durability, humidity resistance and power consumption, other alternatives should be taken into account. The RaspberryPi is connected to two cameras (NOIR V2 with IMX295 lenses) via an Arducam Sync HAT (https://www.arducam.com/) using HDMI cables and CSI-HDMI connectors. These cameras and lenses are common components used with RaspberryPis and the standard camera board shipped with the Arducam splitter does not fit in a standard CCTV casing hence the NOIR V2 with the IMX295 lenses were used instead. The cameras are mounted inside the CCTV housing using a 3D printed piece of plastic. The synchronisation HAT captures the video stream simultaneously as one frame in H.264 format. The camera module is powered by a PiJuice HAT with a 5 Ah onboard battery and two external 50 Ah batteries (see Figure 3.3). Videos were recorded in ~36 fps at 10 minutes length per video. Capture started at 5:30 AM and ended at 19:00 PM as light conditions during the evening and night were too low. Batteries were changed daily. For preprocessing, videos were cropped in the middle of the width (resulting in two 822 x 1232 videos) and converted into MPEG-4 format using a custom python script. The videos were recorded directly onto USB drives (2x 256GB capable of holding over 2 months of footage) and regularly exchanged. The video data was then transferred to a university server (as backup) as well as local drives (for further analysis).

3.6 TRACKING AND IDENTIFICATION PIPELINE

The pipeline was based on three main steps, segmentation, identification, and location (see Figure 3-4, code is available at: https://github.com/jreuka/CBC_pipeline/). A custom trained YOLOv8 (Jocher, Chaurasia and Qiu, 2023) model (see Section 3.6.1) was used as an instance segmentation algorithm including the BOTSORT (Aharon, Orfaig and Bobrovsky, 2022) tracker to segment and track three object classes (body, head and headpiece) on each video frame for the basis of the pipeline. YOLOv8 is one of the most recent versions of the "You Only Look Once" object detection models released by ultralytics (https://www.ultralytics.com/) and was selected as it is a State-of-the-art model in the tasks needed. Furthermore, the codebase and documentation of the ultralytics framework is actively being maintained (alternative methods are listed in Section 1.6). In addition to object detection the algorithm has been extended for other tasks such as segmentation which finds the mask (outline) around an object. The BOTSORT tracker is one of the trackers integrated into YOLOv8. This algorithm designed for multi-object tracking in videos by estimating position, velocity and size to handle occlusions. For this thesis, I focused on times when animals were singly housed or one of the pair housed animals was implanted with a headpiece while the other was still in training and had no implant at the time of interest. The implants are part of the neuroscience projects these animals were involved in, and more information can be found in Glover & Baker (2020). Using absence or presence of the headpiece, I could thus identify the individuals (see Section 3.6.2) and match their location using a manually predefined location map (see Section 3.6.3 and Figure 3-5).

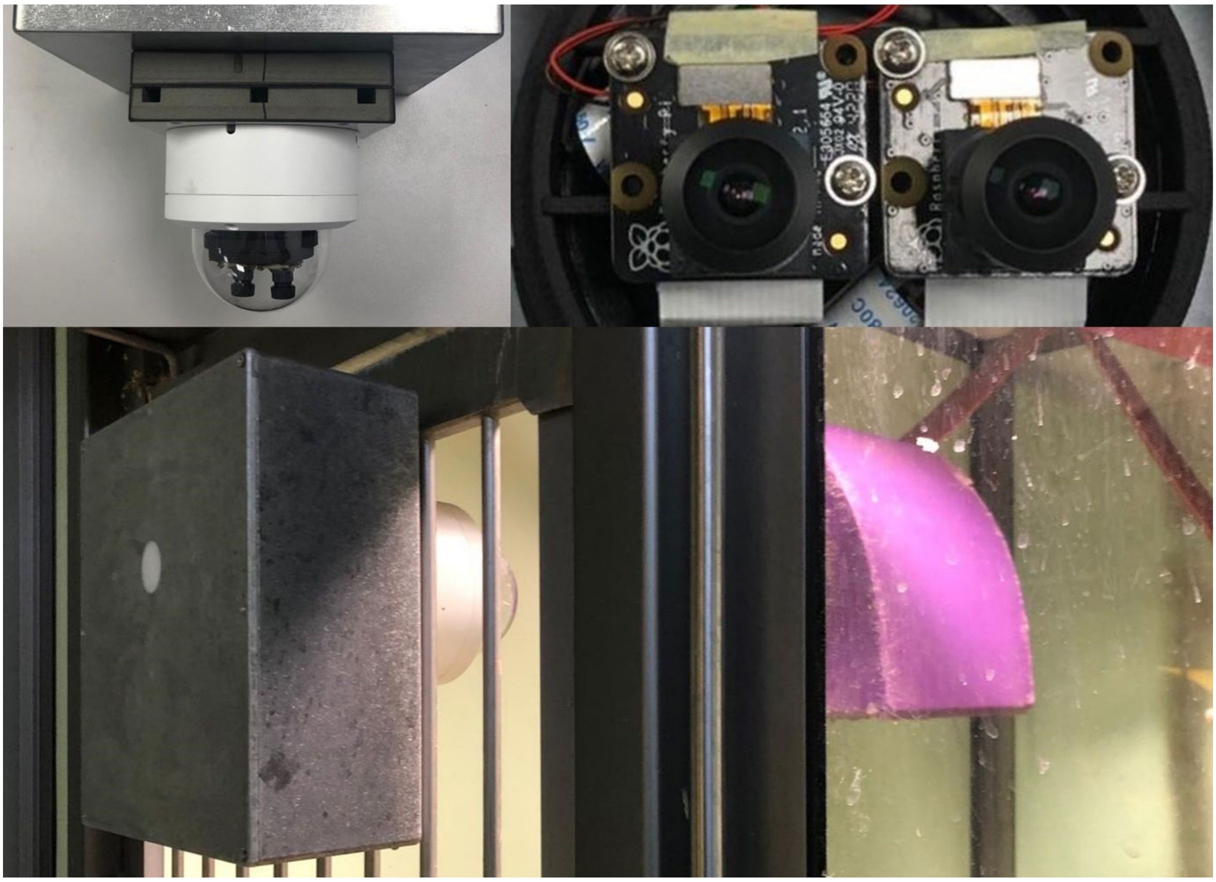

Figure 3-2, Pictures of the camera module. Top left: top view on CCTV-casing and plastic connections with slits for cage bars. Top right: front view on mounted and connected NOIR V2 cameras with IMX295 lenses. Bottom: picture of mounted module within housing area.

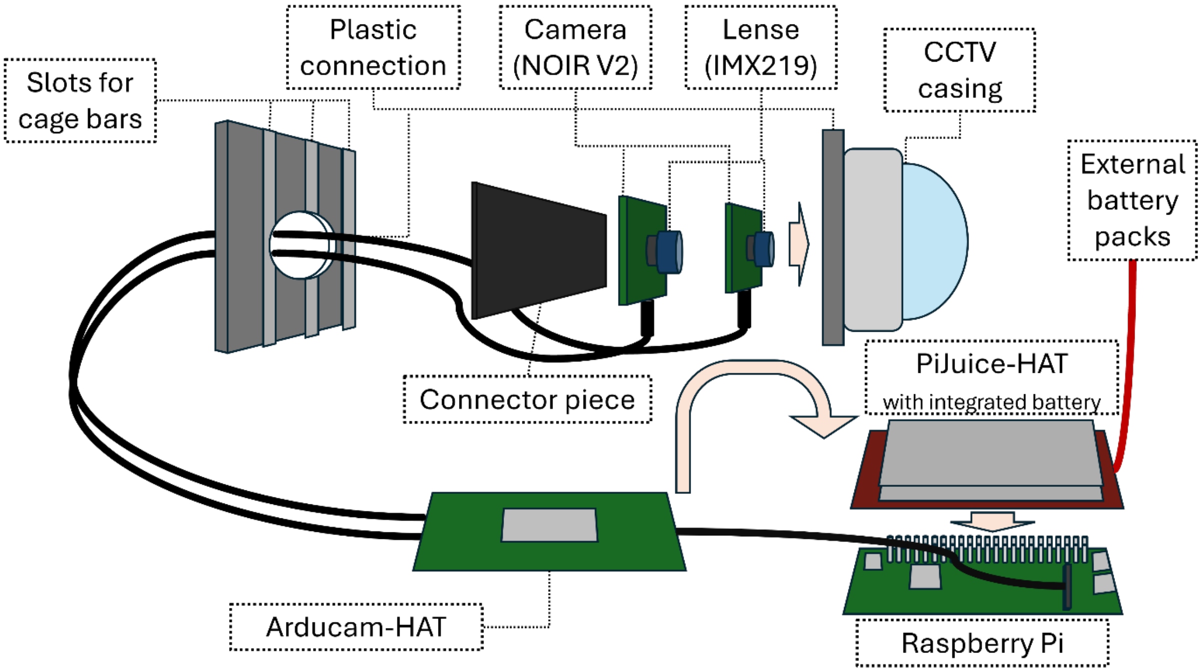

Figure 3-3, Schematic of the camera module setup including all components. Dotted boxes depict the description of the parts used. Black lines indicated cabled connections, red line indicated connection to external power source. Coloured arrows indicate positioning of parts.

Component	**Quantity**	**Price (£)**
Raspberry Pi	1	34
PiJuice-HAT	1	50
Integrated battery (12'000 mAh)	1	25
External battery packs (50'000 mAh)	2	90
Raspberry Pi NOIR V2 Camera	2	50
Arducam IMX219 Wide Angle Camera	2	20
CCTV casing	1	35
		304

Table 3-2, Components and cost of camera module (rounded retail prices at time of purchase)

3.6.1 SEGMENTATION - YOLOV8 BASED ALGORITHM

To obtain masks for objects of interest in the videos I trained a model based on the YOLOv8 segmentation algorithm (Jocher, Chaurasia and Qiu, 2023). The segmentation model was trained, tested, and validated on a set of 1318 images split 70%, 15% and 15% respectively. A custom-made GUI interface (Python 3.10) was used to extract random frames of interest from a selection of videos. 5-10 frames per video with at least one animal present were selected (covering a variety of cage environments and enrichment) and segmentations were labelled using the VGG-VIA (Dutta, Gupta and Zissermann, 2016) image annotator tool. The corresponding labels were then reformatted into the necessary YOLOv8 structure using a custom Python (3.10) script. In addition, background frames (with no animals present, $N = 284$)

were selected with the GUI-interface from two randomly selected videos grouped by cage setups (for each set up of enrichment). The model was trained using the pretrained YOLOv8n model weights until there was no observable improvement over 50 epochs (n-epochs total: 315) and allowing the algorithm to automatically select batch size (28). Model training was conducted on a computer equipped with a Nvidia Geforce RTX3060 GPU and 32GB of RAM and took approximately 1.3 hours to complete. For analysis the number of maximum detections was set to five and I used the BOTSORT tracker to match objects over multiple frames.

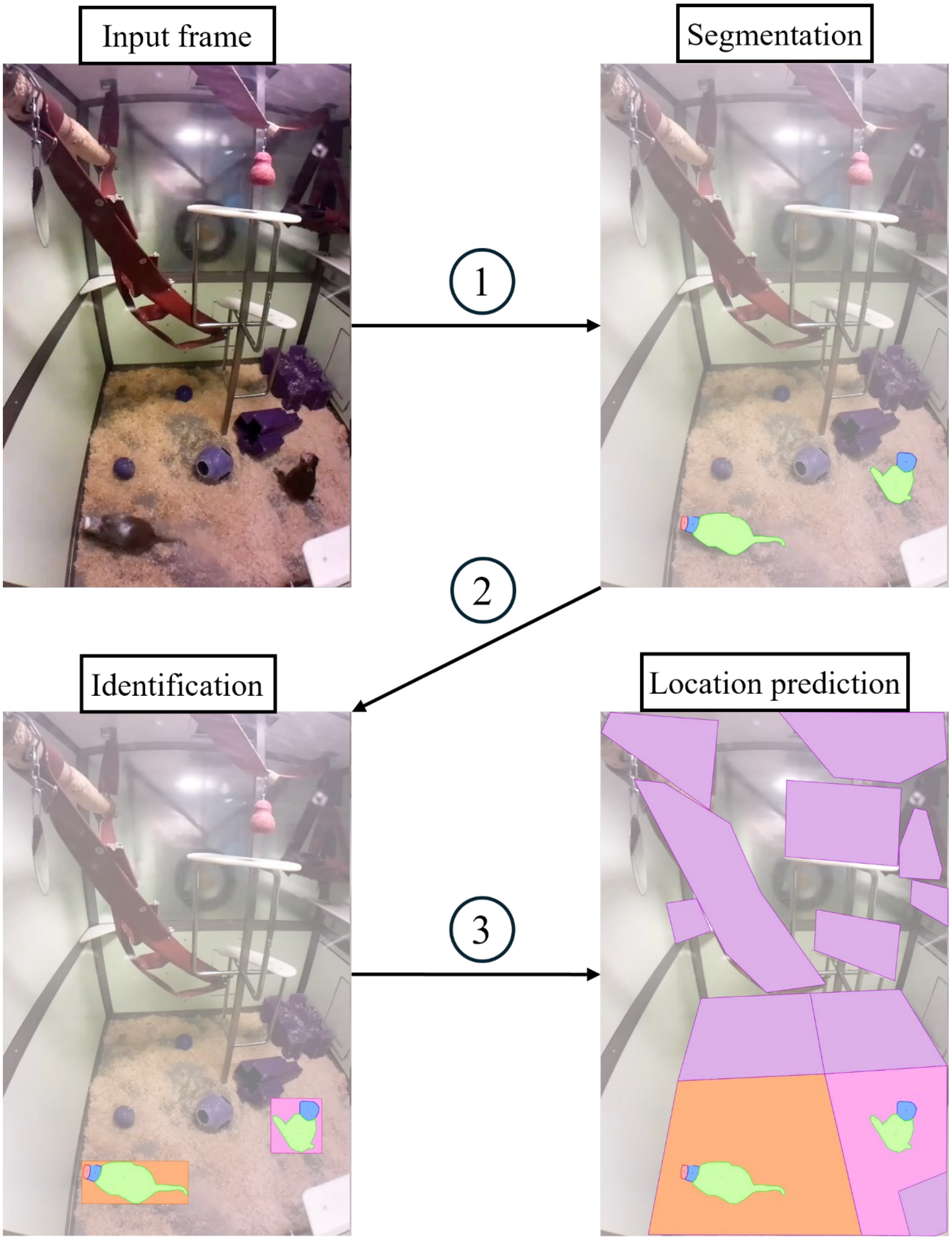

Figure 3-4, Representation of the pipeline; input frame; segmentation masks predicted by the algorithm; the identification of the animals using the overlap of the masks; location prediction using the centroid of the mask and a manually predefined location map.

3.6.2 IDENTIFICATION

The pipeline checks each frame for overlaps between the segmented objects of differing class using the derived polygons from the YOLOv8 segmentation model. In case a body mask was overlapping with both a head and headpiece object, the animal was identified as “Animal 2”; if only two classes overlap it was identified as “Animal 1”. In cases of non-overlapping object’s identity was set to “unsure” (e.g., in case where the head of an animal is occluded). Identities were counted within each track derived with the BOTSORT tracker, and the most frequent occurring identity is assigned to all detections in the track. To avoid multiple detections of the same object (same class and identity), simultaneous occurrences of tracks on frames were considered. Tracks with higher confidence scores given by the segmentation algorithm and presence over a higher number of frames were selected and other occurrences of the same identity were set to “unsure”. In cases where two tracks of the same class were present on the same frame, but only one was identified as either of the animals, the other track was automatically set to the remaining of the two animals.

3.6.3 LOCATION PREDICTION

To predict if animals are present in regions of interest, manually predefined location maps were created by averaging images of 36 random frames from each video, manually inspecting these, and grouping resulting images into cage setups. This was necessary as housing setup was changed regularly by the technical staff for enrichment purposes (changes in environment and enrichment). Frames of cleaning events and low light levels were also excluded from further analysis. The resulting images are again averaged for each setup. I use these pictures as a basis for the manually predefined location maps, labelled using VIA VGG (Dutta, Gupta and Zissermann, 2016) by drawing non-overlapping polygons over locations of interest (see Figure 3.5). At each segmentation step, the centroid of each detection mask was calculated. This centroid was then matched to polygons of the appropriate manually predefined location map.

3.6.4 RESAMPLING AND DETECTION MATCHING

The output of the pipeline is run through an additional python script to resample the data down to a 2 Hz frequency. All predictions were averaged for each individual or in case of location prediction, the most common value was used. This significantly reduced the amount of data and noise, especially for the centroid predictions.

Since videos were recorded in synchrony and the cameras had a large overlap in their vision field, the predictions of both camera streams via the 2 Hz timestamps were matched. Only occurrences where an animal was predicted to be in the same location for both cameras are used in further analysis.

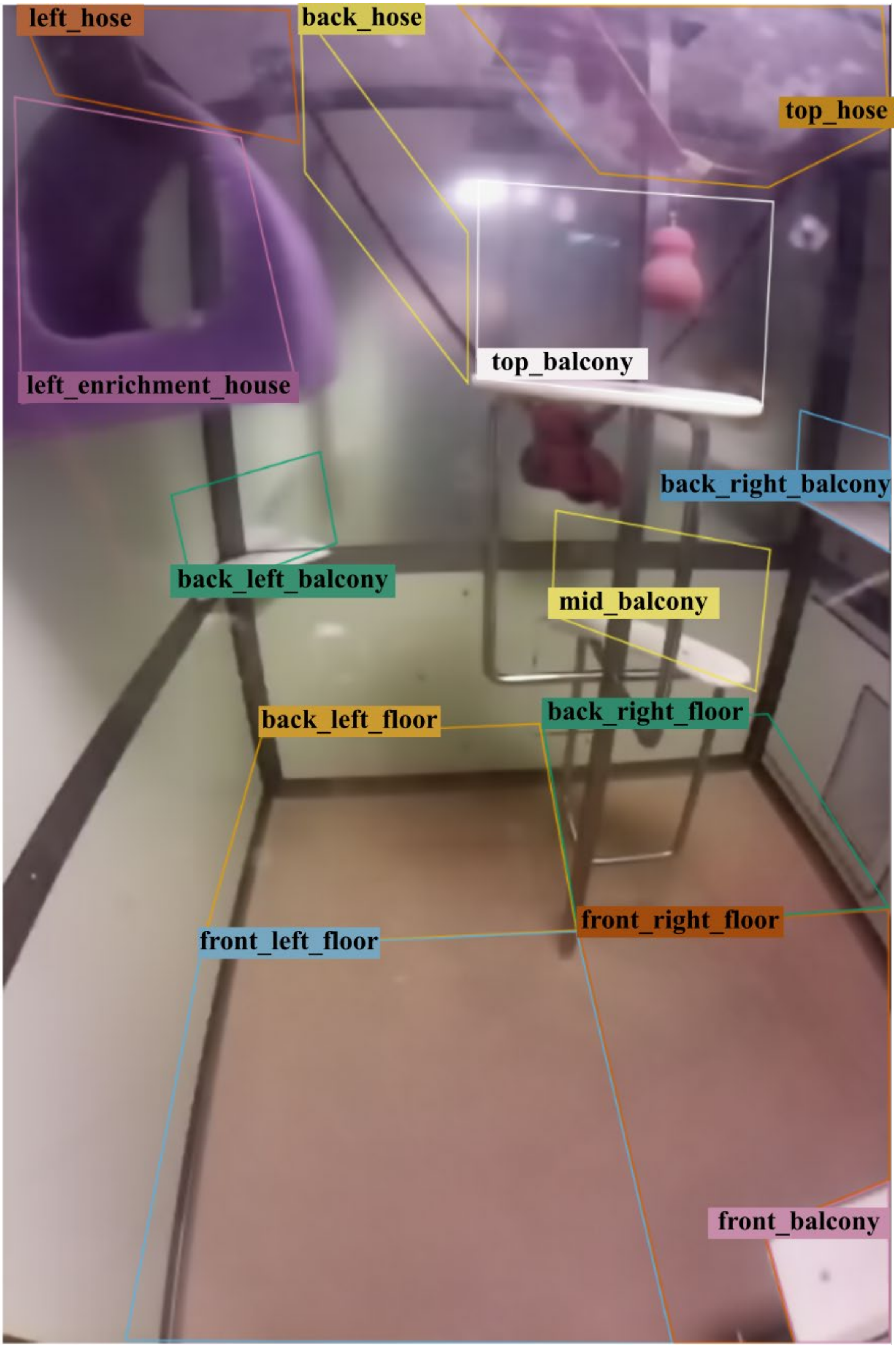

Figure 3-5, Example picture of manually predefined location map, image is an average over multiple average frames taken from each video with the same environmental setup. The coloured boxes with labels represent the locations, everything else is labelled as "other".

3.7 VALIDATION OF THE PIPELINE

I used the inbuilt validation system for the YOLOv8 segmentation model as well as a custom python script to compare manually labelled video data to the pipeline output and get validation scores. The segmentation model was tested on a dataset comprising 198 manually labelled images. Table 3-3 presents the evaluation metrics, including mean average precision (mAP), precision (P), and recall (R) scores. Precision is the ratio of all predicted relevant instances to all retrieved instances and serves as an indication of the model's ability to avoid false positives. Recall is the ratio of relevant predictions to all relevant occurrences, indicating how well the model captures relevant instances. Finally, the mean average precision is a measure of the overall effectiveness, calculated by taking the mean of average values of all precision-recall pairs weighted by the change in recall. This is a commonly used measure for object detection and segmentation algorithms, however, recall values are generally replaced by intersection over union (IoU) values. IoU measures the ratio of the overlap of the true mask with the predicted mask to the union area of the two (Terven, Córdova-Esparza and Romero-González, 2023). Notably, the validation scores exhibit a comparable performance to the baseline Yolo models (e.g. Jocher, Chaurasia and Qiu, 2023). Figure 3-6 shows the Recall-Precision curve of the validation data and Figure 3-7 the confusion matrix of the object classes. The segmentation model scores higher in detecting instances classified as "body" over its performance for "head" or "headpiece" classes. While the lower recall values indicate a reduced capacity to identify existing instances, the higher precision values suggest a lower likelihood of false positive detections.

		Box				**Mask**			
Class	**Instances**	**P**	**R**	**mAP 50**	**mAP 50-95**	**P**	**R**	**mAP 50**	**mAP 50-95**
all	579	0.882	0.776	0.832	0.52	0.851	0.689	0.752	0.372
body	263	0.912	0.909	0.94	0.718	0.905	0.878	0.906	0.546
head	204	0.843	0.74	0.8	0.404	0.811	0.609	0.697	0.309
headpiece	112	0.89	0.679	0.755	0.438	0.837	0.58	0.653	0.26

Table 3-3, Validation metrics, including precision (P), recall (R), and mean average precision (mAP) scores for box and mask detections.

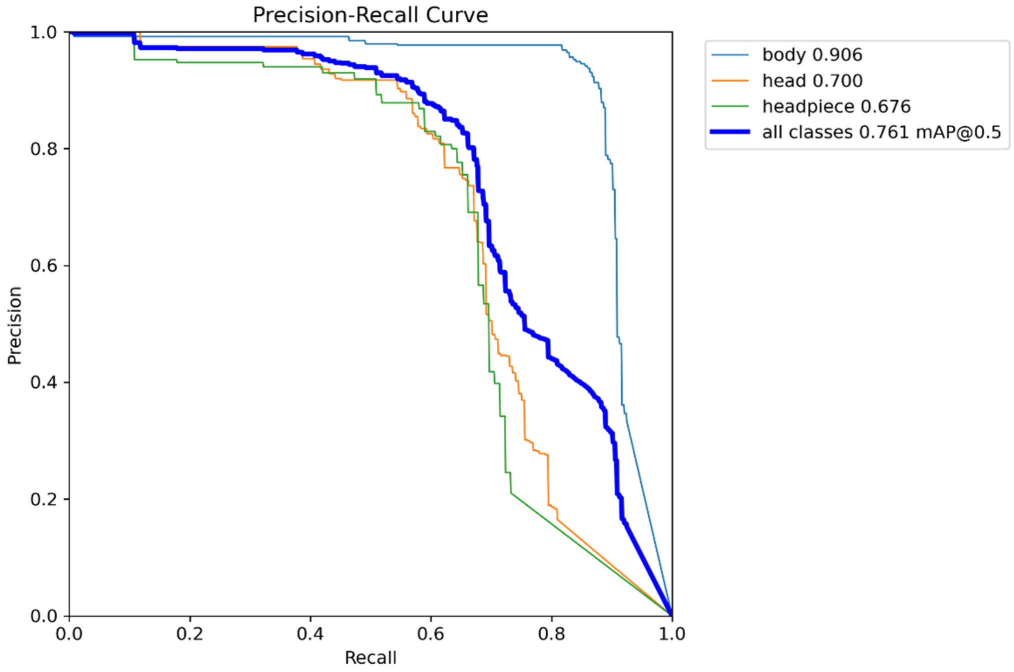

Figure 3-6, Precision-Recall Curve of YOLOv8 validation results. Light blue line is the body mask, yellow line is head mask, green line is headpiece, and all classes in blue.

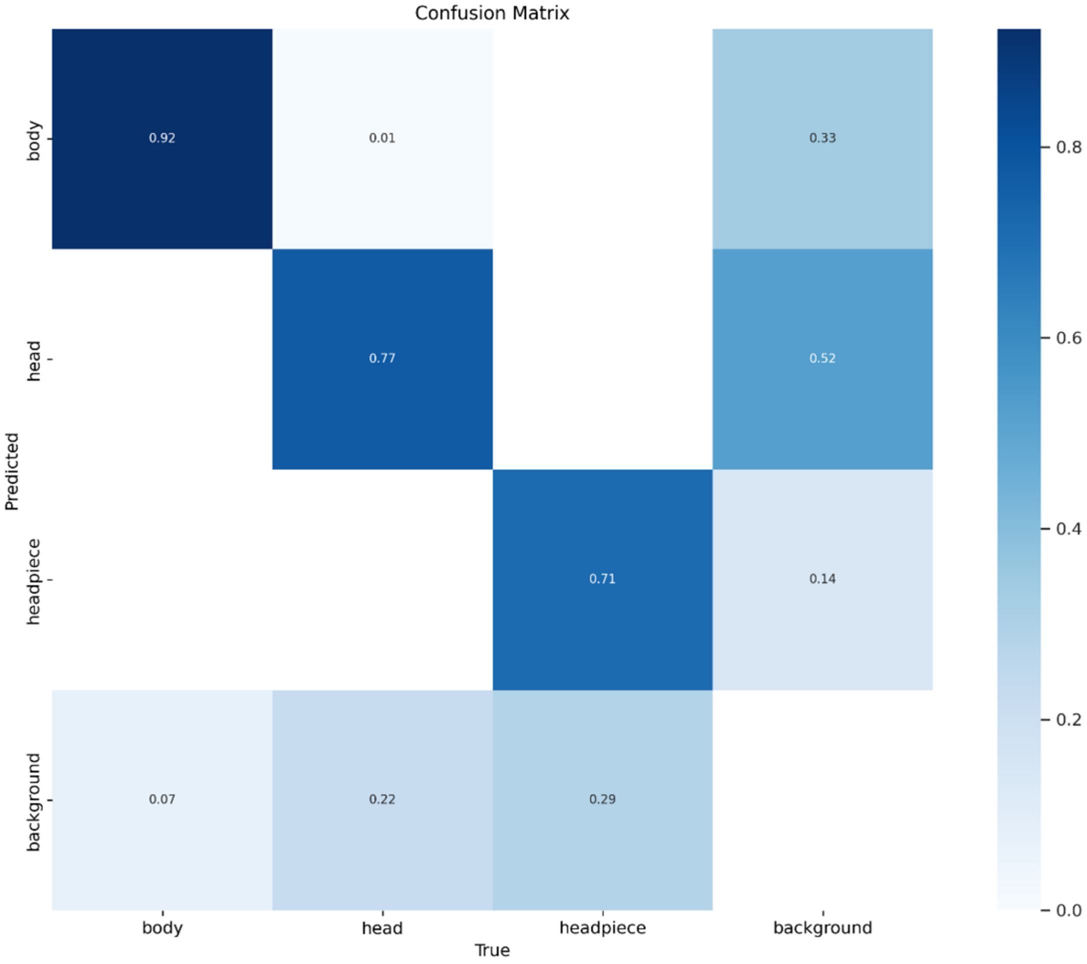

Figure 3-7, Confusion matrix of YOLOv8 validation results. Indicating percentage of predictions and True instances for each class.

A custom-made GUI to label frames on which the pipeline made predictions on animal identity and location was used to validate the identity and location prediction. The 500 frames from video data used in Chapter 5 (see Section 5.3.1) were manually labelled (250 frames randomly selected from all frames with detections, and 250 frames randomly selected from frames with detections where the pipeline has detected at least one occurrence of the "head" class – indicating at least one animal was identifiable) and compared these to the predictions of the pipeline using the statistics program R (Version 4.3.0). In cases of heavy occlusion, the labeller was able to look at the origin videos of the frame to ascertain the identity of the animal. In addition to the location and animal identity, labels were added if a) at least one head was clearly visible (i.e. the animals are easily distinguishable), b) no head is visible (animals can't be distinguishable without more information) or c) one of the animals and/or head is occluded (difficult to distinguish the animals).

