04 University of Plymouth Research Theses

01 Research Theses Main Collection

2023

Towards the Validation of Novel Cognitive and Physiological Indicators of Affective State in Horses (Equus caballus)

Kappel, Sarah

https://pearl.plymouth.ac.uk/handle/10026.1/20644

http://dx.doi.org/10.24382/5012 University of Plymouth

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.



Towards the Validation of Novel Cognitive and

Physiological Indicators of Affective State

in Horses (Equus caballus)

by Sarah Kappel

A thesis submitted to the University of Plymouth in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

OSkappel ne Science

School of Biological and Marine Science March 2023

Copyright statement

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's prior consent.

Acknowledgements

I feel immensely grateful to be able to look back on the past four years as an absolutely enjoyable and extraordinary experience. So many people have helped and supported me through my PhD journey, and I am deeply thankful to everyone one of you.

First and foremost, I would like to thank my four fabulous supervisors. I cannot sufficiently express my thanks for your continuous support, advice and dedication in seeing me through my PhD.

To Dr Carole Fureix, thank you for being a brilliant supervisor, mentor and friend. I would have not been able to work on this project without you. Thanks for always being there in times of struggle, pushing me out of my comfort zone and giving me confidence in my own abilities. You are my inspiration, and I am so grateful for everything you have done for me throughout my PhD journey and beyond.

To Dr Sarah Collins, I feel incredibly lucky you agreed to be my supervisor without hesitation, even though horses were not your 'thing'. I could have not wished for better guidance and support. Above all, thank you for being so enthusiastic about science, always keeping me calm in times of struggle, and often challenging me to take on different perspectives. I have learnt so much from you.

To Prof. Mike Mendl, thank you for your time and invaluable expertise. You have supported me throughout my PhD more than I could have ever hoped for. Being part of the Bristol ABW group has helped me in so many ways, not only in my development as a scientist but also on a personal level. Thank you for making it possible for me to be included in this incredible group.

To Dr Katherine Herborn, I am so grateful you joined my supervisory team. Your advice on the infrared thermography work has been incredibly helpful and I am so thankful I could always ask you for help when I got stuck with R.

The practical elements of my data collection would have been a lot more challenging without Dr Marco Ramirez Montes de Oca. Thank you, Marco, for helping me with the horses, carrying tons of equipment, and keeping me calm when things did not go as planned. Above all, thank you for collaborating with me on the reward study (Chapter 6), and collecting and extracting the IRT data.

To Dr Chanakarn Wongsaengchan, thank you for your immense help with the HGS scoring (Chapter 3) and IRT data extraction. Your hardworking attitude is impressive and I look forward to collaborating on another IRT study with you soon.

My write-up period would have been far less enjoyable if it were not for Shareen Bistre Dabbah. Thank you for spending countless writing 'retreats' with me, cooking up delicious Mexican meals and providing me with coffee *ad libitum*. Also, my deepest thanks for the drawings I included in Chapters 2 and 6.

Acknowledgements

I thank everyone else in the ABW group for all the productive lab meeting discussions, helpful feedback on conference talks and many good chats over cake breaks.

I would have not been able to collect any data if it were not for the three horse facilities that welcomed me and let me use their horses and facilities.

Working on a PhD during the pandemic was challenging but it also opened some unexpected 'zoom' doors to me. I am incredibly grateful to Melanie Hezzell, Feisal Subhan and Richard Mott for taking the time to advise me on various ECG and HRV-related issues. I also thank Jo Murrell and Kieran Bogat for lending me the ECG and PA equipment.

I am deeply thankful to my friends and family for all their support over the past four years. I feel incredibly lucky to share a home with my fantastic housemates Rachel and Megan. Thank you for giving me excuses to pause on work, bake, go for a run and always cheering me up when I needed it. Lockdown life would have been impossible to cope with without you two.

I feel incredibly lucky to have so many good friends, especially Phill, Evie, Nick, Laura, and Steph, who often unknowingly gave me with the strength to carry on with my PhD. I cannot thank Helena and Sam enough for welcoming me at their beautiful Welsh farm on numerous occasions, providing me with a peaceful place for writing up and lots of fun farm activities when I needed to recharge my brain.

As a first-generation academic, I have had many doubts that I would be capable of pursuing a PhD, but the unconditional support and patience from my family has kept me going. I especially thank my mum for always looking out for me, sending me numerous food parcels from Germany and patiently waiting for me to find a 'proper job'.

Finally, I want to thank my dear friend Dr Paul Molnar without whom I would never have considered studying animal welfare. There are no words to express how immensely grateful I am for having become part of your family and spending time on the ranch. Without you, and my two families, none of this would have been possible.

Abstract

Author's declaration

I hereby declare that all works presented in this thesis were prepared solely by the author, Sarah Kappel, except where reference is made to published literature. At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

This study was financed by a studentship from the University of Plymouth.

A version of Chapter 5 has been published in the journal Animal Cognition:

Kappel S, Ramirez Montes De Oca MA, Collins S, Herborn K, Mendl M, Fureix C. Do you see what I see? Testing horses' ability to recognise real-life objects from 2D computer projections. Anim Cogn. 2023 Mar 2:1–13. doi: 10.1007/s10071-023-01761-6

Word count main body of thesis: ~58500

Signed: Saral Kappel

Date: 29/03/2023

Abstract

Sarah Kappel - Towards the Validation of Novel Cognitive and Physiological Indicators of Affective State in Horses (Equus caballus)

Assessing animal emotions reliably is crucial for making accurate decisions about animal welfare, but subjective affective (emotional) states are inaccessible for direct measurement. We know that negative affect modulates attention in humans and some animals. The overarching aim of this thesis is to investigate if the preferential attention allocation towards certain information (*i.e.* attention bias, AB) could be a potential proxy measure of affective valence in horses.

First, I provide an overview of the theoretical background to attention and AB by reviewing the literature on AB in humans and animals (Chapter 2). I examine which methodological aspects need consideration when translating AB tests to animals, before describing my general methods in Chapter 3 and presenting the development of a visual attention bias test in horses using on-screen images of conspecific faces in Chapter 4. For a more informed interpretation of horses' response in the AB test, I further investigated horses' image recognition ability in Chapter 5. As well as cognitive measures, I explored changes in heart rate variability (HRV) and surface temperature via infrared thermography (IRT) as additional measures of affective valence in Chapter 6.

Reviewing the literature, I conclude that AB is a promising novel proxy of affect, but its assessment requires further refinement for reliable application in animals. Testing horses' attention to conspecific faces showing differently valenced facial expressions indicated no significant evidence of AB. This is possibly because artificial stimuli might not be appropriate test stimuli, as image recognition ability might not be generalizable across individual horses. Changes in physiological parameters indicated shifts in arousal but not valence, contributing to my conclusion that measuring emotional valence from indirect parameters like AB, HRV and IRT needs more refinement, which I discuss in my final chapter (Chapter 7).

Table of content

COP	/RIGHT STATEMENTII
ACK	NOWLEDGEMENTSIII
AUTI	IOR'S DECLARATIONV
ABST	RACTVI
TABL	E OF CONTENTVII
LIST	OF FIGURESIX
LIST	OF TABLESXI
Снар	TER 1 GENERAL INTRODUCTION
1.1	ANIMAL WELFARE AND ANIMAL EMOTIONS
1.2	AFFECTIVE STATES
1.3	OVERVIEW OF COMPONENTS OF ANIMAL EMOTIONS, THEIR ASSESSMENT AND LIMITATIONS21
1.4	RESEARCH AIMS AND THESIS OUTLINE
Снар	TER 2 A TTENTION TO EMOTIONAL STIMULI AS AN INDICATOR OF AFFECT IN HUMANS AND NON-HUMAN
ANIM	ALS – A LITERATURE REVIEW
2.1	BACKGROUND AND CHAPTER AIMS
2.2	WHAT IS ATTENTION?
2.3	FUNCTIONS OF ATTENTION
2.4	FORMS OF ATTENTION
2.5	TYPES OF VISUAL ATTENTION AND ITS ASSESSMENT
2.6	TOP-DOWN AND BOTTOM-UP PROCESSES OF ATTENTION
2.7	PROCESSING STAGES OF ATTENTION
2.8	ANXIETY-RELATED ATTENTIONAL BIASES IN RESPONSE TO THREAT IN HUMANS
2.9	AB IN ANIMALS
2.10	Conclusion
Снар	ter 3 General methods
3.1	ETHICAL APPROVAL
3.2	PILOT STUDIES AND CONCEPTUALISING OF EXPERIMENTS
3.3	CALCULATION OF SAMPLE SIZES
3.4	STUDY POPULATION RECRUITMENT AND YARD DESCRIPTIONS
3.5	HORSE WELFARE ASSESSMENT
3.6	EXCLUDED WELFARE FACTORS
Снар	TER 4 DEVELOPING AN AFFECT-RELATED ATTENTION BIAS TEST IN HORSES
4.1	INTRODUCTION

4.2	Methods	
4.3	Results	115
4.4	DISCUSSION	
4.5	CONCLUSION	
Сна	PTER 5 TESTING HORSES' ABILITY TO RECOGNISE REAL-LIFE OBJECTS FROM 2D COI	MPUTER PROJECTIONS
5.1	INTRODUCTION	
5.2	Methods	132
5.3	Results	141
5.4	Discussion	145
5.5	CONCLUSION	149
Сна	PTER 6 I CHANGES IN HEART RATE VARIABILITY AND FACIAL SURFACE TEMPERATU	JRES IN RESPONSE TO
IMMI	EDIATE AND WITHHELD REWARD	150
6.1	INTRODUCTION	151
6.2	Methods	158
6.3	RESULTS	166
6.4	Discussion	171
6.5	CONCLUSION	177
Сна	PTER 7 GENERAL DISCUSSION	181
7.1	How should AB be assessed in animals?	
7.2	IS IT SENSIBLE TO ESTIMATE AFFECTIVE STATE FROM WELFARE MEASUREMENTS?	
7.3	FUTURE RESEARCH QUESTIONS	
7.4	CONCLUDING REMARKS	
REFE	RENCES	
APP	ENDIX A (CHAPTER 2)	
		<u></u>
		3 75
	LIVUR DICHAFTER 3J	
APP	ENDIX C (CHAPTER 4)	
APP	ENDIX D (CHAPTER 5)	228

List of Figures

Figure 1.1. Schema of the Five Domains framework for assessing animal welfare)15
Figure 1.2. The integrative framework to conceptualise animal emotions in terms of their arousal and
valence after Mendl et al. (2010)20
Figure 2.1 Schema of attentional processing stages during early and late cognitive processing
Figure 2.2 A schematic model of the cognitive mechanisms underlying threat processing adapted from Bar-
Haim et al (2007)
Figure 2.3 Behavioural indicators of overt visual attention in (A) humans and (B) primates and (C) horses69
Figure 3.1 Example images of housing conditions in (A) yard 1, (B) yard 2 and (C+D) yard 385
Figure 3.2 Example images of A) application of the pressure algometry device and B) test areas on the
horse's back97
Figure 4.1 Diagram of the AB test set-up108
Figure 4.2 Absolute duration (A) and proportion of time (B) recorded when horses attended to the stimulus
pairs for the first time115
Figure 4.3 Attention duration (raw data presented) to the negative stimulus predicted by welfare score.
Different shapes and colours indicate data points for each yard
Figure 4.4 Graphs to estimated marginal means for Yard (A) and stimulus horse ID (B)117
Figure 5.1 Experimental design comprising the objects discrimination training (ODT) and object recognition
test (ORT)139
Figure 5.2 Proportion of correct responses during each step of the object recognition test (ORT)142
Figure 5.3 Proportion of horses out of the 27 horses responding correctly or incorrectly depending on
whether the preceding trial was refreshed with objects (yes) or not (no)
Figure 6.1 Illustration of the three main theories of how lateralised hemispheric activity during negative
emotion processing might induce asymmetric eye temperatures157
Figure 6.2 Reward schedule and data collection timeline160
Figure 6.3 Illustration of the modified base-apex configuration used for ECG recording
Figure 6.4 Example of a thermal image used for extraction of maximum temperatures (marked by red
triangles) taken in the eyes (left (EI1) and right (EI2)), ear base (left (Li1 and right (Li2)) and hair whorl
(EI3)
Figure 6.5 Mean HR recorded during test phases166
Figure 6.6 Maximum temperatures of the medial canthus of the eye (EyeMax) recorded during test phases at
yard 1 (A) and yard 2 (B)167
Figure 6.7 Maximum ear base temperatures (EarMax) recorded during test phases at yard 1 (A) and yard 2
(B)167
Figure 6.8 Maximum hair whorl temperatures (WhorlMax) recorded during test phases at yard 1 (A) and
yard 2 (B)168

igure 6.9 Change of HRV indices linked to emotional valence measured as (A) RMSSD, (B) HF, (C) SDNN, and	d
(D) LF/HF ratio (log-transformed)16	59
igure 6.10 Differences in asymmetric EyeMax assessed during baseline and reward phases in (A) yard 1 and	d
(B) yard 217	70

List of Tables

Table 1.1 Glossary of keywords used throughout this thesis
Table 2.1 Overview of the most commonly applied paradigms to study anxiety-related AB in humans60
Table 2.2 Overview of commonly applied ABT in animals. 72
Table 3.1 Description of welfare factors, which all equally weighted into the final welfare score87
Table 3.2 Scoring system of the voluntary and forced approach test. 92
Table 3.3 Results of inter-observer reliability testing (weighted kappa) between the three independent
raters
Table 3.4 Overview of facial action units (FAUs) evaluated for horse grimace scale scoring after Dalla Costa et
al. (2014)95
Table 3.5 Information to horses, individual welfare scores and their use in experimental Chapters 4-698
Table 4.1 Ethogram of horse behaviours scored during the attention bias test.
Table 4.2 Example image of stimuli and number of trials per stimulus pair used for AB testing
Table 4.3 Model outputs from the LMER analysis with attention 124
Table 4.4 Model outputs from the glmmTMB analysis with proportion 124
Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent
Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent variable. 125
Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependentvariable.125Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive
Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent variable. 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive stimulus as dependent variable.
Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent variable. 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive 125 stimulus as dependent variable. 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable
Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable 125 Table 5.1 Estimated regression parameters from the GLMM model. 141
Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable 126 Table 5.1 Estimated regression parameters from the GLMM model. 141 Table 5.2 Model estimates of GLMM with 'response' as binary dependent variable (correct/wrong) and 141
Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable 126 Table 5.1 Estimated regression parameters from the GLMM model. 141 Table 5.2 Model estimates of GLMM with 'response' as binary dependent variable (correct/wrong) and predictors with comparator information in square brackets. P-values ≤0.05 are highlighted in bold 144
 Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent variable. 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive stimulus as dependent variable. 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable
 Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent variable. 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive stimulus as dependent variable. 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable126 Table 5.1 Estimated regression parameters from the GLMM model. 141 Table 5.2 Model estimates of GLMM with 'response' as binary dependent variable (correct/wrong) and predictors with comparator information in square brackets. P-values ≤0.05 are highlighted in bold144 Table 6.1 Overview of time-domain and frequency-domain HRV indices and IRT measures and their predicted changes during IR and WR.
 Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent variable. 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive stimulus as dependent variable. 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable. 126 Table 5.1 Estimated regression parameters from the GLMM model. 141 Table 5.2 Model estimates of GLMM with 'response' as binary dependent variable (correct/wrong) and predictors with comparator information in square brackets. P-values ≤0.05 are highlighted in bold 144 Table 6.1 Overview of time-domain and frequency-domain HRV indices and IRT measures and their predicted changes during IR and WR. 162 Table 6.2 Model estimates from LMER analysis of HRV indices. WR withheld reward, IR immeditate reward.
 Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent variable. 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive stimulus as dependent variable. 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable
 Table 4.5 Model outputs from the LMER analysis with attention to the positive stimulus as dependent variable. 125 Table 4.6 Model outputs from the glmmTMB analysis with proportion of time attending to the positive stimulus as dependent variable. 125 Table 4.7 Model outputs from the LMER analysis with total attention duration as dependent variable

1 Chapter 1 | General introduction

2 Summary

3 In this introduction, I describe the theoretical approaches to animal welfare and animal 4 emotions, and the relationship between the two representing the central topic of my thesis. 5 First, I outline how animal welfare, a moral-based concept created by humans to alleviate 6 animal suffering, can be understood according to the three main approaches of determining 7 welfare animals' freedom to express natural behaviour, functioning, and feelings. Given that 8 the focus of this thesis is the development of novel measures of animal emotions, I explain 9 what emotional (affective) states are and how these may be assessed scientifically in animals 10 using behavioural, physiological and cognitive measures. I then justify why I consider the 11 horse (Equus ferus caballus) as a valuable model species for the aims of my thesis. I provide 12 information on horses' ethological background, and discuss how common housing and management conditions relate to horse welfare. Finally, I outline my research aims and how 13 14 I have addressed these in my thesis. For facilitated reading and clarity, I provide a glossary 15 with key terms used in this thesis at the end of this chapter in Table 1.1.

16 **1.1** Animal welfare and animal emotions

17 The scientific assessment of animal welfare originated from societal concerns about the 18 treatment of animals under human care causing animals to suffer (i.e. the experience of 19 unpleasant emotional states, like pain and fear (Harrison, 1964)). Thus far, preventing and alleviating suffering is a core principle of animal welfare science (Broom, 1991; Dawkins, 20 1980). To provide animals with suitable living conditions, accurate scientific measures that 21 22 can indicate the current welfare state and whether/how changes in housing and husbandry 23 improve this state are required. However, there is no 'single measure' of animal welfare 24 (Mason & Mendl, 1993). Identifying reliable welfare indicators first requires clearly defining 25 and conceptualising what welfare is. The answer to this question is, however, rather intricate 26 as moral positions and attitudes towards animals can differ greatly, hence people might apply 27 different criteria in judging what constitutes animal welfare and what animals need to have a 28 good life (Fraser, 1995, 2003, 2007). Fraser et al. (1997) identified three different ways of 29 conceptualising welfare: i) that animals' should have the *freedom* to live naturally through the 30 development and use of their natural adaptations and capabilities, ii) that their physiological

31 and behavioural systems should *function* normally, including good health and growth, and iii) 32 that they should be free from prolonged and intense fear, pain, and other negative *feelings*, 33 and ideally able to experience pleasure. Views on the importance and contribution of these 34 three domains (*i.e.* natural living (freedom), physical health (functioning), and mental state 35 (feelings)) to welfare may be divided (Fraser, 1995, 2007; Fraser et al., 1997). A freedom proponent (i) might perceive welfare optimal when animals' have the opportunity to express 36 37 a full range of natural behaviours (e.g. pasture access for dairy cows), whereas to a 38 veterinarian or farmer, good health and biological functioning (ii) might be the hallmark of 39 animal well-being (e.g. high milk yield in dairy cows). However, a proponent of the feelings 40 approach (iii) might argue the former two are less relevant as long as animals perceive their 41 environment and state as positive or rewarding. In fact, some scientists believe that to be 42 concerned about welfare means to be concerned with suffering and pain (Dawkins, 1980, 43 2008) making it crucial for welfare science to investigate what situations animals perceive as 44 unpleasant, and therefore compromise welfare. More recently, it has been acknowledged 45 that not only the absence of negative experiences but also the experiences of positive 46 emotions is important to good welfare (Boissy et al., 2007; Yeates & Main, 2008).