The performance evaluation of the pipeline shows that predictions made in agreement between left and right camera on animal identity and location were correct in 84.3% of instances. Even under challenging conditions where animals are partly occluded or the head is not visible, the pipeline's prediction are still correct 81.3% of the time. The analysis time for video processing using the pipeline is approximately 110% of the recording time (e.g. ~11 minutes for a 10-minute video) when executed on the same machine (with GPU, described in Section 3.6.1) used for model training and running two batches in parallel. The analysis time is intricately linked to the number of detections and tracks (see Figure 3-8 and Figure 3-9).

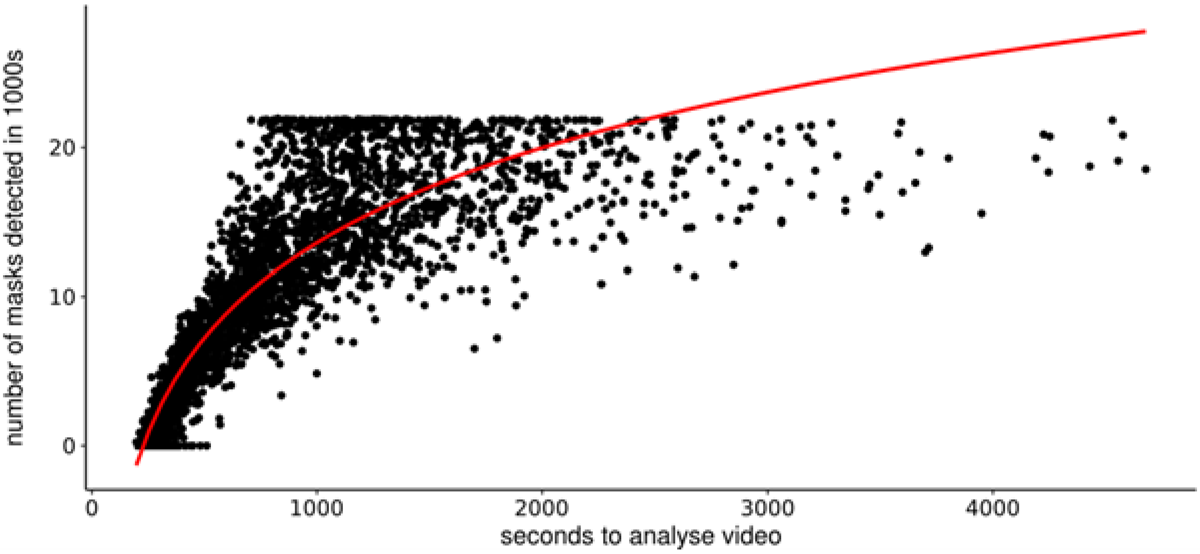

Figure 3-8, Time in seconds taken to analyse videos compared to the number of masks detected in the video in thousands. Red line is a fitted logarithmic curve, this was selected visually to be a good fit.

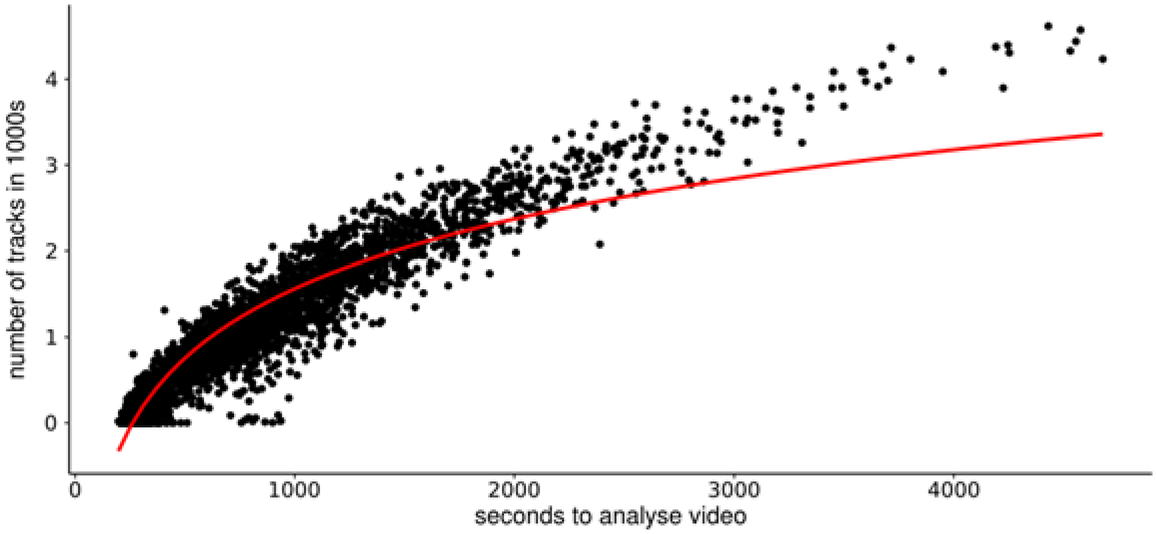

Figure 3-9, Time in seconds taken to analyse videos compared to the number of tracks detected in the video in thousands. Red line is a fitted logarithmic curve, this was selected visually to be a good fit.

3.8 DISCUSSION

This chapter presents the main method developed for this thesis including an in-cage camera module and analysis pipeline and its validation (for an example video see https://youtu.be/m6M8x7O3s7I). Results showed that the pipeline was capable of tracking individuals over multiple frames enabling correct identification of animals at 84.3%, including cases of partial occlusion. The simultaneous recording of two video streams from separate cameras resulted in a halved resolution relative to the potential resolution achievable at the same frame rate. This synchronous recording is leveraged by matching detections from both cameras and exclusively considering detections occurring in both streams. While the presented model demonstrates good precision and recall values for primate detection in videos, it exhibits lower recall values for the "head" and "headpiece" classes. This finding might be attributed to the smaller size of these objects within the frame, compounded by the inherent challenge posed by the low resolution of the obtained recordings. The integration of tracking over multiple frames served to alleviate this issue, provided the model sporadically detected these classes for identification purposes.

The overall accuracy was comparable to previously published identification models for macaques (i.e. Marks *et al.*, 2022 ~ 78 +- 3% with up to 4 individuals), even though the presented model used a smaller number of labelled frames. However, since the identification relies on the implant of one animal only, the applicable timeframe is restricted by the project timelines. On the other hand, this reduces the cost of training and application as there is no need for manually labelling animal identities (Aharon, Orfaig and Bobrovsky, 2022; Mathis *et al.*, 2018; Porto *et al.*, 2013; Weinstein, 2018). This could be done even if both animals are implanted by having different colours for the implants to make them easily distinguishable or aid in creating training sets based on the method presented here. Moreover, my model exhibits robust performance even under conditions of relatively low resolution and instances of occlusion (see Figure 3-10).

Manually predefined location maps were used to predict the location of individuals. This process identified the location of individuals >95% correctly compared to manually labelled data. This approach is fast in processing and takes little manual effort, requiring the creation of a new location map only when there are alterations in the fixed environmental setup, such as changes to the cage or its configuration (i.e. around once a month). A drawback of this approach is that areas in the cage will overlap and given the single-view angle of this setup, not all areas can be clearly distinguished. Moreover, the changes in the environment can make it difficult to compare data over multiple setups, I solved this in following chapters by grouping removable objects (such as hose-enrichments) under the same location label. Especially in bigger enclosures, using multiple synchronized camera modules with different view angles covering all locations in a cage would be highly beneficial.

This pipeline provides a base framework that could help in more detailed behaviour analysis. For instance, the pipeline can identify periods where animals are visible or present in the same location of the cage which would reduce the time to screen the footage manually (as an example doing this for the dataset used in Chapter 5 would reduce the full ~1000h of videos down to ~100h and ~12h respectively). Moreover, since the camera module is comprised in a single unit and self-sufficient with the batteries, it could be fitted in different areas of the cage and potentially be used in combination with other camera systems already in place to cover areas that are otherwise obscured or too far away.

Yolov8 models have been run on Raspberry Pis, integrating the models directly into the camera module could save time and storage space in instances where raw video footage is unnecessary. Additionally, the pipeline holds the potential to include other identification approaches to increase accuracy. The following chapters will make use of the pipeline and showcase its applied potential.

Figure 3-10, Example output of pipeline with heavy occlusion. Blue boxes and lines indicate the detection and segmentation result of the “body” class. These are labelled with the animal identity in blue and location in yellow.

CHAPTER 4. INDIVIDUALITY AND CONSISTENCY IN MOVEMENT-TRAJECTORIES IN LABORATORY-HOUSED RHESUS MACAQUES

4.1 CHAPTER SUMMARY

This chapter's focus is on the individuality and consistency in movement trajectories of laboratory-housed Rhesus macaques. Location tracking is an important tool in primatology and recent research found high reliability in measures related to movement and space use suggesting the existence of spatial personality. Here, a Neural Network is used to classify individual identities based on location trajectories over a time frame of four months. The model's accuracies are suggestive of both consistency and individuality in the data recorded.

4.2 INTRODUCTION

Location tracking has been important to primatologists for decades and is continuing to gain traction in recent years owing to the widespread availability of devices such as GPS (Global Positioning System) collars (Dore *et al.*, 2020; Phillips, Elvey and Abercrombie, 1998). This has been done with Rhesus macaques in the wild using either handheld GPS devices (e.g. Anand, Vaidyanathan and Radhakrishna, 2021; Ganguly and Chauhan, 2018; Kshetri *et al.*, 2023) or GPS collars (Anderson *et al.*, 2017). The data has been leveraged to estimate home ranges (Anderson *et al.*, 2017) and daily distance travelled (Ganguly and Chauhan, 2018; Kshetri *et al.*, 2023), as well as study movement patterns through landscape structures, particularly in relation to crop feeding (Anand, Vaidyanathan and Radhakrishna, 2021).

Rhesus macaque movement has also been evaluated in captive settings employing various methodologies. Beisner and Isbell (2009) utilized manually drawn trajectories to assess movement patterns in adult female Rhesus macaques. They found that higher-ranking individuals tended to follow more direct paths (i.e. fewer changes in direction and fewer stops) within the enclosure. Furthermore, they observed that animals paused more frequently on grass substrates compared to gravel, indicating an influence of ground substrate on the straightness of paths. Hrolenok *et al.* (2018) employed an RFID system to track individual monkeys within a group of six males. Their analysis of tracking data allowed them to identify fleeing behaviours of individual monkeys, which enabled the creation of a dominance network within the group. Finally, Liu *et al.* (2022) utilized video recordings to monitor movement and spatial distribution within the home cage of Rhesus macaques. Their analysis revealed divergent space utilization patterns between two single housed individuals, with one exhibiting a consistent location preference over two years while the other altered its preference between the two years. Furthermore, this study investigated the variations in activity levels throughout the day for these two animals and found a likely association to physiological and husbandry factors.

A recent study by Morita *et al.* (2020) revealed individual patterns in movement trajectories of Japanese macaques. Animals were fitted with Bluetooth collars for 3D location tracking of five individuals. The location data was converted into movement trajectories of two-minute intervals. A movement trajectory is the result of a consecutive sequence of the location of an animal ordered by time (Edelhoff, Signer and Balkenhol, 2016). By employing a machine learning algorithm, animal identities were predicted based on these 3D trajectories. The model of Morita *et al.* (2020) demonstrated the ability to predict identity with a recall score of 0.247-0.577 and an accuracy score of 0.318-0.41 using trajectories from five animals, surpassing chance levels. These results indicate that there is individuality in movement.

A recent meta-analysis revealed high repeatability across time and in different spatial behaviours and referred to this as spatial personality (Stuber, Carlson and Jesmer, 2022). Interestingly, the repeatability of spatial behaviours was almost twice as high as estimates of repeatability in other behaviours, such as foraging (Bell, Hankison and Laskowski, 2009; Stuber, Carlson and Jesmer, 2022). However, the spatial behaviour is heavily influenced by the environment available to an animal and measured differences in behaviour could be simply a reflection of environmental differences in an animal's territory (Spiegel and Pinter-Wollman, 2022).

In this chapter, the pipeline was employed to extract movement trajectories of laboratory-housed macaques and utilize a model constructed following the approach outlined by Morita *et al.* (2020) to identify individuals based on these trajectories. The objective was to discern the presence of individuality within the movement data. Furthermore, the models' stability over a period of three months in predicting animal identity is assessed to indicate temporal consistency in this data. As the animals share the same environment with a limited amount of space, these differences are representative of spatial personality.

4.3 METHODS

4.3.1 MOVEMENT TRAJECTORIES

The pipeline and camera module described in Chapter 3 was used to identify and locate animals within 6616 videos of two pair-housed individuals recorded between 23rd of June and 2nd of December in 2022. With the use of two-minute time bins location data when animals were visible for 50%, 83% and 100% of the time was extracted. This corresponds to 120, 200 and 240 detections within the two-minute period. Table 4-1 summarises the number of movement trajectories extracted for each animal. The extracted data included the normalised X and Y coordinates (all values are scaled to be between 0 and 1 using $x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$) of the animals in the video. Additionally, for the full dataset, a normalised area of the mask was used as a third dimension. As the animals' sizes were approximately the same, the area of the mask is an estimate of the distance of the animals to the camera.

Number of detections	**240**			**200**			**120**		
	A_BM	A_BC	Number of days	A_BM	A_BC	Number days	A_BM	A_BC	Number of days
Full	323	177	110	1509	879	132	5021	3425	139
D1	115	48	40	511	246	52	1682	1065	54
D2	63	36	21	312	185	24	1069	698	24
D3	145	93	49	686	448	56	2270	1662	61

Table 4-1, Number of movement trajectories for Animal BM & BC when visible 100%, 83% and 50% of the two-minute time bin, as well as the number of days the trajectories are taken from. Full indicates the whole sample D1, D2 and D3 are groupings dependent on the date of the trajectory.

4.3.2 BIDIRECTIONAL LONG-SHORT TERM MEMORY NEURAL NETWORK TO CLASSIFY IDENTITY FROM MOVEMENT TRAJECTORIES

A bidirectional LSTM RNN was used (Morita et al., 2020) (see Table 4-2 and Figure 4-1, Representation of movement trajectories and the model architecture). RNN's with their looping connections, where previous inputs can influence current outputs, are designed for use on sequential data. A specific case of a recurrent unit is the LSTM architecture. It has memory Cells with gates that allow the neural network to save and remove information, making it even more effective on time series data (Hochreiter and Schmidhuber, 1997; Schuster and Paliwal, 1997). Here, the model is made of four layers. The first layer, a masking layer, indicates missing data in the input to the following layers in the network. These inputs are then skipped in the proceeding layers as they are not providing any valuable information for the model. The bidirectional layer created the features from the time-series data. The dense layer is a multilayer perceptron. It made predictions transformed into likelihoods of identities in the last SoftMax layer. This bidirectional network predicted the animal's identity based on the movement trajectories. The architecture and model selection were based on Morita *et al.* (2020).

Layer	Output shape	# parameters
Masking	(None, 240, 2)	0
Bidirectional	(None, 256)	134144
MLP	(None, 128)	32896
SoftMax	(None, 2)	258

Table 4-2, Description of layers with their output shapes and number of parameters

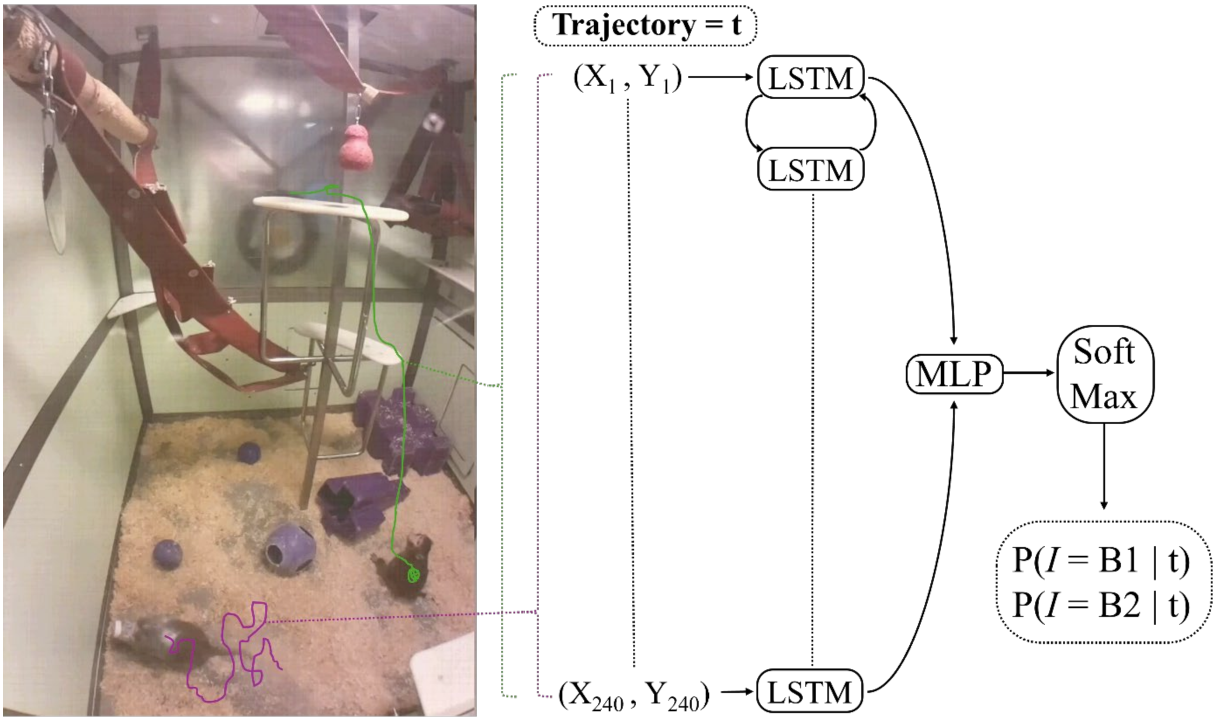

Figure 4-1, Representation of movement trajectories and the model architecture adapted from (Morita et al., 2020); image on the left is an example frame, purple and green are example trajectories for animal BC and BM respectively; on the right is the model architecture with the trajectories represented by x and y coordinates, the bidirectional long-short term network, the multilayer perceptron and the SoftMax layer providing probability values for the identities.

4.3.3 DATA ANALYSIS

The network was trained on the dataset of movement trajectories which was equalised to have the same number of samples from each individual to avoid bias in the data and then split into training, validation and testing data (80%, 10%, 10%). The accuracy for the test set was calculated as the proportion of correct predictions overall predictions after 30 epochs of training. The training and testing step was repeated 30 times and for each repetition, the overall accuracy and recall values for each animal identity were calculated. This was done twice, once with a dataset including only X and Y coordinates once with a dataset including X and Y coordinates and the area of the mask. In order to provide an acceptable sample size for the calculation of confidence intervals, but keeping the number as low as possible due to computational time constraints, 30 repetitions were chosen.

To assess the model's performance over time, the trajectory dataset was then divided into three groups dependent on the date of occurrence (date ranges: D1 = [2022-06-23 – 2022-08-31], D2 = [2022-09-01 – 2022-09-30], D3 = [2022-10-01 – 2022-12-02]; see Table 4-1). A model with identical model architecture (see Figure 4-1 & Table 4-2) was trained on data of D1 movement trajectories for 30 epochs and the accuracy of the test set was calculated. Additionally, the model was tested on the datasets D2 and D3 and the accuracies were calculated. This process was repeated 30 times.

4.4 RESULTS

The mean accuracies and recall values for individual animals for the full dataset (including trajectories from all dates for training) are shown in Table 4-3. The mean accuracies and the 95% confidence intervals calculated with the 30 repeated measures indicate the proportion of correct overall predictions are over 0.5 demonstrating that the model is capable of predicting the identity of the animals based on the movement trajectories better than chance. This is the case for all datasets, however more complete trajectories lead to higher accuracies (see Table 4.3.). Overall, these results indicate the pipelines ability to measure individual differences exhibited in movement trajectories (as shown by Morita *et al.*, 2020).

The recall value was calculated as the proportion of the correct predictions of an identity and all relevant instances. Animals differed in their recall values mainly between individuals but also dependent on the number of detections potentially indicating a loss of information on individuality in these shorter trajectories. The differences between individuals could be an indication of more variability in movement in one of the animals.

Visible	**Mean accuracy**	**CI-95%**	**Recall Animal BM**	**CI-95%**	**Recall Animal BC**	**CI-95%**
100%	0.77	[0.74, 0.79]	0.66	[0.70, 0.88]	0.88	[0.85, 0.91]
83%	0.65	[0.64, 0.65]	0.53	[0.51, 0.54]	0.76	[0.75, 0.77]
50%	0.60	[0.59, 0.61]	0.48	[0.47, 0.49]	0.72	[0.70, 0.74]

Table 4-3, Model performance on dataset including x, y – coordinates only. Mean accuracy and recall scores with 95% confident intervals from 30 repeated training and test steps, including all days of recording using trajectories from two-minute time bins where an animal was visible for 100%, 83%, 50% of the time.

Accuracy and recall values for the full dataset using X and Y coordinates as well as the area of the mask (see Table 4-4) were comparable to the results of the full dataset with only X and Y coordinates. This indicated that the area of the mask was either a bad estimator for the distance to the camera or that the depth information added little information to identify individuals correctly.

Visible	Mean accuracy	CI-95%	Recall Animal BM	CI-95%	Recall Animal BC	CI-95%
100%	0.77	[0.75, 0.79]	0.68	[0.64, 0.71]	0.87	[0.84, 0.89]
83%	0.63	[0.62, 0.64]	0.52	[0.49, 0.54]	0.74	[0.72, 0.76]
50%	0.61	[0.60, 0.61]	0.53	[0.51, 0.54]	0.68	[0.67, 0.69]

Table 4-4, Model performance on dataset including x, y – coordinates and mask area. Mean accuracy and recall scores with 95% confident intervals from 30 repeated training and test steps, including all days of recording using trajectories from two-minute time bins where an animal was visible for 100%, 83%, 50% of the time.

The accuracy values for training on dataset D1 and additional testing on datasets D2 and D3 are shown in Table 4-5. Again, accuracies were above chance level of 0.5 over all datasets, indicating that the differences in movement were consistent over time. More specifically, the model still predicted identity correctly on a higher than chance level given movement trajectories recorded more than two months later.

Visible	Mean accuracy D1	CI_95%	Mean accuracy D2	CI_95%	Mean accuracy D3	CI_95%
100%	0.817	[0.762, 0.871]	0.671	[0.664, 0.678]	0.688	[0.673, 0.702]
83%	0.679	[0.651, 0.706]	0.58	[0.574, 0.586]	0.585	[0.581, 0.588]
50%	0.584	[0.570, 0.598]	0.532	[0.528, 0.536]	0.553	[0.549, 0.556]

Table 4-5, Mean accuracy scores with 95% confident intervals from 30 repeated training and test steps trained on the D1 dataset and tested on datasets D2 and D3.

4.5 DISCUSSION

This chapter found consistent individual differences in location trajectories of two laboratory-housed Rhesus macaques using a non-invasive video-based approach (Morita *et al.*, 2020). These results showed that the pipeline was capable of detecting individual differences and did so consistently over multiple months. The approach used here found these differences based on pixel coordinates that are a two-dimensional representation (image) rather than the three-dimensional location values. Contrary to the argumentation of Morita *et al.* (2020), the method here is used to indicate individual consistency in movement data rather than as a means of identifying individuals. Other more complex model architecture, such as increasing the number of LSTM layers, could potentially increase the accuracy of identity prediction further. All recordings were made during a period without changes in fixed enrichment (see Section 3.4), indicating a stable environment. Hence, the individual differences found are likely to represent spatial personality (Spiegel and Pinter-Wollman, 2022).

As expected, recall values differ between individuals, suggesting that animal BC showed more individuality in its movement trajectories compared to animal BM (Morita *et al.*, 2020). Individuality here indicates the amount of information present in the movement trajectory (i.e. predictability). Hence, this could be an indicator that animal BC was more consistent in its movement. As the total sample size for both individuals was equalised this is not due to an overrepresentation of one of the identities. Additionally, this is unlikely to be caused by a bias in identity predictions simply based on location. If pipeline predictions of identity were biased by the location of an animal or animals show clear location preferences, there would be a bias visible in the movement trajectory data. This does not seem to be the case (see Figure 4-2). However, the causality of changes in recall values are hard to interpret and further detailed analysis would be needed to indicate which underlying pattern in the movement lead to higher predictability rather than the recall itself. Hence, these values are excluded when looking at datasets over time (e.g. datasets D1, D2 and D3).

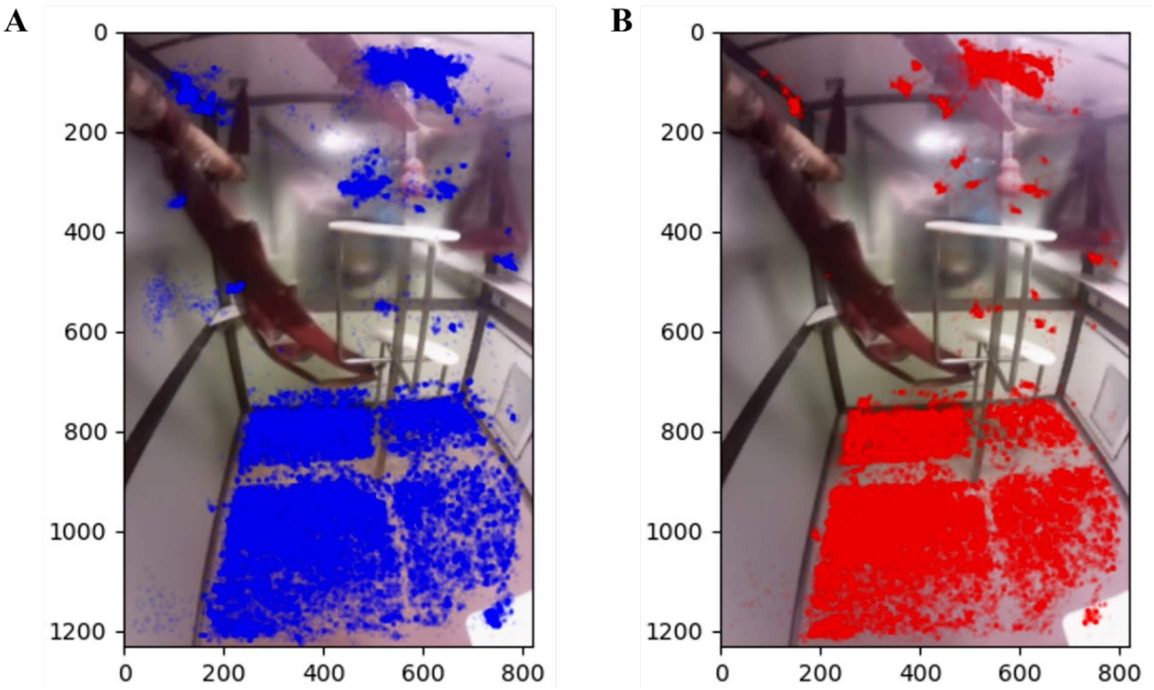

Figure 4-2, Representation of all location detections included for the movement trajectories dataset with at least 83% visibility. A, depicts animal BM’s locations in blue; B, depicts animal BC’s location in red.

The model decreased in accuracy with lower number of detections allowed. This is to be expected as it is likely that these trajectories hold less information. Additionally, these datasets could include part of multiple incomplete movement trajectories as animals were absent for some interval in the two-minute timeframe. Future studies could explore if the length of a trajectory (either by time or number of detections) influences the amount of individuality that can be expressed.

Mean accuracy of identity prediction stays above chance level even when training the model only on the dataset during the first two months of the data but testing it on data of the following month. This indicates the presence of consistency in the information on individuality in the movement trajectories. Furthermore, no degradation of this information is visible even when testing on data over a month apart from the training data set. More fine-grained path analysis could shed light on the specific differences in the individuals.

The capacity of the model to make higher than chance level predictions on animal identity are particularly intriguing as the environment of these animals during this period is limited and stable. The space the animals can use is constrained and so are the physically feasible paths to connect to these spaces. The constraints however are not specific to a given individual and hence should be shared and make the movement trajectories more similar than expected in a natural setting. The dataset is taken in a time period with no changes in the environment, meaning that the available spaces are consistent over time. Hence the individual differences the model is able to detect are unlikely to derive from spatial features in the environment, but rather underlying behavioural differences previously described as spatial personality (Spiegel and Pinter-Wollman, 2022; Stuber, Carlson and Jesmer, 2022). The home cage environment provides a useful setting to explore these concepts and models from different environment set ups should be compared as an environmental control.

This chapter showcased the potential of the pipeline to measure consistency and individuality in laboratory-housed Rhesus macaques. A bidirectional-LSTM model classified identity based on movement trajectories with higher than chance accuracy in a stable environment representative of spatial personalities. This approach highlights the efficacy of the developed pipeline in assessing both consistency and individuality in movement of laboratory-housed Rhesus macaques. Employing a bidirectional Long Short-Term Memory (LSTM) model, identity classification predicted based on movement trajectories yielded accuracy levels surpassing chance. Previously it was argued that individual differences in movement could simply be reflections of differences in the environment (Spiegel and Pinter-Wollman, 2022). As measures in this chapter are derived from a stable environment underlying differences are likely to reflect spatial personality.

CHAPTER 5. CONTEXT DEPENDENT DEVIATION IN SPACE USE BY LABORATORY PAIR-HOUSED RHESUS MACAQUES

5.1 CHAPTER SUMMARY

Animals exhibit consistent individual differences in movement and behaviour that are intricately linked to physiological measures and ethological needs. Automated approaches show promise in aiding the assessment of individual behavioural expressions which are related to animal welfare. With the help of the pipeline developed (see Chapter 3), this chapter compared daily location probabilities to measure variability in space use within individuals across different contexts as a potential indicator for changes in their welfare states showcasing the applicability of automated tools for individual-based welfare measures.

5.2 INTRODUCTION

Rhesus macaques (Macaca mulatta) serve as pivotal subjects in biomedical research due to their genetic and physiological similarities to humans, making them one of the most common non-human primate species utilised (Lear *et al.*, 2022; Mitchell *et al.*, 2018; Treue and Lemon, 2023). Regulations are in place to ensure their rigorous protection and the 3Rs principles are implemented (see Chapter 1). However, accurately evaluating the well-being of non-human primates poses a considerable challenge, given their social complexity and cognitive capacities (Carstens and Moberg, 2000; Hau and Schapiro, 2006; Lear *et al.*, 2022). Even though the benefits of automated systems to measure changes in welfare have been argued for (Eradus and Jansen, 1999; Knaebe *et al.*, 2022), there is a lack of studies on the actual application of such methods. Rhesus macaques are known to change their behaviour in the presence of human observers, masking signs of injury or illness (Gaither *et al.*, 2014). Hence, the use of video recordings is paramount to get more accurate information. As stated in Chapter 1 and Chapter 3, automating the evaluation of such video data can reduce human labour and bias. Furthermore, a bigger amount of data can be analysed and taken into account, assisting in the estimation of individualised levels of behaviour and the identification of deviations from typical behavioural patterns (Carstens and Moberg, 2000; Laudenslager and Boccia, 1996; Lutz *et al.*, 2022; Robinson *et al.*, 2018b). Movement generally describes the change of location. Often behaviours are linked to certain locations, for example foraging in captive housed macaques generally will happen on the ground where the wood shavings and forage are placed . Reinhardt (1992) showed that huddling behaviour is more likely to occur on perches. Chapter 4 discussed the consistency and individuality in space use by animals, indicated by the potential to identify individuals based on their movement trajectories. Hence, this Chapter explores the deviations in space use of individuals as an indicator for changes in their well-being.

The predictions of the pipeline are used to extract daily location probabilities for each individual and to compare them with the use of the Kullback-Leibler divergence (Kullback and Leibler, 1951). This is a statistical measure of distance also called relative entropy, that quantifies the differences between two probability distributions. This approach was selected as a generalised measure of deviations in space use over others as it allows to statistically compare frequency distributions rather than treating it as an alignment issue (e.g. Dynamic Time Warping), with the latter adding difficulties of how to treat missing data. Location distributions over days with different contexts that are likely to have affected the animal involved (e.g. surgery and sedations) were compared. I show that my pipeline is able to detect deviations in behaviour based on space use of each animal. To my knowledge, this is the first study showcasing the use of such automated video analysis methods to measure deviation of everyday home cage behaviour in primates with a particular focus on welfare.

5.3 MATERIAL AND METHODS

5.3.1 VIDEO DATA

Two female Rhesus macaques, both around the age of 6 years at the time of recording (Animal BM ~ 6 years and 1 month, Animal BC ~ 5 years and 11 months), were recorded in their home cage using a custom-made camera module (see Section 3.5). The module recorded 2750 videos of ten-minute length over a period of 36 days (~ 12.4 h per day) (Figure 5-1). For details on husbandry and ethics see Chapter 3.

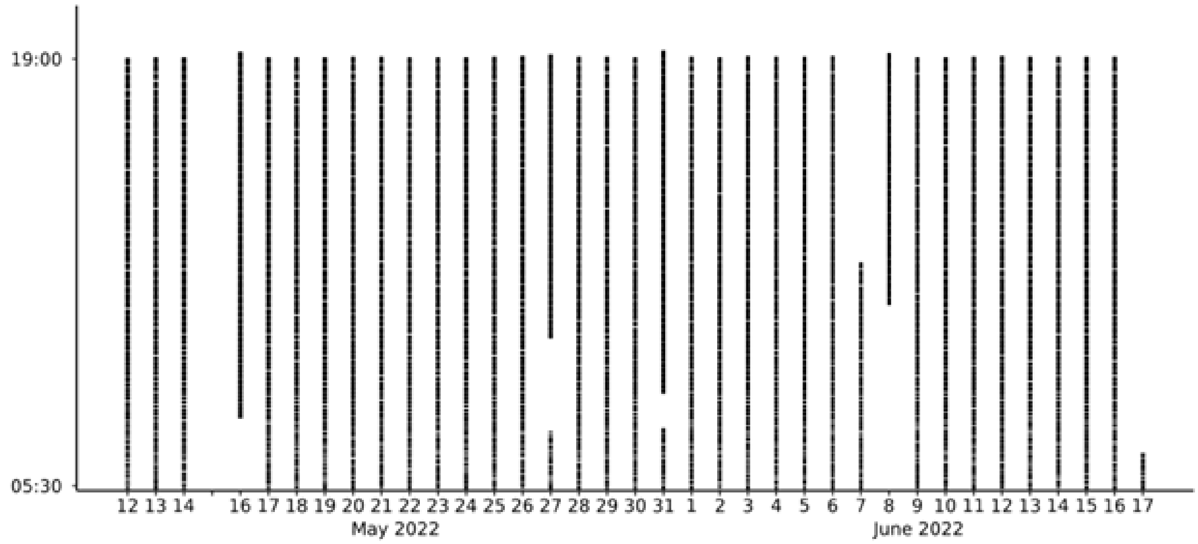

Figure 5-1, Representation of videos collected for each day using black bars from 5:30 to 19:00. Each dot in the bars represents a 10 minute video clip, white spaces indicate missing data.

5.3.2 EVENTS OF IMPACT

Five major events impacting the pair-housed female Rhesus macaques over the time of recording were identified; for a detailed description, see Table 5-1. All procedures and surgeries were part of an ongoing study and performed with full aseptic techniques and under general anaesthesia, additionally animals were under the appropriate postoperative care including pain relief and antibiotic cover. Generally, animals will first complete training on the research task. After successful training animals undergo two surgeries. During the first surgery the animal is implanted with a headpiece for head fixation and to mount electrodes. Additionally, electrodes are implanted in areas of interest (for example biceps brachii or flexor digitorum) and tunnelled subcutaneously to the headpiece. In the second surgery the animal was implanted with electrodes into the pyramidal tract (and potentially other regions of interest) (see Glover and Baker, 2020). Data collection of videos did not interfere with the procedures the animals were involved in.

5.3.3 LOCATION PROBABILITY DISTRIBUTIONS

The pipeline output was down sampled to a 2Hz frequency by taking the most frequent predicted location for each animal over all frames in that interval. Video recordings were made at a rate of ~36 frames per second so each interval consisted of predictions on ~13 frames. Predictions differing in the prediction of the left and right camera for each down sampled interval were excluded. Each animal locations were predicted with matching the centroids of the animal’s mask to a predefined location mask (see Section 3.6.3 and Figure 3-5). Due to an enrichment change on the 20th of May (during the “Jacket” event) two different maps were applied. All removable locations (e.g. swings, ropes, shelves, tubes, and boxes) are grouped within the same location (i.e. "enrichment") to avoid an impact on the location estimation. The location probabilities for each animal on each day with video data available were calculated by counting the detections in each location divided by the total number of detections for that day. The Kullback-Leibler (KL) divergence was used to calculate a measure of difference between two probability distributions. The KL-divergence is also called relative entropy and is a measure of the expected excess of information content (or surprise). The KL-divergence is calculated as follows: $DKL(P\,||Q) = \sum_{x \in X} P(x) \log\left(\frac{P(x)}{Q(x)}\right)$

KL-divergence between each of the daily distributions of event and control days for each animal were compared using the "KL" package (Drost, 2018) and a two-way ANOVA in R (4.3.0).

Event	**Description**	**Number of Days**
"EMG"	EMG-implant and head piece surgery of Animal BC; ➢ Sedation (Animal BC) on the 12th of May: start not indicated but back in home cage at 11PM;	4
"Jacket"	Animal BC was wearing a jacket after the EMG-Implant surgery to aid recovery and keep wounds clean and intact; ➢ Sedation (Animal BC) on the 16th of May: start at 2:20PM and returned to home cage at 16:50PM	4
"Recovery"	Animal BC jacket was removed but still on pain medication;	3
"PT"	Animal BC was sedated for a subcortical electrode implant ➢ Sedation (Animal BC) on the 26th of May: start 8:30AM and returned to home cage at 14:30 PM	
"Fight"	Animal BM was sedated and operated on her left hand (D4 sutured and D5 amputated below the knuckle) after a potential fight with cage mate ➢ Sedation (Animal BM) on the 1st of June: start 19:20PM and returned to home cage the next day	4
"Control"	Days without remarks on special circumstances, Animal BC given diazepam on some days	12

Table 5-1, Description of events and the number of days the event spans

5.4 RESULTS

A two-way ANOVA revealed a significant effect of animal identity ($F(1, 6) = 26.549$, $p <.001$) and events ($F(1, 6) = 94.72$, $p <.001$) but no significant interaction between the effects of event and animal identity on the KL-divergence ($F(1, 6) = 1.56$, $p =.17$). Post-hoc paired t-tests with Tukey correction revealed that the KL-divergence within control days was significantly smaller than KL-divergence during days of the "EMG", "Jacket" and "Recovery" events from "control" days in animal BM ($t(828) = -11.44$, $p < .001$; $t(828) = -11.56$, $p < .001$; $t(828) = -4.28$, $p < .001$, respectively) and animal BC ($t(828) = -8.90$, $p < .001$; $t(828) = -13.01$, $p < .001$; $t(828) =$

-4.40, *p* < .001, respectively), but not compared to other events (see Figure 5-2). Sedation and surgery do not seem to be the main factor impacting the changes in behaviour measured by the approach presented here, as evidenced by the absence of significant alterations in location probabilities during events such as "PT" or "Fight".

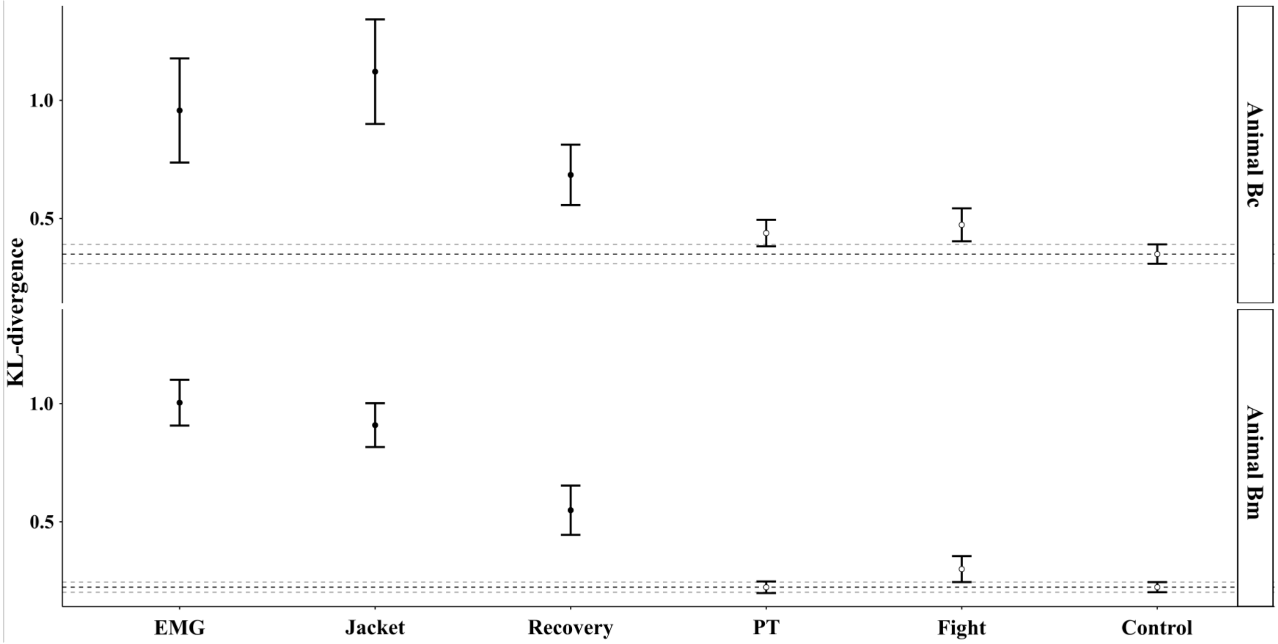

Figure 5-2, Kullback-Leibler (KL) Divergence of event days to control days for each animal. Points represent the mean and error bars represent the 95% confidence intervals. Filled points show that the divergence is significantly different from the control (Post Hoc with Tukey correction). Horizontal lines indicate the control mean (black) and 95% confidence interval of the control (gray).

To estimate the impact of mislabelling an animals' location or identities locations were randomly resampled for the data at different chance levels to induce noise. For this the down sampled output data from the pipeline is taken and for each prediction there is a 40%-0% chance of being replaced by a random location. Changes of the KL-divergence stayed significant even when 40% of the location data were randomly sampled from all locations. Combined with the validation from Section 3.7 that suggested the pipeline has a chance of mislabelling identity or location at less than 20% this indicates that the results are robust and not due to misdetections (see Figure 5-3).

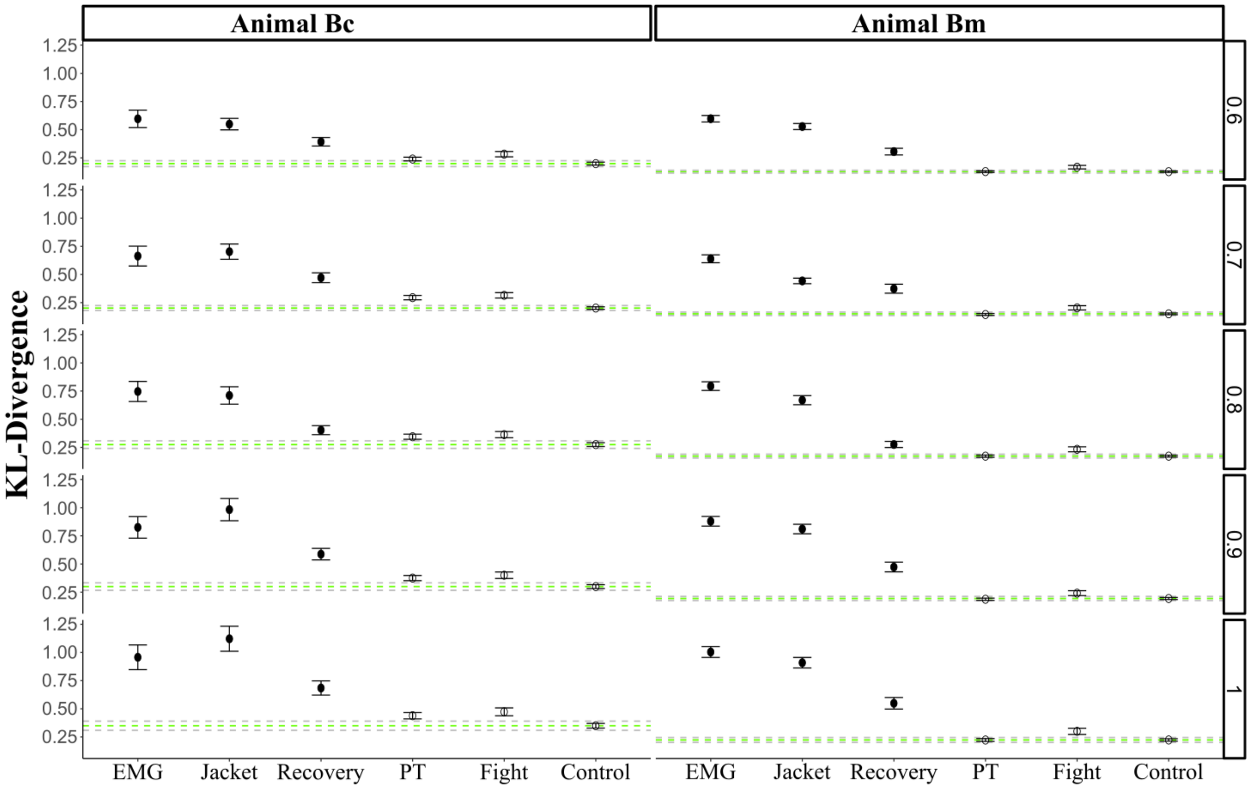

Figure 5-3, Kullback-Leibler (KL) Divergence of event days to control days for each animal with random resampled data. The right facets indicate the percentage of original data contained (e.g., at 0.6 40% location predictions in the data were changed to a randomly chosen location). Points are the mean and error bars represent the 95% Confidence Intervals. Filled points show that the divergence is significantly different from the control (Post Hoc with Tukey correction). Horizontal lines indicate the control mean (green) and 95% confidence interval of the control (gray).

As Animal BC was given Diazepam on seven out of twelve of the control days, I ran an additional test to assess differences in the KL-divergence of control days with Diazepam (*Mdn* = 0.250 [0.226, 0.279]) and control days without Diazepam (*Mdn* = 0.256 [0.228, 0.304]), but there was no significant difference (*W*= 2380, *p* = 0.732).

Daily differences in the KL-divergence were visually inspected (see Figure 5-4). On the days of sedation, there is a divergence in the location probability distributions indicating differences in space use in both animals on these days (12th, 18th, 26th of May and 1st of June) compared to space use on control days. This seems to extend in the case of the 18th of May and the 1st of June, in both these cases the sedation took place in the afternoon, 2PM and 7.20PM respectively, which could be the reason the effects are still recognisable on the following day.

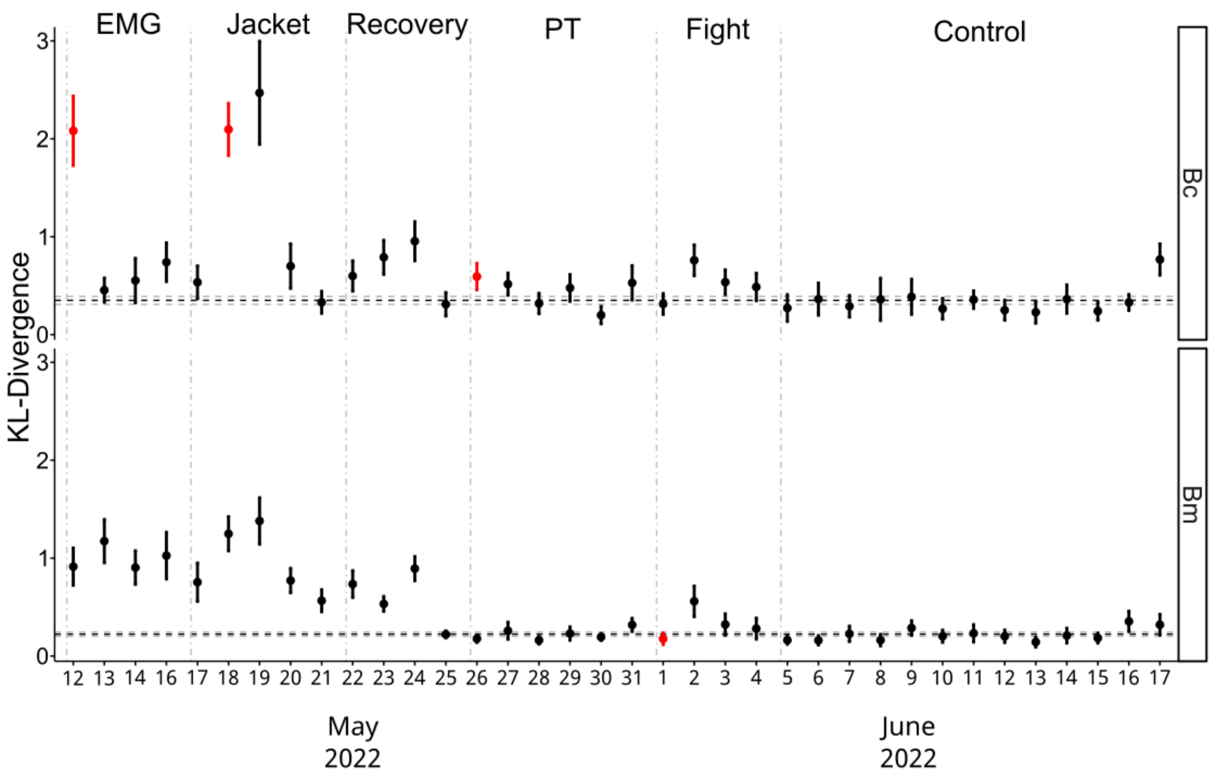

Figure 5-4, Kullback-Leibler divergence of each day compared to control days for each animal. Points are the mean and error bars represent the 95% Confidence Intervals. Red points indicate days of sedation or anaesthesia. Horizontal lines indicate the control mean (black) and 95% confidence interval of the control (gray). Vertical lines indicate the breakpoints of events labelled on top of the graph.

5.5 DISCUSSION

This chapter uses the KL-divergence of the locations probability distribution to provide a daily measure of deviation in space use. This serves as an estimated summary of the overall behaviour patterns of the animal and is not a direct measure of impact severity of a procedure. Animals show individual differences in movement patterns that are consistent over time as shown in Chapter 4. But behavioural patterns can be impacted by welfare states such as post operative pain (Descovich *et al.*, 2019). The results follow my expectations by indicating significant changes in space use after events impacting an animal's welfare. Moreover, these effects are visible in both pair-housed animals independent of which animal underwent a procedure. The results suggest that behavioural changes are caused by an impact on the social dynamics of these two animals rather than just by sedations or post operative effects. The KL-divergence between both individuals show high similarity in trends, as both animals showed changes in their use of space after one of them was fitted with a headpiece and was wearing a jacket. However, other instances of sedation did not lead to significant differences in KL-divergence and the daily effects seem to subside within a day. Rather, the change in the appearance of animal BC, attributable to both the head-implant (a long-term change across all events) and the jacket (a short-term change during "EMG" and "Jacket"), appears to coincide with behavioural shifts in both animals. This alteration, coupled with potential post-surgical behavioural changes in animal BC during recovery, may disrupt the social dynamics between the animals. This could further explain the subsequent fight that happened between the pair after animal BC had recovered. Animal BC was reported to be dominant over Animal BM (according to researchers and animal technicians) prior to any implants, thus the impacts after the EMG surgery could have led to an instability in the social composition of the pair. In addition, the highest daily divergence seems to be on the 19th of May. On this day there are no special remarks in either of the animals' reports, however both animals showed bruising on their cheeks in the afternoon on the 18th of May, which might indicate a scuffle over food that could have further impacted the relationship between the animals and let to changes in their behavioural patterns in the following day. As mentioned in the results, another possible explanation is the sedation of Animal BC that took place in the afternoon that could have an effect on use of space on the next day. Rhesus macaques have a steep hierarchical structure, which are often maintained through aggression (Bernstein and Ehardt, 1985; McCowan, Beisner and Hannibal, 2018; Thierry, 2004), and changes in the hierarchy can lead to stress (Wooddell *et al.*, 2016). Similarly, novelty can lead to stress in Rhesus macaques and hence lead to deviations in behaviour. Identifying these potential causes could lead to changes in procedures by for example habituating animals to the jacket and implant before the procedure or before the reintroduction of the animal into the enclosure.

Applying this measure could in the future help in more specifically identifying events of impact in socially housed laboratory primates. Additionally, while sedations and post operative effects did not show lasting differences in divergence this could be due to appropriate use of care procedures and medication. The assessment of pain after surgical procedures is difficult in animals, and institutions handle this differently (Paterson and Turner, 2022). The proposed method of analysis works on a broad timescale and hence is not suitable to assess pain. Further methods scanning for behavioural changes on a smaller time scale should be applied on the video data recorded in the future. These include algorithms trained on detecting behaviours such as hunching or facial expressions that could be applied to assess pain more specifically (Gris *et al.*, 2024; Poirier and Bateson, 2017). The presented pipeline could be extended with pose estimation algorithms to identify hunched posture (see Section 7.2.2). However, the current proposed cameras do not provide the resolution needed for such facial expression algorithms, but adding additional high-resolution cameras in combination with the current methods is possible (see Section 7.2.1). As Rhesus macaques are known to hide negative behaviours in the presence of human observers, using remote monitoring methods would be advantageous (Anderson and Perona, 2014; Ardoin and Sueur, 2023; Gaither *et al.*, 2014; Hennessy, Chun and Capitanio, 2017; Mathis *et al.*, 2018; Porto *et al.*, 2013; Weinstein, 2018). Furthermore, current existing methods also overlook the impact on the other animal in the enclosure and often focus on signs of pain and physical symptoms of one specific individual (Gris *et al.*, 2024).

Any measure which relies on comparing animal location between days will be adversely affected by the regular changes in placement and types of enrichment items such as swings which is common practice in well-run facilities. The change in cage enrichment on 20th of May was addressed by grouping the mobile enrichment into one category so it did not impact the labels used for the probability distributions, it could however have an effect on the frequency with which animals used the new types of enrichment. As there is a significant difference in space use during the recovery event, that is after this enrichment change, this is less likely to be the explanation for the higher divergence, however it could have an obscuring impact on the derived measure.

Even though Diazepam might have an influence on the behaviour of Rhesus macaques there was no difference found in the KL-divergence of the days Animal BC was dosed compared to days where this wasn't the case (Kumar *et al.*, 1999; Rose *et al.*, 2012). The location probability distributions are a way to condense the vast amount of data collected to an interpretable measure. Rose *et al.* (2012) reported that general measures of movement did not differ in animals dosed with diazepam and only closer inspection could find such differences. Measures suitable for smaller time frames could be applied in the future to see if this method is capable of identifying subtle changes in behaviour such as velocity measurements, which can also be extracted using the pipeline. However, the use of diazepam is monitored heavily and aimed to mainly impact the period of time when animals are performing their task in the lab.

Overall, this chapter showcases the useability of the pipeline in the scope of measuring deviations in behaviour related to welfare. It highlights the importance of consistent observations to infer the causal influences on these changes in patterns and to potentially find levels of normality for each animal.

CHAPTER 6. ASSOCIATING AUTOMATED MEASURES OF MACAQUE BEHAVIOUR TO PERSONALITY TRAITS

6.1 CHAPTER SUMMARY

Behavioural expressions in animals show consistent individual differences often described as personality. This chapter investigated associations between personality components and automated behavioural measures in laboratory-housed Rhesus macaques using the camera system and pipeline developed as part of this thesis. The results underscore the potential and feasibility of automated tools in complementing traditional methods for assessing animal personality, offering insights into behavioural needs and welfare.

6.2 INTRODUCTION

Animal personality describes individual differences that are consistent over time and context (Réale *et al.*, 2007). Chapter 2 reviewed literature on Rhesus macaque personality, how it can be measured and what physiological, welfare, life history and behavioural associations have been found. Generally, personality components are measured using behavioural coding or trait ratings. However, using multiple approaches could be beneficial to overcome drawbacks from specific methods and to get a more comprehensive view of personality (Norman, Rowden and Cowlishaw, 2021).

Non-human primate's welfare and well-being have been associated with individual differences in personality. In general, primates described as more gregarious, active, social, stable and bold showed increased signs of welfare and well-being compared to less socially connected, inactive, anxious and emotional conspecifics (see Robinson and Weiss, 2023). Similarly, ratings in behavioural reactivity to novel situations have shown Rhesus macaques to be a valid model for behavioural inhibition in humans, a trait related to anxiety problems (Wood *et al.*, 2024) and low levels of sociality have been used as a primate model for impaired social behaviour in humans with autism spectrum disorder (Parker, 2022). Welfare has been argued to be a psychological construct that is measured by using an animal's expression of behaviours. The behavioural expression is based on an animal's ability to cope with its environment (Broom, 2001; Broom, 1991) and is a combination of emotional experiences, biological functions and an individual's behavioural needs (Hemsworth *et al.*, 2015). Overall, combining a multitude of measures leads to a more comprehensive picture of the animals' welfare (see Robinson and Weiss, 2023). Getting insight into an animal's personality composition could hence indicate the monitoring intensity required for an animal, such that animals known to be anxious are more likely to be injured and could need more frequent interventions. This could not only help in reducing labour for animal technicians, but also refine the monitoring strategy.

Chapter 2 also described associations of personality components and behaviour, for example Capitanio (1999a) showed that animals rated higher in sociability expressed more affiliative behaviours towards conspecifics than animals with lower ratings. Moreover, Chapter 4 showed the individual variability in movement trajectories (i.e. a measure of locomotion behaviour) of laboratory-housed macaques and how this is indicative of spatial personality. Further studies showed differences in space use and a preference for high-level regions in Rhesus macaques. This preference has been shown to be higher in the evenings than mornings (Clarence *et al.*, 2006) and was influenced by age and rank (Reinhardt, 1992), where older animals preferred fixed rather than mobile platforms and higher ranking animals spent more time in elevated positions. Lutz and Brown (2018) specifically found an effect of temperament and cage location on the use of porches in Cynomolgus macaques. This study rated animals based on their latency to approach during a novel object test as either bold or intermediate temperament. These primates were housed in stacked cages and the usage of a mesh porch was measured. Animals rated as intermediate during the novel object test were less likely to be on the porches when housed in the lower cages.

The technological advances in the last decade provide a variety of tools to measure behaviour and there is an increase in the availability of such methods (see Section 1.6.; Wiltshire *et al.*, 2023). These automated approaches can not only decrease time and labour efforts to measure behaviour, but also reduce observer bias (Anderson and Perona, 2014; Ardoin and Sueur, 2023; Gaither *et al.*, 2014; Mathis *et al.*, 2018; Porto *et al.*, 2013; Weinstein, 2018). New technologies could also be applied to measure temperament in animals (e.g. Bethell *et al.*, 2019b; Ciminelli, 2023). Automating personality analysis would make it feasible to include longer observations periods, more behavioural measures and obtain predictors that are more objective. This has been done in mice automatically scoring behaviours and using factor analysis to get ratings on so called identity domains (Forkosh, 2021).