Unfortunately, subjective experiences are not directly measurable. In humans, verbal report,
albeit an indirect measure, is the gold standard to infer conscious mental states (*i.e. feelings*).
Our (societal) concerns about animal welfare hinge on the possibility that non-human animals
might experience negative feelings that could affect their quality of life, but gaining evidence
of conscious emotional experiences remains problematic (Browning, 2022a; Mendl et al.,
2022). I discuss this issue in more detail in section 1.3.

53 Therefore, the objective assessment of animals' internal states has largely relied on behaviour 54 and physiological measures that address functioning and natural behaviour. Broom's 55 definition of welfare as "the state of the individual as regards its attempts to cope with its 56 environment" aligns with the functioning domain by focussing on measurable aspects that 57 allow us to infer an individual's health and physiological functioning (e.g. immune and brain 58 functioning (Broom, 1986, 1995, 1996)). However, failure to cope will result in reduced 59 mental and bodily functioning; hence subjective states are important to welfare (Broom, 60 1995). With the growing acceptance of the existence of animal emotions, whether these are consciously experienced or not (Bekoff, 2000; Boissy et al., 2007; Browning, 2019, 2022a), the 61

62 scientific evaluation of welfare has shifted towards an approach that prioritises the 63 assessment of emotional (affective) states (Mendl et al., 2009, 2010; Paul et al., 2005). 64 Although there is criticism that emotions cannot be measured directly, and are therefore not 65 available for objective assessment (Dawkins, 2008), scientists have proposed indirect methods by which animals are "asked" how they feel about their living conditions (Dawkins, 66 67 1990, 2008; Duncan, 2005). For instance, animals will push doors (Mason et al., 2001), or 68 perform operant responses (e.g. panel pressing (Søndergaard et al., 2011)), to gain access to 69 resources enriching their environments indicating that they are motivated to obtain highly 70 valued (i.e. rewarding) stimuli, while the lack thereof can result in suffering (Mason et al., 71 2001).

72 In this thesis, I consider animal welfare as a multifaceted subjective state which is determined 73 by an animal's perception of its own circumstances (*i.e.* comprising nutrition, physical 74 environment, health and behaviour), and which can be positioned on a continuous spectrum 75 between good and poor. Measures of welfare provide information about an animal's position 76 on the welfare continuum at that time (Broom, 1995, 1996). I regard the assessment of animal 77 emotions in conjunction with behaviour and physical functioning as indispensable for 78 measuring animal welfare scientifically since behavioural and physical aspects are 79 interrelated with subjective experiences in terms of causes of emotions, their function and components (explained in more detail below; Browning, 2019, 2022a). Mellor's (2017) 'Five 80 81 Domains model', describing the external and internal circumstances (*i.e. domains*) relevant 82 to animal welfare, represents the best theoretical framework for my comprehension of 83 animal welfare (see Figure 1.1). States associated with the first four domains (nutrition, 84 environment, health, behaviour) of the model accumulate into the fifth domain, the mental 85 (emotional) state and all domains combined contribute to animal welfare.



Figure 1.1. Schema of the Five Domains framework for assessing animal welfare. Schema after Mellor(2017).

89 1.2 Affective states

86

90 1.2.1 Terms and definitions

91 The terms *affective state* and *emotions* are often used interchangeably (Paul et al., 2005). 92 However, there is some diversity in the use of the term 'emotions' as researchers from different scientific fields have proposed a variety of definitions (*e.g.* reviewed in Anderson & 93 94 Adolphs, 2014; Kremer et al., 2020; Rolls, 2005). One way of explaining emotions, which I 95 consider most suitable for the purpose of this thesis, uses a functional approach defining 96 emotions as "adaptive processes that enable animals to avoid harm/punishment and seek valuable resources/reward" (Panksepp, 1994; Rolls, 1999, 2005). Accordingly, Mendl & Paul 97 (2020) propose an operational definition of rewards as stimuli/events which animals work to 98 99 access and punishers as stimuli/events which they work to avoid (see also Mendl et al., 2022 100 and Paul & Mendl, 2018). Presence of rewards or the absence of punishers, and associated 101 predictors thereof, elicit positive affective states. In contrast, the absence of rewards or the 102 presence of punishers, and associated predictors thereof, induce negative affective states 103 (Mendl & Paul, 2020).

The term 'affective state' can further be conceived as an umbrella label for both short-lasting transient emotions and longer-lasting mood states (Mendl et al., 2010; Paul et al., 2005). While emotions occur in immediate response to a stimulus (*e.g.* the sight of a predator elicits instantaneously the emotion fear), mood arises from the accumulation of emotions without the presence of a specific stimulus (*i.e.* mood states are not attached to external triggers;

Mendl et al., 2010). Moods can be understood as an integrative function of an individual's acute emotional experiences over time (Nettle & Bateson, 2012), or the 'running mean' of affect (Mendl et al., 2010). For example, in a high predator-risk environment, a negative mood state (*i.e.* anxiety) may arise from frequent experiences of the short-lasting negative emotion such as fear (see sections 1.2.2. and 1.2.3. for further details).

1 use affective states and emotions interchangeably in my thesis following the use of cited
authors, and specify in the text when differentiating between stimulus-driven short-term
emotions and longer-lasting mood states.

117 **1.2.2** Functions of affective states

118 Following the definition of emotions as intrinsic processes enabling animals to avoid 119 punishment and seek reward (Panksepp, 1994; Rolls, 1999, 2005), the main function of 120 emotional states can be seen as potentiating behaviours that are appropriate for 121 environmental challenges (de Waal, 2011; Mendl et al., 2010; Panksepp, 2011). Hence, 122 emotions coordinate the approach/avoidance of potentially rewarding/punishing stimuli and 123 increase chances of survival by prioritising the allocation of (energetic) resources towards the 124 currently most important fitness-relevant behaviours (Bateson, 2016). Mendl & Paul 2020 125 proposed a framework for integrating animal affect and decision-making. Broadly speaking, 126 there are two potential control mechanisms for decision-making called model-free (habitual) 127 or model-based (goal directed) control. The former refers to decision-making that is based on 128 past experiences (*i.e.* less flexible and responsive to changes in the environment). In contrast, 129 decision-making control in the latter is more flexible than habitual responses as the stimulus 130 value is flexible and rapidly adaptable to new circumstances. For instance, for a predator, food 131 X might become less valuable after having eaten a lot of food X, which influences the selected 132 action (*i.e.* approach/avoid more food X) in ambiguous situations (*e.g.* rustle in the grass could 133 be more food X). In addition to the feedback of the internal status (*i.e.* I am full of food X), 134 memory of the same situational circumstances (rustle previously predicted food X) and incoming sensory information (e.g. rustling sounds like food X) are integrated to make a 135 136 decision about an appropriate response action.

Furthermore, the accumulation of specific emotions can lead to shifts in cognitive processing resulting in specific response profiles (*i.e.* optimistic or pessimistic responses) based on the probability of expected outcomes, thereby changing individuals' threshold of sensitivity to

rewards or punishment (Nettle & Bateson, 2012). With regards to the previous example, living in a high-threat level environment, a negative bias would result in a greater likelihood of anticipating that the rustle in the grass is signalling a negative outcome (*e.g.* predator attack). Hence, a negative judgement would in this case increase the animal's chance of detecting actual, or potential, danger more quickly (Mendl et al., 2010). Consequently, these emotion-induced shifts in cognitive processing (*i.e.* cognitive biases) are important adaptations for survival.

147 Emotions can thus be regarded as 'sources of information' in the guidance of behaviour. 148 Moreover, this implies that the capacity to experience (negative) emotions has positive 149 survival value (Nettle & Bateson, 2012). It can be questioned why longer-lasting mood states 150 developed when short-term emotions are sufficient for coordinating adaptive behavioural 151 responses. Nettle & Bateson (2012) suggested that it is beneficial to carry over emotions from 152 one time to another, or over from one situation to the next, because environmental cues are 153 probabilistically associated with fitness-relevant information, i.e. involve a degree of 154 uncertainty. For an individual to respond appropriately to uncertain cues (e.g. whether the rustling in the grass signals the presence of prey, predators, or just wind), it has to estimate 155 156 the likelihood for the signal to indicate the fitness-relevance of the stimulus before mobilising 157 the appropriate behavioural response (*e.g.* approach or flight).

158 Two factors are important for an appropriate response; the probability of the event (*i.e.* how 159 likely is the rustling in the grass a sign of prey, not a predator), and the relative cost of the 160 possible outcomes (Bateson et al., 2011; Nettle & Bateson, 2012). If misjudging a truly 161 negative outcome is associated with death rather than losing a few calories, it would be of 162 importance for the individuals' survival to judge the noise as a potential punishment, and 163 even more so when an individual experienced frequent short-term emotional experiences 164 associated with punishment in the past. It would therefore be of evolutionary advantage if 165 the experience of an anxious mood state, resulting from frequent fear-inducing events, led to adjustments (i.e. lowering) of an individual's threshold of expecting future outcomes to be 166 167 punishing (and increase the threshold of expecting rewarding events) based on prior 168 experience. Hence, emotions not only have the function of evoking appropriate acute responses but also influence the adjustment of expectations about future experiences 169 170 (Bateson, 2016; Nettle & Bateson, 2012). Adjustments in anxiety-like responses can be

characterized by physiological, behavioural and cognitive changes preparing the organism for
detecting and dealing with imminent threats to survival (Bateson et al., 2011).

173 Chapter 2 addresses the difference between the normal adaptive fear response and its 174 dysfunction, resulting in (human) anxiety disorders with regard to the consequences of 175 cognitive processing associated with attention, which is the central topic of this thesis. Not 176 only negative but also positive emotions are useful in the adjustment of expectations of 177 ambiguous outcomes. Individuals experiencing rewarding events frequently may perceive 178 their environment more favourably (Mendl et al., 2010), which could, for instance, have 179 buffering effects against negative events (Reefmann et al., 2012). In social species such as the 180 modern domesticated horse, positive emotions associated with the presence of conspecifics 181 (e.g. social security) can attenuate stress response to an emotionally challenging situation 182 (e.g. confrontation with a fear-inducing object (Ricci-Bonot et al., 2021)).

183 **1.2.3 The two-dimensional framework**

184 Studies of animal emotions can be based on different theoretical frameworks. The basic 185 (discrete) emotion theory suggests that in mammals, discrete emotional states (e.g. fear, 186 rage, panic, pleasure) are conserved in specific hard-wired neural circuits (*i.e.* unconditioned 187 responses (Ekman, 2016; Panksepp, 1982, 2011)), which generate particular adaptive 188 emotional responses facilitating animal survival in life-challenging circumstances (Panksepp, 189 1982). For example, the 'fear system' coordinates emotion-specific reactions (i.e. fight or 190 flight response) under hypothalamic control to prevent actual, or potential, bodily harm. 191 However, Mendl et al. (2010) argued that the basic emotion theory lacks an overarching 192 structure of states that can incorporate a wide range of possible emotions. Proponents of the 193 'core affect' model (e.g. Posner et al., 2005; Russell, 2003), which build upon the two-194 dimensional framework, therefore suggest categorising emotions according to their valence 195 (*i.e.* pleasantness/unpleasantness continuum) and arousal (intensity) in a two-dimensional 196 space instead of using the discrete emotion categorisation (Posner et al., 2005; Russell, 2003). 197 Subjective experiences that can be described according to their valence and arousal are 198 referred to as core affect (Mendl et al., 2010; Russell, 2003).

However, since an understanding of the subjective emotional experiences presents the basis of core affect and the discrete emotion approach, these conceptual frameworks are only useful when private experience can be directly measured or reported (Mendl et al., 2010).

202 Neither option is available when considering animal emotions. Therefore, Mendl et al. (2010) 203 proposed an integrative two-dimensional framework combining both existing concepts to 204 distinguish animal emotions. A schema of this integrative framework is shown in Figure 1.2. 205 The vertical axis indicates arousal and the horizontal axis represents valence, the latter being 206 particularly relevant to animal welfare (Paul et al., 2005). Positively valenced affective states 207 are located in the right half of the framework (quadrants Q1 and Q2). Negatively valenced-208 affective states are positioned on the opposite side (quadrants Q3 and Q4). The framework 209 further provides a functional perspective on the adaptive value of different emotional states. 210 States of positive high arousal (located in Q1 in Figure 1.2) are associated with reward-211 seeking/appetitive motivational states. Exposing an animal to a reward will induce Q1 states 212 of positive-arousal. The behavioural, physiological and/or cognitive responses the animal 213 shows in presence of the reward can be inferred to be markers of Q1 states. In contrast, states 214 of negative low arousal promote low activity and conservation of energy (located in Q3), and 215 the emotional components in response to reward absence can be accessed via physiological, 216 behaviour and/or cognitive measures as markers of Q3 states. While affective states within 217 the Q1-Q3 axis are associated with acquiring fitness-enhancing rewards, i.e. are part of the 218 reward acquisition system (RAS, see Figure 1.2), states located in Q4-Q2 are associated with the punishment-avoidance system (PAS). Any emotion can be distinguished based on its 219 220 combination of valence and arousal (Posner et al., 2005). For instance, fear is associated with 221 negative valence and high arousal, while relaxation is associated with positive valence and 222 low arousal (see Figure 1.2). Hence, positioning emotional experiences in the core affect 223 space provides a 'common currency' to describe subjective states (Cabanac, 1992; Mendl et 224 al., 2010; Spruijt et al., 2001), and the integrative framework by Mendl et al. (2010) offers a 225 predictive structure that aims to help to outline these.

When considering animal welfare, identifying the direction of affective valence is of particular
importance, because the experience of negative emotional states has moral implications
(Paul et al., 2005; Webb et al., 2019).



229

Figure 1.2. The integrative framework to conceptualise animal emotions in terms of their arousal and valence after Mendl et al. (2010). Emotional states can be distinguished based on their function as part of the reward acquisition system (RAS; states along the Q1-Q3 axis) or punishment avoidance system (PAS; states along the Q2-Q4 axis). The model provides an empirical way of specifying an animal's location in core affect space by presenting animals with known punishers/rewards generating particular states. Measuring physiological, behavioural and/or cognitive responses in animals in these states can be used as markers of affect.

237 The causal relationship between emotions and the environment is still unclear (Mendl et al., 238 2010). Following the functional definition, affective states emerge from the environment as 239 a product of imminent rewarding/punishing stimuli (short-term emotions) and the 240 accumulation thereof (longer-term mood states). On the other hand, core affect may influence how the environment is perceived. Mendl et al. (2010) suggested that both might 241 242 be true, meaning that short-term (discrete) emotions and longer mood states have a 243 bidirectional relationship. Position in the core affect space, therefore, may represent a 244 common currency, enabling decision-making under trade-off between discrete emotions (e.g. 245 fear) and sensations/motivations (e.g. hunger) arising from physiological needs related to 246 homeostasis. For instance, in a fearful state, an animal may weigh out its chances for 247 searching for food safely differently than in a non-fearful state.

As previously explained, the accumulation of positively valenced emotions can result in positive mood states (*i.e.* 'happiness') and greater expectation of rewarding events (*i.e.* optimistic bias), while frequent negative emotions may lead to shifts of comparatively more

negative mood states and pessimistic bias (Mendl et al., 2009). By assessing the environmental conditions animals are provided with and how these are perceived by the animals (*i.e.* rewarding/punishing), the valence of the underlying mood state can be inferred. However, to determine the position of an animal's affective state within the framework, an objective assessment of the emotional components accessible to quantitative measurement is required.

257 1.3 Overview of components of animal emotions, their assessment and 258 limitations

259 Affective states comprise behavioural, physiological, and cognitive components, and at least 260 in humans, conscious subjective experiences (Frijda, 1988; Paul et al., 2005; Scherer, 1984). 261 As suffering is a major concern for animal welfare, assessing affective valence is important 262 (Mendl et al., 2009; Mendl & Paul, 2004). However, it is impossible to directly measure the 263 subjective component of emotions, in either humans or animals. Verbal self-report is the 'gold 264 standard' indicator of felt affective states in humans, who can verbally describe their 265 subjective experiences, but this approach is unsuitable for non-verbal beings (e.g. infants, 266 non-speaking autistic people) and animals (Paul et al., 2005). Therefore, there is presently no 267 definite answer to whether animals *feel* (*i.e.* consciously experience) anything at all, and there 268 may never be (Dawkins, 2015, 2017; Weary et al., 2017). Moreover, if we knew animals do 269 have feelings, it further remains unknown what kind of experience they have, *i.e. how* they 270 feel (Paul et al., 2020). Nevertheless, given the considerable anatomical, behavioural and 271 physiological similarities between humans and animals, it is very likely that non-human 272 animals, at least some, may also experience emotions consciously (Low, 2012). Therefore, it 273 has been argued that the exploration of affective states in animals, independently of whether 274 they may consciously experience emotions or not, is a crucial pursuit in animal welfare 275 research (Browning, 2019; Paul et al., 2005, 2020).

276 1.3.1 Behavioural components

An essential step for using animal behaviour as an indicator of emotional state is to verify which behavioural measures reliably indicate positive and negative affective states. At the simplest experimental level, this can be achieved by observing animals' spontaneous behaviours, *e.g.* in relation to their social and living conditions, and making an assumption about animals' affective state (Mendl et al., 2022). For instance, behaviours such as freezing,

282 aggression, or attack, are generally accepted indicators of negative affect, as these are 283 responses provoked by stimuli associated with fitness-threatening events (e.g. fear in presence of a predator, aggression during food competition, pain induced by thermal damage 284 285 (Mendl et al., 2010; Paul et al., 2005)). Conversely, behaviours that only occur in absence of 286 fitness-challenging stimuli (i.e. activities which reduce when energy resources are low (Mandel et al., 2013)), such as affiliative social interactions and animal play, are conceived to 287 288 reflect positively valenced affective states (e.g. pleasure associated with play (Boissy et al., 289 2007; Held & Špinka, 2011)). However, interpreting behaviours as exclusively positive or 290 negative is not always straightforward. Play, for example, reduces under poor welfare 291 conditions (*i.e.* during periods of hunger or predation), but there are counter-examples where 292 the opposite relationship can be observed (Ahloy-Dallaire et al., 2018). In cats, restricting 293 maternal care (*i.e.* maternal separation, rationed feeding) enhances object play in kittens 294 (Bateson et al., 1990; Bateson & Young, 1981). In horses, higher levels of adult social play are 295 positively correlated with the level of chronic stress (Hausberger et al., 2012). This lack of 296 specificity limits the use of behaviours as indicators of emotional state (Mason & Mendl, 297 1993), and more so when the affective significance of behaviour is dependent on the 298 expectation of observers and their interpretation of the situational context (Tuyttens et al., 299 2014). Another challenge in interpreting animal behaviour may be difficulties in judging at 300 which level of behaviour change welfare might be at risk (Barnett & Hemsworth, 1990).

301 Testing animals' responses to distinct stimuli likely to elicit particular emotions might 302 eliminate some specificity issues. Behavioural measures explicit to negative emotional states 303 might be gained from exposing animals to 'punishers' (*i.e.* unpleasant stimuli/events inducing 304 negative affective states such as hunger or pain) and measuring the resulting response 305 (Mason & Mendl, 1993). For instance, studies on fear (i.e. the emotional response to the 306 perception of actual or potential danger (Boissy, 1995)) have used a variety of experimental 307 designs in which (farm) animals' are deliberately exposed to fear-inducing stimuli (reviewed 308 in Forkman et al., 2007). The novel object test investigates animals' reactions to an unfamiliar 309 object, whereby avoidance of the item is indicative of fear, while approach and exploration 310 suggest a positive emotional response. However, how novel stimuli are evaluated is also 311 dependent on the animals' needs and expectations (i.e. based on previous experience, and 312 predictability of the stimulus) as well as their coping possibilities within their environment

(Desire et al., 2002). Desire et al. (2002) pointed out that predictability and controllability of 313 314 negative stimuli influence animals' responses to these events. Predicted occurrences of 315 shocks are less likely to cause ulcers in rats than unpredicted shocks (Weiss, 1972). 316 Furthermore, the directionality of behavioural changes associated with emotions can be 317 inconsistent (Mason & Mendl, 1993). For instance, an increased locomotor activity could be related to escape or exploratory motivation, hence be associated with different affective 318 319 states (Archer, 1973). On the other hand, contrasting changes in behaviours have been 320 reported in putatively same-valenced situations. Exploration behaviour can increase or 321 decrease with increasing fear (Archer, 1973). Behaviours that can be linked to both negatively 322 and positively reinforcing situations are likely to be associated with arousal rather than 323 valence (Barnett & Hemsworth, 1990; Rushen, 1991).

324 Facial expressions are behavioural signals that can be linked to specific emotions (Darwin, 325 1965; Ekman, 2003). For example, pain has been associated with particular facial patterns (i.e. 326 'pain-face') and species-specific grimace scales helping to assess pain levels based on facial 327 expressions have been developed in a range of animals (e.g. mice (Mus musculus) - Langford 328 et al., 2010; sheep (Ovis aries) - McLennan et al., 2016; pigs (Sus scrofa domesticus) - Di 329 Giminiani et al., 2016; horses - Dalla Costa et al., 2014a). However, facial expressions are also 330 of importance for social communication and therefore may be influenced by the presence or 331 absence of an 'audience' (i.e. other conspecifics). Hence, some facial expressions may have a 332 communicative function rather than a one-to-one correspondence with the actual emotional 333 state, which might reduce the reliability of these expressions as a measure of affect (Paul et 334 al., 2005). For instance, in presence of adult horses, foals 'snap' (perform rapid up-and-down 335 movements of the mandibles) which is thought to signal a submissive social status (Crowell-336 Davis et al., 1985), hence might not directly reflect emotional state. Moreover, expressions 337 with strong communicative functions (such as snapping) are unlikely to occur in absence of 338 an audience (*i.e.* dominant horse).

For the purpose of this thesis, I applied behavioural measures in support of the interpretation of cognitive measures (*i.e.* attention and attention disengagement) described in Chapter 2. In the general methodological section, I further detail how I intended to use the rating of facial expressions (*i.e.* horse grimace scale) as a potential marker of emotional states in horses, and aimed to evaluate horse welfare based on activity budgets derived on behaviour

observations. The data of these approaches were, however, not considered for furtheranalysis for reasons explained in Chapter 3.

346 1.3.2 Physiological components

347 Physiological indicators linked to animal emotions traditionally derive from changes observed 348 in states of physiological stress likely to be associated with negative affect as stress is 349 generally observed in aversive situations (Paul et al., 2005). The main approach to determine 350 stress relies on measures of the activity of the hypothalamic-pituitary-adrenal (HPA) axis, in 351 particular by assessing levels of 'stress hormones', such as vasopressin, adrenocorticotrophic 352 hormone (ACTH), or glucocorticoids (e.g. cortisol (Selye, 1976)). For example, stress during 353 transport increases plasma cortisol levels in horses (e.g. Fazio & Ferlazzo, 2003; Schmidt et 354 al., 2010), with longer transport durations leading to higher cortisol levels (Fazio & Ferlazzo, 355 2003). However, using endocrine parameters can have methodological limitations when 356 invasive sampling methods are required (*e.g.* blood cortisol), which can confound results by 357 altering animals' emotional state during the sampling procedure (Broom & Johnson, 1993). 358 Regarding cortisol, collecting samples from urine, faeces or hair can be a non-invasive 359 alternative. Mason et al. (2001) found that urine cortisol levels were elevated in mink 360 (*Neovison vison*) denied access to a swimming pool indicating that preventing these animals 361 from swimming causes stress. Cortisol levels in these animals were similar to those deprived 362 of food revealing that inhibiting mink from strongly motivating activities (swimming), may 363 result in an emotional state similar to frustration (Mason et al., 2001). However, not only 364 aversive situations but also positive situations (e.g. courtship, mating, prey catching) 365 temporarily increase HPA activity (Broom, 2001). Hence, changes in HPA activity are more 366 likely to reflect the activity of coping systems (*i.e.* physiological processes that respond to 367 environmental challenges (Broom, 2001)), for example, to prepare animals for action (be it 368 catching prey or avoiding predation) by mobilising energy, and therefore reflect emotional 369 arousal rather than valence.

Assessing autonomic nervous system (ANS) activity is another way to measure emotions physiologically. The ANS involuntarily regulates physiological processes for bodily homeostasis and is a key component in emotion regulation (Kreibig, 2010). ANS activity can be measured non-invasively by assessing its control of cardiac activity. The heart receives innervation from two branches of the ANS, the sympathetic (SNS) and the parasympathetic

375 nervous system (PNS). Fluctuations in time intervals between successive heartbeats (inter-376 beat intervals, IBIs) evident as heart rate variability (HRV (Billman et al., 2015)) is the result of 377 the antagonistic influences of the PNS and SNS on cardiac activity. Higher HRV is associated 378 with good health and greater ANS flexibility which enhances the body's ability to respond to 379 stress (Friedman & Thayer, 1998; Thayer et al., 1997). In contrast, reduced HRV caused by a shift towards SNS dominance has been associated with a range of clinical diseases (e.g. 380 381 cardiovascular diseases, diabetes) as well as psychological stress, in both humans and farm 382 animals (e.g. cattle (Bos taurus), pigs, horses (von Borell et al., 2007)).

Far fewer studies have investigated the effect of positive situations on cardiac activity in animals. In cows, talking in a gentle voice and neck stroking appears to have lowering effects on heart rate and increases HRV suggesting positive (relaxing) effects of these types of human-animal interactions (Lange et al., 2021). Similarly, brushing increases HRV in sheep (Tamioso et al., 2018) and feeding results in higher HRV in pigs (Zebunke et al., 2011).

388 Emotion-induced autonomic responses can further be assessed *via* infrared thermography 389 (IRT; reviewed in Stewart et al., 2005; Travain & Valsecchi, 2021). This approach is based on 390 the measurement of radiated electromagnetic energy (*i.e.* photons of $3-12 \mu m$ wavelength) 391 to detect surface temperature changes, which are controlled by autonomic (SNS) activity 392 (Stewart et al., 2005). IRT recordings are non-invasive and have furthermore the advantage 393 that emotion-induced changes can be recorded without, or little, animal handling and in real-394 time (Stewart et al., 2005). For instance, Stewart et al. (2010a) found that castrating calves 395 led to a drop in eye temperature immediately after the procedure (<2 min). This was likely 396 due to peripheral vasoconstriction as part of the sympathetic response to pain (Stewart et al., 397 2010) causing stress-induced hyperthermia (SIH), a phenomenon in which blood flow to 398 internal organs and the skeletal musculature is prioritised over peripheral areas (e.g. skin, 399 eyes) causing core temperature to rise and surface temperature in peripheral areas to 400 decrease (Vianna & Carrive, 2005). Pharmacologically induced sympathetic activation with 401 epinephrine resulted in a similar drop in eye temperature in calves proving that eye 402 temperature is mediated by the SNS (Stewart et al., 2010b). However, in horses, eye IRT 403 results and cortisol (blood and saliva samples) were found to be significantly positively 404 correlated (*i.e.* higher cortisol was associated with higher eye temperatures) following show

jumping competitions further suggesting that changes in eye temperature might also belinked to the activation of the HPA axis (Valera et al., 2012).

In addition to physiological stress, psychological stress (*i.e.* stress not caused by actual (or potential) tissue damage or physical exertion as was the case in the previous two examples) has been reported to cause changes in surface temperature changes in a variety of animals. In pheasants (*Phasianus colchicus*), head temperature dropped immediately prior to an aggressive encounter with another conspecific and increased rapidly after the attack indicating that mental stress associated with social conflict can cause temperature variations measurable *via* IRT (Knoch et al., 2022).

414 Positive emotions might also induce physiological changes reflected in surface temperature 415 changes. For instance, a significant rise in eye temperature was found in dogs (Canis lupus 416 familiaris) whilst receiving food (Travain et al., 2016). In chickens (Gallus gallus domesticus), 417 positive (e.g. food (Moe et al., 2012)), as well as negative situations (e.g. airpuff directed at 418 hen or its chicks - Edgar et al., 2011; negative handling - Edgar et al., 2013; Herborn et al., 2015), can result in a drop in comb temperature. Hence, IRT results are generally thought to 419 420 detect changes in states of emotional arousal rather than valence, unless thermal proxies are 421 combined with other physiological or behavioural indicators (Travain & Valsecchi, 2021).

422 Preliminary IRT findings observed in chickens and cows recently suggested that asymmetries 423 in bilateral facial features (e.g. eyes, ears, nostrils) might be linked to emotional valence 424 (Telkanranta et al., 2019). According to the emotional lateralisation theory, positive emotions 425 are mainly processed in the left hemisphere, while the right hemisphere is dominant in the 426 processing of negative emotions (Davidson, 1992; Davidson et al., 1990). This asymmetric 427 hemispheric activation can be observed in a range of animal species (reviewed in Leliveld et 428 al., 2013) as lateralised motoric responses (e.g. asymmetric tail wagging in dogs (Quaranta et 429 al., 2007)), lateralised ear movement in sheep (Reefmann et al., 2009), or lateralised sensory 430 processing (e.g. preferential eye or nostril use in response to emotive stimuli found in cows 431 (Phillips et al., 2015), horses (De Boyer Des Roches et al., 2008; Larose et al., 2006), and 432 chickens (Vallortigara et al., 2001)). Consistent with right hemisphere processing due to the 433 cross-over of neurons from the left eye to the right hemisphere, a left eye preference was 434 found in cattle (Phillips et al., 2015) and horses (De Boyer Des Roches et al., 2008; Larose et 435 al., 2006) exposed to fear-inducing situations (e.g. novel objects, people or situations).

436 Lateralised sensory processing is thus believed to be a useful indicator of emotional valence 437 (Leliveld et al., 2013), although the modulation of asymmetric cerebral activity by emotional 438 state is still largely unexplored. One study investigated the links between eye preference 439 when viewing novel objects and underlying affect used dairy cows (putatively in negative 440 affective state associated with lameness), but the findings were inconclusive (Kappel et al., 2017). Whether emotion-induced hemispheric lateralisation is assessable as asymmetric 441 442 temperature profiles in facial areas such as eyes or nostrils is largely unexplored. The initial 443 evidence in chickens (Telkanranta et al., 2019) and calves (Ramirez Montes De Oca, 2021) 444 supports the theory that emotional valence may be inferred from asymmetric facial 445 temperature profiles. For instance, temperature changes in the inner corner of the eye 446 (medial canthus) might be associated with lateralised hemispheric activity due to increased 447 vascular activity mediated by the SNS and contralateral brain hemisphere. In horses, 80-90% 448 of the optical nerve axons decussate in the optic chiasm from the left eye to the right 449 hemisphere and vice versa (Herron et al., 1978; Schmidt et al., 2019). Hence, higher 450 temperatures of the left than the right medial canthus might be the result of right hemispheric 451 dominance (Ramirez Montes De Oca, 2021). If no laterality effects occur, eye temperatures 452 might still increase during emotionally arousing situations as a reflection of an increase in ANS 453 activity, hence changes in HRV and eye temperature might be related.

454 In this thesis, I therefore investigated whether HRV and IRT can be used as measures of 455 emotional valence in horses, which is described in Chapter 6.

456 1.3.3 Cognitive components

457 Finally, emotions comprise cognitive components and their assessment can provide valuable 458 insights into emotional states in animals (Paul et al., 2005). As described previously, emotions 459 influence, and can be influenced by, cognitive processes, in particular cerebral functions 460 associated with interpretation, decision-making, memory, or attention (reviewed in Paul et 461 al., 2005). In humans, negative affect disorders (e.g. depression, anxiety) lead to pessimistic 462 judgement of ambiguous information, negative expectations of future events (MacLeod & 463 Byrne, 1996; Mathews & MacLeod, 1994), heightened attention to threat-relevant information (Bar-Haim et al., 2007), and preferential recall of negative memories (Denny & 464 465 Hunt, 1992; Watkins et al., 1992). Harding et al. (2004) were the first to show that rats kept in housing-induced chronic mild stress were more likely to respond to ambiguous cues (i.e. 466

467 auditory cues intermediate to two conditioned stimulus tones associated with reward (S+) or 468 punishment (S-)) as if these represented S-, compared to rats kept in predictable housing. 469 These findings provided some initial support for the hypothesis that long-term negative affect 470 (mood) associated with unpredictable housing conditions can induce a pessimistic judgement 471 bias in animals, similar to mood-congruent biases found in humans. Since this seminal study, 472 the judgement bias test (JBT) has been adapted for a wide range of animal species (reviewed 473 in Baciadonna & McElligott, 2015; Gygax, 2014). Study findings have largely shown that JBT 474 can be applied to measure animal emotions and that positive affect (induced e.g. through 475 comparatively better welfare conditions or pharmacological manipulations) results in a more 476 positive interpretation of uncertainty while negative affective leads to pessimistic judgement 477 of ambiguous information (reviewed in Lagisz et al., 2020; Neville et al., 2020), with no 478 significant evidence of publication biases in the literature as explored in Lagisz and colleagues' 479 meta-analysis and systematic review. Nonetheless, reported JBT findings do not always 480 support this prediction as some studies have generated the opposite or null results (reviewed 481 in Mendl et al., 2009; Gygax 2014; Lagisz et al., 2020). For instance, brief owner absence likely 482 to induce separation-related anxiety in pet dogs did not lead to pessimistic judgment bias 483 (Müller et al., 2012). In sheep, the release from restraint and isolation stress resulted in these 484 animals responding more optimistically instead of more pessimistically in a JBT (Doyle et al., 485 2010). The authors proposed contrast effects (*i.e.* treated sheep did not perceive the dog 486 presented as negative stimulus during testing as aversive as the control sheep) or 'relief' 487 effects following the restraint and isolation inducing positive emotions in the treated as 488 possible explanations for observing findings opposite to their initial predictions. On the other 489 hand, animals' JBT responses might not be exclusively caused by underlying affective state. 490 Chronic food restriction likely to cause the negative emotional state associated with hunger 491 tended to induce optimistic bias in sheep (Verbeek et al., 2014), possibly as a reflection of 492 increased motivation to find food rather than shifts in emotional state. Differences between 493 responses of male and female rats (Barker et al. 2016) and goats (Briefer & McElligott, 2013) 494 further suggest that sex also influences JBT outcomes. Moreover, personality trait has been 495 reported to be a predictor of judgment bias (e.g. in dogs - Barnard et al., 2018; pigs - Asher et al., 2016). Therefore, the role of other biological modulators (e.g. personality, social status, 496

497 sex) need to be considered as potentially confounding factors when testing animals'498 responses in JBT (Whittaker & Barker, 2020).

499 The application of JBTs has certain practical limitations. For instance, time-consuming training 500 periods to condition the positive and negative stimuli, selection biases of test subjects (*i.e.* 501 animals in negative affect are less likely to meet learning criterion, Mendl et al., 2010), 502 practical restrictions for on-farm testing, and learning associated with repeated exposure to 503 ambiguous test cues (*i.e.* animals learn that novel cues are not reinforced), further reduce the 504 feasibility of the judgment bias paradigm (Crump et al., 2018). Developing new cognitive tests 505 that rely on no, or little, training that are suitable for a large number of animals and practically 506 applicable within a farm setting are therefore needed. Measuring cognitive processes 507 involving attention to threat have been suggested as an alternative test of emotional state, 508 since this would allow the testing of emotion-modulated innate mechanisms and therefore 509 eliminate training requirements (Bethell, 2009; Crump et al., 2018). In Chapter 2, I provide an 510 in-depth overview of attentional mechanisms, their assessment, and practical approaches to 511 measure emotion-driven biases in attention in humans and animals, before describing the 512 development of an attention bias test in horses in Chapter 4.