This chapter explored the association of automated measures of behavioural frequencies to ratings in personality using the Hominoid Personality Questionnaire (Weiss, 2017; Weiss *et al.*, 2009) and approach latencies in a novel object task (Massey, 2023). I expected to find animals that scored higher in activity to have a higher frequency of location changes and a higher frequency of expressing high velocity. Moreover, I expected animals that are rated higher in confidence, lower in anxiety and have a lower neophobia score to show higher frequencies of spending time in front of the cage (in areas opposing corridors where animals are potentially closer to humans). Animals rated higher in dominance are expected to spend more time in higher locations. Lastly, the proportion of detections in the same location as the cage mate are expected to be higher with individuals scoring high in friendliness. Overall, this chapter showcased the potential to automate measures of personality in laboratory-housed Rhesus macaques.

6.3 METHODS

6.3.1 HOMINOID PERSONALITY QUESTIONNAIRE

To assess personalities of Rhesus macaques the Hominoid Personality Questionnaire was used (HPQ; Weiss, 2017; Weiss *et al.*, 2009). This questionnaire is suitable for macaques housed in captivity and derived components related to welfare and well-being (Robinson *et al.*, 2021b). A recent review found the HPQ to be one of the most commonly used measure for primate personality (Norman, Rowden and Cowlishaw, 2021). The questionnaire consists of 54 items represented by an adjective and a descriptor. For example, the item *fearful*: "*FEARFUL: Subject reacts excessively to real or imagined threats by displaying behaviours such as screaming, grimacing, running away or other signs of anxiety or distress.*" Shorter versions of the HPQ have been validated (Hopper, Cronin and Ross, 2018; Robinson *et al.*, 2018b), however I was advised to use the full questionnaire to get the most amount of personality variation possible (personal communication with L. Robinson). A total of five raters scored eleven individuals, four animals were rated by multiple observers (see Table 6-1). The inter-rater reliability was calculated using intraclass correlations in the 'irr' package (Gamer *et al.*, 2012) in R (4.3.3) for all animals that were rated by multiple observers (see Table 6-1).

The scores ranging from 1-7 for the personality domains of Confidence, Openness, Dominance, Friendliness, Anxiety and Activity were calculated using the loadings available from another publication (Robinson *et al.*, 2021b). While interrater reliability was in a similar range to previous studies (Robinson *et al.*, 2021b), only one rater (EW) rated all of the individuals and since interrater reliability is rather low (< 0.7) and ratings varied widely across different raters, only ratings of EW were used (Polit, 2014). EW was a senior animal technician for multiple years at the time caring for the animals daily and having observed these animals in their home cage for much more than 30 minutes each. I additionally calculated Cronbach's α reliability coefficient to assess internal consistency of each component using the loadings of the trait ratings from a single rater and from multiple raters (see table 6.2.; Cronbach, 1951). Cronbach's α is a measure of variability of the item ratings and the overall variability of the personality trait score. It indicates the internal consistency of the item ratings. Cronbach α values above .7 are considered acceptable reliability (George and Mallery, 2003). Again, the higher Cronbach α values from single rater scores suggest a better representation of the personality components when using the single rater over multiple rater's assessments. The Cronbach's α reliability coefficient and the confidence intervals were calculated using the 'psych' package (Revelle and Revelle, 2015). The confidence intervals presented are derived using Duhachek and Iacobucci (2004), this method adjusts the standard errors with the serial correlation and does not assume the data to be normally distributed.

Animal ID	Number of raters	ICC	CI-95%	p-value
BC	2	0.59	[0.39, 0.74]	< .001
BM	2	0.57	[0.36, 0.73]	< .001
CS	2	0.19	[-0.09, 0.43]	.087
MB	3	0.40	[0.23, 0.56]	< .001

Table 6-1, Inter-rater reliability, confidence intervals and number of raters for each animal where multiple observers were present.

	multiple raters		**single rater (EW)**	
Component	**α**	**CI-95%**	**α**	**CI-95%**
Confidence	0.7	[0.45, 0.96]	0.86	[0.72, 1.01]
Openness	0.69	[0.41, 0.96]	0.85	[0.71, 0.99]
Dominance	0.89	[0.8, 0.97]	0.9	[0.81, 0.99]
Friendliness	0.48	[0.07, 0.88]	0.75	[0.51, 0.99]
Activity	0.89	[0.8, 0.98]	0.82	[0.64, 1.01]
Anxiety	0.73	[0.53, 0.94]	0.78	[0.53, 1.03]
Total	0.89	[0.82, 0.96]	0.89	[0.8, 0.99]

Table 6-2, Cronbach's α and confidence intervals calculated using Duhachek and Iacobucci (2004)

6.3.2 NOVELTY TESTING – NEOPHOBIA SCORE

The Rhesus macaques observed for this thesis were all bred at the Medical Research Council's Centre for Macaques (CFM) except for animal MB who was bred at the UK Health Security Agency (UKHSA) before being transported to the CBC for use in scientific studies. While at CFM, seven of the animals were rated on their temperament using novelty testing, as part of another research project, (Massey, 2023), and their scores were included in this chapter.

For the novelty test, animals were separated from other animals in the colony after feeding to make sure hunger was not a motivation to engage with the new stimulus. The test consisted of five phases, the familiar food phase (FF), the novel food phase (NF), two novel object phases (NO-1 and NO-2) and a phase where the animal was left alone after the novelty tests for two minutes. For the FF phase raisins were used that are part of the training routine in CFM. These were placed on a wooden shelf and the animal was left alone for five minutes. Following the FF phase the raisins were removed, and six pieces of novel food (i.e. coloured sweets) was placed on the shelf and the animal was left alone for two minutes. After this, remaining NF stimuli were removed and replaced by a novel object (NO-1 was a toy with no eyes). After another two minutes the first novel object was removed and the second object was placed in front of the cage (NO-2, a toy with eyes). Both novel objects were attached with a carabiner and could not be removed. Again, the animal was left alone for two minutes before the NO-2 was removed and the last phase began in which the animal was left with no stimulus for two minutes. The detailed procedure and ethogram can be found in Massey (2023). The neophobia score was derived by z-scoring the variables 'behind visual barrier', 'escape attempts', 'freeze' and the latency variables for the FF, NF, NO-1, and NO-2 phases, normalising them to a mean of 0 and a standard deviation of 1. This was done on the data provided by David Massey for all animals that had been scored and were present for behavioural recordings (see Table 6-4). A PCA on the normalised scores was conducted to retain one component for neophobia (Massey, 2023) using the 'principal' function in the 'psych' package (Revelle and Revelle, 2015). In my subsample of animals, 'escape attempts' as well as 'Latency NF' showed little loadings on my neophobia component (< 0.4) the component was recalculated without these behavioural measure (see Table 6-3). All remaining variables had salient loadings (> 0.4) and the derived component explained 75% of the total variance of the new neophobia component used in this chapter.

Variable	**Loading**	**h^2**	**Loading**	**h^2**
Behind visual barrier	0.81	0.65	0.83	0.69
Latency (FF)	0.95	0.9	0.96	0.92
Latency (NF)	0.01	0		
Latency (NO-1)	0.9	0.81	0.88	0.77
Latency (NO-2)	0.9	0.82	0.88	0.77
Freeze	0.75	0.57	0.78	0.61
Escape attempts	0.17	0.03		

Table 6-3, Loadings of the variables into the principal component. Loadings on the right side depict the component used to calculate the neophobia score. FF, NF, NO-1, and NO-2 describe the phases of the novel object test, familiar food, novel food, novel object without eyes, novel object with eyes respectively. h^2 is the amount of variance in the variable explained by the retained component.

6.3.3 AUTOMATED BEHAVIOURAL MEASURES

Using three of the camera systems in the CBC, I identified periods between 12AM-14 PM and 17-19 PM where footage was available for the full two hours for each period on 10 occasions (the afternoon and evening periods do not need to be on the same day). This was the case for seven individuals from which six had been scored on the personality questionnaire and five had been scored in temperament at the CFM (see Table 6-4). The pipeline (see Section 3.6) was used to analyse locations and velocities of these seven individuals (see Table 6-4). Using R (4.3.3) the proportion of overall detections of an animal in an elevated position and in the front part of the cage for each afternoon and evening period was extracted. Detections that did not correspond to the location the animal was previously in were identified, and the proportion of location changes over all detections was calculated for each period. Similarly, the detections where an animal was in the same location as their cage mate were identified to calculate the frequency of being in social proximity. One animal (CF) was singly housed at the time of recording, so the social proximity measure did not apply. This measure is an estimate of tolerance as it does not provide information on which individual started or ended the social contact. Instances of high velocities were defined based on the scores falling in the third quartile of all velocity measures and were again measured as proportions over all detections of an afternoon or evening period. Since all measures are proportions of all detections within a given period, all scores fall between 0 and 1.

6.3.4 CORRELATION ANALYSIS OF BEHAVIOUR AND TEMPERAMENT

To investigate the relationship between behavioural measures and temperament scores correlation tests were used. Correlation analysis was deemed appropriate as it allowed for the exploration of potential associations between variables without assuming causality, aligning with the complex and multifaceted nature of primate personality and behaviour. Correlation coefficients larger than .4 indicate moderate and values smaller than .4 weak correlation (Akoglu, 2018). Two correlation tests were conducted separately for correlating behavioural measures with personality ratings and correlations of behavioural measures and the derived neophobia score from the temperament test conducted at the CFM before animals came to the CBC in Newcastle.

Animal ID	Automated behavioural measures (CBC)	Personality trait ratings from HPQ (CBC)	Novelty testing data (CFM)
MB	✓ (162.4)	✓	
CS	✓ (133.6)	✓	✓
CF	✓(132.9)	✓	✓
CI		✓	✓
CA		✓	✓
BM	✓ (551.4)	✓	✓
BC	✓ (362.6)	✓	
BZ	✓ (898.2)	✓	✓
BL	✓ (297.5)		✓

Table 6-4, Data availability for automated behavioural measures, personality trait ratings using the HPQ and novelty testing data from (Massey, 2023). The brackets in the automated behavioural measures column depict the number of minutes the animal was detected by the camera module during the time of observation.

6.4 RESULTS

Automated behavioural measures were obtained from seven individuals of which six were rated using the HPQ (see Figure 6-1) and five were tested on neophobia at the CFM (see Table 6-4). The correlations of the personality components rated using the HPQ with the automated behavioural measures are presented in Table 6-5. Moderate positive correlations were found between confidence, openness, friendliness, and activity ratings with the proportions of high velocity detections, frequency of location changes, proportion of detections where animals were in proximity with their cage mate and frequency of detections individuals were located in the front of the cage. Friendliness and activity also showed a moderate correlation with increased frequency of detection in the front of the cage. While significant, neither the proportion of detections in an elevated location nor the anxiety trait ratings showed a correlation coefficient larger than 0.2.

The neophobia score estimated from the novel object tests conducted at the CFM showed weak correlations with the proportion of high velocity detections and location changes as well as a decrease in the proportion of detections in an elevated location (see Table 6-6).

	Confidence	**Openness**	**Dominance**	**Friendliness**	**Activity**	**Anxiety**
High velocity	0.41 **	0.29	0.09 **	0.21	0.28	-0.07 **
Location change	0.37 **	0.29 **	0.12	0.14	0.28	-0.01
Elevated	-0.26 **	-0.02 **	0.09	-0.04	-0.02 **	-0.09
Front	0.03 **	0.27	0.14 **	0.19	0.28	0.02 *
Social proximity	0.46 **	0.35	0.51	0.09	0.28	0.22 **

Table 6-5, Correlation coefficients of personality components and automated behavioural measures (*N* = 204) derived using the spearman's rank correlation test, significance levels were calculated using Holm adjustment, * < .05, **< .01.

The automated behavioural frequency measures are shown in Figure 6-2. A three-way ANOVA was performed to analyse the effect of the type of behaviour, the time of day and the animal's identity on the frequency of occurrence of the behaviour. There was no significant interaction between all three (*F*(1, 19) = 0.60, *p* =.539), however there was a significant interaction between the type of measure and animal identity on the frequency of occurrence (*F*(1, 19) = 3.95, *p* <.001). Additionally, there was a significant effect of animal identity on the measured frequency of behaviours (*F*(1, 5) = 8.38, *p* <.001) but no effect of time of day on the recording on frequency of behaviours (*F*(1, 1) = 1.19, *p* =276).

	Neophobia score
High velocity	0.35 **
Location change	0.31 *
Elevated	-0.43 **
Front	-0.19
Social proximity	-0.15

Table 6-6, Correlation coefficients of neophobia score and automated behavioural measures derived using the spearman's rank correlation test, significance levels were calculated using Holm adjustment, * <.05, **< .01.

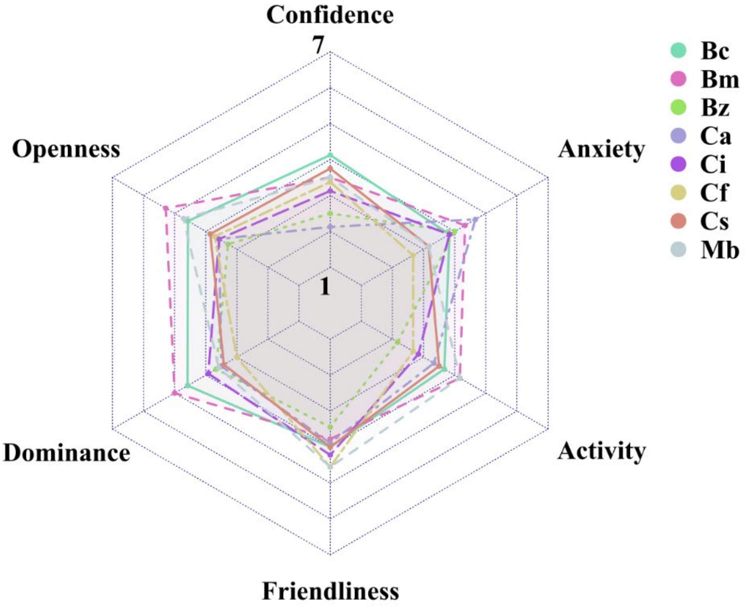

Figure 6-1, Personality trait ratings derived from the HPQ, the webbed lines scale from 1-7 coloured lines depict the score for each animal.

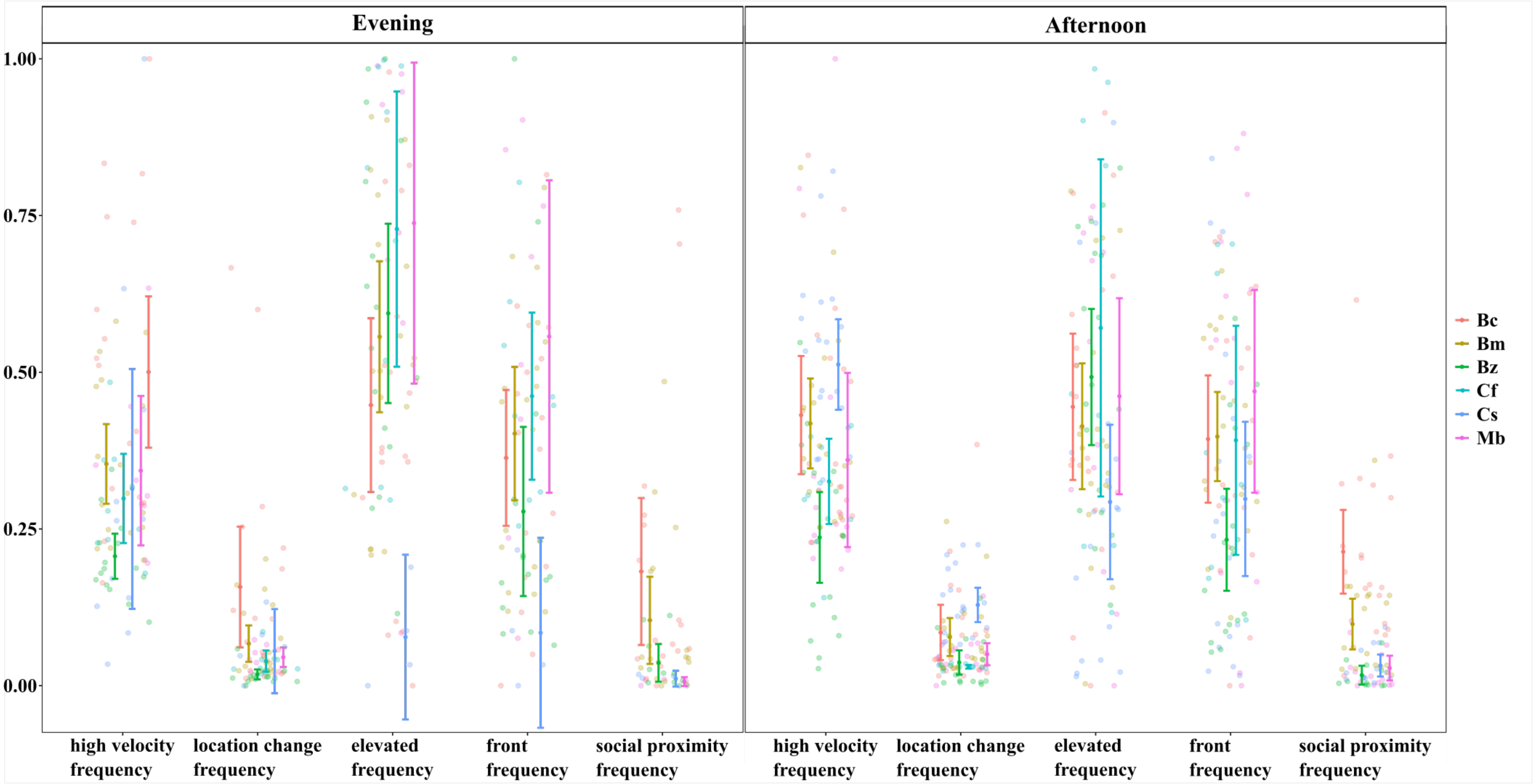

Figure 6-2, Automated behavioural frequency measures for each animal for the evening and afternoon period. Coloured dots depict single measurements and coloured lines depict the mean and 95% confidence intervals, colours are unique for a specific animal identity.

6.5 DISCUSSION

This chapter found significant correlations between automated measures of behaviour and personality ratings derived from the HPQ as well as the neophobia score measured at the breeding facility at least a year prior to the behavioural measures. Contrary to what was expected, higher ratings in activity showed no correlation with the frequency of location changes and detections with high velocities. Furthermore, animals that scored higher on the friendliness trait were not detected more often in proximity to their cage mate. This could potentially mean that our behavioural measures are not good representations of the personality traits we assessed. Nevertheless, we do see weak and moderate correlations of four of the behavioural measures and the confidence trait. There is a weak negative correlation of confidence ratings and the frequency an animal was in an elevated position in the cage. This correlation is negative when compared to the neophobia score. The latter is unexpected as elevated positions are generally locations of safety for (semi-)arboreal species. A possible explanation for this is that most of the top halves of the pens are made of glass and some areas of the floor might provide a better barrier from visual contact. Confidence also shows a moderate correlation with animals moving at high velocity in the cage as well as a weak effect in the location change frequency. This could be indicative that confident animals spend more time exploring and making use of their environment. The neophobia score also showed weak significant positive correlation with frequencies of high velocity and location changes. While this could indicate an animal to be ‘jumpy’ it is unexpected as these behavioural measures also showed a moderate positive correlation to the confidence and openness ratings. Massey (2023) used the neophobia score as a descriptor for a tendency to avoid novelty. However, all animals were already acclimated to the CBC and their cages for more than two months at time of recording. The low correlations with behavioural scores might indicate that since the animals were used to their environment, the extracted behavioural measures might be better predictors of animal personality than the neophobia scores. Possibly animals could be assessed for behaviours and performance during training for a new task to find more representative measures for this.

As mentioned in Section 6.3.1 the inter-rater reliability for animals that were rated by multiple observers was relatively low and only ratings from EW were used to calculate the personality scores. EW is one of the technicians working at the CBC and the ratings are probably more reflective of the animal’s behaviour in the home cage. The other raters for these animals were researchers and the personality ratings might be influenced more by how the animal behaved in the laboratory while training and conducting the tasks rather than behaviour in the home cage.

Mouchet and Dingemanse (2021) found that measures of personality in laboratory- and field-based tests reflected expressions of distinct characteristics and do not always lead comparable outcomes. This could similarly be the case here in the comparison of home cage- and laboratory-based view of the animal.

Overall, animals scored relatively close (never more than two points apart on a seven-point scale) in all personality ratings (see Figure 6-1). This could be because these animals represent a biased sample as they were selected for use in research by the CFM. Nevertheless, the behavioural measures monitored using the pipeline still show relevant correlations to the personality measures. Furthermore, the behavioural frequencies differed depending on the behaviour scored and between individuals. Also, a main effect on behavioural frequency was found between animals but not between afternoon or evening recordings. This suggests that the presented approach is sensitive to individual differences in behavioural expression and could be used for personality assessment.

Forkosh *et al.* (2019) presented a mathematical framework to describe individual differences from automated longitudinal measures derived from group-housed mice. This study used 60 behavioural readouts. Here I present five example behaviours that fit within the framework presented by Forkosh *et al.* (2019) showing the potential of future studies to create more of these readouts and automate measures of personality in Rhesus macaques.

This chapter showed the potential of the created pipeline to measure behaviours associated with personality trait ratings. The results underline the potential and feasibility of complementing current methods to identify levels of behavioural expression as well as behavioural needs with the use of automated tools. These could not only give inside to the composition of primate personality but help in the monitoring of changes in the expression of behaviour that could occur after animals are moved to a new institution. Overall, this could impact the refinement and increase welfare for animals housed in captivity.

CHAPTER 7. GENERAL DISCUSSION

7.1 SUMMARY OF KEY FINDINGS

This thesis described and showcased an in-cage camera module and pipeline to automate monitoring of Rhesus macaques within their home environment. The long-term data collected from this is leveraged to identify individual differences and consistency in movement and space use. This showcases the capacity for automated methods to measure interindividual differences in captive primates consistently. The application of these methods could hence not only impact colony management but advance research both in behavioural ecology by exploring measures of personality (see Chapter 6) and by aiding in the phenotyping of animal models for biomedical research (e.g. Silverman *et al.*, 2022). Further, the work focuses on deviations in these measures in contexts of periods likely to affect welfare. The capacity to monitor the animals on a high temporal resolution is leveraged to focus on intraindividual differences in behaviour indicating changes in behaviour on an individual level. This could not only refine welfare through the identification of impactful events but refine research by providing more nuanced measures of impact in biomedical experiments within the home cage environment without bias through observer presence. Overall, this showcases how long-term monitoring data can be applied, automated and analysed to complement current procedures and help to refine animal research and promote more detailed well-being assessments.

7.2 LIMITATIONS AND FUTURE DIRECTIONS

7.2.1 HARDWARE

The component that builds the basis for the presented camera module is a Raspberry Pi (https://www.raspberrypi.com/), a low-cost single-board computer. There is a variety of products that could serve as an alternative and recent studies have proved their applicability in animal research (e.g. Curtin and Matthews, 2019; Zualkernan *et al.*, 2022). One of the main limitations I encountered is humidity. Even if this is kept constant throughout the unit (see Section 3.4) the camera system is exposed to liquids, especially during cleaning events. Throughout the two and a half years of recordings I lost two SD cards and three USB sticks as they were corrupted. While I can't certainly pinpoint the cause to be humidity exposure it is likely to be a contributing factor in the deterioration of the materials. In one of the cages, the casing was also heavily exposed to the urine of the animals as there was a perch right above the camera module. This led to the corrosion of the plastic fitting in the first CCTV-casing used.

The future versions of the module included a CCTV-casing constructed only of metal to avoid this. Alternatively, the perch had been elongated in another enclosure which somewhat protected the camera.

Another issue I encountered was that the newer version of the Arducam HAT that combined the inputs of two cameras did not start up correctly when attached to the modified NOIR V2 cameras. I found a work around by using a simple HDMI switch adapter that allowed me to start the camera while being connected to the stereo-camera shipped with the Arducam HAT and was then switched to the NOIR V2 cameras when recording started using the GPIO pins and a relay module (see Figure 7-1). However, this added more electrical parts within the module making it more vulnerable.

As only one camera system was used in each cage to measure behaviours of multiple animals and the locations where I could mount the system were limited, not all areas of the cage are in view of the camera. Since I mostly used frequencies of behaviours or calculated proportions between timepoints where an animal was actually visible this should not directly impact measures used in this thesis. However, adding multiple viewing angles would give a more complete picture of the actual behavioural repertoires observed within the animal. While employing a multi camera system introduces the issue of synchronising the recordings, it could help to track the animals throughout the full environment available to them. Furthermore, it could automatically identify the times where an animal is out of the cage (i.e. in the laboratory) compared to just not visible, which could aid in creating datasets for individual recognition that are not based on the presence or absence of a head implant. Moreover, I did not record footage during the night as light levels were too low. Future versions of the camera system could include an infra-red light to study nocturnal behaviours and monitor sleep (Fruth, Tagg and Stewart, 2018). Again, this would lead to a more complete picture when looking for deviations from normality.

Two of the camera modules have since been given to the CFM to test in their enclosures (see Figure 7-2). However, only the housing has been mounted so far for a brief period to test if the full casing is robust as animals now have unrestricted access to the full camera module while in the CBC the aluminium box was located outside their cage door (see Figure 3-2). The camera system developed as part of this thesis could serve as an additional module that is moveable to cover areas that are otherwise difficult to observe. As mentioned before this could help in tracking animals throughout the full enclosure or help to gather close-up footage in areas were animals are often observed to enable easier identification (e.g. Witham, 2018).

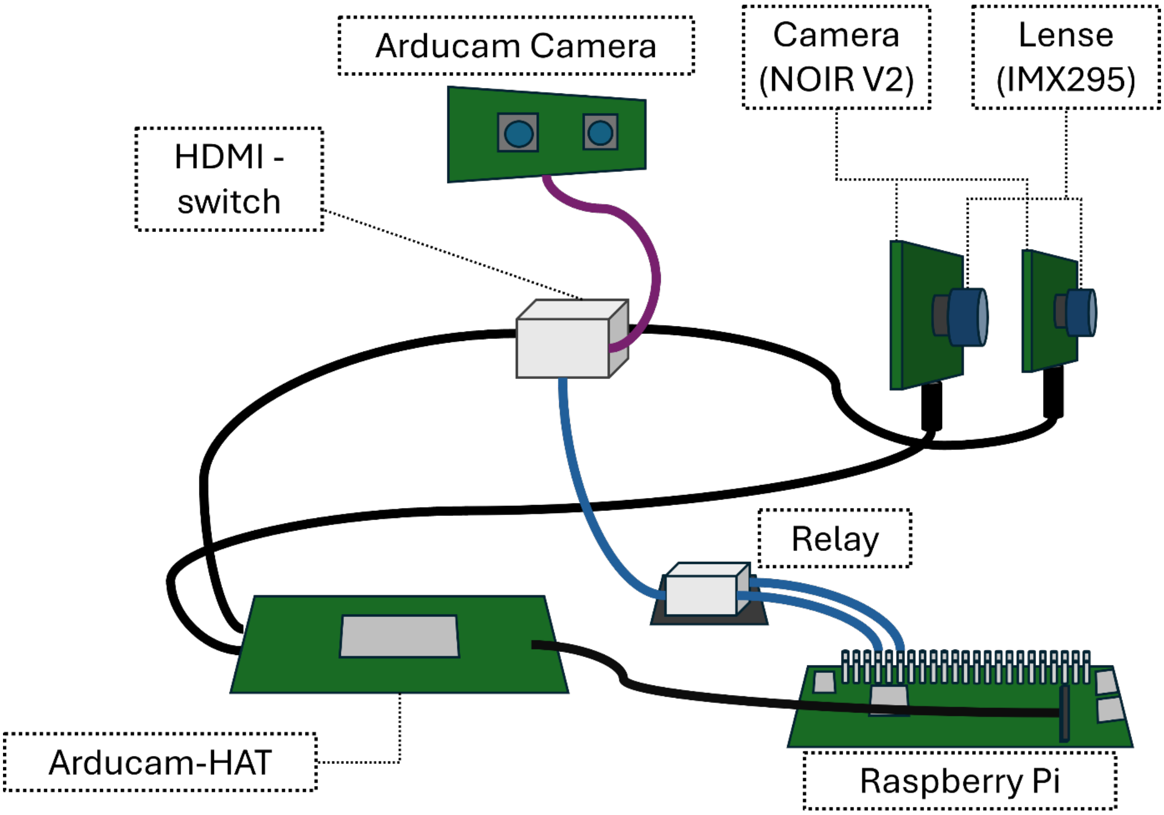

Figure 7-1, Representation of the camera system including the relay to switch between the cameras when starting up the Raspberry Pi and starting the recording. Dotted boxes depict the description of the parts used. Black lines indicate cabled connections, blue line indicate the connection to the relay switch and the purple line is the connection to the camera that is active only on startup (before relay is activated).

Figure 7-2, Picture of camera casing being tested at the CFM. Red arrow points at the camera casing.

The field of electronics is consistently evolving, and new technologies and cheaper alternatives might be available soon. An example for this are RGB-D cameras (colour channels red, green and blue; D stands for depth). These cameras emit structured patterns of light onto a scene and analyse the distortion of these patterns to determine the depth of objects, allowing for the creation of detailed 3D models alongside the traditional RGB image. The additional channel for depth did not disturb the usage of current deep learning architectures (see Shao *et al.*, 2017) and has already been applied to measure behaviours in mice (Gerós, Magalhães and Aguiar, 2020) and social interactions in humans (Coppola *et al.*, 2017).

7.2.2 SOFTWARE

The initial version of the pipeline was based on Yolact (Bolya *et al.*, 2019), another segmentation algorithm. However, in January 2023 YOLOv8 (Jocher, Chaurasia and Qiu, 2023) was released. I adapted the pipeline to the new algorithm as it ran much faster and is more user-friendly providing easy ways of extracting all of the segmentation information and having the option to include integrated tracking algorithms. About a month before submitting this thesis (21.02.2024) the new version of YOLO has been released (Wang, Yeh and Liao, 2024) and shows potential to further optimising detection and segmentation tasks.

The initial thought for having a camera system that used two cameras in synchrony was to get estimates in 3D. In stereovision approaches, 3D estimation relies on the principle of triangulation, where corresponding points in the left and right images are matched to compute the depth information. By analysing the positional disparity between these matched points, the system can infer the relative distances of objects from the cameras. This process generates a disparity map or depth map, providing a spatial representation of the scene's geometry, essential for various applications such as object detection, scene reconstruction, and depth-based segmentation. For this the camera need to be calibrated, this often involves presenting a board depicting shapes (e.g. checker or CharuCo boards). Recalibration of a stereovision setup is crucial, particularly after camera movement, as it ensures accurate alignment of the camera pair, enabling precise depth perception and spatial reconstruction in the captured imagery. As the developed camera module was mounted on the door, cameras would have needed to be recalibrated frequently which was not feasible as it would have interfered too much with the animals. Other calibration approaches such as using the dimensions of the interior of the environment could be applied. I tested a software alternative using Global-Local Path Networks (Kim et al., 2022) to estimate depth from 2D images. This approach utilizes a dual-path architecture that combines global contextual information with local detailed features. The global path captures broad scene context to infer coarse depth estimates, while the local path focuses on fine-grained details for accurate depth refinement. Vertical cut depth mechanisms further enhance performance by efficiently integrating information across multiple scales, resulting in robust depth estimation from single images. For an example frame, the depth estimate and the constructed point cloud from that estimate see Figure 7-3. While this approach yielded measures of the expected distance between the animal and the camera (not quantitatively validated), this slowed down the video processing (0.5-1 second per frame) and is hence not a sensible approach for long term monitoring.

Advancing the pipeline outlined in this thesis doesn't solely rely on acquiring 3D information. There's potential to enhance the analysis of behaviour captured by refining the measures, allowing for a more detailed examination of behaviour at a finer scale. As mentioned in Chapter 4, movement trajectories are commonly used in ecological studies. Features such as velocity profiles and turn angles could be estimated in the future. This could obtain a more detailed perspective on an individual’s movement. Such measures have previously been related to welfare factors, such as diazepam treatment (Rose et al., 2012). Other approaches have shown the potential of pixel based classification of behaviour (see Marks et al., 2022). Applying such approaches to monitor even simple behaviours like a hunched posture and inactivity can help to identify signs of welfare (Hennessy et al., 2014; Perera et al., 2011; Willard et al., 2014). Moreover, pose estimation algorithms have gained a lot of traction with different approaches being presented to capture key point locations in 2D and 3D on single and multiple animals (see

Section 1.6 for a summary). As part of my thesis, I trained a DeepLabCut model for pose estimation for multiple individuals (Lauer et al., 2022). I labelled 22 key points on the animals as well as two fixed key points in the environment on 1280 images. The body landmarks included the middle of the back, chest, and hip, the left and right ankles, knees, hips, shoulders, elbows, wrists, ears and eyes, the tail base, the tail end, and the nose (see Figure 7-4). While mean squared errors for training and testing are acceptable for the task (see Table 7-1) visually inspecting the output indicated issues with identifying individuals (see mismatch in identities between Figure 7-4 A and B) and tracking them consistently.

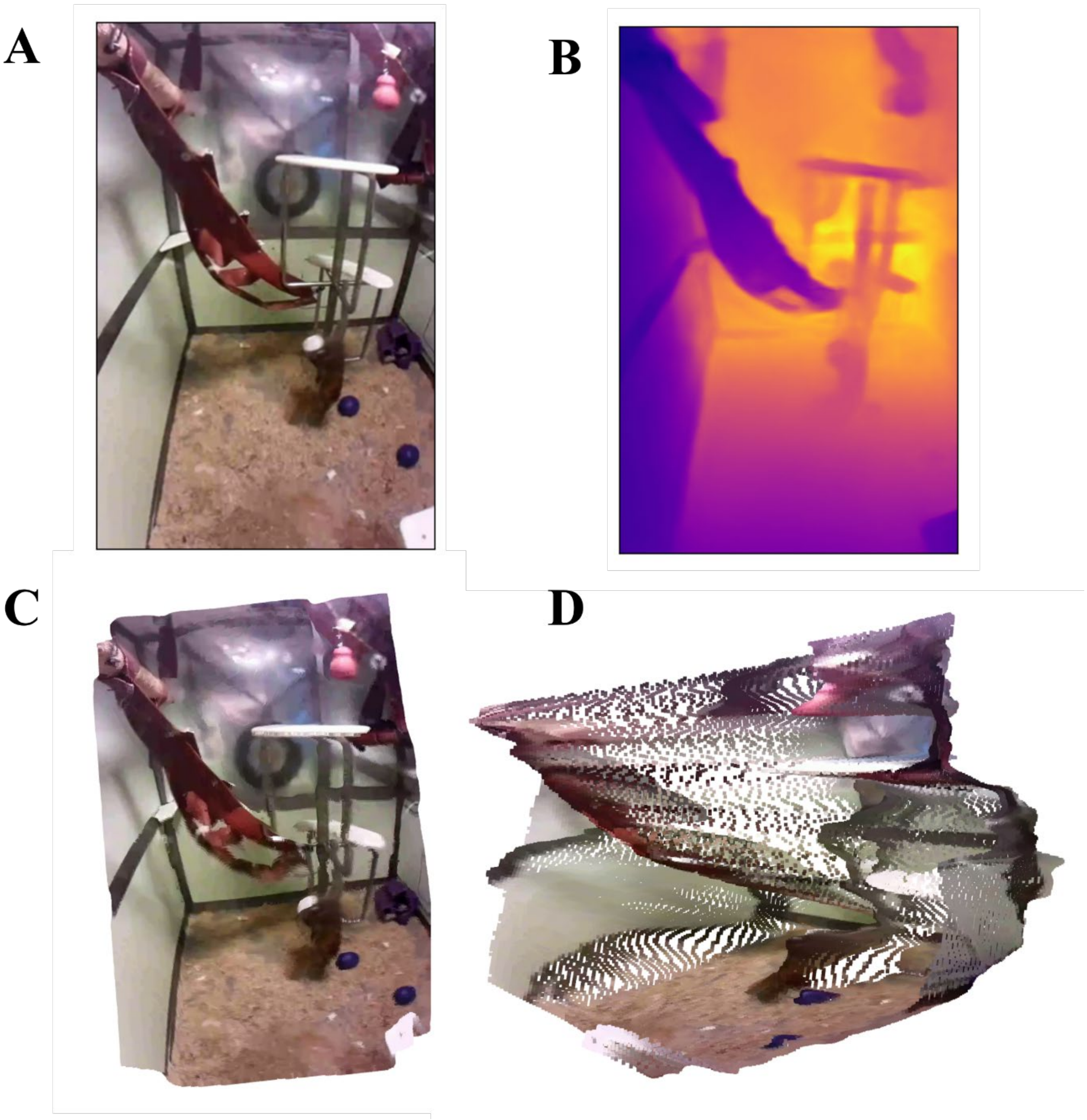

Figure 7-3, Example output from a Global Local Path Network to predict depth from a 2D image. Graphic A depicts the example frame from footage recorded using the camera module; B is the heatmap for the depth estimation, yellow indicates points further away, blue points closer to the camera; C point cloud front view; D point cloud side view, points are represented in a 3D representation of the initial frame.

Training iterations:	**%Training dataset**	**Train Error (pixels)**	**Test Error (pixels)**
100000	95	2.92 (2.87)	7.79 (6.07)
200000	95	2.37 (2.36)	6.32 (5.15)

Table 7-1, Mean squared error for the key point prediction on body (N = 22) and in the environment (N = 2). Numbers are given as the overall error and the error with a threshold of > 0.6 in brackets.

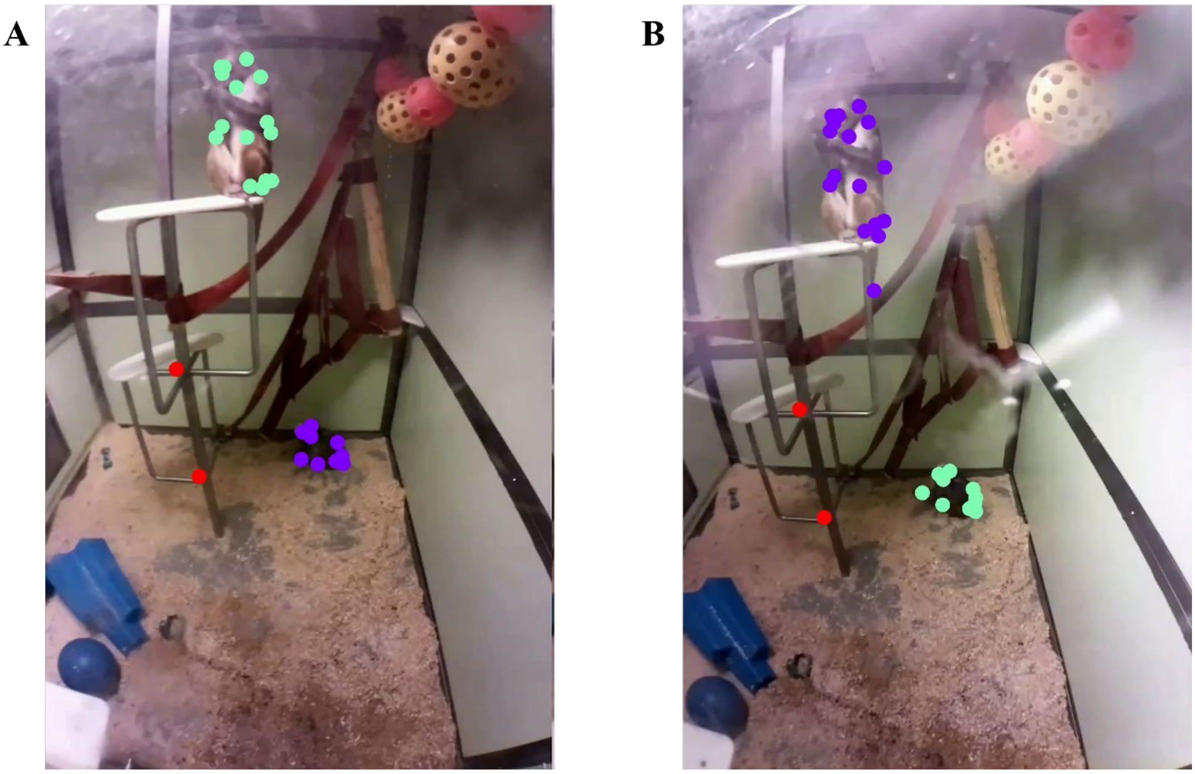

Figure 7-4, Example images of the DeepLabCut model output; A, is the left frame; B, the right frame. Red dots indicate the environment landmarks, green and purple indicate identity of an animal and the dots are the body landmarks.

Pose estimation algorithms lead to a vast amount of data tracking multiple body parts in each frame to gather information on the animals pose and behaviour. Such big datasets have previously been described as 'big behaviour' (von Ziegler, Sturman and Bohacek, 2021). The resulting large datasets open the potential to use deep learning approaches that do not need feature construction or even labelling, this is also called unsupervised learning and have been applied to measure primate behaviour (e.g. Matsumoto et al., 2023; Mimura et al., 2023). A more novel approach uses pose detection data to animate an artificial avatar of a monkey. This can not only help in analysing behaviours, but also lead to stimuli that can be used in an experimental setting leading to results comparable to genuine videos (Martini et al., 2024).

Overall, these automated tools can not only help increase welfare by providing closer insights on the behavioural needs of individual animals but could also refine research procedure. For example, aiding in finding the correct timings to terminate experiments in dosing studies by taking home cage behaviour as an additional indicator. Moreover, tracking an animal's behaviour at a more detailed level could help identify changes in behaviour. As an example researchers use primates as models for stroke and spinal injury (Zaaimi *et al.*, 2012), often behavioural tasks are needed to keep track of changes in motor output. Applying pose estimation to and extract angular changes in the limbs could help in identifying changes in laterality pre and post lesions in the home cage and not just in the laboratory setting.

7.2.3 AUTOMATED MEASURES OF BEHAVIOUR AND PERSONALITY

This thesis showed the potential of automating personality trait ratings by correlating automated behavioural measures to the rating of animals using the Hominoid Personality Questionnaire (HPQ; Weiss *et al.*, 2009; Whitham and Washburn, 2017). Such automated methods could in the future aid longitudinal studies to track adaptions in behaviour across major stages (such as living in a breeding colony and being moved to laboratory housing). This provides valuable insights into potential selective pressures and mechanisms for the composition of personality (Stamps and Groothuis, 2010).

While I associated measures of both personality ratings from raters in the CBC as well as neophobia scores derived from the CFM with the automated behavioural measures only in four of these animals do I have both scores. Comparing personality measurements from these different contexts to test for stability could help to shed light individual behavioural expression (see Robinson and Weiss, 2023). Macaques now housed in the CBC are going to be tested with a novel object test again using a similar procedure but different stimuli. Also, inter-rater reliability scores obtained from two or three raters for each animal were low. I argued that this might be due to the different perspectives these raters had on the animals as one rater was working as an animal technician (probably basing the ratings on home cage behaviours) and the other two were researchers that might have a laboratory-based view on the animals. This could also be further explored to potentially use behaviours during tasks and training as another stage-dependent context of personality expression.

Overall, this thesis is a proof of concept that currently available methods can be adapted to measure behavioural patterns in laboratory-housed Rhesus macaques. I showed examples of behaviours that can be extracted based on this method and show that they detect individual differences, are consistent and are correlated with personality traits. These methods can be extended to cover a higher variety and more detailed behaviours that can be monitored longitudinally at different time scales.

7.3 COMPUTER VISION AND WELFARE MEASURES

The use of novel technologies such as machine and deep learning approaches based on video data to monitor animal welfare has so far mostly been applied in farm animals and laboratory rodents (Valletta et al., 2017).

In farm animals, computer vision techniques have been applied to monitor gait in pigs, cows and chickens. Stavrakakis et al. (2014) used camera-based motion capture on walking pigs that identified differences in joint flexion and stance-related angular changes in joints linked to a variety of orthopaedic health issues. Approaches to track hoof location or the arching of the back were shown to be capable of tracking characteristics related to lameness in cows (Barney et al., 2023; Kang, Zhang and Liu, 2020; Song et al., 2007). In chickens, computer vision methods were used to extract step frequency, step length, speed and lateral body oscillation which showed high correlations to lameness scores with potential to automate lameness detection in broiler chickens (Aydin, 2017). Further, tail biting can be a major welfare concern in the husbandry of pigs. Low tail posture is an early warning sign for higher rate of tail biting. D'Eath et al. (2018) presented a model using RGB-D cameras to automate pig tail posture as an early warning sign for tail biting risk. The use of enrichment can have positive effects on welfare of the animals, however assessing the actual preferences for specific enrichment can be time consuming. Chen et al. (2020) automatically monitored the frequency of use of a specific enrichment item in pen housed pigs. Selecting toys that animals are more likely to engage with and that might keep them engaged more frequently can increase health and welfare and incidents of negative behaviours have shown to decrease with more engaging enrichment.

In laboratory rodents, RFID tracking in combination with computer vision techniques have led to the automated assessment of unforced behaviour in the home cage over long time intervals. This has led to better understanding of mouse mutant phenotyping and detecting changes due to aging, sex, and disease (Bains et al., 2023; Bains et al., 2018). Additionally, the application of deep learning models on videos and images was effective in identifying pain in rodents via tracking facial expressions or body posture (Fried et al., 2020; Tuttle et al., 2018).

Overall, automated methods such as computer vision have shown to be valuable tools for animal welfare. The camera technology is readily available and cheap in comparison to housing costs. Furthermore, video recordings are non-invasive and is less biased. Hence, the application of such methods is feasible and highly encouraged, especially for animals in captivity.

7.4 BRIDGING THE GAP – CALL FOR APPLICATION

Primates used in research are heavily monitored and a variety of measures are retrieved on a daily basis (e.g. weights, task-performance). Such measures are already used to detect deviations from normality that could be indicative of changes in welfare levels. The husbandry of animals in research is expensive and labour intense, however, especially regarding video surveillance, there is room for improvement. This thesis presented a camera system to monitor animals within their home-cage. This method has no external cables, making it suitable to be mounted within an animal's cage to avoid occlusions from bars and glass panels. Moreover, I showcased how currently available methods can be adapted to derive automated measures form video recordings with little human labour. These methods are getting more user friendly over time. I strongly argue that such implementations of video monitoring should become part of the standard procedures for animals housed in captivity. There is a gap between the development of new methods and the actual application of these on an everyday basis and this thesis is a proof of concept for an approach to bridge this gap. It showcased the feasibility of monitoring animals at a high temporal resolution (multiple hours per day) on a longitudinal scale (multiple months to years).

Complementing the current standards of monitoring with automatically derived daily behavioural measures could help in specifically identifying levels of normality and instances of deviations, further identifying the intricacies of individual variability and personality (Robinson and Weiss, 2023). This is not only a step towards better refinement for animal research but could aid in strengthening the research being conducted (Bethell *et al.*, 2019b; Buchanan-Smith, 2006; Jennings, Prescott and Refinement, 2009; Prescott and Lidster, 2017; Würbel, 2002). Finally, the application of automated methods can aid in the decision-making process for husbandry procedures such as what to feed and which enrichment to use (see Ciminelli, 2023).

7.5 SUMMARY OF FUTURE STEPS

The integration of multiple cameras within a single enclosure represents a promising avenue for enhancing animal monitoring systems. This approach ensures comprehensive coverage of all areas within the habitat, minimizing instances of unrecorded animal activity. Additionally, the incorporation of infrared lights into the existing lighting infrastructure enables continuous 24/7 surveillance, including nocturnal behaviours (Brando and Buchanan-Smith, 2018). Such advancements facilitate caretakers and researchers in delivering optimal care and gaining novel insights into individual behavioural needs.

One of the major issues I faced with the pose estimation was the correct and consistent identification and tracking of the animals, this can be addressed with the use of YOLOv8. the pipeline output can retrospectively identify body landmarks based on bounding boxes or masks, or through adapting YOLOv8 as a fully integrative approach for pose estimation (Jocher, Chaurasia and Qiu, 2023). This leads to more fine scaled measures of behaviour that could be beneficial for both welfare and research purposes (see Section 7.3).

The methodology proposed in Chapter 5 for detecting deviations from normal behaviour should be extended to longer observation periods and potentially varying temporal resolutions (e.g., hourly measurements instead of daily). Although labour-intensive, this approach should cross-reference automated measures with laboratory journals and husbandry data, yielding deeper insights into normal levels of behaviour for each individual and the causes for divergences.

Lastly, as similar methods are being applied in the breeding colonies and the system developed in this thesis is also being tested (see Ciminelli, 2023), standardised methods and measures should be applied to make comparison between housing contexts for an individual possible. Moreover, after the change of institutions animals should be heavily monitored as this could provide more specific insights into the expression of personality (Robinson and Weiss, 2023).

As the technology keeps growing and getting more accessible, it is high time that these methods are applied to monitor animals in captivity. Especially regarding the 3Rs and in the light of recent refinements it seems unethical to not try and capture as much data as we can. This becomes even more important as it comes with low levels of invasiveness (adding a camera to the enclosure) and with the potential of getting further insights into means to refine husbandry and experimental procedures by comparing behavioural changes to standard levels such as during standard husbandry events (e.g. feeding) or over time. The collected data can additionally be highly informative to the research the animals are involved in. Focusing on illness specific behaviours within the home-cage such as motor impairment in stroke models or measuring sociality and anxiety related behaviours for models of mental diseases. Monitoring methods applied to breeding facilities could also help in phenotyping animal models for neurodevelopmental disorders (Silverman *et al.*, 2022).

In the end there is little to no arguments of why such methodology should not be applied on a wider scale. And while there is no ‘one fits all’ solution for automated analysis yet, at least the video monitoring should be integrated into standard procedures and policies.

7.6 CONCLUSION

This thesis presented a design for an in-cage camera system allowing long-term video capture with little occlusion and an automated pipeline to detect and track animals using these recordings. It showed the applicability of adapting current methodologies to measure consistent individual differences in behaviour in laboratory pair-housed Rhesus macaques. Furthermore, the presented method is capable to measure deviations in space use associated with common research procedures. Some general frequencies of behaviour were related to measures of personality indicating its potential to automate measures of temperament. Overall, this thesis underscored the viability of an automated method to enhance existing husbandry practices for laboratory Rhesus macaques, emphasizing individual welfare. It also exemplified the integration of contemporary techniques to bridge the divide between technological progress and behavioural research, together with welfare protocols.

REFERENCES

Aguilar-Moreno, M. and Graña, M. 'Computational Ethology: Short Review of Current Sensors and Artificial Intelligence Based Methods'. *International Conference on Engineering Applications of Neural Networks*: Springer, 17-27.

Aharon, N., Orfaig, R. and Bobrovsky, B.-Z. (2022) 'BoT-SORT: Robust associations multi-pedestrian tracking', *arXiv preprint arXiv:2206.14651.*

Akoglu, H. (2018) 'User's guide to correlation coefficients', *Turkish journal of emergency medicine,* 18(3), pp. 91-93.

Altmann, J. (1974) 'Observational study of behavior: sampling methods', *Behaviour,* 49(3-4), pp. 227-266.

Altschul, D. M., Robinson, L. M., Coleman, K., Capitanio, J. P. and Wilson, V. A. D. (2019) 'An Exploration of the Relationships Among Facial Dimensions, Age, Sex, Dominance Status, and Personality in Rhesus Macaques (Macaca mulatta)', *International Journal of Primatology,* 40(4-5), pp. 532-552.

Anand, S., Vaidyanathan, S. and Radhakrishna, S. (2021) 'The role of landscape structure in primate crop feeding: Insights from rhesus macaques (Macaca mulatta) in northern India', *International Journal of Primatology,* 42(5), pp. 764-780.

Anderson, C. J., Heard, D. J., Andreu, M. A., Hostetler, M. E. and Johnson, S. A. (2017) 'Winter Home Range and Habitat Selection of a Rhesus Macaque Group (Macaca mulatta) at Silver Springs State Park', *Florida Scientist,* 80(4), pp. 159-164.

Anderson, David J. and Perona, P. (2014) 'Toward a Science of Computational Ethology', *Neuron,* 84(1), pp. 18-31.

Anderson, J. R. and Chamove, A. S. (1980) 'Self-aggression and social aggression in laboratory-reared macaques', *Journal of Abnormal Psychology,* 89(4), pp. 539.

Ardoin, T. and Sueur, C. (2023) 'Automatic Identification of Stone-Handling Behaviour in Japanese Macaques Using LabGym Artificial Intelligence', *arXiv preprint arXiv:2310.07812.*

Ares, E. 2022. Animal Welfare (Sentience) bill. House of Commons library.

Aydin, A. (2017) 'Development of an early detection system for lameness of broilers using computer vision', *Computers and Electronics in Agriculture,* 136, pp. 140-146.

Bachevalier, J., Machado, C. J. and Kazama, A. (2011) 'Behavioral outcomes of late-onset or early-onset orbital frontal cortex (areas 11/13) lesions in rhesus monkeys.', *Annals of the New York Academy of Sciences December,* 1, pp. 71-86.

Bailey, J. and Taylor, K. (2016) 'Non-human primates in neuroscience research: the case against its scientific necessity', *Alternatives to Laboratory Animals,* 44(1), pp. 43-69.

Bain, M., Nagrani, A., Schofield, D., Berdugo, S., Bessa, J., Owen, J., Hockings, K. J., Matsuzawa, T., Hayashi, M. and Biro, D. (2021) 'Automated audiovisual behavior recognition in wild primates', *Science Advances,* 7(46), pp. eabi4883.

Bains, R. S., Forrest, H., Sillito, R. R., Armstrong, J. D., Stewart, M., Nolan, P. M. and Wells, S. E. (2023) 'Longitudinal home-cage automated assessment of climbing behavior shows sexual dimorphism and aging-related decrease in C57BL/6J healthy mice and allows early detection of motor impairment in the N171-82Q mouse model of Huntington’s disease', *Frontiers in Behavioral Neuroscience,* 17.

Bains, R. S., Wells, S., Sillito, R. R., Armstrong, J. D., Cater, H. L., Banks, G. and Nolan, P. M. (2018) 'Assessing mouse behaviour throughout the light/dark cycle using automated in-cage analysis tools', *Journal of Neuroscience Methods,* 300, pp. 37-47.

Baker, K. C., Bloomsmith, M. A., Oettinger, B., Neu, K., Griffis, C., Schoof, V. and Maloney, M. (2012) 'Benefits of pair housing are consistent across a diverse population of rhesus macaques', *Applied Animal Behaviour Science,* 137(3), pp. 148-156.

Bala, P. C., Eisenreich, B. R., Yoo, S. B. M., Hayden, B. Y., Park, H. S. and Zimmermann, J. (2020) 'Automated markerless pose estimation in freely moving macaques with OpenMonkeyStudio', *Nature communications,* 11(1), pp. 4560.

Ballesta, S., Reymond, G., Pozzobon, M. and Duhamel, J.-R. (2014) 'A real-time 3D video tracking system for monitoring primate groups', *Journal of neuroscience methods,* 234, pp. 147-152.

Bardi, M. and Huffman, M. A. (2006) 'Maternal Behavior and Maternal Stress Are Associated with Infant Behavioral Development in Macaques.', *Developmental Psychobiology January,* 48(1), pp. 1-9.

Barney, S., Dlay, S., Crowe, A., Kyriazakis, I. and Leach, M. (2023) 'Deep learning pose estimation for multi-cattle lameness detection', *Scientific Reports,* 13(1), pp. 4499.

Barr, C. S., Dvoskin, R. L., Schwandt, M. L., Lindell, S. G., Kling, M. A., Gold, P. W., Higley, J. D., Heilig, M., Suomi, S. J. and Goldman, D. (2008) 'CRH haplotype predicts CSFCRH, HPA axis activity, temperament, and alcohol consumption in rhesus macaques', *Biological Psychiatry,* 63(7), pp. 108S-108S.

Bateson, M. and Martin, P. (2021) *Measuring Behaviour: An Introductory Guide.* 4 edn. Cambridge: Cambridge University Press.

Bateson, P., Johansen-Berg, H. and Jones, D. (2011) 'Review of research using non-human primates', *London: Biotechnology and Biological Sciences Research Council, Medical Research Council and Wellcome Trust,* 286.

Bauman, M. D., Iosif, A. M., Ashwood, P., Braunschweig, D., Lee, A., Schumann, C. M., Van de Water, J. and Amaral, D. G. (2013) 'Maternal antibodies from mothers of children with autism alter brain growth and social behavior development in the rhesus monkey', *Translational Psychiatry,* 3, pp. 12.

Baxter, A., Capitanio, J. P., Bales, K. L. and Kinnally, E. L. (2021) 'Biobehavioral organization shapes the immune epigenome in infant rhesus Macaques (Macaca mulatta)', *Brain Behavior and Immunity,* 96, pp. 256-270.

Bayne, K. (2005) 'Potential for unintended consequences of environmental enrichment for laboratory animals and research results', *Ilar j,* 46(2), pp. 129-39.

Bayne, K. and Novak, M. (1998) 'Behavioral disorders', *Nonhuman primates in biomedical research*: Elsevier, pp. 485-500.

Bayne, K. and Würbel, H. (2014) 'The impact of environmental enrichment on the outcome variability and scientific validity of laboratory animal studies', *Rev Sci Tech,* 33(1), pp. 273-80.

Beisner, B. A. and Isbell, L. A. (2009) 'Movement ecology in a captive environment: the effects of ground substrate on movement paths of captive rhesus macaques, Macaca mulatta', *Animal Behaviour,* 78(5), pp. 1269-1277.

Beisner, B. A. and McCowan, B. (2014) 'Signaling context modulates social function of silent bared-teeth displays in rhesus macaques (Macaca mulatta).', *American Journal of Primatology,* 76(2), pp. 111-121.

Bell, A. M., Hankison, S. J. and Laskowski, K. L. (2009) 'The repeatability of behaviour: a meta-analysis', *Animal behaviour,* 77(4), pp. 771-783.

Benton, C. G., West, M. W., Hall, S. M., Marko, S. T. and Johnson, J. C. (2013) 'Effect of short-term pair housing of juvenile rhesus macaques (Macaca mulatta) on immunologic parameters', *Journal of the American Association for Laboratory Animal Science,* 52(3), pp. 240-246.

Bernstein, I. S. and Ehardt, C. L. (1985) 'Agonistic aiding: kinship, rank, age, and sex influences', *American Journal of Primatology,* 8(1), pp. 37-52.