513 As mentioned in section 1.3.2, lateralised brain activity during emotional processing has been 514 described in humans (Davidson, 1995) and other vertebrates (reviewed in Leliveld et al., 2013; 515 Siniscalchi et al., 2021) suggesting that the direction of emotional valence (i.e. negative-516 positive) might be inferred from hemispheric dominance. Although both hemispheres are 517 involved in emotional processing, right-hemispheric dominance has been linked to the 518 processing of negatively valence emotions (e.g. fear, aggression), while positively valenced 519 emotions (e.g. emotions associated with food) induce left-hemispheric dominance (Silberman 520 & Weingartner, 1986; Leliveld et al., 2013). One approach to evaluate affective valence in 521 animals through cerebral lateralisation is by presenting animals with emotional stimuli (e.g. 522 unfamiliar objects are likely to induce fear, Forkman et al., 2007) to observe asymmetric 523 behavioural responses reflecting prevalent activation of the left or right hemisphere 524 (Siniscalchi et al., 2010). For instance, a eye side bias when viewing unfamiliar (threatening) 525 stimuli has been described in cows (Robins & Phillips, 2010) and horses (Larose et al., 2008) as a reflection of right hemisphere dominance, given the high degree of decussation in optic 526 527 neurons (nerves from the eyes projecting to the contra-lateral side of the brain resulting in

visual information from the left eye dominantly processed in the right hemisphere and vice versa, Herron et al., 1978). In chapter 6, I explore whether lateralised brain activity is also reflected in asymmetric facial temperature profiles in horses supposedly experiencing differently valenced emotional states, hence whether differences in temperatures across bilateral facial areas could be used as a novel proxy of emotional valence in this species as previous work on cows and chickens has suggested (see 1.3.2).

534

535 **1.3.4** The domestic horse as model species for affective studies

536 The domestic horse is a special case with regard to their roles in human-animal relationships 537 (e.g. companion, working animal, sport competitors), which ultimately affects how these 538 animals are cared for by humans (Hockenhull et al., 2010). Before the industrial revolution, 539 horses were mainly kept as work animals for agricultural purposes (Harris, 1999). In today's 540 Western countries, horses are mostly considered companion animals. There are about 847 541 000 horses in the UK (as of 2019, British Equestrian Trade Association), used in equine 542 activities such as sport competitions (e.g. show jumping, eventing, and dressage), racing, or 543 leisure. The housing requirements for horses have changed little with domestication, yet they 544 are predominantly kept, similar to other farm animals, in intensive conditions imposing 545 restrictions to their ethological needs (e.g. social, physical, access to roughage) presenting 546 risks to animal welfare (Hockenhull & Creighton, 2014a, 2015; Minero & Canali, 2009; Waran, 547 2007). In the following, I describe species-specific needs and conflicting management-related 548 constraints of modern horses to illustrate that this species presents an important and valid 549 model for the study of animal welfare and cognition.

550 **1.3.4.1** The behavioural ecology of free-roaming horses

551 The domestic horse is a member of the family Equidae and a subspecies of the wild horse 552 (Equus ferus) which evolved as grazing, group-living animals adapted to unpredictable 553 environmental conditions, prompting migration, and enabling constant, but regularly 554 changing, forage supply (Goodwin, 2002). That is, horses have developed a degree of 555 behavioural flexibility to adapt to changing living conditions, a capacity that has played an 556 important role in their adjustment to confinement during domestication (Waran, 2007, p. 3). 557 However, the modern horse still shares behavioural and physiological traits with its pre-558 domestication ancestors (Waring, 2003), as evidenced by the fact that this species can still

adapt to free-roaming living conditions. For instance, Dartmoor, Exmoor, or New Forest
ponies (Tyler, 1972), are feral horse breeds that reverted to live in wild conditions proving
that certain survival traits have never been modified by domestication.

562 Unmanaged, horses form stable social groups (bands), often of less than 10 animals per band 563 (Waring, 2003). Bands maintain a home range with territorial boundaries of $0.8 - 78 \text{km}^2$ 564 (Waring, 2003, p. 236) and individuals may move up to 25-80 k per day (McGreevy, 2012). 565 Social groups can be of different organisations, with family bands comprising one stallion, 566 multiple adult females and their offspring, bachelor bands of multiple young males, or female 567 unisex groups (Waring, 2003). Adult band members form stable social hierarchies, which are 568 matriarchal in the absence of a stallion (Waran, 2007). Juveniles (males and females) tend to 569 disperse from their natal bands by the age of 2-3 years. Social dominance is displayed in a 570 variety of activities (e.g. feeding, drinking, selecting resting sites), although group activities 571 can be initiated by any member of the group (Waring, 2003). As a prey species, group 572 coherence is important for anti-predator behaviour under wild conditions, hence group 573 stability is fostered to ensure survival (Goodwin, 2002). To maintain stable relationships, 574 horses express a range of social interactions fostering attachment, including affiliative behaviours such as allogrooming, social play (mostly in foals), or spending time in close 575 576 proximity (Costa et al., 2019). Aggressive encounters are rarely seen among established band 577 members, but group integrity is defended against intruders which can elicit aggression (Fureix 578 et al., 2012).

Horses' sensory abilities facilitate early predator detection, *e.g.* near-panoramic and scotopic vision allows quick reactions to potential threats, triggering rapid flight responses. Visual communication is the primary form of social interaction in horses, which is highly adapted to perceive subtle changes in facial expression and body language of conspecifics and serves, along with social cohesion, as social signals in response to a potential predator attack (Goodwin, 2002).

Horses feed up to 16 hours per day and as preferential grazers, they eat a wide range of grasses, shrubs and trees (Goodwin, 2002). Their diet is affected by seasonal changes. For instance, during summer, the diet of New Forest ponies comprises mainly grasses (60%), while this drops to 30% in the winter, and gorse, heather, shrubs and holly form most of their diet

while the animals forage in more sheltered habitats such as woodland (Hansen, 1976; Putman
et al., 1987; Tyler, 1972).

591 These naturally evolved behaviours enable horses to live in unpredictable and changing 592 environmental conditions that are very different from the settings of intensive management 593 systems in which most horses are kept today.

594 **1.3.4.2** The "gilded cage": management-related risks to horse welfare

595 Many horse owners believe that their animals live 'like kings' as they are provided with 596 purpose-built spaces of 'comfort', shielding them from any harm or distress, receive a 597 balanced diet, and are exercised regularly (Hockenhull & Furtado, 2021). Hence, horses 598 apparently require human intervention to live a good life (Hockenhull & Furtado, 2021). Horse 599 owners often do not recognise that the repercussions of this apparent luxurious life can have 600 consequences for animal welfare when housing conditions do not align with horses' 601 ethological needs (Fletcher et al., 2021; Horseman et al., 2016). For instance, animals may 602 experience frustration when prevented from performing highly motivated behaviour patterns 603 (Goodwin, 2002; Hockenhull & Creighton, 2014a, 2014b). As a result, abnormal (repetitive) 604 behaviours (e.g. redirected behaviour activities such as wood-chewing and bed eating; (P. 605 McGreevy et al., 1995), aggression (Hockenhull & Creighton, 2015), and stereotypies (i.e. 606 frustration-induced repetitive behaviours, associated with repeated attempts to cope and/or 607 CNS dysfunction (Mason, 2006)) are highly prevalent in stabled horses. A previous survey 608 reported prevalences of abnormal behaviour for dressage, eventing, and endurance horses 609 of 32.5, 30.8 and 19.5%, respectively (McGreevy et al., 1995a). In a more recent survey on 610 leisure horses, 82% of horses were reported to express at least one form of handling problem 611 (Hockenhull & Creighton, 2015).

612 Most UK leisure horses (approx. 75%) are stabled for at least 9h per day, and some (5%) for 613 up to 21-24h per day (Hockenhull & Creighton, 2015; Wylie et al., 2013). Given the trend away 614 from winter-only towards all-year around indoor housing (Anon-BETA, 2006), the time horses 615 spend indoors is likely to be even higher today. Globally, constant single stall housing (i.e. 616 >20h/day confinement) remains the most common form of housing for sport and riding 617 school horses (Lesimple et al., 2016, 2019), often varying in design and opportunity for social 618 contact. As a gregarious species, horses are highly motivated to interact with conspecifics 619 (Søndergaard et al., 2011). Animals will work harder to access a paddock with a conspecific

than accessing an empty paddock (Lee et al., 2011) and perform more conditioned panel presses enabling them varying degrees of access to a social partner (*i.e.* full, head or muzzle contact) than access to an empty area (Søndergaard et al., 2011). Social isolation has consequences for physiological and psychological well-being. For example, horses may not lie down due to a lack of social security resulting in a lack of REM sleep (only possible in a recumbent position), which increases the risk of drowsiness (Fraser, 2009) and can lead to episodes of sudden collapse (Williams et al., 2008).

627 Some studies have suggested that visual and tactile contact will alleviate the negative effects 628 of social isolation (e.g. expression of abnormal behaviours (Cooper et al., 2000)) and horses 629 do, in fact, work equally hard for social access, regardless of whether the interaction allows 630 head, muzzle, or full social contact (Søndergaard et al., 2011). However, there is also evidence 631 that it may be frustrating for horses to perceive conspecifics without being able to engage in 632 full physical contact (Hockenhull & Creighton, 2014a), indicating that full access to 633 conspecifics is overall more beneficial to psychological well-being. Moreover, a change from 634 full to restricted social access can be detrimental to horse welfare. Visser et al. (2008) found 635 that moving group-housed horses to individual stables induced stress-related behaviours (e.g. 636 neighing, pawing, snorting) and 67% of individually housed animals performed one or more 637 stereotypic behaviours (STs) 12 weeks later.

638 Opportunity for social interactions with conspecifics is furthermore important to horses 639 establishing stable social bonds and performance of affiliative behaviours (Christensen et al., 640 2002). Free-ranging horses frequently perform mutual grooming between pairs, e.g. 641 reciprocal nuzzling of the withers (Wolter et al., 2018). Imitating this behaviour through 642 manual massage has been found to have a positive (i.e. lowering) effect on heart rate 643 indicating that this type of interaction is beneficial to the functioning of the peripheral 644 nervous system (McBride et al., 2004). However, stable designs prevent full-body contact 645 between horses. Furthermore, the interruption of established bonds when horses move 646 between yards, stall neighbours change, or new animals are introduced to stable groups, can 647 have negative consequences for horses and trigger aggression (Fureix et al., 2012; Hockenhull 648 & Creighton, 2015). Under naturalistic conditions, aggression is low in pastured horses 649 (Sigurjónsdóttir & Haraldsson, 2019). Aggression between horses forced to stay close 650 together in adjacent stalls may occur if animals do not choose or have a bond with their

neighbours. In this case, a lack of solid walls can mean that there is no escape from unwantedor agonistic interactions (Hockenhull & Furtado, 2021).

653 Stabling also imposes access restrictions to other resources, including food. Under natural 654 conditions, horses spend the majority (over 60 %) of their time grazing (Goodwin, 2002). In 655 modern practices, however, food is often provided in small hard feed rations (*i.e.* high energy 656 concentrate), one or multiple times per day alongside limited amounts of roughage (hay). 657 'Meal'-feeding, provision of concentrate, and low hay diets have been associated with a high 658 prevalence of gastric ulcers indicating that the feed is not suitable for the equine digestive 659 system (McGreevy et al., 1995b; Murray et al., 1996). In a survey of 201 horses used for sport 660 and leisure 83.5% of animals were affected by gastric ulcerations of varying degrees of clinical 661 severity (*i.e.* having 1-2 localised lesions (Luthersson et al., 2009)). Pain associated with ulcers 662 can trigger the development of oral STs (reviewed in Waran, 2002), although this form of 663 behavioural response might be to alleviate pain by increasing saliva production which buffers 664 stomach acidity (Nicol et al., 2002). In addition, prolonged phases of food unavailability may 665 further cause frustration when the motivation to feed is not fulfilled, which can also lead to 666 abnormal behaviour and STs. McGreevy et al. (1995b) found that thoroughbred horses fed a 667 limited amount of forage (<6.8 kg of hay) per day had a higher risk of performing abnormal 668 repetitive behaviours (i.e. weaving, a repetitive left-right swaying movement involving the 669 front feet and head) and wood-chewing (i.e. parts of wooden stable features are chewed and 670 possibly ingested). In contrast, constant availability of high-fibre roughage may reduce these 671 behaviours as animals have less time to perform these behaviours, by decreasing the 672 motivation to eat through increased gut-fill, or other feedback mechanisms associated with 673 (Cooper & Mason, 1998; McGreevy et al., 1995b, 1995c). Offering varied forage types further 674 reduces abnormal behaviours (McGreevy et al., 1995b; Ninomiya et al., 2004). Horses prefer 675 straw compared to other bedding types (e.g. shavings, rubber mats (Baumgartner et al., 2020; 676 Mills et al., 2000)) and the use of bedding types other than straw has been linked to increased 677 weaving (locomotor ST (McGreevy et al., 1995b)). It is suggested that digesting straw might 678 ameliorate behaviour problems associated with low-fibre diets (McGreevy et al., 1995b), 679 although straw eating can result in intestinal impaction and colic (Hayes (1978) in McGreevy 680 1995b). Horse owners perceive routine as important to their horses and see predictability as 681 central to keeping their horses calm and relaxed (Furtado et al., 2021a). However, heightened

anticipation of food during established feeding times can be stressful, indicated by the fact that most abnormal behaviours are observed right before food delivery (Hockenhull & Creighton, 2014b; Peters et al., 2012). However, delaying hay feeding can increase investigative behaviours, likely associated with motivation (Nicol & Guilford, 1991), and may induce frustration (Ninomiya et al., 2004). Instead, multiplying hay feeding times and locations (*i.e.* feed less more often, in different areas) can enhance normal active behaviours, thereby seemingly contributing to feeding satisfaction (Ninomiya et al., 2004).

689 Limited exercise, confinement and a non-suitable diet can result in equine obesity (Furtado 690 et al., 2021b; Robin et al., 2015), which is one of the most severe welfare concerns in UK 691 leisure horses and is thought to affect between 60% of the UK equine population (Robin et 692 al., 2015). Equine obesity can have severe health consequences, including the development 693 of insulin resistance and laminitis, a potentially painful and sometimes fatal condition (Geor, 694 2008). When horses are provided with outdoor access, enclosures are often not of adequate 695 space, may lack enrichment (e.g. trees, shrubs), and it is increasingly common to turn out 696 horses alone (Hockenhull & Furtado, 2021). Social isolation during turnout can reduce grazing 697 time and increase activity and vigilant behaviours (Houpt & Houpt, 1988; Singer et al., 1999) 698 suggesting the lack of companions heightens attention to potential threats. If horses are 699 turned out together, adequate space for the number of animals needs to be available as 700 overcrowding can otherwise elicit aggression, especially when resources (e.g. water troughs, 701 shelter, hay feeder) are limited (Benhajali et al., 2008). In fact, resource guarding has been 702 described in managed horses (Hockenhull & Furtado, 2021; Jonckheer-Sheehy & Houpt, 703 2015), a behaviour not observed under conditions where resources are plentiful (Fureix et al., 704 2012). Although keeping horses together best fulfils their ethological needs, managed groups 705 are often larger than naturally occurring social groups and rarely mirror the naturalistic 706 compositions of bands (Van Dierendonck, 2006).

Beyond management-related factors that can compromise horse welfare, the type of work (*i.e.* disciplines such as dressage, jumping, hacking) and riding techniques can further influence equine well-being by inducing undesirable postures during riding, leading to back pain (Lesimple et al., 2010, 2012, 2013). Chronic back problems linked to vertebral problems are highly prevalent in riding school horses (reported in French horses by Fureix et al., 2010a; Lesimple et al., 2010, 2012, 2013), which were used as the study population for this thesis,
Chapter 1 | General introduction

713 and likely cause physical pain with negative consequences to psychological well-being. 714 However, the assessment of pain is difficult (Sneddon et al., 2014), especially in prey species 715 like cattle or sheep, which tend to be 'stoic', *i.e.* a developed behavioural adaption to mask 716 vulnerability for predator protection (Carbone, 2020; Dwyer, 2004; Salzer et al., 2021). There 717 are different approaches to detecting pain in horses, using either facial expression (e.g. the 718 horses grimace scale (Dalla Costa et al., 2014) or protocols combining body posture and 719 behaviour (e.g. composite scale (Dyson et al., 2018; van Loon & Van Dierendonck, 2015)). 720 Moreover, certain behaviours such as bucking or rearing have been associated with 721 musculoskeletal pain (Barstow & Dyson, 2015; Jonckheer-Sheehy et al., 2012). Chronic pain 722 or discomfort can also increase aggressiveness towards humans, but often this might be 723 mistaken as 'bad temper' (Fureix et al., 2010). Examining the spine of French riding school 724 horses, Fureix et al. (2010) found that most horses (73%) were severely affected by vertebrae 725 problems and animals affected by vertebral problems were more prone to be aggressive in 726 response to a human.

727 **1.3.4.3** The horse as a model for cognitive studies associated with negative affect

728 In humans, psychological disorders (e.g. depression, anxiety) are associated with alterations 729 in cognitive processing, including cerebral functions involved in decision-making, attention 730 and memory (reviewed in Paul et al., 2005). Hence, affective experiences can alter how 731 certain stimuli or situations are perceived, and ultimately an individual's perception of its own 732 state (i.e. welfare). As described in section 1.3.3., the assessment of cognitive biases can 733 facilitate measuring emotional states in humans and other animals (Mendl et al., 2009; Paul 734 et al., 2005), including horses (Briefer Freymond et al., 2014; Henry et al., 2017; Hintze & 735 Schanz, 2021). In fact, the cognitive consequences of inadequate housing conditions and 736 resulting welfare issues have been described in different ways. For instance, Henry et al. 737 (2017) reported that restricted riding school horses kept in single stalls, with limited roughage 738 access, and ridden by inexperienced riders, displayed pessimistic responses in a spatial 739 judgement bias task (JBT), whereas horses living under more naturalistic conditions 740 performed optimistically. Their findings indicated that negative housing conditions can induce negative mood, which was further supported by their findings that restricted horses 741 742 presented more signs of poor welfare (e.g. abnormal behaviours, aggressiveness towards humans) compared to naturalistically housed horses (Henry et al., 2017). In contrast, 743

Chapter 1 | General introduction

providing access to pasture and conspecifics appears to induce optimistic bias in horses(Löckener et al., 2016).