Bethell, E. J., Cassidy, L. C., Brockhausen, R. R. and Pfefferle, D. (2019a) 'Toward a Standardized Test of Fearful Temperament in Primates: A Sensitive Alternative to the Human Intruder Task for Laboratory-Housed Rhesus Macaques (Macaca mulatta)', *Frontiers in Psychology,* 10, pp. 17.

Bethell, E. J., Cassidy, L. C., Brockhausen, R. R. and Pfefferle, D. (2019b) 'Toward a Standardized Test of Fearful Temperament in Primates: A Sensitive Alternative to the Human Intruder Task for Laboratory-Housed Rhesus Macaques (Macaca mulatta)', *Frontiers in Psychology,* 10.

Birn, R. M., Shackman, A. J., Oler, J. A., Williams, L. E., McFarlin, D. R., Rogers, G. M., Shelton, S. E., Alexander, A. L., Pine, D. S., Slattery, M. J., Davidson, R. J., Fox, A. S. and Kalin, N. H. (2014) 'Evolutionarily conserved prefrontal-amygdalar dysfunction in early-life anxiety', *Molecular Psychiatry,* 19(8), pp. 915-922.

Blanco Negrete, S., Labuguen, R., Matsumoto, J., Go, Y., Inoue, K.-i. and Shibata, T. (2021) 'Multiple monkey pose estimation using openpose', *bioRxiv*, pp. 2021.01. 28.428726.

Blaszczyk, M. B. (2020) 'Primates got personality, too: Toward an integrative primatology of consistent individual differences in behavior', *Evolutionary Anthropology: Issues, News, and Reviews,* 29(2), pp. 56-67.

Bliss-Moreau, E., Toscano, J. E., Bauman, M. D., Mason, W. A. and Amaral, D. G. (2010) 'Neonatal Amygdala or Hippocampus Lesions Influence Responsiveness to Objects.', *Developmental Psychobiology July,* 52(5), pp. 487-503.

Blomquist, G. E., Hinde, K. and Capitanio, J. P. (2021) 'Inheritance of Hormonal Stress Response and Temperament in Infant Rhesus Macaques (Macaca Mulatta): Nonadditive and Sex-Specific Effects', *Behavioral Neuroscience*, pp. 12.

Bolig, R., Price, C. S., O'Neill, P. L. and Suomi, S. J. (1992) 'Subjective assessment of reactivity level and personality traits of rhesus monkeys', *International Journal of Primatology,* 13(3), pp. 287-306.

Bolya, D., Zhou, C., Xiao, F. and Lee, Y. J. 'Yolact: Real-time instance segmentation'. *Proceedings of the IEEE/CVF international conference on computer vision*, 9157-9166.

Bonini, L. (2019) 'Refinement techniques in non-human primate neuroscientific research', *Annali dell'Istituto Superiore di Sanità,* 55(4), pp. 409-412.

Boyce, W. T. (2004) 'Social Stratification, Health, and Violence in the Very Young.', *Annals of the New York Academy of Sciences*, pp. 47-68.

Boyce, W. T., O'Neill-Wagner, P., Price, C. S., Haines, M. and Suomi, S. J. (1998) 'Crowding Stress and Violent Injuries Among Behaviorally Inhibited Rhesus Macaques.', *Health Psychology May,* 17(3), pp. 285-289.

Brando, S. and Buchanan-Smith, H. M. (2018) 'The 24/7 approach to promoting optimal welfare for captive wild animals', *Behavioural Processes,* 156, pp. 83-95.

Braundmeier, A. and Fazleabas, A. (2009) 'The non-human primate model of endometriosis: research and implications for fecundity', *Molecular human reproduction,* 15(10), pp. 577-586.

Brent, L. J. N., Semple, S., MacLarnon, A., Ruiz-Lambides, A., Gonzalez-Martinez, J. and Platt, M. L. (2014) 'Personality Traits in Rhesus Macaques (Macaca mulatta) Are Heritable but Do Not Predict Reproductive Output', *International Journal of Primatology,* 35(1), pp. 188-209.

Brommer, J. E. and Class, B. (2017) 'Personality from the perspective of behavioral ecology', *Personality in nonhuman animals*: Springer, pp. 73-107.

Broom, D. (2001) 'Coping, stress and welfare', *Coping with challenge: Welfare in animals including humans*, pp. 1-9.

Broom, D. M. (1991) 'Animal welfare: concepts and measurement', *Journal of Animal Science,* 69(10), pp. 4167-4175.

Buchanan-Smith, H. M. (2006) 'Primates in laboratories: standardisation, harmonisation, variation and science', *ALTEX: Alternatives to Animal Experimentation,* (Special Issue), pp. 115-119.

Camus, S. M. J., Blois-Heulin, C., Li, Q., Hausberger, M. and Bezard, E. (2013) 'Behavioural Profiles in Captive-Bred Cynomolgus Macaques: Towards Monkey Models of Mental Disorders?', *PLOS ONE,* 8(4), pp. e62141.

Capitanio, J. P. (1999a) 'Personality dimensions in adult male rhesus macaques: Prediction of behaviors across time and situation', *American Journal of Primatology: Official Journal of the American Society of Primatologists,* 47(4), pp. 299-320.

Capitanio, J. P. (1999b) 'Personality dimensions in adult male rhesus macaques: Prediction of behaviors across time and situation', *American Journal of Primatology,* 47(4), pp. 299-320.

Capitanio, J. P. (2002) 'Sociability and responses to video playbacks in adult male rhesus monkeys (Macaca mulatta)', *Primates,* 43(3), pp. 169-177.

Capitanio, J. P., Abel, K., Mendoza, S. P., Blozis, S. A., McChesney, M. B., Cole, S. W. and Mason, W. A. (2008) 'Personality and serotonin transporter genotype interact with social context to affect immunity and viral set-point in simian immunodeficiency virus disease', *Brain Behavior and Immunity,* 22(5), pp. 676-689.

Capitanio, J. P., Blozis, S. A., Snarr, J., Steward, A. and McCowan, B. J. (2017a) 'Do "Birds of a Feather Flock Together" or Do "Opposites Attract"? Behavioral Responses and Temperament Predict Success in Pairings of Rhesus Monkeys in a Laboratory Setting', *American Journal of Primatology,* 79(1), pp. 11.

Capitanio, J. P., Blozis, S. A., Snarr, J., Steward, A. and McCowan, B. J. (2017b) 'Do “birds of a feather flock together” or do “opposites attract”? Behavioral responses and temperament predict success in pairings of rhesus monkeys in a laboratory setting', *American Journal of Primatology,* 79(1), pp. e22464.

Capitanio, J. P., Del Rosso, L. A., Calonder, L. A., Blozis, S. A. and Penedo, M. C. T. (2012) 'Behavioral Effects of Prenatal Ketamine Exposure in Rhesus Macaques Are Dependent on MAOA Genotype', *Experimental and Clinical Psychopharmacology,* 20(3), pp. 173-180.

Capitanio, J. P., Hawkley, L. C., Cole, S. W. and Cacioppo, J. T. (2014) 'A Behavioral Taxonomy of Loneliness in Humans and Rhesus Monkeys (Macaca mulatta)', *Plos One,* 9(10), pp. 11.

Capitanio, J. P., Mendoza, S. P. and Baroncelli, S. (1999) 'The relationship of personality dimensions in adult male rhesus macaques to progression of simian immunodeficiency virus disease', *Brain Behavior and Immunity,* 13(2), pp. 138-154.

Capitanio, J. P., Mendoza, S. P. and Bentson, K. L. (2004) 'Personality characteristics and basal cortisol concentrations in adult male rhesus macaques (Macaca mulatta)', *Psychoneuroendocrinology,* 29(10), pp. 1300-1308.

Capitanio, J. P., Mendoza, S. P. and Cole, S. W. (2011) 'Nervous temperament in infant monkeys is associated with reduced sensitivity of leukocytes to cortisol's influence on trafficking', *Brain Behavior and Immunity,* 25(1), pp. 151-159.

Capitanio, J. P., Miller, L. A., Schelegle, E. S., Mendoza, S. P., Mason, W. A. and Hyde, D. M. (2011) 'Behavioral Inhibition Is Associated With Airway Hyperresponsiveness but not Atopy in a Monkey Model of Asthma', *Psychosomatic Medicine,* 73(4), pp. 288-294.

Capitanio, J. P. and Widaman, K. F. (2005) 'Confirmatory factor analysis of personality structure in adult male rhesus monkeys (Macaca mulatta)', *American Journal of Primatology,* 65(3), pp. 289-294.

Careaga, M., Murai, T. and Bauman, M. D. (2017) 'Maternal immune activation and autism spectrum disorder: from rodents to nonhuman and human primates', *Biological psychiatry,* 81(5), pp. 391-401.

Carlsson, H. E., Schapiro, S. J., Farah, I. and Hau, J. (2004) 'Use of primates in research: a global overview', *American Journal of Primatology: Official Journal of the American Society of Primatologists,* 63(4), pp. 225-237.

Carstens, E. and Moberg, G. P. (2000) 'Recognizing pain and distress in laboratory animals', *Ilar Journal,* 41(2), pp. 62-71.

Carter, A. J., Feeney, W. E., Marshall, H. H., Cowlishaw, G. and Heinsohn, R. (2013) 'Animal personality: what are behavioural ecologists measuring?', *Biological Reviews,* 88(2), pp. 465-475.

Castellano Bueno, J. (2022) *Multidisciplinary assessment of cumulative experience in laboratory rhesus macaques.* Newcastle University.

Ceniceros, L. C., Capitanio, J. P. and Kinnally, E. L. (2021) 'Prenatal Relocation Stress Enhances Resilience Under Challenge in Infant Rhesus Macaques', *Frontiers in Behavioral Neuroscience,* 15, pp. 14.

Chamove, A. S., Eysenck, H. J. and Harlow, H. F. (1972) 'Personality in monkeys - Factor-analyses of Rhesus social behavior', *Quarterly Journal of Experimental Psychology,* 24(NOV), pp. 496-&.

Champoux, M., Bennett, A., Shannon, C., Higley, J. D., Lesch, K. P. and Suomi, S. J. (2002a) 'Serotonin transporter gene polymorphism, differential early rearing, and behavior in rhesus monkey neonates', *Molecular Psychiatry,* 7(10), pp. 1058-1063.

Champoux, M., Hibbeln, J. R., Shannon, C., Majchrzak, S., Suomi, S. J., Salem Jr, N. and Higley, J. D. (2002b) 'Fatty acid formula supplementation and neuromotor development in rhesus monkey neonates', *Pediatric Research,* 51(3), pp. 273-281.

Champoux, M., Higley, J. D. and Suomi, S. J. (1997) 'Behavioral and physiological characteristics of Indian and Chinese-Indian hybrid rhesus macaque infants', *Developmental Psychobiology,* 31(1), pp. 49-63.

Champoux, M., Kriete, M. F., Eckhaus, M. A. and Suomi, S. J. (1997) 'Behavioral and physical concomitants of congenital hydrocephalus in a rhesus macaque (Macaca mulatta) neonate', *Contemporary Topics in Laboratory Animal Science,* 36(1), pp. 56-61.

Chatfield, K. and Morton, D. (2018) 'The use of non-human primates in research', *Ethics Dumping: Case Studies from North-South Research Collaborations*, pp. 81-90.

Chen, C., Zhu, W., Oczak, M., Maschat, K., Baumgartner, J., Larsen, M. L. V. and Norton, T. (2020) 'A computer vision approach for recognition of the engagement of pigs with different enrichment objects', *Computers and Electronics in Agriculture,* 175, pp. 105580.

Chun, K. and Capitanio, J. P. (2016) 'Developmental consequences of behavioral inhibition: a model in rhesus monkeys (Macaca mulatta)', *Developmental Science,* 19(6), pp. 1035-1048.

Chun, K., Miller, L. A., Schelegle, E. S., Hyde, D. M. and Capitanio, J. P. (2013) 'Behavioral Inhibition in Rhesus Monkeys (Macaca mulatta) Is Related to the Airways Response, but Not Immune Measures, Commonly Associated with Asthma', *Plos One,* 8(8), pp. 8.

Ciminelli, G. (2023) *Automated Methodologies to Assess Welfare of Rhesus Macaques (Macaca mulatta) in a Breeding Colony.* Newcastle University.

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', *Journal of the American Association for Laboratory Animal Science,* 45(5), pp. 31-34.

Clarke, A. S. and Snipes, M. (1998) 'Early behavioral development and temperamental traits in mother- vs peer-reared rhesus monkeys', *Primates,* 39(4), pp. 433-448.

Cloninger, C. R. (1994) 'Temperament and personality', *Current Opinion in Neurobiology,* 4(2), pp. 266-273.

Coe, C. L., Lubach, G. R., Crispen, H. R., Shirtcliff, E. A. and Schneider, M. L. (2010) 'Challenges to Maternal Wellbeing during Pregnancy Impact Temperament, Attention, and Neuromotor Responses in the Infant Rhesus Monkey', *Developmental Psychobiology,* 52(7), pp. 625-637.

Coleman, K. (2012) 'Individual differences in temperament and behavioral management practices for nonhuman primates', *Applied Animal Behaviour Science,* 137(3-4), pp. 106-113.

Coleman, K., Dahl, R. E., Ryan, N. D. and Cameron, J. L. (2003) 'Growth hormone response to growth hormone-releasing hormone and clonidine in young monkeys: Correlation with behavioral characteristics', *Journal of Child and Adolescent Psychopharmacology,* 13(3), pp. 227-241.

Coleman, K., Lutz, C. K., Worlein, J. M., Gottlieb, D. H., Peterson, E., Lee, G. H., Robertson, N. D., Rosenberg, K., Menard, M. T. and Novak, M. A. (2017) 'The Correlation Between Alopecia and Temperament in Rhesus Macaques (Macaca mulatta) at Four Primate Facilities', *American Journal of Primatology,* 79(1), pp. 10.

Coleman, K., Tully, L. A. and McMillan, J. L. (2005) 'Temperament correlates with training success in adult rhesus macaques', *American Journal of Primatology,* 65(1), pp. 63-71.

Coppola, C., Cosar, S., Faria, D. R. and Bellotto, N. 'Automatic detection of human interactions from RGB-D data for social activity classification'. *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 28 Aug.-1 Sept. 2017, 871-876.

Corcoran, C. A., Pierre, P. J., Haddad, T., Bice, C., Suomi, S. J., Grant, K. A., Friedman, D. P. and Bennett, A. J. (2012) 'Long-term effects of differential early rearing in rhesus macaques: Behavioral reactivity in adulthood', *Developmental Psychobiology,* 54(5), pp. 546-555.

Cronbach, L. J. (1951) 'Coefficient alpha and the internal structure of tests', *Psychometrika,* 16(3), pp. 297-334.

Curtin, B. H. and Matthews, S. J. 'Deep learning for inexpensive image classification of wildlife on the Raspberry Pi'. *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*: IEEE, 0082-0087.

D'Eath, R. B., Jack, M., Futro, A., Talbot, D., Zhu, Q., Barclay, D. and Baxter, E. M. (2018) 'Automatic early warning of tail biting in pigs: 3D cameras can detect lowered tail posture before an outbreak', *PLOS ONE,* 13(4), pp. e0194524.

Davidson, R. J., Kalin, N. H. and Shelton, S. E. (1993) 'Lateralized response to diazepam predicts temperamental style in rhesus monkeys', *Behavioral Neuroscience,* 107(6), pp. 1106-1110.

Dawkins, M. (2012) *Animal suffering: the science of animal welfare.* Springer Science & Business Media.

Delaloye, S. and Holtzheimer, P. E. (2014) 'Deep brain stimulation in the treatment of depression', *Dialogues in clinical neuroscience,* 16(1), pp. 83-91.

Descovich, K. A., Richmond, S. E., Leach, M. C., Buchanan-Smith, H. M., Flecknell, P., Farningham, D. A. H., Witham, C., Gates, M. C. and Vick, S.-J. (2019) 'Opportunities for refinement in neuroscience: Indicators of wellness and post-operative pain in laboratory macaques', *ALTEX - Alternatives to animal experimentation,* 36(4), pp. 535-554.

Dingemanse, N. J. and Wolf, M. (2010) 'Recent models for adaptive personality differences: a review', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 365(1560), pp. 3947-3958.

Dore, K. M., Hansen, M. F., Klegarth, A. R., Fichtel, C., Koch, F., Springer, A., Kappeler, P., Parga, J. A., Humle, T., Colin, C., Raballand, E., Huang, Z.-P., Qi, X.-G., Di Fiore, A., Link, A., Stevenson, P. R., Stark, D. J., Tan, N., Gallagher, C. A., Anderson, C. J., Campbell, C. J., Kenyon, M., Pebsworth, P., Sprague, D., Jones-Engel, L. and Fuentes, A. (2020) 'Review of GPS collar deployments and performance on nonhuman primates', *Primates,* 61(3), pp. 373-387.

Doudet, D. P., Hommer, D. M. D., Higley, D. J. P., Andreason, P. J. M. D., Moneman, R. P., Suomi, S. J. P. and Linnoila, M. M. D. P. (1995) 'Cerebral Glucose Metabolism, CSF 5-HIAA

Levels, and Aggressive Behavior in Rhesus Monkeys.', *American Journal of Psychiatry December,* 152(12), pp. 1782-1787.

Drost, H.-G. (2018) 'Philentropy: information theory and distance quantification with R', *Journal of Open Source Software,* 3(26), pp. 765.

Duhachek, A. and Iacobucci, D. (2004) 'Alpha's Standard Error (ASE): An Accurate and Precise Confidence Interval Estimate', *Journal of Applied Psychology,* 89(5), pp. 792-808.

Dutta, A., Gupta, A. and Zissermann, A. (2016) 'VGG image annotator (VIA)', *URL: http://www. robots. ox. ac. uk/~ vgg/software/via.*

Edelhoff, H., Signer, J. and Balkenhol, N. (2016) 'Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns', *Movement ecology,* 4(1), pp. 1-21.

Elfenbein, H. A., Del Rosso, L., McCowan, B. and Capitanio, J. P. (2016) 'Effect of Indoor Compared with Outdoor Location during Gestation on the Incidence of Diarrhea in Indoor-Reared Rhesus Macaques (Macaca mulatta)', *Journal of the American Association for Laboratory Animal Science,* 55(3), pp. 277-290.

Eradus, W. J. and Jansen, M. B. (1999) 'Animal identification and monitoring', *Computers and Electronics in Agriculture,* 24(1-2), pp. 91-98.

Erickson, K., Gabry, K. E., Schulkin, J., Gold, P., Lindell, S., Higley, J. D., Champoux, M. and Suomi, S. J. (2005) 'Social withdrawal behaviors in nonhuman primates and changes in neuroendocrine and monoamine concentrations during a separation paradigm', *Developmental Psychobiology,* 46(4), pp. 331-339.

Erwin, J. (1979) 'Strangers in a strange land: abnormal behavior or abnormal environments?', *Captivity and behavior.*

Forkosh, O. (2021) 'Animal behavior and animal personality from a non-human perspective: Getting help from the machine', *Patterns,* 2(3).

Forkosh, O., Karamihalev, S., Roeh, S., Alon, U., Anpilov, S., Touma, C., Nussbaumer, M., Flachskamm, C., Kaplick, P. M., Shemesh, Y. and Chen, A. (2019) 'Identity domains capture individual differences from across the behavioral repertoire', *Nature Neuroscience,* 22(12), pp. 2023-2028.

Fox, A. S., Harris, R. A., Rosso, L. D., Raveendran, M., Kamboj, S., Kinnally, E. L., Capitanio, J. P. and Rogers, J. (2021) 'Infant inhibited temperament in primates predicts adult behavior, is heritable, and is associated with anxiety-relevant genetic variation', *Molecular Psychiatry,* 26(11), pp. 6609-6618.

Fox, A. S., Oler, J. A., Birn, R. M., Shackman, A. J., Alexander, A. L. and Kalin, N. H. (2018) 'Functional Connectivity within the Primate Extended Amygdala Is Heritable and Associated with Early-Life Anxious Temperament', *Journal of Neuroscience,* 38(35), pp. 7611-7621.

Fox, A. S., Oler, J. A., Shackman, A. J., Shelton, S. E., Raveendran, M., McKay, D. R., Converse, A. K., Alexander, A., Davidson, R. J., Blangero, J., Rogers, J. and Kalin, N. H. (2015) 'Intergenerational neural mediators of early-life anxious temperament', *Proceedings of the National Academy of Sciences of the United States of America,* 112(29), pp. 9118-9122.

Fox, A. S., Oler, J. A., Shelton, S. E., Nanda, S. A., Davidson, R. J., Roseboom, P. H. and Kalin, N. H. (2012) 'Central amygdala nucleus (Ce) gene expression linked to increased trait-like Ce metabolism and anxious temperament in young primates', *Proceedings of the National Academy of Sciences of the United States of America,* 109(44), pp. 18108-18113.

Fox, A. S., Shelton, S. E., Oakes, T. R., Davidson, R. J. and Kalin, N. H. (2008) 'Trait-Like Brain Activity during Adolescence Predicts Anxious Temperament in Primates', *Plos One,* 3(7), pp. 9.

Fox, A. S., Souaiaia, T., Oler, J. A., Kovner, R., Kim, J. M. H., Nguyen, J., French, D. A., Riedel, M. K., Fekete, E. M., Rabska, M. R., Olsen, M. E., Brodsky, E. K., Alexander, A. L., Block, W. F., Roseboom, P. H., Knowles, J. A. and Kalin, N. H. (2019) 'Dorsal Amygdala Neurotrophin-3 Decreases Anxious Temperament in Primates', *Biological Psychiatry,* 86(12), pp. 881-889.

Franks, B., Reiss, D., Cole, P., Friedrich, V., Thompson, N. and Higgins, E. T. (2013) 'Predicting how individuals approach enrichment: Regulatory focus in cotton-top tamarins (Sanguinus oedipus)', *Zoo biology,* 32(4), pp. 427-435.

Freeman, H. D. and Gosling, S. D. (2010) 'Personality in nonhuman primates: a review and evaluation of past research', *American journal of primatology,* 72(8), pp. 653-671.

Fried, N. T., Chamessian, A., Zylka, M. J. and Abdus-Saboor, I. (2020) 'Improving pain assessment in mice and rats with advanced videography and computational approaches', *PAIN,* 161(7).

Fruth, B., Tagg, N. and Stewart, F. (2018) 'Sleep and nesting behavior in primates: A review', *American Journal of Physical Anthropology,* 166(3), pp. 499-509.

Gaither, A. M., Baker, K. C., Gilbert, M. H., Blanchard, J. L., Liu, D. X., Luchins, K. R. and Bohm, R. P. (2014) 'Videotaped behavior as a predictor of clinical outcome in rhesus macaques (Macaca mulatta)', *Comparative Medicine,* 64(3), pp. 193-199.

Gamer, M., Lemon, J., Gamer, M. M., Robinson, A. and Kendall's, W. (2012) 'Package 'irr'', *Various coefficients of interrater reliability and agreement,* 22, pp. 1-32.

Ganguly, I. and Chauhan, N. S. (2018) 'Daily behavioural activity patterns and reproductive ecology of urban rhesus macaques (Macaca mulatta) in human dominated landscape and its implication in management', *Journal of Applied and Natural Science,* 10(4), pp. 1101-1108.

Gardner, J. (2013) 'A history of deep brain stimulation: Technological innovation and the role of clinical assessment tools', *Social studies of science,* 43(5), pp. 707-728.

Gartner, M. C. and Weiss, A. (2018) 'Studying primate personality in zoos: Implications for the management, welfare and conservation of great apes', *International Zoo Yearbook,* 52(1), pp. 79-91.

George, D. and Mallery, P. (2003) 'SPSS for Windows step by step: answers to selected exercises', *A simple guide and reference,* 63(1), pp. 1461-1470.

Gerós, A., Magalhães, A. and Aguiar, P. (2020) 'Improved 3D tracking and automated classification of rodents' behavioral activity using depth-sensing cameras', *Behavior Research Methods,* 52(5), pp. 2156-2167.

Gibboni, R. R., Zimmerman, P. E. and Gothard, K. M. (2009) 'Individual differences in scanpaths correspond with serotonin transporter genotype and behavioral phenotype in rhesus monkeys (Macaca mulatta)', *Frontiers in Behavioral Neuroscience,* 3, pp. 11.

Glover, I. S. and Baker, S. N. (2020) 'Cortical, corticospinal, and reticulospinal contributions to strength training', *Journal of Neuroscience,* 40(30), pp. 5820-5832.

Golub, M. S. and Hogrefe, C. E. (2015) 'Fetal Iron Deficiency and Genotype Influence Emotionality in Infant Rhesus Monkeys', *Journal of Nutrition,* 145(3), pp. 647-653.

Golub, M. S., Hogrefe, C. E. and Germann, S. L. (2007) 'Iron deprivation during fetal development changes the behavior of juvenile rhesus monkeys', *Journal of Nutrition,* 137(4), pp. 979-984.

Golub, M. S., Hogrefe, C. E., Germann, S. L., Tran, T. T., Beard, J. L., Crinella, F. M. and Lonnerdal, B. (2005) 'Neurobehavioral evaluation of rhesus monkey infants fed cow's milk formula, soy formula, or soy formula with added manganese', *Neurotoxicology and Teratology,* 27(4), pp. 615-627.

Golub, M. S., Hogrefe, C. E. and Unger, E. L. (2012) 'Influence of prenatal iron deficiency and MAOA genotype on response to social challenge in rhesus monkey infants', *Genes Brain and Behavior,* 11(3), pp. 278-290.

Golub, M. S., Hogrefe, C. E., Wjdaman, K. F. and Capitanio, J. P. (2009) 'Iron deficiency anemia and affective response in Rhesus monkey infants', *Developmental Psychobiology,* 51(1), pp. 47-59.

Gosling, S. D. (2001) 'From mice to men: what can we learn about personality from animal research?', *Psychological bulletin,* 127(1), pp. 45.

Gottlieb, D. H., Capitanio, J. P. and McCowan, B. (2013) 'Risk factors for stereotypic behavior and self-biting in rhesus macaques (Macaca mulatta): Animal's history, current environment, and personality', *American Journal of Primatology,* 75(10), pp. 995-1008.

Gottlieb, D. H., Del Rosso, L., Sheikhi, F., Gottlieb, A., McCowan, B. and Capitanio, J. P. (2018a) 'Personality, environmental stressors, and diarrhea in Rhesus macaques: An interactionist perspective', *American Journal of Primatology,* 80(12), pp. e22908.

Gottlieb, D. H., Del Rosso, L., Sheikhi, F., Gottlieb, A., McCowan, B. and Capitanio, J. P. (2018b) 'Personality, environmental stressors, and diarrhea in Rhesus macaques: An interactionist perspective', *American Journal of Primatology,* 80(12), pp. 14.

Gottlieb, D. H., Maier, A. and Coleman, K. (2015) 'Evaluation of environmental and intrinsic factors that contribute to stereotypic behavior in captive rhesus macaques (Macaca mulatta)', *Applied Animal Behaviour Science,* 171, pp. 184-191.

Gris, V. N., Crespo, T. R., Kaneko, A., Okamoto, M., Suzuki, J., Teramae, J.-n. and Miyabe-Nishiwaki, T. (2024) 'Deep Learning for Face Detection and Pain Assessment in Japanese macaques (Macaca fuscata)', *Journal of the American Association for Laboratory Animal Science*.

Haley, G., Landauer, N., Renner, L., Weiss, A., Hooper, K., Urbanski, H., Kohama, S., Neuringer, M. and Raber, J. (2009) 'Circadian activity associated with spatial learning and memory in aging rhesus monkeys', *Experimental neurology,* 217(1), pp. 55-62.

Haley, G. E. a. b., Eghlidi, D. H. a. b., Kohama, S. G. b., Urbanski, H. F. a. b. c. and Raber, J. a. b. d. (2012a) 'Association of microtubule associated protein-2, synaptophysin, and apolipoprotein E mRNA and protein levels with cognition and anxiety levels in aged female rhesus macaques.', *Behavioural Brain Research June,* 232(1), pp. 1-6.

Haley, G. E. a. b., McGuire, A. a., Berteau-Pavy, D. a., Weiss, A. b., Patel, R. a., Messaoudi, I. b. c., Urbanski, H. F. a. b. d., Raber, J. a. b. e. and lowast (2012b) 'Measures of anxiety, amygdala volumes, and hippocampal scopolamine phMRI response in elderly female rhesus macaques.', *Neuropharmacology January,* 62(1), pp. 385-390.

Hamel, A. F., Lutz, C. K., Coleman, K., Worlein, J. M., Peterson, E. J., Rosenberg, K. L., Novak, M. A. and Meyer, J. S. (2017) 'Responses to the Human Intruder Test Are Related to Hair Cortisol Phenotype and Sex in Rhesus Macaques (Macaca mulatta)', *American Journal of Primatology,* 79(1), pp. 10.

Harlow, H. F. and Zimmermann, R. R. (1959) 'Affectional response in the infant monkey: Orphaned baby monkeys develop a strong and persistent attachment to inanimate surrogate mothers', *Science,* 130(3373), pp. 421-432.

Hau, J. and Schapiro, S. (2006) 'Non-human primates in biomedical research', *Scandinavian Journal of Laboratory Animal Science,* 33(1), pp. 9-12.

Hau, J. and Schapiro, S. J. (2007) 'The welfare of non-human primates', *The welfare of laboratory animals*: Springer, pp. 291-314.

Heinz, A., Higley, J. D., Gorey, J. G., Saunders, R. C., Jones, D. W., Hommer, D., Zajicek, K., Suomi, S. J., Lesch, K. P., Weinberger, D. R. and Linnoila, M. (1998) 'In vivo association between alcohol intoxication, aggression, and serotonin transporter availability in nonhuman primates', *American Journal of Psychiatry,* 155(8), pp. 1023-1028.

Hemsworth, P. H., Mellor, D. J., Cronin, G. M. and Tilbrook, A. J. (2015) 'Scientific assessment of animal welfare', *New Zealand Veterinary Journal,* 63(1), pp. 24-30.

Hennessy, M. B., Chun, K. and Capitanio, J. P. (2017) 'Depressive-like behavior, its sensitization, social buffering, and altered cytokine responses in rhesus macaques moved from outdoor social groups to indoor housing', *Social Neuroscience,* 12(1), pp. 65-75.

Hennessy, M. B., McCowan, B., Jiang, J. and Capitanio, J. P. (2014) 'Depressive-like behavioral response of adult male rhesus monkeys during routine animal husbandry procedure', *Frontiers in behavioral neuroscience,* 8, pp. 309.

Herrington, J. A., Del Rosso, L. A. and Capitanio, J. P. (2016) 'Biobehavioral Consequences of Prenatal Exposure to a Matrilineal Overthrow and Relocation in Captive Infant Rhesus (Macaca mulatta) Monkeys', *American Journal of Primatology,* 78(9), pp. 895-903.

Hinde, K. and Captanio, J. P. (2010) 'Lactational Programming? Mother's Milk Energy Predicts Infant Behavior and Temperament in Rhesus Macaques (Macaca mulatta)', *American Journal of Primatology,* 72(6), pp. 522-529.

Hinde, K., Skibiel, A. L., Foster, A. B., Rosso, L. D., Mendoza, S. P. and Capitanio, J. P. (2015) 'Cortisol in mother's milk across lactation reflects maternal life history and predicts infant temperament', *Behavioral Ecology,* 26(1), pp. 269-281.

Hochreiter, S. and Schmidhuber, J. (1997) 'Long short-term memory', *Neural computation,* 9(8), pp. 1735-1780.

Hollands, C. (1986) 'The animals (scientific procedures) act 1986', *Lancet (London, England),* 2(8497), pp. 32-33.

Hopkins, W. D. and Latzman, R. D. (2017) 'Future Research with Captive Chimpanzees in the United States', *Handbook of primate behavioral management.*

Hopper, L. M., Cronin, K. A. and Ross, S. R. (2018) 'A multi-institutional assessment of a short-form personality questionnaire for use with macaques', *Zoo biology,* 37(5), pp. 281-289.

Hopper, L. M., Price, S. A., Freeman, H. D., Lambeth, S. P., Schapiro, S. J. and Kendal, R. L. (2014) 'Influence of personality, age, sex, and estrous state on chimpanzee problem-solving success', *Animal cognition,* 17, pp. 835-847.

Howell, S., White, D., Ingram, S., Jackson, R., Larin, J., Morales, P., Garcia, A. P., Hicks, C., Hopper, K. and Wagner, J. (2012) 'A bio-behavioral study of chronic idiopathic colitis in the rhesus macaque (Macaca mulatta)', *Applied Animal Behaviour Science,* 137(3-4), pp. 208-220.

Hrolenok, B., Balch, T., Byrd, D., Roberts, R., Kim, C., Rehg, J. M., Gilliland, S. and Wallen, K. 'Use of position tracking to infer social structure in rhesus macaques', *Proceedings of the Fifth International Conference on Animal-Computer Interaction*, Atlanta, Georgia, USA: Association for Computing Machinery, Article 15.

Isakova, I. A., Baker, K., DuTreil, M., Dufour, J., Gaupp, D. and Phinney, D. G. (2007) 'Age- and Dose-Related Effects on MSC Engraftment Levels and Anatomical Distribution in the Central Nervous Systems of Nonhuman Primates: Identification of Novel MSC Subpopulations That Respond to Guidance Cues in Brain.', *Stem Cells,* 25(12), pp. 3261-3270.

Isakova, I. A., Baker, K. C., Dufour, J. and Phinney, D. G. (2017) 'Mesenchymal Stem Cells Yield Transient Improvements in Motor Function in an Infant Rhesus Macaque with Severe Early-Onset Krabbe Disease.', *Stem Cells Translational Medicine Open Access January,* 6(1), pp. 99-109.

Isakova, I. A., Lanclos, C., Bruhn, J., Kuroda, M. J., Baker, K. C., Krishnappa, V. and Phinney, D. G. (2014) 'Allo-Reactivity of Mesenchymal Stem Cells in Rhesus Macaques Is Dose and Haplotype Dependent and Limits Durable Cell Engraftment In Vivo', *Plos One,* 9(1), pp. 16.

Jennings, M., Prescott, M. J. and Refinement, J. W. G. o. (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', *Laboratory Animals,* 43(1_suppl), pp. 1-47.

Jiang, J., Kanthaswamy, S. and Capitanio, J. P. (2013) 'Degree of Chinese ancestry affects behavioral characteristics of infant rhesus macaques (Macaca mulatta)', *Journal of Medical Primatology,* 42(1), pp. 20-27.

Jocher, G., Chaurasia, A. and Qiu, J. (2023) *YOLO by Ultralytics*. Available at: https://github.com/ultralytics/ultralytics (Accessed: 16/07/2023 2023).

Johnston, J. M., Cohen, Y. E., Shirley, H., Tsunada, J., Bennur, S., Christison-Lagay, K. and Veeder, C. L. (2016) 'Recent refinements to cranial implants for rhesus macaques (Macaca mulatta)', *Lab Anim (NY),* 45(5), pp. 180-6.

Kalin, N. H., Fox, A. S., Kovner, R., Riedel, M. K., Fekete, E. M., Roseboom, P. H., Tromp, D. P. M., Grabow, B. P., Olsen, M. E., Brodsky, E. K., McFarlin, D. R., Alexander, A. L., Emborg, M. E., Block, W. F., Fudge, J. L. and Oler, J. A. (2016) 'Overexpressing Corticotropin-Releasing Factor in the Primate Amygdala Increases Anxious Temperament and Alters Its Neural Circuit', *Biological Psychiatry,* 80(5), pp. 345-355.

Kalin, N. H., Larson, C., Shelton, S. E. and Davidson, R. J. (1998) 'Asymmetric frontal brain activity, cortisol, and behavior associated with fearful temperament in rhesus monkeys', *Behavioral Neuroscience,* 112(2), pp. 286-292.

Kalin, N. H. and Shelton, S. E. (1989) 'Defensive behaviors in infant rhesus monkeys: environmental cues and neurochemical regulation', *Science,* 243(4899), pp. 1718-1721.

Kalin, N. H., Shelton, S. E. and Davidson, R. J. (2007) 'Role of the primate orbitofrontal cortex in mediating anxious temperament', *Biological Psychiatry,* 62(10), pp. 1134-1139.

Kalin, N. H., Shelton, S. E., Davidson, R. J. and Kelley, A. E. (2001) 'The primate amygdala mediates acute fear but not the behavioral and physiological components of anxious temperament', *Journal of Neuroscience,* 21(6), pp. 2067-2074.

Kang, X., Zhang, X. D. and Liu, G. (2020) 'Accurate detection of lameness in dairy cattle with computer vision: A new and individualized detection strategy based on the analysis of the supporting phase', *Journal of Dairy Science,* 103(11), pp. 10628-10638.

Kay, D. B., Marsiske, M., Suomi, S. J. and Higley, J. D. (2010) 'Exploratory factor analysis of human infant temperament in the rhesus monkey', *Infant Behavior & Development,* 33(1), pp. 111-114.

Kemp, C., Thatcher, H., Farningham, D., Witham, C., MacLarnon, A., Holmes, A., Semple, S. and Bethell, E. J. (2017) 'A protocol for training group-housed rhesus macaques (Macaca mulatta) to cooperate with husbandry and research procedures using positive reinforcement', *Applied Animal Behaviour Science,* 197, pp. 90-100.

Kim, D., Ka, W., Ahn, P., Joo, D., Chun, S. and Kim, J. (2022) 'Global-local path networks for monocular depth estimation with vertical cutdepth', *arXiv preprint arXiv:2201.07436.*

King, J. E. and Figueredo, A. J. (1997) 'The five-factor model plus dominance in chimpanzee personality', *Journal of research in personality,* 31(2), pp. 257-271.

Kinnally, E. L. (2008) 'Genetic and environmental factors that impact infant temperamental features and parameters of the serotonergic system associated with adult impulsivity in rhesus macaques (Macaca mulatta)', *Dissertation Abstracts International: Section B: The Sciences and Engineering,* 68(8-B).

Kinnally, E. L., Gonzalez, M. N. and Capitanio, J. P. (2018) 'Paternal line effects of early experiences persist across three generations in rhesus macaques.', *Developmental Psychobiology December,* 60(8), pp. 879-888.

Kinnally, E. L., Tarara, E. R., Mason, W. A., Mendoza, S. P., Abel, K., Lyons, L. A. and Capitanio, J. P. (2010) 'Serotonin transporter expression is predicted by early life stress and is associated with disinhibited behavior in infant rhesus macaques.', *Genes, Brain, & Behavior February,* 9(1), pp. 45-52.

Kinnally, E. L., Whiteman, H. J., Mason, W. A., Mendoza, S. P. and Capitanio, J. P. (2008) 'Dimensions of response to novelty are associated with social engagement and aggression in adult male rhesus macaques (Macaca mulatta)', *Journal of Comparative Psychology,* 122(2), pp. 195-203.

Knaebe, B., Weiss, C. C., Zimmermann, J. and Hayden, B. Y. (2022) 'The promise of behavioral tracking systems for advancing primate animal welfare', *Animals,* 12(13), pp. 1648.

Kohn, J. N., Snyder-Mackler, N., Barreiro, L. B., Johnson, Z. P., Tung, J. and Wilson, M. E. (2016) 'Dominance rank causally affects personality and glucocorticoid regulation in female rhesus macaques', *Psychoneuroendocrinology,* 74, pp. 179-188.

Koolhaas, J. M. (2008) 'Coping style and immunity in animals: Making sense of individual variation', *Brain, Behavior, and Immunity,* 22(5), pp. 662-667.

Kovner, R., Souaiaia, T., Fox, A. S., French, D. A., Goss, C. E., Roseboom, P. H., Oler, J. A., Riedel, M. K., Fekete, E. M., Fudge, J. L., Knowles, J. A. and Kalin, N. H. (2020) 'Transcriptional Profiling of Primate Central Nucleus of the Amygdala Neurons to Understand the Molecular Underpinnings of Early-Life Anxious Temperament', *Biological Psychiatry,* 88(8), pp. 638-648.

Kraemer, G. W., Moore, C. F., Newman, T. K., Barr, C. S. and Schneider, M. L. (2008) 'Moderate level fetal alcohol exposure and serotonin transporter gene promoter polymorphism affect neonatal temperament and limbic-hypothalamic-pituitary-adrenal axis regulation in monkeys', *Biological Psychiatry,* 63(3), pp. 317-324.

Kshetri, S., Khanal, L., Tandon, S. and Kyes, R. C. (2023) 'Seasonal Variation in the Activity Budget and Daily Path Length of Semi-Provisioned Rhesus Macaques (Macaca mulatta) in Daunne Forest, Nawalpur, Nepal', *Indonesian Journal of Primatology,* 2(01), pp. 27-36.

Kullback, S. and Leibler, R. A. (1951) 'On information and sufficiency', *The annals of mathematical statistics,* 22(1), pp. 79-86.

Kumar, R., Palit, G., Singh, J. and Dhawan, B. (1999) 'Comparative behavioural effects of benzodiazepine and non-benzodiazepine anxiolytics in rhesus monkeys', *Pharmacological research,* 39(6), pp. 437-444.

Labuguen, R., Matsumoto, J., Negrete, S. B., Nishimaru, H., Nishijo, H., Takada, M., Go, Y., Inoue, K.-i. and Shibata, T. (2021) 'MacaquePose: a novel “in the wild” macaque monkey pose dataset for markerless motion capture', *Frontiers in behavioral neuroscience,* 14, pp. 581154.

Lambeth, S. P., Hau, J., Perlman, J. E., Martino, M. and Schapiro, S. J. (2006) 'Positive reinforcement training affects hematologic and serum chemistry values in captive chimpanzees (Pan troglodytes)', *Am J Primatol,* 68(3), pp. 245-56.

Laudenslager, M. L. and Boccia, M. L. (1996) 'Some observations on psychosocial stressors, immunity, and individual differences in nonhuman primates', *American Journal of primatology,* 39(4), pp. 205-221.

Laudenslager, M. L., Rasmussen, K. L., Berman, C. M., Lilly, A. A., Shelton, S. E., Kalin, N. H. and Suomi, S. J. (1999) 'A preliminary description of responses of free-ranging rhesus monkeys to brief capture experiences: Behavior, endocrine, immune, and health relationships', *Brain Behavior and Immunity,* 13(2), pp. 124-137.

Lauer, J., Zhou, M., Ye, S., Menegas, W., Schneider, S., Nath, T., Rahman, M. M., Di Santo, V., Soberanes, D., Feng, G., Murthy, V. N., Lauder, G., Dulac, C., Mathis, M. W. and Mathis, A. (2022) 'Multi-animal pose estimation, identification and tracking with DeepLabCut', *Nature Methods,* 19(4), pp. 496-504.

Laule, G. E., Bloomsmith, M. A. and Schapiro, S. J. (2003) 'The use of positive reinforcement training techniques to enhance the care, management, and welfare of primates in the laboratory', *J Appl Anim Welf Sci,* 6(3), pp. 163-73.

Lear, A., Baker, S. N., Clarke, H. F., Roberts, A. C., Schmid, M. C. and Jarrett, W. (2022) 'Understanding them to understand ourselves: the importance of NHP research for translational neuroscience', *Current Research in Neurobiology,* 3, pp. 100049.

Li, C., Xiao, Z., Li, Y., Chen, Z., Ji, X., Liu, Y., Feng, S., Zhang, Z., Zhang, K. and Feng, J. (2023) 'Deep learning-based activity recognition and fine motor identification using 2D skeletons of cynomolgus monkeys', *Zoological Research,* 44(5), pp. 967.

Li, N., Ren, Z., Li, D. and Zeng, L. (2020) 'Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming', *animal,* 14(3), pp. 617-625.

Li, X., Xu, F., Xie, L., Ji, Y., Cheng, K., Zhou, Q., Wang, T., Shively, C., Wu, Q. and Gong, W. (2013) 'Depression-like behavioral phenotypes by social and social plus visual isolation in the adult female Macaca fascicularis', *PLoS One,* 8(9), pp. e73293.

Lidow, M. S. (1998) 'Nonhuman primate model of the effect of prenatal cocaine exposure on cerebral cortical development', in Harvey, J.A. and Kosofsky, B.E. (eds.) *Cocaine: Effects on the Developing Brain Annals of the New York Academy of Sciences*. New York: New York Acad Sciences, pp. 182-193.

Linden, J. B. (2018) 'Factors affecting hair cortisol, rank, and aggression in a large newly formed social group of captive rhesus macaques (Macaca mulatta)', *Dissertation Abstracts International Section A: Humanities and Social Sciences,* 79(3-A(E).

Linden, J. B., Capitanio, J. P., McCowan, B. and Isbell, L. A. (2018) 'Coping style and cortisol levels in infancy predict hair cortisol following new group formation in captive rhesus macaques (Macaca mulatta)', *American Journal of Primatology,* 80(12), pp. 13.

Liu, M.-S., Gao, J.-Q., Hu, G.-Y., Hao, G.-F., Jiang, T.-Z., Zhang, C. and Yu, S. (2022) 'MonkeyTrail: A scalable video-based method for tracking macaque movement trajectory in daily living cages', *Zoological Research,* 43(3), pp. 343.

Loyer, C. and Ha, J. (2017) 'Applications of Research in Nonhuman Animal Personality', *Personality in Nonhuman Animals*: Springer, pp. 307-318.

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', *American Journal of Primatology: Official Journal of the American Society of Primatologists,* 60(1), pp. 1-15.

Lutz, C. K. and Baker, K. C. (2023) 'Using Behavior to Assess Primate Welfare', in Robinson, L.M. and Weiss, A. (eds.) *Nonhuman Primate Welfare: From History, Science, and Ethics to Practice*. Cham: Springer International Publishing, pp. 171-205.

Lutz, C. K. and Brown, T. A. (2018) 'Porches as enrichment for singly housed cynomolgus macaques (Macaca fascicularis)', *Journal of the American Association for Laboratory Animal Science,* 57(2), pp. 134-137.

Lutz, C. K., Coleman, K., Hopper, L. M., Novak, M. A., Perlman, J. E. and Pomerantz, O. (2022) 'Nonhuman primate abnormal behavior: Etiology, assessment, and treatment', *American journal of primatology,* 84(6), pp. e23380.

Machado, C. J. and Bachevalier, J. (2006) 'The impact of selective amygdala, orbital frontal cortex, or hippocampal formation lesions on established social relationships in rhesus monkeys (Macaca mulatta)', *Behavioral Neuroscience,* 120(4), pp. 761-786.

Machado, C. J., Emery, N. J., Capitanio, J. P., Mason, W. A., Mendoza, S. P. and Amaral, D. G. (2008) 'Bilateral neurotoxic amygdala lesions in rhesus monkeys (Macaca mulatta): Consistent pattern of behavior across different social contexts', *Behavioral Neuroscience,* 122(2), pp. 251-266.

MacKay, J. R. D. and Haskell, M. J. (2015) 'Consistent Individual Behavioral Variation: The Difference between Temperament, Personality and Behavioral Syndromes', *Animals,* 5(3), pp. 455-478.

Maddali, H. T., Novitzky, M., Hrolenok, B., Walker, D., Balch, T. and Wallen, K. (2014) 'Inferring social structure and dominance relationships between rhesus macaques using RFID tracking data', *arXiv preprint arXiv:1407.0330*.

Maestripieri, D. (2000) 'Measuring temperament in rhesus macaques: consistency and change in emotionality over time', *Behavioural Processes,* 49(3), pp. 167-171.

Maestripieri, D., Schino, G., Aureli, F. and Troisi, A. (1992) 'A modest proposal: displacement activities as an indicator of emotions in primates', *Animal Behaviour,* 44(5), pp. 967-979.

Magden, E. R., Haller, R. L., Thiele, E. J., Buchl, S. J., Lambeth, S. P. and Schapiro, S. J. (2013) 'Acupuncture as an adjunct therapy for osteoarthritis in chimpanzees (Pan troglodytes)', *J Am Assoc Lab Anim Sci,* 52(4), pp. 475-80.

Maness, N. J., Schouest, B., Singapuri, A., Dennis, M., Gilbert, M. H., Bohm, R. P., Schiro, F., Aye, P. P., Baker, K., Van Rompay, K. K. A., Lackner, A. A., Bonaldo, M. C., Blair, R. V., Permar, S. R., Coffey, L. L., Panganiban, A. T. and Magnani, D. (2019) 'Postnatal Zika virus infection of nonhuman primate infants born to mothers infected with homologous Brazilian Zika virus', *Scientific Reports,* 9, pp. 9.

Maninger, N., Capitanio, J. P., Mendoza, S. P. and Mason, W. A. (2003) 'Personality influences tetanus-specific antibody response in adult male rhesus Macaques after removal from natal group and housing relocation', *American Journal of Primatology,* 61(2), pp. 73-83.

Marks, M., Jin, Q., Sturman, O., von Ziegler, L., Kollmorgen, S., von der Behrens, W., Mante, V., Bohacek, J. and Yanik, M. F. (2022) 'Deep-learning-based identification, tracking, pose estimation and behaviour classification of interacting primates and mice in complex environments', *Nature machine intelligence,* 4(4), pp. 331-340.

Martini, L. M., Bognár, A., Vogels, R. and Giese, M. A. (2024) 'MacAction: Realistic 3D macaque body animation based on multi-camera markerless motion capture', *bioRxiv*, pp. 2024.01. 29.577734.

Mason, G. J. and Latham, N. (2004) 'Can't stop, won't stop: is stereotypy a reliable animal welfare indicator?'.

Massey, D. A. (2023) *Delaying the Weaning of Captive Rhesus Macaques (Macaca mulatta) Destined for Use in Neuroscience Research: An Assessment of the Costs and Benefits.* Newcastle University.

Mather, J. A. and Anderson, R. C. (1993) 'Personalities of octopuses (Octopus rubescens)', *Journal of Comparative Psychology,* 107(3), pp. 336.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W. and Bethge, M. (2018) 'DeepLabCut: markerless pose estimation of user-defined body parts with deep learning', *Nature neuroscience,* 21(9), pp. 1281-1289.

Mathis, M. W. and Mathis, A. (2020) 'Deep learning tools for the measurement of animal behavior in neuroscience', *Current opinion in neurobiology,* 60, pp. 1-11.

Matsumoto, J., Kaneko, T., Kimura, K., Negrete, S. B., Guo, J., Suda-Hashimoto, N., Kaneko, A., Morimoto, M., Nishimaru, H. and Setogawa, T. (2023) 'Three-dimensional markerless motion capture of multiple freely behaving monkeys for automated characterization of social behavior', *bioRxiv*, pp. 2023.09. 13.556332.

Matthews, S. G., Miller, A. L., Plötz, T. and Kyriazakis, I. (2017) 'Automated tracking to measure behavioural changes in pigs for health and welfare monitoring', *Scientific reports,* 7(1), pp. 17582.

McClintick, M. N. and Grant, K. A. (2016) 'Aggressive temperament predicts ethanol self-administration in late adolescent male and female rhesus macaques', *Psychopharmacology,* 233(23-24), pp. 3965-3976.

McCormack, K. a. b. c., Newman, T. K. d., Higley, J. D. e., Maestripieri, D. b. f. and Sanchez, M. M. b. c. g. (2009) 'Serotonin transporter gene variation, infant abuse, and responsiveness to stress in rhesus macaque mothers and infants.', *Hormones & Behavior April,* 55(4), pp. 538-547.

McCowan, B., Beisner, B. and Hannibal, D. (2018) 'Social management of laboratory rhesus macaques housed in large groups using a network approach: A review', *Behavioural processes,* 156, pp. 77-82.

McCowan, B., Beisner, B. A., Capitanio, J. P., Jackson, M. E., Cameron, A. N., Seil, S., Atwill, E. R. and Fushing, H. (2011) 'Network Stability Is a Balancing Act of Personality, Power, and Conflict Dynamics in Rhesus Macaque Societies', *Plos One,* 6(8), pp. 8.

McCrae, R. R. and Weiss, A. (2007) 'Observer ratings of personality', *Handbook of research methods in personality psychology*, pp. 259-272.

McFarland, D. (1989) 'Problems of animal behaviour', *(No Title).*

Mimura, K., Matsumoto, J., Mochihashi, D., Nakamura, T., Hirabayashi, T., Higuchi, M. and Minamimoto, T. (2023) 'Unsupervised decomposition of natural monkey behavior into a sequence of motion motifs', *bioRxiv*, pp. 2023.03.04.531044.

Mitchell, A. S., Thiele, A., Petkov, C. I., Roberts, A., Robbins, T. W., Schultz, W. and Lemon, R. (2018) 'Continued need for non-human primate neuroscience research', *Current biology,* 28(20), pp. R1186-R1187.

Molendijk, M. L. and de Kloet, E. R. (2015) 'Immobility in the forced swim test is adaptive and does not reflect depression', *Psychoneuroendocrinology,* 62, pp. 389-391.

Morita, T., Toyoda, A., Aisu, S., Kaneko, A., Suda-Hashimoto, N., Matsuda, I. and Koda, H. (2020) 'Animals exhibit consistent individual differences in their movement: A case study on location trajectories of Japanese macaques', *Ecological Informatics,* 56, pp. 101057.

Morton, F. B., Lee, P. C., Buchanan-Smith, H. M., Brosnan, S. F., Thierry, B., Paukner, A., de Waal, F. B. M., Widness, J., Essler, J. L. and Weiss, A. (2013) 'Personality Structure in Brown Capuchin Monkeys (Sapajus apella): Comparisons With Chimpanzees (Pan troglodytes), Orangutans (Pongo spp.), and Rhesus Macaques (Macaca mulatta)', *Journal of Comparative Psychology,* 127(3), pp. 282-298.

Mouchet, A. and Dingemanse, N. J. (2021) 'A quantitative genetics approach to validate lab-versus field-based behavior in novel environments', *Behavioral Ecology,* 32(5), pp. 903-911.

NC3Rs (2017) 'NC3Rs Guidelines: Non-human primate accommodation, care and use.', *National Centre of Replacement, Refinement, & Reduction of Animals in Research*, pp. 1–52.

Nettle, D. (2006) 'The evolution of personality variation in humans and other animals', *American Psychologist,* 61(6), pp. 622.

Neudecker, V., Perez-Zoghbi, J. F., Coleman, K., Neuringer, M., Robertson, N., Bemis, A., Glickman, B., Schenning, K. J., Fair, D. A., Martin, L. D., Dissen, G. A. and Brambrink, A. M. (2021) 'Infant isoflurane exposure affects social behaviours, but does not impair specific cognitive domains in juvenile non-human primates.', *BJA: British Journal of Anaesthesia February,* 126(2), pp. 486-499.

Norman, M., Rowden, L. J. and Cowlishaw, G. (2021) 'Potential applications of personality assessments to the management of non-human primates: a review of 10 years of study', *PeerJ,* 9, pp. e12044.

O'Shea, K. and Nash, R. (2015) 'An introduction to convolutional neural networks', *arXiv preprint arXiv:1511.08458*.

Office, U. H. (2022) *Statistics of scientific procedures on living animals, Great Britain: 2021*. https://www.gov.uk/government/statistics/statistics-of-scientific-procedures-on-living-animals-great-britain-2021. Available at: https://www.gov.uk/government/statistics/statistics-of-scientific-procedures-on-living-animals-great-britain-2021 (Accessed: January 2024).

Oler, J. A., Fox, A. S., Shelton, S. E., Christian, B. T., Murali, D., Oakes, T. R., Davidson, R. J. and Kalin, N. H. (2009) 'Serotonin Transporter Availability in the Amygdala and Bed Nucleus of the Stria Terminalis Predicts Anxious Temperament and Brain Glucose Metabolic Activity', *Journal of Neuroscience,* 29(32), pp. 9961-9966.

Oler, J. A., Fox, A. S., Shelton, S. E., Rogers, J., Dyer, T. D., Davidson, R. J., Shelledy, W., Oakes, T. R., Blangero, J. and Kalin, N. H. (2010) 'Amygdalar and hippocampal substrates of anxious temperament differ in their heritability', *Nature,* 466(7308), pp. 864-868.

Olsson, I. A. S., Silva, S. P. d., Townend, D. and Sandøe, P. (2017) 'Protecting animals and enabling research in the European Union: An overview of development and implementation of directive 2010/63/EU', *ILAR journal,* 57(3), pp. 347-357.

Parker, K. J. (2022) 'Leveraging a translational research approach to drive diagnostic and treatment advances for autism', *Molecular psychiatry,* 27(6), pp. 2650-2658.

Parker, K. J., Garner, J. P., Oztan, O., Tarara, E. R., Li, J., Sclafani, V., Del Rosso, L. A., Chun, K., Berquist, S. W., Chez, M. G., Partap, S., Hardan, A. Y., Sherr, E. H. and Capitanio, J. P. (2018) 'Arginine vasopressin in cerebrospinal fluid is a marker of sociality in nonhuman primates.', *Science Translational Medicine May,* 10, pp. 439.

Paterson, E. A. and Turner, P. V. (2022) 'Challenges with assessing and treating pain in research primates: A focused survey and literature review', *Animals,* 12(17), pp. 2304.

Paukner, A., Capitanio, J. P. and Blozis, S. A. (2020) 'A new look at neurobehavioral development in rhesus monkey neonates (Macaca mulatta)', *American Journal of Primatology,* 82(5), pp. 11.

Perera, T. D., Dwork, A. J., Keegan, K. A., Thirumangalakudi, L., Lipira, C. M., Joyce, N., Lange, C., Higley, J. D., Rosoklija, G. and Hen, R. (2011) 'Necessity of hippocampal neurogenesis for the therapeutic action of antidepressants in adult nonhuman primates', *PloS one,* 6(4), pp. e17600.

Phillips, K. A., Elvey, C. R. and Abercrombie, C. (1998) 'Applying GPS to the study of primate ecology: A useful tool?', *American Journal of Primatology,* 46(2), pp. 167-172.

Pickard, J., Buchanan-Smith, H., Dennis, M., Flecknell, P., Joannides, A., Lemon, R., Prescott, M. and Schultz, W. (2013) 'Review of the assessment of cumulative severity and lifetime experience in non-human primates used in neuroscience research', *London, UK: Animal Procedures Committee*.

Poirier, C. and Bateson, M. (2017) 'Pacing stereotypies in laboratory rhesus macaques: Implications for animal welfare and the validity of neuroscientific findings', *Neuroscience & Biobehavioral Reviews,* 83, pp. 508-515.

Polit, D. F. (2014) 'Getting serious about test–retest reliability: a critique of retest research and some recommendations', *Quality of Life Research,* 23, pp. 1713-1720.

Pomerantz, O. and Capitanio, J. P. (2021) 'Temperament Predicts the Quality of Social Interactions in Captive Female Rhesus Macaques (Macaca mulatta)', *Animals,* 11(8), pp. 18.

Poole, T. (1997) 'Happy animals make good science', *Lab Anim,* 31(2), pp. 116-24.