746 However, using horses' responses in JBT to infer mood is not always straightforward. Horses 747 trained with negative reinforcement (i.e. handler used gentle pressure on the halter and 748 pulling on the lead rope to get them to perform different exercises) approached the 749 intermediate cue faster than horses trained using positive reinforcement (i.e. received food 750 whilst performing the exercises; Briefer Freymond et al., 2014). This suggests a more 751 optimistic responses bias in the former, contrary to the predictions, possibly because 752 negatively reinforced horses were more motivated to receive food than positively reinforced 753 horses that already received food rewards during training (Freymond Briefer et al., 2014). In 754 addition, other research suggests that poor welfare can alter cognitive processes in horses, 755 as well as their ability to cope with their environment (Fureix et al., 2010, 2012, 2015; 756 reviewed in Hausberger et al. (2019)). For instance, negative emotional states, particularly 757 short-term states such as fear, have been associated with impairments in learning, although 758 individual characteristics (e.g. temperament (i.e. fearfulness - Valenchon et al. (2013)) and 759 emotionality (*i.e.* propensity of reacting to challenging situations with high behavioural 760 reactions - Lansade et al. (2017)) can also influence involved cognitive processes (reviewed in 761 Hausberger et al., 2019). Most relevant to this thesis, there is some evidence that horses' 762 attention to environmental stimuli can be modulated by negative affect (Rochais et al., 2016). 763 Horses monitored their surroundings more the less they were affected by spine problems, 764 indicating that putative pain associated with vertebral disorders reduces overall attention 765 (Rochais et al., 2016). These findings are in line with other observations suggesting that 766 negative mood (*i.e.* putative depressive-like states) reduces alertness in horses (Fureix et al., 767 2012), which has also been linked with physical damage such as body lesions (Popescu & 768 Diugan, 2013).

769 **1.4 Research aims and thesis outline**

Finding reliable indicators of animal emotion is a central research goal to accurately assess animal welfare. There is scientific evidence that negative affect modulates attention to threat in humans and other animals. Hence, AB might be a valuable indicator of emotional state. The overarching aim of this thesis is therefore to develop and validate a visual AB test in horses allowing measuring AB in this species.

Chapter 1 | General introduction

In chapter 2, I first review the existing literature on attention bias thereby drawing heavily from human studies to provide an overview of the theoretical background to attentional processes and anxiety-induced changes. I then examine experimental approaches to measuring attention bias in animals to discuss methodological aspects that need consideration when translating AB tests to animals.

780 Chapter 3 summarises the general methods applied in this thesis and describes my781 approaches to developing the experimental designs used in the subsequent chapters.

In chapter 4, I present the development of a visual AB test in horses. In a follow-up study
(Chapter 5), I investigate horses' response to on-screen images, insights that help to further
understand horses' behaviour in the AB test.

In chapter 6, I explore additional measures of emotional state comprising behavioural and
physiological indices, with a particular focus on the potential relationship of cardiac activity

787 (HRV) and asymmetric facial temperature patterns as indicators of affective valence.

In the final chapter, Chapter 7, I discuss the findings of this thesis and evaluate how AB assessed in Chapter 4 might relate to the behavioural and physiological indices assessed in the subsequent chapters. To conclude, I identify methodological limitations and give future directions on how the here developed methods may be improved to use similar applications in future studies of animal emotions.

793 Table 1.1 Glossary of keywords used throughout this thesis.

Term	Meaning
Affective state	Overarching term for short-term emotions and longer lasting mood states. These can be described in terms of their valence (see below) and arousal. In humans, affective states comprise feelings alongside physiological, neurological, behavioural and cognitive components. In animals, the term is used without implication of a conscious experience.
Animal welfare	A multi-faceted subjective state determined by an animal's perception of its own circumstances influenced by factors associated with nutrition, the physical environment, health and behaviour. Welfare state can be positioned on a continuous spectrum between good and poor.

Term	Meaning			
Attention	Accognitive function involving a set of evolved cerebral processes enabling the selective processing of certain information. This results in the prioritised processing of internal or external information allowing normal functioning of subsequent cognitive and behavioural processes.			
Attention bias	Describes the preferential allocation of attention towards certain information over other.			
Arousal	Refers to how activated or energised the individual is during the emotional experience.			
Core affect	Term derived from two-dimensional framework; subjective states that can be described within the 'Core affect space' of valence and arousal.			
Feelings	Subjective component of affective states that is consciously experienced.			
Mood, mood states	Longer-lasting 'free-floating' subjective experiences that are not attached to specific rewarding or punishing events/stimuli and derives from the accumulation of short-term emotions.			
<i>Positive affective state(s)</i>	Emotional states that are associated with reward-seeking/appetitive motivational states and (<i>i.e.</i> Q1 and Q2 states in the integrative framework by Mendl et al. (2010), Figure 1.2). States elicited by rewards (see below) or predictors thereof.			
Negative affective state(s)	Emotional states that are associated with punsihment motivational states (<i>i.e.</i> Q3 and Q4 states). States that are elicited by punishers (see below) or predictors thereof.			
Punishers	Stimulus/event an animal will work to escape from or avoid; induce negative affective states.			
Rewards	Stimulus/event an animal will work for; elicit positive affective states			
Short-term emotions	Subjective experiences elicited by rewarding or punishing stimuli/events.			
Stimulus/Stimuli	Valenced events, objects or situations that animals associate with rewards or punishers.			
Valence	Pleasant or unpleasantness of the emotional experience.			

795 Chapter 2 | Attention to emotional stimuli as an indicator of affect in humans

796 and non-human animals – a literature review

797 Summary

The overarching goal of this chapter is to give an in-depth overview of the concept of attention 798 799 and attention bias (AB), and explain why AB could be considered a novel cognitive proxy of 800 animal affective state. Furthermore, the conclusions I draw from this literature review 801 rationalise my experimental approach to developing a visual AB test in horses (Chapter 4). I 802 first describe some key concepts, definitions, functions, types and stages of attention termed 803 in the human literature, to provide a better understanding of attentional processes, and of 804 the relationship between AB to threat and human anxiety. In the second part of this review, I 805 present an up-to-date overview of existing AB studies conducted on animals. I discuss 806 experimental discrepancies in existing animal AB studies. From this, I conclude that AB to 807 threat could potentially be a useful indicator of affective valence in animals, but experimental 808 refinements, including the identification of appropriate species-specific test stimuli, are 809 needed to validate this paradigm.

810 **2.1**

Background and chapter aims

811 There is a general acceptance that animals can experience emotional states (de Vere & Kuczaj, 812 2016), even though the level of conscious processing of affect is still unclear (Paul et al., 2020). 813 Finding ways to reliably measure these internal states is central to animal welfare science to 814 prevent suffering and promote positive experiences (Boissy et al., 2007; Dawkins, 1990; Paul 815 et al., 2005). Behavioural and physiological components of emotions can be measured 816 quantitatively, but the acquisition, interpretation and understanding of these types of indices 817 concerning animal emotions are not always straightforward (Mason & Mendl, 1993; Mendl 818 et al., 2009; Paul et al., 2005).

In humans, cognition and emotions are interdependent (reviewed by Brosch et al., 2013; Liu et al., 2009). Affective states (*i.e.* stimulus-driven short-lasting emotions and stimulusindependent longer-lasting mood states) modulate cognition, possibly to enable appropriate adaptive responses to environmental conditions. That is, emotions determine how we perceive the world around us and are important for successful (*i.e.* normal) cognitive functioning involved in the execution of daily activities, such as social interactions and

decision-making (Brosch et al., 2013). On the other hand, cognitive appraisal (*i.e.* the evaluation of the relevance of objects, events or situations concerning an individual's needs or goals), elicits adaptive affective responses that activate resources enabling the organism to cope with the situation (Brosch et al., 2013). Thus, emotions also arise from cognitive activity thereby assigning emotional values to stimuli and guiding behaviour responses (*e.g.* approaching positive/rewarding stimuli or avoiding negative/punishing stimuli; Paul et al. 2005).

832 The modulating effects of emotions on cognitive functions are hypothesized to have adaptive 833 survival value (*i.e.* preparing the organism to detect and deal with threats likely to occur in 834 the environment, Bateson et al., 2011; Nettle & Bateson, 2012). It is therefore possible that 835 affect-driven cognitive biases, at least to some extent, also occur in non-human animal 836 species (Paul et al., 2020). Harding et al., (2004) were the first to show that negative affect 837 can induce a bias in decision-making in animals. They found that rats kept under stressful 838 unpredictable housing conditions (therefore putatively experiencing a negatively-valenced 839 emotional state) were more likely to respond to ambiguous test cues, intermediate to two 840 conditioned cues associated with positive or negative outcomes, as if the stimulus signalled a 841 negative outcome compared to rats in less stressful conditions. Since this first evidence of 842 'pessimistic' bias in an animal model, experimental variations of the seminal judgement bias 843 task (JBT) have been applied to various other species (reviewed in e.g. Gygax, 2014). Overall, 844 JBT findings confirm that animals in comparatively better conditions (induced by housing or 845 pharmacological interventions), therefore putatively more positive affect, show more 846 'optimistic' judgements of ambiguity than those in relatively worse conditions (Lagisz et al., 847 2020; Neville et al., 2020).

848 An important advantage of assessing cognitive components of emotions is that is it possible 849 to make a priori predictions about their direction in relation to affective valence (i.e. 850 negative/positive states). Behavioural and physiological indices, in contrast to cognitive indices, may be better measures of emotional arousal (*i.e.* the intensity) as they are not free 851 852 from interpretation problems with regard to emotional valence (Paul et al., 2005). However, 853 the application of JBTs has several limitations, including time and labour constraints 854 associated with animal training, selection bias of test animals (only individuals that learn the 855 discrimination step can subsequently be tested in the JBT), or unsuitability of test setups in

applied settings. Moreover, animals' responses in the JBT may be influenced by non-valencerelated confounds, such as arousal, motivation, distraction, and general activity (Crump et al., 2018). As a result, animals' responses in JBTs may not always correspond with the expected mood-congruent shifts in decision-making. For instance, chronic food-restriction induced optimistic judgement bias in sheep, possible because hungry animals were motivated to search for food and were, therefore, more likely to approach the test probes (Verbeek et al., 2014). Therefore, new experimental paradigms free from these constraints are needed.

863 Recently, quantifying attentional biases in response to emotional stimuli has been proposed 864 as an alternative measure of affect-related changes in cognition (Bethell, 2009; Bethell et al., 865 2012; reviewed in Crump et al., 2018). Attention bias (AB) refers to the preferential allocation 866 of attention to certain information over another (MacLeod et al., 1986). In humans, there is 867 ample evidence that anxiety modulates attention to threatening information (Bar-Haim et al., 868 2007; Cisler & Koster, 2010; Mathews & MacLeod, 2002). An advantage of using attention-869 based tests in animal studies is that experimental designs target innate tendencies of sensory 870 information processing (e.g. visual or auditory attention) thereby avoiding experimental 871 limitations associated with animal training required in JBTs. This means that measuring AB 872 may be more practical in applied settings and be may suitable for testing larger numbers of 873 animals. Moreover, understanding how attention is modulated by affect in animals might 874 further provide insights into other affect-induced forms of cognitive biases (Mendl et al., 875 2010). For instance, affect-driven shifts in attention towards stimulus-related (e.g. negative, 876 threatening) information, and subsequent changes in memory processes, may contribute to 877 the development of pessimistic judgement bias (Mendl et al., 2010).

878 However, to date, relative few studies have investigated affect-driven ABs in animals (Crump 879 et al., 2018). A review by Crump et al. (2018) comprised 12 studies conducted on eight species 880 (e.g. starlings (Sturnus vulgarisnus vulgaris), sheep (Ovis aries), rhesus macaques (Macaca 881 mulatta), cattle (Bos taurus), chimpanzees (Pan troglodytes), orange-winged amazon 882 (Amazona amazonica), Guinea baboon (Papio paio), brown rat (Rattus norvegicus) using five 883 different test designs. Even in the relatively small set of studies, there is a noticeable diversity 884 in experimental designs and approaches to how AB was quantified and interpreted in animals. 885 In human studies, contradictory findings are suggesting that anxiety shifts attention to or away from threat (Bar-Haim et al., 2007). The inconsistencies in human findings likely result 886

from differences in experimental tasks targeting different aspects of attentional processing (Bar-Haim et al., 2007; Cisler & Koster, 2010). It is therefore fundamental to understand what AB looks like (*i.e.* what are the measurable characteristics) and how different paradigms measure affect-induced changes in attention (*i.e.* different paradigms might address different stages of attention, hence how we understand AB might be influenced by the chosen approach to measure it), before applying similar attention bias tests (ABTs here after) in animals.

894 The aims of this chapter are threefold. First, I describe some of the key concepts (i.e. 895 definitions, functions, types and stages of attention) from the human literature to facilitate a 896 deeper understanding of the different approaches to quantifying attention. Second, I review 897 paradigms of anxiety-related attention biases in humans to support later comparisons of ABTs 898 applied in animal studies. Third, I provide an up-to-date overview of existing AB studies 899 conducted on animals. I review differences in experimental designs, with a particular focus 900 on what stages of attention were targeted and how AB was measured. I further evaluate the 901 influence of test stimuli characteristics, including the type, emotionality (i.e. valence, 902 intensity) and presentation types. With this, I want to highlight knowledge gaps needing 903 further investigation to validate AB as an indicator of emotional valence. Most importantly, 904 the conclusion I have drawn from this review will explain my approach to developing a visual 905 AB test in horses presented in Chapter 4.

For an extensive review of ABTs in humans, I refer elsewhere (*e.g.* Bar-Haim et al., 2007; Cisler
& Koster, 2010; Yiend, 2010). However, I draw heavily from human studies to explain
terminology for a deeper understanding of the scientific concepts of attention and biases in
attentional processing.

910 For reviewing the animal literature, I searched "Web of Science" for animal AB studies 911 published to-date following Crump et al. (2018) using the search terms "attention bias animal welfare". This resulted in 88 findings in June 2022. I screened all abstracts for relevance (i.e. 912 913 mentioning assessment of attention bias or attention to emotional stimuli in animals) and 914 only experimental studies were considered for my review (29 publications). I found additional 915 references in previous reviews (Crump et al., 2018; van Rooijen et al., 2017; Winters et al., 916 2015) or through references provided in other AB studies. In total, I evaluated 50 animal AB 917 publications, although eight of these addressed AB indirectly (*i.e.* testing visual attention

918 using emotional stimuli, but not specifically referring to their results as AB). The full list of919 these studies is given in Table A1 in Appendix A.

920 **2.2 What is attention?**

921 Cerebral mechanisms by which an individual acquires, processes, stores and acts on 922 information from the environment encompass cognition (Shettleworth, 1998). Attention is, 923 alongside learning, memory and decision-making, a core component of cognition and 924 attentional processes are involved in or precede other cognitive processes, *e.g.* long-term and 925 working memory (Chun et al., 2011).

926 The concept of attention is a long-term subject of scientific study and was first described by 927 William James (1890) as "taking the possession by the mind, in clear and vivid form, of one out 928 of what seem several objects or trains of thought". However, the term 'attention' is used so 929 widely that there is no consensus about a common definition. Some scientists have argued 930 that attention is a 'catch-all' term for how the brain controls information processing, possibly 931 because attention is part of many different cognitive processes (Chen et al., 2020; Hommel et 932 al., 2019). Generally, definitions of attention are consistent in that they refer to attention as 933 a cerebral process which enhances information processing by allocating processing resources 934 based on a current goal (Broadbent, 1958; Carrasco, 2011; Desimone & Duncan, 1995). This 935 means that attention involves selective processes by which certain information (e.g. objects, 936 locations) is selected to receive more sensory (e.g. visual) and cognitive processing than 937 others (Carrasco, 2011; Findlay & Gilchrist, 2003). Accordingly, attention is related to the 938 limited processing capacities of the brain due to the high-energy cost of neuronal activity 939 (Carrasco, 2011). To emphasise that attention is not an agent acting on the brain but a subset 940 of brain processes that accomplice selective cerebral processing, others suggest defining 941 attention as "a set of evolved brain processes that lead to adaptive and effective behavioural 942 selection" (Krauzlis et al., 2021).

To combine existing definitions, I refer to attention in my thesis as "a cognitive function involving a set of evolved processes that enables selective processing of certain information resulting in the prioritised processing of subsequent cognitive, physiological and behavioural processes".

947 2.3 Functions of attention

948 The central function of attention is to prioritise certain information. There are two possible 949 reasons why not all environmental information impacting the sensory system can be 950 processed at an equal level. The central nervous system (CNS) has a limited capacity to sustain 951 a high level of information processing, meaning that high numbers of, and complex stimuli 952 can only be attended to for a limited period of time (Dukas & Clark, 1995; Eysenck & Keane, 953 2002). Moreover, cortical computation requires a high level of energy resources, but these 954 are limited (Carrasco, 2011). Therefore, stimulus selection is required to reduce information 955 processing, and attention facilitates the decision of which stimuli are important. Selection 956 mechanisms are involved in early processing stages such as sensory processing (i.e. at a 957 perceptual level) as well as later processing stages involving memory and decision-making 958 (Chun et al., 2011). Stimuli salience is determined by two different mechanisms, either 959 involving characteristics associated with the stimulus itself (*e.g.* colour, location, suddenness) 960 or characteristics dependent on the motivational state of the organisms (i.e. processes 961 respectively referred to as *bottom-up* and *top-down* attention, see section 2.6). The ability to 962 selectively attend to specific stimuli has fitness-enhancing behavioural functions during 963 interactions with the environment (Carrasco, 2011; Findlay & Gilchrist, 2003). Stimulus 964 selection is followed by behavioural action (*e.g.* attending to the water bottle in front of you 965 enables you to pick it up), which is particularly important for organisms that can actively move 966 through their environment (i.e. are not sessile or drift; Krauzlis et al., 2021). Possessing 967 selective orienting behaviours for the early detection of prey or predators is crucial for the 968 survival of an organism (Krauzlis et al., 2021). Support for this theory is the fact that 969 subcortical circuits (*i.e.* optic tectum, superior colliculus in mammals) responsible for stimulus 970 detection, as well as approach-avoidance and decision-making, are preserved across diverse 971 vertebrate species (*e.g.* fish, amphibians, reptiles, birds, mammals; Knudsen, 2020).