Porto, S. M., Arcidiacono, C., Anguzza, U. and Cascone, G. (2013) 'A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns', *Biosystems Engineering,* 115(2), pp. 184-194.

Prescott, M. J. and Lidster, K. (2017) 'Improving quality of science through better animal welfare: the NC3Rs strategy', *Lab Anim (NY),* 46(4), pp. 152-156.

Qin, D., Chu, X., Feng, X., Li, Z., Yang, S., Lü, L., Yang, Q., Pan, L., Yin, Y. and Li, J. (2015) 'The first observation of seasonal affective disorder symptoms in Rhesus macaque', *Behavioural brain research,* 292, pp. 463-469.

Raper, J. P. D., Alvarado, M. C. P. D., Murphy, K. L. B. V. M. C. V. A. D. P. and Baxter, M. G. P. D. (2015) 'Multiple Anesthetic Exposure in Infant Monkeys Alters Emotional Reactivity to an Acute Stressor.', *Anesthesiology November,* 123(5), pp. 1084-1092.

Réale, D., Reader, S. M., Sol, D., McDougall, P. T. and Dingemanse, N. J. (2007) 'Integrating animal temperament within ecology and evolution', *Biological reviews,* 82(2), pp. 291-318.

Reinhardt, V. (1992) 'Space utilization by captive rhesus macaques', *Animal technology: journal of the Institute of Animal Technology,* 43(1).

Rendina, D. N., Lubach, G. R. and Coe, C. L. (2016) 'Gestational Timing of Prenatal Disturbance and Fetal Sex Determine the Developmental Outcomes', *Neonatology,* 109(4), pp. 314-320.

Revelle, W. and Revelle, M. W. (2015) 'Package 'psych'', *The comprehensive R archive network,* 337(338).

Robinson, L. M., Coleman, K., Capitanio, J. P., Gottlieb, D. H., Handel, I. G., Adams, M. J., Leach, M. C., Waran, N. K. and Weiss, A. (2018a) 'Rhesus macaque personality, dominance, behavior, and health', *American Journal of Primatology,* 80(2), pp. 8.

Robinson, L. M., Coleman, K., Capitanio, J. P., Gottlieb, D. H., Handel, I. G., Adams, M. J., Leach, M. C., Waran, N. K. and Weiss, A. (2018b) 'Rhesus macaque personality, dominance, behavior, and health', *American journal of primatology,* 80(2), pp. e22739.

Robinson, L. M., Waran, N. K., Handel, I. and Leach, M. C. (2021a) 'Happiness, welfare, and personality in rhesus macaques (Macaca mulatta)', *Applied Animal Behaviour Science,* 236, pp. 10.

Robinson, L. M., Waran, N. K., Handel, I. and Leach, M. C. (2021b) 'Happiness, welfare, and personality in rhesus macaques (Macaca mulatta)', *Applied Animal Behaviour Science,* 236, pp. 105268.

Robinson, L. M. and Weiss, A. (2023) 'Primate Personality and Welfare', in Robinson, L.M. and Weiss, A. (eds.) *Nonhuman Primate Welfare: From History, Science, and Ethics to Practice.* Cham: Springer International Publishing, pp. 395-411.

Roelfsema, P. R. and Treue, S. (2014) 'Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research', *Neuron,* 82(6), pp. 1200-1204.

Rogers, J., Raveendran, M., Fawcett, G. L., Fox, A. S., Shelton, S. E., Oler, J. A., Cheverud, J., Muzny, D. M., Gibbs, R. A., Davidson, R. J. and Kalin, N. H. (2013) 'CRHR1 genotypes, neural circuits and the diathesis for anxiety and depression', *Molecular Psychiatry,* 18(6), pp. 700-707.

Rogers, J., Shelton, S. E., Shelledy, W., Garcia, R. and Kalin, N. H. (2008) 'Genetic influences on behavioral inhibition and anxiety in juvenile rhesus macaques', *Genes Brain and Behavior,* 7(4), pp. 463-469.

Rommeck, I., Capitanio, J. P., Strand, S. C. and McCowan, B. (2011) 'Early Social Experience Affects Behavioral and Physiological Responsiveness to Stressful Conditions in Infant Rhesus Macaques (Macaca mulatta)', *American Journal of Primatology,* 73(7), pp. 692-701.

Rose, C., De Heer, R., Korte, S., Van der Harst, J., Weinbauer, G. and Spruijt, B. (2012) 'Quantified tracking and monitoring of diazepam treated socially housed cynomolgus monkeys', *Regulatory Toxicology and Pharmacology,* 62(2), pp. 292-301.

Roseboom, P. H., Nanda, S. A., Fox, A. S., Oler, J. A., Shackman, A. J., Shelton, S. E., Davidson, R. J. and Kalin, N. H. (2014) 'Neuropeptide Y Receptor Gene Expression in the Primate Amygdala Predicts Anxious Temperament and Brain Metabolism', *Biological Psychiatry,* 76(11), pp. 850-857.

Roughan, J., Wright-Williams, S. and Flecknell, P. (2009) 'Automated analysis of postoperative behaviour: assessment of HomeCageScan as a novel method to rapidly identify pain and analgesic effects in mice', *Laboratory animals,* 43(1), pp. 17-26.

Ruhde, A. A., Baker, K. C., Russell-Lodrigue, K. E., Blanchard, J. L. and Bohm, R. P. (2020) 'Trio housing of adult male rhesus macaques (Macaca mulatta): Methodology and outcome predictors', *Journal of Medical Primatology,* 49(4), pp. 188-201.

Rushen, J., Chapinal, N. and de Passile, A. (2012) 'Automated monitoring of behavioural-based animal welfare indicators', *Animal Welfare,* 21(3), pp. 339-350.

Salguero, F. J., White, A. D., Slack, G. S., Fotheringham, S. A., Bewley, K. R., Gooch, K. E., Longet, S., Humphries, H. E., Watson, R. J. and Hunter, L. (2021) 'Comparison of rhesus and cynomolgus macaques as an infection model for COVID-19', *Nature Communications,* 12(1), pp. 1260.

Schapiro, S. J. and Hau, J. (2023) 'Research Benefits of Improving Welfare in Captive Primates', *Nonhuman Primate Welfare: From History, Science, and Ethics to Practice*: Springer, pp. 445-462.

Schapiro, S. J., Nehete, P. N., Perlman, J. E. and Sastry, K. J. (2000) 'A comparison of cell-mediated immune responses in rhesus macaques housed singly, in pairs, or in groups', *Applied Animal Behaviour Science,* 68(1), pp. 67-84.

Schapiro, S. J., Perlman, J. E. and Boudreau, B. A. (2001) 'Manipulating the affiliative interactions of group-housed rhesus macaques using positive reinforcement training techniques', *American Journal of Primatology: Official Journal of the American Society of Primatologists,* 55(3), pp. 137-149.

Schapiro, S. J., Perlman, J. E., Thiele, E. and Lambeth, S. (2005) 'Training nonhuman primates to perform behaviors useful in biomedical research', *Lab Anim (NY),* 34(5), pp. 37-42.

Schino, G., Perretta, G., Taglioni, A. M., Monaco, V. and Troisi, A. (1996) 'Primate displacement activities as an ethopharmacological model of anxiety', *Anxiety,* 2(4), pp. 186-191.

Schino, G., Troisi, A., Perretta, G. and Monaco, V. (1991) 'Measuring anxiety in nonhuman primates: Effect of lorazepam on macaque scratching', *Pharmacology Biochemistry and Behavior,* 38(4), pp. 889-891.

Schlaepfer, T. E., Bewernick, B. H., Kayser, S., Mädler, B. and Coenen, V. A. (2013) 'Rapid effects of deep brain stimulation for treatment-resistant major depression', *Biological psychiatry,* 73(12), pp. 1204-1212.

Schneider, M. L., Coe, C. L. and Lubach, G. R. (1992) 'Endocrine activation mimics the adverse effects of prenatal stress on the neuromotor development of the infant primate', *Developmental Psychobiology,* 25(6), pp. 427-439.

Schneider, M. L., Kraemer, G. W. and Suomi, S. J. (1991) 'The effects of vestibular-proprioceptive stimulation on motor maturation and response to challenge in rhesus monkey infants', *Occupational Therapy Journal of Research,* 11(3), pp. 135-154.

Schneider, M. L., Moore, C. F., Barnhart, T. E., Larson, J. A., DeJesus, O. T., Mukherjee, J., Nickles, R. J., Converse, A. K., Roberts, A. D. and Kraemer, G. W. (2005) 'Moderate-Level Prenatal Alcohol Exposure Alters Striatal Dopamine System Function in Rhesus Monkeys.', *Alcoholism: Clinical & Experimental Research September,* 29(9), pp. 1685-1697.

Schneider, M. L., Moore, C. F. and Becker, E. F. (2001) 'Timing of Moderate Alcohol Exposure During Pregnancy and Neonatal Outcome in Rhesus Monkeys (Macaca mulatta).', *Alcoholism: Clinical & Experimental Research August,* 25(8), pp. 1238-1245.

Schneider, M. L., Moore, C. F., Suomi, S. J. and Champoux, M. (1991) 'Laboratory assessment of temperament and environmental enrichment in rhesus monkey infants (Macaca mulatta)', *American Journal of Primatology,* 25(3), pp. 137-155.

Schneider, M. L. and Suomi, S. J. (1992a) 'Laboratory assessment of temperament and postrotary nystagmus responses in rhesus monkey infants (macaca mulatta)', *Physical and Occupational Therapy in Pediatrics,* 12(1), pp. 37-52.

Schneider, M. L. and Suomi, S. J. (1992b) 'Neurobehavioral assessment in rhesus monkey neonates (Macaca mulatta): developmental changes, behavioral stability, and early experience', *Infant Behavior and Development,* 15(2), pp. 155-177.

Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D. and Carvalho, S. (2019) 'Chimpanzee face recognition from videos in the wild using deep learning', *Science Advances,* 5(9), pp. eaaw0736.

Schofield, D. P., Albery, G. F., Firth, J. A., Mielke, A., Hayashi, M., Matsuzawa, T., Biro, D. and Carvalho, S. (2023) 'Automated face recognition using deep neural networks produces robust primate social networks and sociality measures', *Methods in Ecology and Evolution,* 14(8), pp. 1937-1951.

Schuster, M. and Paliwal, K. K. (1997) 'Bidirectional recurrent neural networks', *IEEE transactions on Signal Processing,* 45(11), pp. 2673-2681.

Scott, J. T. and Bourne, J. A. (2022) 'Modelling behaviors relevant to brain disorders in the nonhuman primate: Are we there yet?', *Progress in Neurobiology,* 208, pp. 102183.

Shackman, A. J., Fox, A. S., Oler, J. A., Shelton, S. E., Davidson, R. J. and Kalin, N. H. (2013) 'Neural mechanisms underlying heterogeneity in the presentation of anxious temperament', *Proceedings of the National Academy of Sciences of the United States of America,* 110(15), pp. 6145-6150.

Shackman, A. J., Fox, A. S., Oler, J. A., Shelton, S. E., Oakes, T. R., Davidson, R. J. and Kalin, N. H. (2017) 'Heightened extended amygdala metabolism following threat characterizes the early phenotypic risk to develop anxiety-related psychopathology', *Molecular Psychiatry,* 22(5), pp. 724-732.

Shao, L., Cai, Z., Liu, L. and Lu, K. (2017) 'Performance evaluation of deep feature learning for RGB-D image/video classification', *Information Sciences,* 385-386, pp. 266-283.

Shively, C. A. (2017) 'Depression in captive nonhuman primates: theoretical underpinnings, methods, and application to behavioral management', *Handbook of Primate Behavioral Management*: CRC Press, pp. 115-125.

Sih, A., Bell, A. M., Johnson, J. C. and Ziemba, R. E. (2004) 'Behavioral syndromes: an integrative overview', *The quarterly review of biology,* 79(3), pp. 241-277.

Silverman, J. L., Thurm, A., Ethridge, S. B., Soller, M. M., Petkova, S. P., Abel, T., Bauman, M. D., Brodkin, E. S., Harony-Nicolas, H. and Wöhr, M. (2022) 'Reconsidering animal models used to study autism spectrum disorder: Current state and optimizing future', *Genes, brain and behavior,* 21(5), pp. e12803.

Simpson, E. A., Robinson, L. M. and Paukner, A. (2019) 'Infant rhesus macaque (Macaca mulatta) personality and subjective well-being', *Plos One,* 14(12), pp. 23.

Sloan, E. K., Capitanio, J. P., Tarara, R. P. and Cole, S. W. (2008) 'Social temperament and lymph node innervation', *Brain Behavior and Immunity,* 22(5), pp. 717-726.

Song, X., Leroy, T., Vranken, E., Maertens, W., Sonck, B. and Berckmans, D. (2007) 'Automatic detection of lameness in dairy cattle (1): Vision-based trackway analysis in cow's locomotion', *Precision livestock farming'07*: Wageningen Academic, pp. 277-283.

Spiegel, O. and Pinter-Wollman, N. (2022) 'Dissecting how behavior and environment shape spatial personalities: a comment on Stuber et al', *Behavioral Ecology,* 33(3), pp. 489-491.

Stamps, J. and Groothuis, T. G. G. (2010) 'The development of animal personality: relevance, concepts and perspectives', *Biological Reviews,* 85(2), pp. 301-325.

Stavrakakis, S., Guy, J. H., Warlow, O. M. E., Johnson, G. R. and Edwards, S. A. (2014) 'Walking kinematics of growing pigs associated with differences in musculoskeletal conformation, subjective gait score and osteochondrosis', *Livestock Science,* 165, pp. 104-113.

Stevenson-Hinde, J., Stillwell-Barnes, R. and Zunz, M. (1980) 'Subjective assessment of rhesus monkeys over four successive years', *Primates,* 21(1), pp. 66-82.

Stevenson-Hinde, J. and Zunz, M. (1978) 'Subjective assessment of individual rhesus monkeys', *Primates,* 19, pp. 473-482.

Story, L. and Kennedy, S. (2004) 'Animal studies in endometriosis: a review', *ILAR journal,* 45(2), pp. 132-138.

Stuber, E. F., Carlson, B. S. and Jesmer, B. R. (2022) 'Spatial personalities: a meta-analysis of consistent individual differences in spatial behavior', *Behavioral Ecology,* 33(3), pp. 477-486.

Suarez-Jimenez, B., Hathaway, A., Waters, C., Vaughan, K., Suomi, S. J., Noble, P. L., Pine, D. S., Fox, N. A. and Nelson, E. E. (2013) 'Effect of Mother's Dominance Rank on Offspring Temperament in Infant Rhesus Monkeys (Macaca mulatta)', *American Journal of Primatology,* 75(1), pp. 65-73.

Sueur, C., Forin-Wiart, M.-A. and Pelé, M. (2020) 'Are they really trying to save their buddy? the anthropomorphism of animal epimeletic behaviours', *Animals,* 10(12), pp. 2323.

Sullivan, E. C., Hinde, K., Mendoza, S. P. and Capitanio, J. P. (2011) 'Cortisol Concentrations in the Milk of Rhesus Monkey Mothers Are Associated With Confident Temperament in Sons, But Not Daughters', *Developmental Psychobiology,* 53(1), pp. 96-104.

Sullivan, E. C., Mendoza, S. P. and Capitanio, J. P. (2011) 'Similarity in Temperament Between Mother and Offspring Rhesus Monkeys: Sex Differences and the Role of Monoamine Oxidase-A and Serotonin Transporter Promoter Polymorphism Genotypes', *Developmental Psychobiology,* 53(6), pp. 549-563.

Suomi, S. J., Novak, M. A. and Well, A. (1996) 'Aging in rhesus monkeys: Different windows on behavioral continuity and change', *Developmental Psychology,* 32(6), pp. 1116-1128.

Sussman, A. F., Ha, J. C., Bentson, K. L. and Crockett, C. M. (2013) 'Temperament in Rhesus, Long-Tailed, and Pigtailed Macaques Varies by Species and Sex', *American Journal of Primatology,* 75(4), pp. 303-313.

Talbot, C. F., Maness, A. C., Capitanio, J. P. and Parker, K. J. (2021) 'The factor structure of the macaque social responsiveness scale-revised predicts social behavior and personality dimensions', *American Journal of Primatology,* 83(5), pp. 10.

Terven, J., Córdova-Esparza, D.-M. and Romero-González, J.-A. (2023) 'A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas', *Machine Learning and Knowledge Extraction,* 5(4), pp. 1680-1716.

Thiele, A. and Bellgrove, M. A. (2018) 'Neuromodulation of attention', *Neuron,* 97(4), pp. 769-785.

Thierry, B. (2004) '12 Social epigenesis', *Macaque societies: A model for the study of social organization,* 41, pp. 267.

Treue, S. and Lemon, R. (2023) 'The indispensable contribution of nonhuman primates to biomedical research', *Nonhuman Primate Welfare: From History, Science, and Ethics to Practice*: Springer, pp. 589-603.

Tromp, D. P. M., Fox, A. S., Oler, J. A., Alexander, A. L. and Kalin, N. H. (2019) 'The Relationship Between the Uncinate Fasciculus and Anxious Temperament Is Evolutionarily Conserved and Sexually Dimorphic', *Biological Psychiatry,* 86(12), pp. 890-898.

Truelove, M. A., Martin, J. E., Langford, F. M. and Leach, M. C. (2020) 'The identification of effective welfare indicators for laboratory-housed macaques using a Delphi consultation process', *Scientific Reports,* 10(1), pp. 20402.

Tuttle, A. H., Molinaro, M. J., Jethwa, J. F., Sotocinal, S. G., Prieto, J. C., Styner, M. A., Mogil, J. S. and Zylka, M. J. (2018) 'A deep neural network to assess spontaneous pain from mouse facial expressions', *Molecular Pain,* 14, pp. 1744806918763658.

Ueno, M., Hayashi, H., Kabata, R., Terada, K. and Yamada, K. (2019) 'Automatically detecting and tracking free-ranging Japanese macaques in video recordings with deep learning and particle filters', *Ethology,* 125(5), pp. 332-340.

Ueno, M., Kabata, R., Hayashi, H., Terada, K. and Yamada, K. (2022) 'Automatic individual recognition of Japanese macaques (Macaca fuscata) from sequential images', *Ethology,* 128(5), pp. 461-470.

Urano, E., Okamura, T., Ono, C., Ueno, S., Nagata, S., Kamada, H., Higuchi, M., Furukawa, M., Kamitani, W. and Matsuura, Y. (2021) 'COVID-19 cynomolgus macaque model reflecting human COVID-19 pathological conditions', *Proceedings of the National Academy of Sciences,* 118(43), pp. e2104847118.

Valletta, J. J., Torney, C., Kings, M., Thornton, A. and Madden, J. (2017) 'Applications of machine learning in animal behaviour studies', *Animal Behaviour,* 124, pp. 203-220.

VandeBerg, J. L. and Williams-Blangero, S. (1997) 'Advantages and limitations of nonhuman primates as animal models in genetic research on complex diseases', *Journal of medical primatology,* 26(3), pp. 113-119.

Vandeleest, J. J. and Capitanio, J. P. (2012) 'Birth Timing and Behavioral Responsiveness Predict Individual Differences in the Mother-Infant Relationship and Infant Behavior During Weaning and Maternal Breeding', *American Journal of Primatology,* 74(8), pp. 734-746.

Vandeleest, J. J., Mendoza, S. P. and Capitanio, J. P. (2013) 'Birth timing and the mother-infant relationship predict variation in infant behavior and physiology.', *Developmental Psychobiology December,* 55(8), pp. 829-837.

Vazire, S., Gosling, S. D., Dickey, A. S. and Schapiro, S. J. (2007) 'Measuring personality in nonhuman animals', *Handbook of research methods in personality psychology*, pp. 190-206.

Veazey, R. S. and Lackner, A. A. (2017) 'Nonhuman primate models and understanding the pathogenesis of HIV infection and AIDS', *ILAR journal,* 58(2), pp. 160-171.

Vermeire, T., Epstein, M., Badin, R., Flecknell, P., Hoet, P., Hudson-Shore, M., Jones, D., Krätke, R., Langermans, J. and Prescott, M. (2017) 'Final opinion on the need for non-human primates in biomedical research, production and testing of products and devices (update 2017)', *Sci Comm Heal Environ Emerg Risks*.

Vitale, A., Manciocco, A. and Alleva, E. (2009) 'The 3R principle and the use of non-human primates in the study of neurodegenerative diseases: the case of Parkinson's disease', *Neuroscience & Biobehavioral Reviews,* 33(1), pp. 33-47.

von Borell, C., Kulik, L. and Widdig, A. (2016) 'Growing into the self: the development of personality in rhesus macaques', *Animal Behaviour,* 122, pp. 183-195.

von Ziegler, L., Sturman, O. and Bohacek, J. (2021) 'Big behavior: challenges and opportunities in a new era of deep behavior profiling', *Neuropsychopharmacology,* 46(1), pp. 33-44.

Walker, C. K., VandeVoort, C. A., Li, C.-S., Chaffin, C. L. and Capitanio, J. P. (2018) 'Adiposity and weight gain during pregnancy associate independently with behavior of infant rhesus monkeys (Macaca mulatta).', *Developmental Psychobiology September,* 60(6), pp. 629-638.

Wang, C.-Y., Yeh, I.-H. and Liao, H.-Y. M. (2024) 'YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information', *arXiv preprint arXiv:2402.13616*.

Webster, J. (1995) 'Animal welfare: a cool eye towards Eden'.

Weinstein, B. G. (2018) 'A computer vision for animal ecology', *Journal of Animal Ecology,* 87(3), pp. 533-545.

Weinstein, T. A. R. and Capitanio, J. P. (2008) 'Individual differences in infant temperament predict social relationships of yearling rhesus monkeys, Macaca mulatta', *Animal Behaviour,* 76, pp. 455-465.

Weinstein, T. A. R. and Capitanio, J. P. (2012) 'Longitudinal Stability of Friendships in Rhesus Monkeys (Macaca mulatta): Individual- and Relationship-Level Effects', *Journal of Comparative Psychology,* 126(1), pp. 97-108.

Weiss, A. (2017) 'Exploring factor space (and other adventures) with the Hominoid Personality Questionnaire', *Personality in nonhuman animals*: Springer, pp. 19-38.

Weiss, A. and Adams, M. J. (2013) 'Differential behavioral ecology', *Animal personalities: Behavior, physiology and evolution*, pp. 96-123.

Weiss, A., Adams, M. J. and Johnson, W. (2011) 'The big none: No evidence for a general factor of personality in chimpanzees, orangutans, or rhesus macaques', *Journal of Research in Personality,* 45(4), pp. 393-397.

Weiss, A., Adams, M. J., Widdig, A. and Gerald, M. S. (2011) 'Rhesus Macaques (Macaca mulatta) as Living Fossils of Hominoid Personality and Subjective Well-Being', *Journal of Comparative Psychology,* 125(1), pp. 72-83.

Weiss, A., Inoue-Murayama, M., Hong, K. W., Inoue, E., Udono, T., Ochiai, T., Matsuzawa, T., Hirata, S. and King, J. E. (2009) 'Assessing chimpanzee personality and subjective well-being in Japan', *American Journal of Primatology: Official Journal of the American Society of Primatologists,* 71(4), pp. 283-292.

Whitham, W. and Washburn, D. A. (2017) 'A history of animal personality research', *Personality in nonhuman animals*: Springer, pp. 3-16.

Wichmann, T., Bergman, H. and DeLong, M. R. (2018) 'Basal ganglia, movement disorders and deep brain stimulation: advances made through non-human primate research', *Journal of Neural Transmission,* 125, pp. 419-430.

Willard, S. L., Hemby, S. E., Register, T. C., McIntosh, S. and Shively, C. A. (2014) 'Altered expression of glial and synaptic markers in the anterior hippocampus of behaviorally depressed female monkeys', *Neuroscience Letters,* 563, pp. 1-5.

Willette, A. A., Coe, C. L., Colman, R. J., Bendlin, B. B., Kastman, E. K., Field, A. S., Alexander, A. L., Allison, D. B., Weindruch, R. H. and Johnson, S. C. (2012) 'Calorie restriction reduces psychological stress reactivity and its association with brain volume and microstructure in aged rhesus monkeys', *Psychoneuroendocrinology,* 37(7), pp. 903-916.

Willette, A. A. a., Lubach, G. R. a., Knickmeyer, R. C. b., Short, S. J. a., Styner, M. b. c., Gilmore, J. H. b. and Coe, C. L. a. (2011) 'Brain enlargement and increased behavioral and cytokine reactivity in infant monkeys following acute prenatal endotoxemia.', *Behavioural Brain Research May,* 219(1), pp. 108-115.

Willner, P. and Mitchell, P. (2002) 'The validity of animal models of predisposition to depression', *Behavioural pharmacology,* 13(3), pp. 169-188.

Wiltshire, C., Lewis-Cheetham, J., Komedová, V., Matsuzawa, T., Graham, K. E. and Hobaiter, C. (2023) 'DeepWild: Application of the pose estimation tool DeepLabCut for behaviour tracking in wild chimpanzees and bonobos', *Journal of Animal Ecology,* 92(8), pp. 1560-1574.

Witham, C. L. (2018) 'Automated face recognition of rhesus macaques', *Journal of neuroscience methods,* 300, pp. 157-165.

Wolf, M. and Weissing, F. J. (2010) 'An explanatory framework for adaptive personality differences', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 365(1560), pp. 3959-3968.

Wood, E. K., Champoux, M., Lindell, S. G., Barr, C. S., Suomi, S. J. and Higley, J. D. (2020a) 'Neonatal temperament and neuromotor differences are predictive of adolescent alcohol intake in rhesus monkeys (Macaca mulatta)', *American Journal of Primatology,* 82(11), pp. 10.

Wood, E. K., Halter, C. M., Byrne, E., Baron, Z. D., Forvil, M., Marett, L., Smith, E., Hafen, E., Hepworth, E. and Johnson, M. (2024) 'Behavioral inhibition in a translational nonhuman primate model: A pilot study of Kagan's behavioral inhibition paradigm modified for use in infant rhesus macaques (Macaca mulatta)', *Developmental Psychology.*

Wood, E. K., Halter, C. M., Gabrielle, N., Capitanio, J. P. and Higley, J. D. (2021a) 'Stress-induced plasma cortisol concentrations in infancy are associated with later parenting behaviors in female rhesus macaques (Macaca mulatta).', *Developmental Psychobiology July,* 63(5), pp. 1098-1108.

Wood, E. K., Higley, J. D., Champoux, M., Marsiske, M., Olsen, J. A., Suomi, S. J. and Kay, D. B. (2021b) 'Multi-group multi-time point confirmatory factor analysis of the triadic structure of temperament: A nonhuman primate model', *Developmental Psychobiology,* 63(1), pp. 65-73.

Wood, E. K., Hunter, J. N., Olsen, J. A., Almasy, L., Lindell, S. G., Goldman, D., Barr, C. S., Suomi, S. J., Kay, D. B. and Higley, J. D. (2021c) 'Parental genetic contributions to neonatal temperament in a nonhuman primate (Macaca mulatta) model', *Developmental Psychobiology,* 63(5), pp. 997-1005.

Wood, E. K., Kruger, R., Cash, E., Lindell, S. G., Schwandt, M. L., Barr, C. S., Suomi, S. J. and Higley, J. D. (2020b) 'Early life temperamental anxiety is associated with excessive alcohol intake in adolescence: A rhesus monkey (Macaca mulatta) model', *Addiction Biology,* 25(6), pp. 10.

Wooddell, L. J., Kaburu, S. S., Rosenberg, K. L., Meyer, J. S., Suomi, S. J. and Dettmer, A. M. (2016) 'Matrilineal behavioral and physiological changes following the removal of a non-alpha matriarch in rhesus macaques (Macaca mulatta)', *PLoS One,* 11(6), pp. e0157108.

Wooddell, L. J., Simpson, E. A., Murphy, A. M., Dettmer, A. M. and Paukner, A. (2019) 'Interindividual differences in neonatal sociality and emotionality predict juvenile social status in rhesus monkeys', *Developmental Science,* 22(2), pp. 13.

Workman, K. P., Healey, B., Carlotto, A. and Lacreuse, A. (2019) 'One-year change in cognitive flexibility and fine motor function in middle-aged male and female marmosets (Callithrix jacchus)', *American journal of primatology,* 81(2), pp. e22924.

Würbel, H. (2002) 'Behavioral phenotyping enhanced – beyond (environmental) standardization', *Genes, Brain and Behavior,* 1(1), pp. 3-8.

Zaaimi, B., Edgley, S. A., Soteropoulos, D. S. and Baker, S. N. (2012) 'Changes in descending motor pathway connectivity after corticospinal tract lesion in macaque monkey', *Brain,* 135(7), pp. 2277-2289.

Zhang, B., Suarez-Jimenez, B., Hathaway, A., Waters, C., Vaughan, K., Noble, P. L., Fox, N. A., Suomi, S. J., Pine, D. S. and Nelson, E. E. (2012) 'Developmental changes of rhesus monkeys in response to separation from the mother', *Developmental Psychobiology,* 54(8), pp. 798-807.

Zijlmans, D. G. M., Vernes, M. K., Sterck, E. H. M. and Langermans, J. A. M. (2021) 'The utility of voluntary weighing in captive group-living rhesus macaques (Macaca mulatta)', *Animal Welfare,* 30(3), pp. 249-258.

Zualkernan, I., Dhou, S., Judas, J., Sajun, A. R., Gomez, B. R. and Hussain, L. A. (2022) 'An IoT system using deep learning to classify camera trap images on the edge', *Computers,* 11(1), pp. 13.