Attention further facilitates sensory processing throughout the cortex by modulating selected
information, *e.g.* by changing the characteristic of perceived objects (Carrasco et al., 2004).
That is, attention provides the organism with an optimised representation of sensory input
that emphasises relevant details (Carrasco, 2011), thereby enhancing the processing of
information most relevant to ongoing goals and behaviours (Pashler et al., 2001).

Attention is also responsible for suppressing irrelevant stimuli (which enhances the processing of important information; Steinman & Steinman, 1998). This could be understood as a noise reduction in external signals that co-exist with the signal of interest, *i.e.* letting specific information pass through 'the filter' while distracting information remains outside 'the filter' (Carrasco, 2011). For instance, when focussing on a specific local area (*e.g.* the words in the sentence you are reading right now), the strength of stimuli outside the locus (*e.g.* the words in the next paragraph) is actively suppressed.

984 Moreover, context-dependent adjustments (e.g. affect-modulated motivations) to the 985 internal state of the animal are important during attentional processes (Mendl & Paul, 2020). 986 For instance, an animal might take a greater risk to obtain food when hungry, but prioritise 987 the avoidance of harm in absence of hunger. Hence, accurately accessing risks and benefits is 988 vital and the ability to accurately do so would be enhanced by incorporating more relevant 989 information. In this case, attention can be considered as a set of brain functions that may have emerged to ensure that an animal can access information readily to select and guide 990 991 behavioural responses (Krauzlis et al., 2021). This might also include context-dependent 992 learning based on previous outcomes. Previous experiences influence attentional 993 mechanisms resulting in a dynamic function of attention to regulate and select sensory 994 processing in interaction with its environment (Krauzlis et al., 2021).

Attentional processes are further important components for other cognitive functions, such as memory formation and decision-making (Mendl, 1999; Mendl & Paul, 2020). For example, attentional shifts and subsequent narrowing of what information is processed or later recalled may influence response speed and errors in decision-making tasks (Mendl, 1999). Hence, how attention is deployed may have significant consequences on decision-making in animals (Mendl & Paul, 2020).

1001 **2.4 Forms of attention**

Attention can be categorised in various ways, for instance, according to the targets of attentional processes. Targets of internal (endogenous) origin comprise thoughts or memories, while exogenous (external) targets are perceived through the sensory system (*e.g.* visual, auditory or tactile information; Chun et al., 2011). Exogenous attention can entail all sensory modalities (*i.e.* visual, auditory, olfactory, gustatory, tactile stimuli), but in my review,

1007 I will only focus on visual attention as the bulk of human and animal AB studies use1008 behavioural correlates associated with visual attention.

1009 **2.5 Types of visual attention and its assessment**

1010 As the primary visual system (retina) is constantly stimulated by an overwhelming amount of 1011 information, visual attention enables filtering relevant information out of irrelevant noise. In 1012 other words, attention turns looking into seeing (Carrasco, 2011; Findlay & Gilchrist, 2003). In 1013 fact, less than 1% of the information projected via the optic nerves to the brain reaches 1014 attentive scrutiny (Anderson et al., 2005). As a result, humans fail to notice or recollect most 1015 of the visual information of a typical scene due to inattentional blindness (Mack & Rock, 1016 1998). The most famous demonstration of this phenomenon may be the "gorilla in the room" 1017 experiment by Simons & Chabris (1999). Subjects were asked to mentally count the number 1018 of passes made between teammates of two basketball teams. While keeping exact count 1019 thereby focusing their attention on a specific event (ball tossing), most subjects missed seeing 1020 a person dressed in a gorilla costume walk across the scene. One possible explanation for 1021 inattentional blindness is that unexpected events may be consciously perceived, but 1022 immediately forgotten (Wolfe, 1999).

1023 Only by prioritising relevant information and simultaneously ignoring irrelevant noise, we are 1024 able to interpret the world around us (Carrasco, 2011). For this, selective visual attention is a 1025 fundamental process whereby stimuli that are irrelevant to ongoing behaviours or goals are 1026 ignored so that neural resources are prioritised for attending to stimuli more relevant to an 1027 individual's goals. There are different forms of selective attention (reviewed in Carrasco, 2011 1028 and Chun et al., 2011). Spatial attention, in which overt attention is directed to a specific 1029 location, can be assessed as eye movements reflecting spatial fixation of foveal acuity (*i.e.* the 1030 ability to discriminate fine details; Chan & Courtney, 1996). In humans, foveal acuity is 1031 restricted, which is why eye movement is deployed to view prioritised locations, and overt 1032 attention can be inferred from eye fixation on a very specific focal point (Findlay & Gilchrist, 1033 2003).

However, the location of the visual field attended is not necessarily the same as the attention
focus, since it is possible to fixate your eyes on one object whilst paying attention to another.
Moreover, attention can be deployed to a relevant location in the periphery without direct

1037 eye movement, which is called *covert* orientation (Carrasco, 2011; Findlay & Gilchrist, 2003). 1038 This type of attention is, for example, deployed in humans, whilst searching for objects, such 1039 as your favourite peanut butter in the supermarket aisle, whereby attention is guided by 1040 feature-based cues (e.g. the familiar blue label of the peanut butter jar) or the saliency of 1041 features (e.g. in a row of red-labelled jars, a blue-labelled peanut butter jar draws more 1042 attention). Covert attention enables processing advantages for localised parts of the visual 1043 field without overt eye movement (Findlay & Gilchrist, 2003). The relationship between 1044 covert and overt attention is debated, *i.e.* whether both are deployed independently and co-1045 occur because both are driven by similar visual input or closely linked (reviewed in (Findlay & 1046 Gilchrist, 2003).

1047 **2.6 Top-down and bottom-up processes of attention**

1048 Two types of attention mechanisms can be distinguished (Corbetta & Shulman, 2002; Egeth 1049 & Yantis, 1997; Yantis & Egeth, 1999). Top-down (i.e. goal-driven) mechanisms are guided by 1050 internal motivational states. For instance, in search of the favourite peanut butter, top-down 1051 attention mechanisms will prime the visual search for the familiar blue-labelled jar. This form 1052 of 'selective' or 'sustained' attention is voluntarily modulated through cortical structures 1053 (Steinman & Steinman, 1998; Caresco, 2011). Cognitive processing of top-down attention is 1054 tied to late processing stages of visual information. Duration of attention allows perceiving 1055 details about a stimulus, and top-down mechanisms prime attention based on expectancy 1056 and moves attention to a specific location (e.g. to attend the specific shelf location where 1057 peanut butter is usually stocked).

1058 In contrast, bottom-up (stimulus-driven) mechanisms are responsible for noticing specific 1059 object features, such as a blue-labelled peanut butter jar "popping out" in a row of red-1060 labelled jars. In this case, attention is involuntarily captured and controlled by the salient 1061 attribute of the stimulus not necessarily relevant to the observers' perceptual goals (Yantis & 1062 Egeth, 1999). Here, attention is guided by events in the sensory periphery and tied to the early 1063 processing stages of visual attention (Lockhofen et al., 2021). This attention mechanism is 1064 characterised by a transient, rapid engagement with new stimuli such as the sudden 1065 appearance of signals outside the attentional focus resulting in a shift of attention to the 1066 stimuli (Steinman & Steinman, 1998).

1067 In any act of (visual) attention, a combination of both modes of attentional control might 1068 occur as both complement each other and are closely connected *via* common neural 1069 pathways (Yantis & Egeth, 1999).

1070 2.7 Processing stages of attention

1071 Considering the function of visual attention to prioritize and filter information, the following 1072 task-dependent stages of attention can be defined: a) attention engagement in preparation 1073 to detect a target, b) orienting attention to a specific location (which Posner (1980) defines 1074 as the alignment of attentional resources with a source of external or internal information), 1075 c) locking attention on its locus (*i.e.* attention maintenance), d) suppression of irrelevant 1076 information, and e) attention disengagement in preparation for attending new information 1077 (Steinman & Steinman, 1998). Figure 2.1 provides a simplified schema of early and later 1078 processing stages of attention.



1079



1081

1082 The neurophysiological basis of visual attention is still not fully understood, and therefore 1083 sometimes speculative or controversial (Carrasco, 2011; Posner & Petersen, 1990). However, 1084 three fundamental neuronal networks comprising both cortical and subcortical brain areas 1085 have been identified (Posner & Petersen, 1990; reviewed in Carrasco, 2011). Broadly, alerting 1086 (attention engagement) is a state of high sensitivity to incoming stimuli and is, in humans, 1087 associated with the frontal and parietal regions of the right hemisphere. Orienting describes 1088 the selection of information from sensory input and is associated with posterior brain areas 1089 (e.g. superior parietal lobe, temporal parietal junction and front eye fields). Executive control refers to the mechanisms involved in conflict-solving among appropriate responses (i.e. 1090 1091 maintenance, disengagement), which are processed by the anterior cingulate and the lateral 1092 prefrontal cortex (reviewed in Carrasco, 2011).

2.8 Anxiety-related attentional biases in response to threat in humans

1094 The following paragraphs provide an overview of the assessment and major findings of 1095 anxiety-related AB to threat in humans. I first describe one of the key theories of how anxiety 1096 might modulate attentional processes proposed by Bar Haim et al. (2007), before outlining 1097 the most common experimental designs to measure ABs.

1098 2.8.1 AB to threat as a normal adaptive response

1099 Before considering the pathological aspects of anxiety-related ABs, it should be noted that, 1100 from an evolutionary perspective, the preferential processing of threatening information is a 1101 normal adaptive mechanism enabling early detection of potential threats (Ohman & Mineka, 1102 2001; Paul et al., 2005). Prioritising attentional resources for threat detection is important 1103 because the cost of missing potential threat signals is much greater than the cost of a false 1104 alarm (Steimer, 2002). Biologically relevant threat cues (e.g. the sight of deadly predators, 1105 height drop) are automatically detected and can trigger fear responses before conscious 1106 perception (LeDoux, 1996; Ohman & Mineka, 2001). Hence, the primary function of fear and 1107 anxiety is to avoid danger and protect the body from harm (Lang et al., 2000; Mathews & 1108 MacLeod, 1994; Ohman & Mineka, 2001).

1109 Although fear and anxiety share similarities, both states have clear differences in terms of 1110 their time course and response pattern (Steimer, 2002). Both are associated with high 1111 arousal, the expectancy of threat to well-being and specific behavioural patterns which have 1112 the function to cope with an aversive or unexpected situation. They overlap in the underlying 1113 brain and behavioural mechanisms, which is why fear and anxiety result in similar adaptive or 1114 defensive behaviours (Craig et al. (1995) cited in Steimer, 2002). However, fear is a basic 1115 emotion evoked by a specific (imminent) threat or danger resulting in defensive behaviour or 1116 escape (MacFarland (1987) cited in Steimer, 2002). In contrast, anxiety persists in absence of 1117 a fear-inducing stimulus, hence is *mood* state detached from any immediately triggering 1118 event, which also appears to prepare the organism for unknown threats (Craig et al., 1995). 1119 The origins of anxiety might therefore be linked to uncertainty or lack of control resulting in 1120 maladaptive responses, such an overprotective fear reactions, or the complete avoidance of 1121 stimuli or situations (Mogg & Bradley, 2016). However, distinguishing when an anxiety 1122 response is normal (*i.e.* makes adaptive sense from an evolutionary perspective) and when it 1123 becomes a dysfunction, or in case of humans a clinical disorder, is difficult to identify

1124 (Wakefield, 1992; Bateson et al., 2011; Nettle & Bateson, 2012). Wakefield (1992) proposed 1125 to define (mental) dysfunction as a term describing the failure of an internal mechanism to 1126 perform the function it was evolutionary selected for resulting in harmful consequences 1127 (Wakefield, 1992). In humans, clinical anxiety is mainly diagnosed based on the level of 1128 suffering and impact on quality of life, but distress does not necessarily mean a malfunction 1129 from an evolutionary perspective (since symptoms of anxiety such as restlessness, increased 1130 heart rate and processing priority of potential threat prepares the organism to detect and 1131 respond to threat, Bateson et al., 2011). Bateson et al. (2011) proposed a 'signal detection 1132 model' which explains that the threshold for an anxiety response should vary with the 1133 probability of threats occurring in an individual's environment and its vulnerability if these 1134 threats occur. According to this framework, the symptoms of generalised anxiety disorder 1135 (GAD) might be classified as a mental disorder if the threshold for mobilising an anxiety 1136 response is low and leads to the proportion of false alarms being high (Bateson et al., 2011). 1137 Consequently, anxiety responses are exhibited in non-threat situations, in fact any situation 1138 could trigger an anxiety response resulting in an impairment of daily functioning.

1139 **2.8.2** Theories of anxiety-related attention biases to threat

There are many theories (reviewed in Bethell, 2009; Bar-Haim et al., 2007) describing how anxiety (*i.e.* trait and state anxiety) might modulate threat-selective attentional processing. It goes beyond the scope of this thesis to explain each theoretical framework in depth. Instead, we briefly describe the most recent framework by Bar-Haim et al. (2007) for a better understanding of which attentional stages common AB tasks address, and how these are modulated by anxiety.

1146 The integrative model proposed by Bar-Haim et al. (2007) is illustrated in Figure 2.2. An 1147 environmental stimulus is first categorised by the pre-attentive evaluation system (PTES) as 1148 having either a high or low threat value. In case of a high threat value (*e.g.* the sight of a grizzly 1149 bear), information is forwarded to the resource allocation system (RAS), which triggers the 1150 physiological fear response (e.g. increase in HR, release of stress hormones), interrupts on-1151 going task, and allocates processing resources to the stimulus. If the stimulus threat value is 1152 still perceived as high, the signal is sent to the guided threat evaluation system (GTES), which 1153 analyses the context of the threat signal, in combination with prior learning and memory. If 1154 the stimulus is perceived as a low threat at this stage (e.g. taxidermy grizzly bear), the

precursory systems relax and return to a non-threat state. If the stimulus threat value is still considered as high (*e.g.* the grizzly bear is truly real), the goal evaluation system (GES) interrupts all on-going goals and orientation to threat is induced.

1158 Different forms of anxiety disorders may lead to malfunctions at different stages of the 1159 proposed model (Bar-Haim et al., 2007). For example, types of phobias are characterized by 1160 a persistent fear reaction in response to a specific object, situation, or signals thereof. The 1161 phobic stimulus (e.g. a butterfly for lepidopterophobic subjects) almost invariably provokes 1162 an immediate anxiety response, which in Bar-Haim et al.'s model may be explained by a 1163 specific bias in the PTES, resulting in an alert response of the RAS. Furthermore, phobic subjects may recognize that their fear is irrational (a butterfly is harmless), indicating that the 1164 1165 GTES is functioning properly, but the overriding mechanism responsible for relaxing the alert 1166 state imposed by the RAS may be dysfunctional. In contrast, individuals with post-traumatic stress disorder (PTSD) selectively attend to trauma-related information stored in memory 1167 1168 following a traumatic experience. A vigorous alarm response to trauma-related stimuli may 1169 be enforced by altered functions of the PTES and RAS, or when the GTES fails to integrate the context and coping resources causing the experience of reliving the trauma (Bar-Haim et al., 1170 2007). 1171



1172

Figure 2.2 A schematic model of the cognitive mechanisms underlying threat processing adapted from Bar-Haim et al (2007). A stimulus is perceived and evaluated by early attention stages (*i.e.* PTES, RAS). These systems estimate the stimulus threat value and activate fear responses as well as later attention processing stages (*i.e.* GTE, GES). High threat value leads to a disruption of on-going activities and redirection of all processing resources to an appropriate stimulus response.

1179 It is important to know that different mood disorders lead to specific patterns in cognitive 1180 biases as they modulate different aspects of cognitive processing. For instance, there is less 1181 consistent support that depression induces AB to threat, but strong evidence that depression 1182 causes pessimistic judgement and negative memory bias (Mathews & MacLeod, 1994; but see 1183 (Peckham et al., 2010) for evidence that depression might cause AB measured in the dot-1184 probe paradigm). In turn, these two latter types of biases are less likely to occur in anxiety 1185 (Fox et al., 2001).

1186 **2.8.3** Methods to quantify anxiety-related ABs to threat in humans

1187 There is an overall agreement that anxiety is associated with biases in attention to threat (Bar-1188 Haim et al., 2007). Threat-related AB is observed across several different experimental tasks 1189 and different types of anxious populations (*i.e.* individuals with clinical disorders and non-1190 clinical individuals reporting high levels of anxiety; Bar-Haim et al., 2007; Cisler & Koster, 1191 2010). However, there are significant procedural differences in testing and measuring AB, 1192 resulting in contrasting findings, *i.e.* whether anxiety induces AB towards or away from threat, 1193 likely because different tests target different attention processes (Bar-Haim et al., 2007; 1194 Driver, 2001).

1195 Test designs can diverge in many variables, e.g. how AB is generally quantified. A within-1196 subject AB is inferred from the difference in the attention allocation of highly anxious people 1197 (clinically anxious or high trait anxiety) concerning threat-related versus relatively neutral 1198 stimuli. In contrast, a *between-subject* AB is noted when there are significant differences in 1199 the patterns of attention prioritisation in anxious versus non-anxious subjects. In some 1200 instances, between-subject AB but not within-subject AB, or vice versa, have been reported 1201 (reviewed in Bar-Haim et al., 2007). Both types of comparisons have found supporting 1202 evidence that anxiety is associated with threat-related ABs. However, divergence in anxiety-1203 induced attention allocation patterns might occur when findings of within-subject and 1204 between-subject AB are compared (Bar-Haim et al., 2007). For example, anxious subjects may 1205 attend preferentially to threat-related cues relative to neutral cues (*i.e.* within-subject AB), 1206 but not significantly more than non-anxious controls (*i.e.* between-subject AB; Kyrios & lob, 1207 1998). Importantly, non-anxious controls, particularly when compared to clinically anxious 1208 subjects, may considerably vary in terms of non-clinical anxiety-levels. In this case, between-1209 subject comparisons may be less reliable.

1210 Moreover, study populations may differ in that psychopathology populations (*i.e.* clinically 1211 anxious subjects) or high-anxious subjects (*e.g.* individuals who self-report high anxiety levels) 1212 are compared with non-clinical controls (selected from the general population), or low-1213 anxious subjects. Since clinically anxious subjects have generally more severe symptoms, it 1214 would be expected (under the assumption that the cognitive phenomenon is sensitive to the 1215 emotional intensity) that they show greater ABs than high-anxious subjects. However, the 1216 meta-analysis by Bar-Haim et al. (2007) found no significant difference in the combined effect 1217 sizes of the two populations. It remains under debate as to how state (*e.g.* induced through 1218 emotional manipulation) and trait anxiety (*i.e.* individual characteristics of being anxious) 1219 affect threat-related AB; Bar-Haim et al., 2007).

1220 Tasks can differ in the type of test stimuli used, *e.g.* word- or image-based tasks. Word-based 1221 tasks are more common in humans, but the use of word stimuli has been criticised because 1222 anxious people might be more familiar with threat words due to higher exposure (e.g. 1223 thinking or talking about threat) than non-anxious people. Image-based stimuli mostly depict 1224 facial expressions as the recognition of facial expressions is highly adaptive (Morris et al., 1225 1998), with negative faces (e.g. angry or fearful faces) conveying a threat. However, Bar-Haim 1226 et al.'s (2007) meta-analysis suggested that both words and pictures were equally effective in 1227 reflecting biases in anxious subjects.

1228 Stimulus specificity, *i.e.* the specific type of information the stimuli convey, can affect 1229 attention. Emotional information is processed preferentially due to its critical role in survival 1230 (Ohman & Mineka, 2001). Presented under the threshold for awareness (*i.e. via* subliminal 1231 presentation), threat-related information such as fearful (Amting et al., 2010) or threatening 1232 faces (Bradley et al., 1999) are detected faster than neutral or happy faces, which is why the 1233 latter two are typically used as controls (Bar-Haim et al., 2007). Likewise, animal stimuli (e.g. 1234 snakes, spiders) that are survival relevant are processed more readily (Ohman & Mineka, 1235 2001). Also, the intensity of the emotional material, as well as the stimulus relevance to the 1236 subject, influences that subject's attention. Specific anxiety disorders cause AB to stimuli 1237 congruent with the type of concerns (e.g. social phobia and social cues; spider phobia and 1238 spider cues; Pergamin-Height et al., 2015; Watts et al., 1986).

Another difference in study design is stimulus exposure time. Most studies use presentation
times of 500ms or longer, which allows conscious (supraliminal) processing of stimuli.

However, studies using shorter (subliminal) exposure times, resulting in test conditions that preclude conscious processing, have also reported significant biases towards threat in association with anxiety, indicating that automatic processing of threat is mediated by anxiety (Bar-Haim et al., 2007).

Lastly, three main AB paradigms are applied to measure threat-related ABs in anxious subjects. These tasks quantify attention through reaction times (RT) whilst subjects are requested to perform a secondary task (*e.g.* locate a probe, name the print colour of a target word), and biases in attention are inferred from RT to emotional, relative to neutral, cues.

1249 In the emotional Stroop task (reviewed MacLeod, 1991; Williams et al., 1996), the subject is 1250 required to name the colour of target words. The Stroop effect generally refers to the 1251 difference in colour-naming latency between congruent (e.q., the word 'red' written in red1252 ink) and incongruent (e.g., the word 'red' in green ink) stimuli (Stroop, 1935). The delay in 1253 latency to name incongruent words reflects an inability to ignore irrelevant information (the 1254 content of words) in presence of the target dimension (the colour of words). Threat-related 1255 AB is inferred by longer latencies to name the colour of words with threatening content (*e.g.* 1256 cancer, wound, disease) compared to words with neutral content (e.g. plate, table; Macleod, 1257 1991). Mathews & MacLeod (1985) observed that GAD is associated with longer RTs to name 1258 words associated with a physical threat (e.q. fatal, mutilated) as well as a social threat (e.q. 1259 inferior, lonely) than control words. The delay in response time might reflect different 1260 anxiety-modulating effects on attention. Threat words may be more distracting (*i.e.* capture 1261 attention more, thereby distracting from the actual task to name the colour) because more 1262 attentional resources are demanded to process threat instead of neutral words. However, the 1263 use of this paradigm has been criticised as it might tap into later processing stages unrelated 1264 to attention. That is, threat and non-threat words may be processed at the same speed, but 1265 the response to threat may be impaired due to the emotional content intensifying negative 1266 affect or avoidance of processing threat-related cues (de Ruiter & Brosschot, 1994). Hence, 1267 delayed reaction time may reflect more effortful avoidance of processing threat cues rather 1268 than attention capture (de Ruiter & Brosschot, 1994; MacLeod et al., 1986). Moreover, AB 1269 observed in the emotional Stroop task seems to be highly specific to the stimuli used, and 1270 whether these match the subject's concerns. For example, Watts et al. (1986) observed that 1271 spider-phobic subjects were slower in responding to words related to their fears (e.g. hairy,

1272 crawl), but showed little disruption in colour naming in negative emotional words (*e.g.* fear,
1273 death, grief). This shows that the Stroop effect is not caused by emotionality *per se* (*i.e.*1274 positive/negative word content), but can depend on the degree of relatedness of words to
1275 subjects' concerns (Williams et al., 1996).

1276 In the **dot-probe task**, subjects are required to respond to a neutral cue (probe) which is 1277 shown in the location of one of two stimuli (words or images) displayed simultaneously side-1278 by-side on a screen before the probe cue. This paradigm measures spatial attention 1279 allocation. The distribution of attention whilst viewing the test stimuli is inferred by the 1280 response latency to locate the probe, with faster responses to probes presented in the 1281 attended relative to the unattended location. AB is evident in faster detection of probes on 1282 congruent trials (where threat information primes probe location, *i.e.* probe occurs on the 1283 same side as the threat cue) than incongruent trials where the probe is on the opposite side 1284 to the threat. Often, the average latencies of congruent trials are subtracted by the average 1285 latencies of incongruent trails to calculate an AB score (van Rooijen et al., 2017).

1286 MacLeod et al. (1986), who designed this paradigm, found that anxious subjects appeared to 1287 have an attention shift toward negative information, having a shorter latency to respond to 1288 probes occurring in the congruent location. Conversely, non-anxious subjects showed a faster 1289 response to probes in incongruent trials, *i.e.* indicating attention in the latter was directed 1290 away from negative material (MacLeod et al., 1986). Their observations confirmed that 1291 anxious people orientate towards threat but this could be due to facilitated (faster) attention 1292 engagement or difficulty in disengaging from threat (*i.e.* have longer dwelling time due to 1293 difficulties to reallocate attention).

1294 To address this question, the exogenous spatial cues paradigm was developed by Posner 1295 (1980). In this paradigm, subjects attend to a neutral cue (e.g. "+" in Posner, 1980) appearing 1296 in one of two locations followed by a target presented at the cued location on a majority of 1297 the trials (valid trials), but at the alternative location on remaining trials. The ability to detect 1298 the target is typically faster on valid trials than on invalid trials because attentional 1299 engagement is already primed for the target location (generally known as cue validity effect 1300 (Fox et al., 2001)). In contrast, slower responses on invalid trials reflect difficulties in 1301 disengagement from the cued location (Bar-Haim et al., 2007). In a modified version of this 1302 test, the emotional spatial cueing paradigm (Fox et al., 2001), the attended cue is either a

1303 threat or a neutral stimulus. A biased attention engagement and delayed disengagement are 1304 respectively inferred by faster responses to targets in valid trials (with a faster response when 1305 primed by a threat cue relative to a neutral cue) and by delayed reactions to invalid trials. 1306 Thus, unlike the dot-probe task, this paradigm allows distinguishing between different 1307 processing stages of attention (Crump et al., 2018). Fox et al. (2001) found that, compared to 1308 non-anxious subjects, anxious subjects (high and low trait anxious subjects) were slower in 1309 responding to the target primed by threat cues (words or faces) on invalid trials (*i.e.* took 1310 longer to reallocate their attention) relative to neutral or positive-cued trials. This suggests 1311 anxiety-induced difficulties of attentional disengagement resulting in longer attention 1312 dwelling times toward threat cue location and henceforth delayed response to locate the 1313 target (Fox et al., 2001).

1314 The visual search task allows the assessment of spatial attention allocation, but this paradigm 1315 is less commonly applied in anxiety-related AB studies. Subjects are asked to find a target in 1316 a matrix of distracting stimuli, *e.g.* the word spider in embedded in a matrix of neutral words 1317 or vice versa when testing the effect of negative stimuli as distractors (Cisler & Koster, 2010). 1318 AB is inferred from faster response times to find a threat stimulus in a neutral matrix relative 1319 to latency to find a neutral stimulus in neutral matrices. A quicker response time to a threat 1320 indicates that the individual's attention is *drawn* to the threat stimulus. In contrast, AB can 1321 also be inferred from longer response latencies to detect a neutral stimulus in a threat matrix 1322 relative to latencies to detect neutral stimuli in a neutral matrix, indicating that attention is 1323 captured by the threat stimulus. Spider phobics are faster to find spider stimuli relative to 1324 beetle and butterfly stimuli, and these biases were greater relative to non-spider phobic 1325 controls (Rinck et al., 2005). However, this type of task has been criticised because visual 1326 search is strongly influenced by the intensity of the emotional stimuli as well as their valence 1327 (i.e. high-arousal happy faces are detected faster than low-arousal negative faces, Lundqvist 1328 et al., 2014).

The so far presented paradigms quantify attention based on subjects' response to a secondary task (*i.e.* reactions times RT). However, RT measures have limitations. For instance, response speed might be biased by approach- and avoidance-related motoric response patterns when acting upon emotional stimuli (Rooijen et al., 2017). Subjects asked to push a lever away from their body in response to negative images and pull the level towards them in response to

positive images (*i.e.* actions congruent with the general action tendencies to avoid negativity
and approach positivity) responded faster than subjects asked to perform the reverse actions
(*i.e.* negative/pull and positive/push; Marsh et al., 2005).

1337 Alternatively, eye-tracking technology allows relatively direct and continuous measurement 1338 of eye-movements associated with overt visual attention (Armstrong & Olatunji, 2012). The 1339 advantages of measuring overt attention via eye-tracking are that this approach provides 1340 both temporal and spatial parameters of attention allocation. That means that the continuous 1341 tracking of eye movements offers important advances in characterizing the time course and 1342 components of attentional bias. In contrast, RT-based tasks such as the emotional spatial 1343 cueing task only provide one measure of the speed of orienting (Armstrong & Olatunji, 2012). 1344 Eye tracking can be employed in free viewing or visual search tasks. In the **free viewing tasks**, 1345 subjects are instructed to view images on a screen usually showing neutral and emotional (*i.e.* 1346 negative) image pairs. The sum and duration of looking intervals at a specific stimulus as well 1347 as latency and location of fixation are potential measures of overt attention.

1348 Some studies used time intervals as the only basis for quantifying attention bias (*e.g.* Buckner 1349 et al., 2010), assuming that early fixation duration reflects facilitated detection of threat, 1350 whereas fixation duration in a subsequent window reflects maintenance of attention to 1351 threat. Parameters can be divided into event-related measures (e.g. latency, direction, and 1352 duration of initial fixation) reflecting specific eye movements. These indicators are typically 1353 used to study hypervigilance (Armstrong & Olatunji, 2012). In contrast, epoch-related 1354 measures (*i.e.* data here referring to specific time windows or 'epochs') are extracted fixation 1355 duration from specific gaze events, and are used for measuring attentional avoidance 1356 (Armstrong & Olatunji, 2012) or difficulties in disengaging (*i.e.* attentional maintenance) 1357 derived from longer/more frequent fixations to stimuli after first orientation, Gregory et al., 1358 2019).

1359 In the visual search task, tracked eye movements can be used to measure location latency to 1360 detect threat stimuli as a proxy of exogenous orienting towards a threat target. A meta-1361 analysis by Armstrong & Olatunji (2012) found that relative to non-anxious controls, anxious 1362 subjects showed increased vigilance for threat during free viewing and visual search (assessed 1363 *via* eye movement tracking to infer attention), and showed difficulty disengaging from threat 1364 in visual search tasks, but not during free viewing tasks.

Eye tracking can further be used alongside RT measures, for instance, in a dot-probe task. Powers et al. (2019) used the dot-probe task and found that women with PTSD have significantly longer dwelling time toward angry faces than women without PTSD.

Using eye tracking, research has suggested that in depression, similar to anxiety, AB is characterised by difficulty in disengaging attention from negative stimuli once attention is captured (Sanchez et al., 2013). In contrast, positive stimuli are less attended to in depressed

1371 subjects compared to healthy controls (e.g. Kellough et al., 2008; Sears et al., 2010).

1372

1374 Table 2.1 Overview of the most commonly applied paradigms to study anxiety-related AB in humans.

	Emotional Stroop task	Dot-probe task	Emotional spatial cueing paradigm	Visual search task
Experimental design	Neutral trials Flower Death Tree Murder time Chair Kill	Congruent trial + + + + + + + + + + + + +	Valid trial + + + time * + + *	000 000 000
Stimulus type	word (images less common)	words, images	words, images	words, images
AB measure	AB to threat: Latency to name colour of threat words > latency to name neutral words	AB to threat: RT to locate probe on congruent trials < RT on incongruent trials AB away from threat: RT to locate probe on congruent trials > RT on incongruent trials	AB to threat: RT to locate probe on valid trials < RT on invalid trials AB away from threat: RT to locate probe on valid trials > on invalid trials	AB to threat: RT to detect threat target in matrix of neutral distractors < RT to locate neutral image/words in matrix of threat images/words
Targeted attention stage	Later attentional stages, distraction	Attention orientation and/or disengagement	Attention orientation and/or disengagement	Spatial orientation and/or distraction

1376 **2.9 AB in animals**

1377 Since the perception of emotional stimuli (*i.e.* reward/threat-related signals) is crucial to 1378 individuals' survival, animals should have evolved adaptions to focus attention to these cues 1379 (van Rooijen et al., 2017), similarly to humans. Perceiving emotionally valenced social cues is 1380 vital for social-living animals, directly affecting individual survival (van Rooijen et al., 2017). 1381 For instance, the display of fearful facial expressions warns conspecifics of potential danger 1382 thereby increasing survival chances. Attentional prioritization of threat stimuli should occur 1383 since the rapid detection of negative stimuli (*e.g.* predator cues) within the environment is 1384 essential for these cues to be beneficial in responding to threatening situations (van Rooijen 1385 et al., 2017). Indeed, emerging scientific evidence supports this, and attention allocation 1386 appears to be mediated by the observers' emotional state similar to AB described in anxious 1387 humans (Crump et al., 2018). For instance, modulating effects of emotions on attention have 1388 also been described when animals experience unsuitable housing conditions. Denying captive 1389 starlings access to water baths reduces birds' ability to fly due to poor feather conditions, 1390 hence their ability to escape predator attacks, and ultimately induces anxiety (Brilot et al., 1391 2009a). This can alter starlings' attention to threat as birds without water bath access 1392 perceived a conspecifics alarm call (an ambiguous cue of threat according to the authors) 1393 more negatively, as shown by increased vigilant behaviours compared to birds with water 1394 bath access (Brilot & Bateson, 2012). Moreover, pharmacologically induced emotional states 1395 have been associated with AB. In sheep, drug administration to induce anxiety-like (with m-1396 chlorophenylpiperazine) or depression-like states (with para-chlorophenylalanine) altered 1397 animals' attention to a threat stimulus (Lee et al., 2016; Monk et al., 2018b). When presented 1398 with a live dog and the image of a conspecific (*i.e.* positive stimulus), 'depressed' sheep 1399 attended more to the threat than control sheep (receiving saline treatment), whereas 1400 'anxious' sheep paid more attention to the conspecific image rather than the treat compared 1401 to controls, possibly indicating AB away from threat in anxious sheep (Monk et al., 2018b). 1402 That 'anxious' and 'depressed' sheep displayed AB in opposing directions in this test suggests 1403 that it might be possible to discriminate between different states of negative valence using 1404 ABTs (Monk et al., 2018b).

1405 Although the number of AB studies conducted in animals is still relatively small, the 1406 experimental approaches to assessing and quantifying attentional biases are very diverse.

1407 Since methodological differences contribute significantly to contradictory findings in human 1408 AB studies (Bar-Haim et al., 2007; Yuan et al., 2019), I believe it is essential to draw from this 1409 knowledge when translating AB tests in animals. The potential of affect-driven AB as an 1410 indicator of animal welfare has been reviewed by Crump et al. (2018) concluding that AB 1411 might not be suitable to infer general well-being. With this review, I aim to evaluate 1412 experimental differences in study designs, as different ABTs may not always measure the 1413 same attentional processes, and I argue that research findings may diverge as a result of these 1414 differences, making the validation of AB as a cognitive marker of animal emotion difficult.

1415 **2.9.1 Defining attention and AB in animals**

1416 Attention and vigilance are at times used interchangeably. In human psychology, vigilance can 1417 describe a state of alertness towards certain stimuli (e.g. task-dependent stimuli like a colour, 1418 or word content) that allows the brain to better process and respond to these stimuli (Dukas 1419 & Clark, 1995). Accordingly, vigilance and attention are used interchangeably to refer to 1420 cognitive functions that enhance the ability to process certain types of information (Dukas & 1421 Clark, 1995). Concerning animals, attention and vigilance might address two distinct 1422 phenomena. In behavioural ecology, vigilance manifests in a state of alertness, e.g. in 1423 dangerous situations (e.g. predator attack, Lima & Dill, 1990) or during fear events (Paul et al 1424 2005)). However, vigilance behaviour (i.e. scanning of the environment to detect predators, Beauchamp et al., 2021) might also occur in the absence of actual threatening stimuli (Crump 1425 1426 et al., 2018). That is, vigilance is measured differently compared to attention with the former 1427 being reflected in increased monitoring behaviours (e.g. head up posture, increased 1428 locomotion) without direct eye contact (*i.e.* overt attention) with the threat stimulus (*e.g.* 1429 assessed in Brilot & Bateson, 2012; Lee et al., 2016, 2018; Monk et al., 2018a,b). Attention to 1430 a (threat) stimulus, on the other hand, usually implies that an animal is making direct eye 1431 contact with the stimulus or related cues (e.g. Lee et al. (2018) defined attention to threat as 1432 the time cows looked in the direction of the threat location; vigilance was defined as the head 1433 is at shoulder height or higher). However, Bethell et al. (2012) followed a different approach 1434 by defining macaques' first eye gaze towards one of two possible stimuli as 'initial vigilance' 1435 and referring to the total time animals looked at either stimulus as 'overall vigilance'. Thus, it 1436 seems necessary to clearly define how attention and vigilance are understood, and hence 1437 measured, as the terms might not mean the same to different researchers. Hereafter, I use

the term 'vigilance' to describe non-directed monitoring/scanning behaviour and use'attention' when visual focus is directed to the (threat) stimulus.

The human definition of AB as the differential allocation of attention towards one stimulus compared to others has also been used in animals (Crump et al., 2018). Moreover, the term affect-driven attention bias (ADAB) was coined by Crump et al. (2018) when animals' affective state is considered as a causative source of attention to emotive stimuli. However, I also evaluated AB studies that did not manipulate (or validate) the emotional state of test animals, hence I will use AB and specify in the text where ADAB was investigated.

1446 2.9.2 AB test designs

1447 AB has been investigated in a relatively small number of animal studies. Yet, there is great 1448 methodological variability between approaches in testing and quantifying AB. Table 2.2 1449 provides a summary of commonly applied animal ABTs. As I will outline below, some 1450 experimental designs in animals are analogues to human paradigms, while others measure 1451 preferential attention allocation by observing animals' spontaneous responses to emotional 1452 stimuli. Evaluating different ABT designs I aim to identify associated experimental limitations 1453 and moderating effects on AB results to be taken into consideration when studying AB in 1454 animals.

1455 **2.9.2.1** Tasks requiring learning (modified paradigms of human ABTs)

1456 Non-human primates (e.g. chimpanzees (Pan troglodytes) and gorillas (Gorilla gorilla gorilla) 1457 - Leinwand et al., 2022; bonobos (Pan paniscus) - Kret et al., 2016; rhesus macaques (Macaca 1458 mulatta) - Lacreuse et al., 2013; Parr et al., 2013) have been trained in a modified dot-probe 1459 task requiring the animals to contact a probe (e.g. yellow dot in Lacreuse et al., 2013) on a 1460 touchscreen following the brief display of two emotional stimuli (*e.g.* neutral and negative 1461 faces images of conspecifics) shown side-by-side on the screen. Similar to what can be 1462 observed in humans (see Table 2.1 for human dot-probe task summary), rhesus monkeys 1463 show AB towards negative (threatening) facial expressions of conspecifics demonstrated in 1464 faster RTs when the probe appears in the negative face location than when it appeared behind 1465 a neutral face (Lacreuse et al., 2013). Likewise, bonobos were significantly faster in 1466 responding to probes replacing emotional pictures than neutral pictures indicating that 1467 animals' attention was biased towards the location of emotional images (Kret et al., 2016). 1468 Interestingly, in contrast to humans and rhesus monkeys (Lacreuse et al., 2013), bonobos

showed AB towards positive social stimuli (*e.g.* images showing affiliative and protective behaviours) but not negative emotional scenes (*e.g.* aggression), possibly because positive social interactions are pivotal for bonobos, hence might have higher adaptive value for this species, and are therefore attentionally prioritised in contrast to negative social cues (Kret et al., 2016).

1474 An advantage of the modified task is that it allows, as in humans, investigation of the 1475 directionality of AB (*i.e.* toward or away from threat). Moreover, it might be possible to tap 1476 into different attentional processing states by modulating the length of stimulus presentation. 1477 Cassidy et al. (2021) tested female long-tailed macaques (Macaca fascicularis) with threat-1478 neutral face pairs and observed AB to threat at baseline and AB away from threat following 1479 anaesthesia (likely to induce stress), but only when the stimuli were presented for 100 ms and 1480 not 1000 ms. Hence, AB to threat appeared to be only detectable during automatic attention 1481 capture (*i.e.* early attention processing). The longer stimulus presentation time might have 1482 encouraged animals' to anticipate the probe location as one animal reached toward the 1483 screen before the stimuli disappeared and the probe was shown (Cassidy et al., 2021).

Using a task applicable to humans and animals also allows valuable results to be obtained for comparative studies (reviewed in van Rooijen et al., 2017). For instance, when Lacreuse et al. (2013) compared the performances of humans and rhesus monkeys, they found that both species showed AB towards negative conspecific faces, but only humans showed AB away from non-social images (*e.g.* images of guns, syringes, car crashes). These findings indicate attentional differences between biologically relevant (*i.e.* innate) and conditioned cues, which I further discuss in section 2.9.3.1.

1491 The emotional Stroop-like task has also been adapted to primates. In Allritz et al. (2016) 1492 chimpanzees were first trained to discriminate between two stimulus frame colours (i.e. 1493 yellow and blue) by touching the rewarded frame on a touch-screen. Once the discrimination 1494 was learnt, emotional images showing humans with putative differently valenced 1495 relationships to the animals (i.e. familiar caretaker, veterinarian or stranger) were embedded 1496 into the frames. AB was inferred from animals' accuracy and latency to touch the reward 1497 frame colour when these contained the emotional images. Chimpanzees performed less 1498 accurately and responded slower in trials showing images of the veterinarian than in all other

1499 conditions, indicating that negative stimulus content resulted in response slowing suggesting 1500 interference of the emotional stimuli on task performance (*i.e.* Stroop-like effects). Similar 1501 observations were made by Hopper et al. (2021) who used the same test design in 1502 chimpanzees, gorillas, and Japanese macaques (Macaca fuscata) using images of positive 1503 (preferred food), negative (a snake), or neutral (human-made objects) contents. Animals 1504 made more errors in selecting the correct frame when it embedded positive or negative 1505 images as compared to neutral images further supporting the idea that emotional stimuli can 1506 alter primates' cognitive processes by capturing attention. Interestingly, positive Stroop-like 1507 effects (induced by food cues) were observed in apes and monkeys, but only apes showed 1508 Stroop-like effects with snake images suggesting a species-specific difference in attentional 1509 responses to these stimuli.

1510 Bethell et al. (2016) simplified the emotional-Stroop-like paradigm to a response-slowing task 1511 using only one coloured target stimulus (frame) that animals were trained to touch, hence 1512 requiring no discrimination learning. Rhesus macaques were conditioned to touch a grey-1513 coloured square, in which conspecifics' face images (*i.e.* direct gaze as the negative stimulus 1514 or averted gaze as the neutral stimulus) were embedded during the test condition. Monkeys 1515 were less likely and slower to touch the target surrounding direct gaze images, but only 1516 following a stressful handling procedure indicating that stress (*i.e.* negative mood) mediated 1517 animals' attention to social threat.

1518 Visual search tasks expect subjects to detect a target among distracters (*e.g.* locating a snake 1519 image among flower images, Shibasaki & Kawai, 2009) as quickly and correctly as possible. 1520 Shibasaki & Kawai (2009) reported that Japanese macaques responded faster to snake images 1521 (threat as target) among flower images (neutral distractors) than vice versa (i.e. neutral target 1522 and threat distractors) suggesting rapid detection, hence attentional prioritization, of threat 1523 cues. In another version of this paradigm (Marzouki et al., 2014), baboons (Papio papio) were 1524 trained to locate T-shaped targets among L-shaped distractors. Baboons responded more 1525 slowly in trials that followed the spontaneous expression of negatively (*i.e.* stereotypy, fear 1526 screaming, resting; behaviours categorised as negative by the authors), rather than neutrally 1527 or positively valenced behaviours (*i.e.* play, allogrooming, copulation), indicating that 1528 negative mood associated with negative social interactions can influence baboons' cognitive

performance (inferred from response speed) in this attention-requiring task (Marzouki et al.,2014).

1531 Although these adapted paradigms can provide valuable insights into how animals might 1532 prioritise attention to different emotional stimuli, similarly to humans, there are significant 1533 limitations to assessing AB in (non-primate) animals this way. Most crucially, these types of 1534 tests demand learning to perform a cognitive task before AB can be measured. However, 1535 training-associated challenges (e.g. time constraints, trainability of animals) might restrict the 1536 number of animals available for AB testing. For instance, in Hopper et al. (2021) roughly half 1537 (10/19) of the animals learnt the required discrimination step. Furthermore, using on-screen 1538 stimuli, animals' ability to perceive and recognise image contents as representations of real-1539 life objects/individuals also needs to be considered (see Chapter 5 for a more in-depth 1540 discussion on this topic and testing horses' ability to recognise artificial images). This could be 1541 why only primates have been tested in these AB tasks so far.

1542 **2.9.2.2** ABTs measuring spontaneous attention behaviours

1543 Looking time paradigms, originally developed to study perception and cognition in pre-verbal 1544 infants (Frantz, 1958; Spelke, 1985), are grounded in the assumption that animals' 1545 spontaneous looking direction reveals what (e.g. objects, situations) is most interesting to 1546 them (Winters et al., 2015). Hence, free viewing tasks may present a more practical approach 1547 to testing AB as they often require little (*i.e.* habituation to the experimental setup, training 1548 animals to sit by a target for testing, Howarth et al., 2021) or no training. Crump et al. (2018) 1549 distinguished between two types of looking time paradigms, namely dual (i.e. visual 1550 comparison tasks or preferential looking time tasks) and single presentation tasks. In the 1551 former, the simultaneous presentation of two stimuli introduces processing competition 1552 (Desimone & Duncan, 1995). AB is determined based on looking duration elicited by each 1553 stimulus. Bethell et al. (2012) developed a preferential-looking time task in rhesus macaques 1554 in which animals were presented with aggressive-neutral image pairs of unfamiliar 1555 conspecifics' faces. Whilst macaques exhibited an AB toward aggressive faces during a period 1556 of enrichment (*i.e.* positive condition), negative faces were less attended following health 1557 checks (*i.e.* negative condition) demonstrating that changes in emotional state can modulate 1558 macaques' attention to negative stimuli. Despite the seemly easy testing procedure, the use 1559 of free-looking tasks has been criticised because of challenges in interpreting animals'

1560 attentional responses to stimuli (Winters et al., 2015). More specifically, attention towards a 1561 stimulus is generally interpreted as greater interest, but this could be driven by a curiosity for 1562 novelty (*i.e.* putatively positive) but also reflect vigilance (*i.e.* increases attention to threat). 1563 Moreover, a lack of attention might be the result of active avoidance or no interest (Raoult & 1564 Gygax, 2019). Bethell et al. (2012) interpreted the change in macaques' attentional responses 1565 as a sustained avoidance of the aggressive face following the health check, but it is not 1566 completely clear why macaques attended to the neutral faces significantly longer during 1567 negative compared to the positive condition as they, in principle, could have also directed 1568 their gaze away from both screens (which was not assessed).

1569 In single stimulus paradigms, animals are confronted with one emotional cue at a time, hence 1570 deciding what animals are attending to might be more straightforward. Food-deprived sheep 1571 interacted (*i.e.* sniffed, licked) with a familiar food bucket for longer than over-fed sheep (*i.e.* 1572 receiving 110% to 150% of the theoretical maintenance requirements) suggesting heightened 1573 attention to food-related cues in the supposedly more hungry sheep (Verbeek et al., 2014). 1574 In cattle (Lee et al., 2018) and sheep (Lee et al., 2016; Monk et al., 2018a,b, 2019, 2020), AB 1575 to threat was tested by exposing animals to a live dog (presented for 3 or 10s) while hay was 1576 available in the test area. Orientation behaviours (e.g. head turn towards the threat location, 1577 forward ear position) and signs of vigilance (e.g. head up, increased latency to feed, and 1578 increased locomotion) were measured to infer vigilance and attention to threat (dog). Sheep 1579 receiving pharmacological treatments inducing anxiety-like states exhibited heightened 1580 vigilance to the threat location (despite the removal of the dog), whereas drug-induced 1581 depression-like sheep showed reduced vigilance to the threat location, with significant 1582 differences between these groups but no significant alternations to control sheep (saline 1583 treatment; Lee et al. 2016). Using a similar test protocol where hay was replaced by 1584 conspecifics' images as a positive cue to control for drug-induced differences in feeding 1585 motivation, Monk et al. (2018b) observed that 'depressed' and 'anxious' sheep differed 1586 significantly from controls with both being significantly more vigilant (i.e. head-up posture, 1587 recorded regardless of head orientation) than control sheep. While 'depressed' sheep were 1588 more vigilant to the dog location more than controls, 'anxious' sheep showed the opposite 1589 direction of attentional response as they paid more attention to the conspecifics' image, and 1590 hence were less vigilant towards the dog location, than controls (Monk et al., 2018b). Hence,

1591 'depressed' and 'anxious' sheep showed contrasting attentional responses Lee et al.'s (2016) 1592 findings. Monk et al. (2018b) proposed this was likely due to substituting food with 1593 conspecific's images as a positive stimulus. Anxiety might have enhanced strong flocking 1594 instincts induced by the threatening situation causing 'anxious' sheep to pay more attention 1595 to the companion's image during the threatening situation (Monk et al., 2018b). It should be 1596 noted that although these studies used two stimuli, these were of differing qualities (*i.e.* real 1597 dog paired with hay (e.g. in Lee et al. 2016, 2018; Monk et al. (2018a) or a conspecific' image 1598 (Monk et al. 2018b, 2019, 2020) and were furthermore not presented for the same length of 1599 time (*i.e.* hay/conspecific image remained in the test area, whereas the dog was removed 1600 after 3 or 10 s). The observations resulting from this test set-up might not reflect AB in a 1601 narrow sense (Raoult & Gygax, 2019), i.e. given the lack of stimulus competition considered 1602 as important for AB testing in humans (Desimone & Duncan, 1995). Therefore, I regard these 1603 studies as single stimulus designs, which are more likely to reflect animals' level vigilance 1604 rather than AB.

1605 In animals, selective attention (*i.e.* attention to a specific stimulus or location) is mostly 1606 inferred from behaviours such as orientation responses whereby animals' sensory organs 1607 (e.g. eyes, ears, or nose) are directed towards the source of information (e.g. sheep looking 1608 towards the threat location in Lee et al. (2016)). Attention can be quantified via spatial (i.e. 1609 what stimulus is attended?) and/or temporal (i.e. how long is the stimulus attended?) 1610 assessment. In species with frontally placed eyes, overt visual attention can be inferred 1611 directly from gaze (see Figure 2.3 A&B). In animals with laterally placed eyes (e.g. ungulates 1612 like sheep, cows, or horses) direct gaze is more difficult to assess, as visual input is 1613 predominantly monocular, hence each eye may perceive different visual information at the 1614 same time (Tyrrell et al., 2014). In these species, the direction of spatial attentional is usually 1615 inferred from the direction of head turn (illustrated in Figure 2.3 C), often in consideration with ear positions (Rochais et al., 2016; Wathan et al., 2016). 1616

1617



1618

Figure 2.3 Behavioural indicators of overt visual attention in (A) humans and (B) primates and (C) horses. In species with predominately monocular vision, like horses, head turn direction is the most noticeable indicator of attentional focus. In species with mainly binocular vision (*e.g.* humans, primates), attentional focus can be shifted without noticeable head turn due to their ability to move the eyes and pupils. Drawings by Shareen Bistre Dabbah.

1624

1625 However, one important caveat of inferring visual attention in animals with laterally placed eyes this way is that the direction of the head turn might not automatically indicate the centre 1626 1627 of attention (Raoult & Gygax, 2018). In fact, an animal might turn its head to the right (see 1628 Figure 2.3C) to inspect something more closely using its left eye. Preferential monocular inspection has been reported in cows (Phillips et al., 2015), horses (Larose et al., 2006) and 1629 1630 sheep (Versace et al., 2007) as animals were found likely to inspect negative stimuli (e.g. novel objects, unfamiliar person) using their left eye. Moreover, it could be that prey species divide 1631 their visual attention within the horizontal pane (i.e. discriminating between cues near and 1632 1633 far), hence they might attend to cues in a distance whilst feeding with their head down 1634 seeming to look at the ground. Automated eye-tracking technology could help to understand 1635 what animals are really looking at but this technology has only been validated in very few 1636 species with laterally placed eyes (e.g. peahens - Yorzinski et al., 2013; starlings - Tyrrell et al., 1637 2014). Raoult & Gygax (2018) utilised an automated tracking system to record head and ear 1638 positions in sheep, but they had significant difficulties in judging what sheep were looking at

when simultaneously presenting two videos of putatively opposing emotive situations (*i.e.*negative – dog, positive - conspecifics).

1641 Given difficulties in determining attentional focus in animals, there are obvious limitations as 1642 to what stages of attention (*i.e.* initial attention capture, engagement, and disengagement) 1643 can be measured and ultimately how AB can be quantified. If eye movements are recorded 1644 directly, it might be possible to assess distinctive attentional stages. For instance, in Bethell 1645 et al.'s (2012) study, macaques' direction and latency of the first gaze towards each stimulus, 1646 the duration of the initial gaze and latency to disengage from the initial gaze were recorded. 1647 This level of detail in attentional processing might not be attainable from observing animals 1648 with laterally placed eyes given the methodological constraints of recording direct gaze in 1649 these animals.

1650 **Recording behaviour tendencies,** such as approach and avoidance, is another method to 1651 assess attention to emotional stimuli, grounded in the theory that motivation to move 1652 towards or away from cues is likely to reflect stimulus valence (Lang et al., 1990; Raoult & 1653 Gygax, 2018). Indeed, horses (Equus caballus) were more likely to approach the images of 1654 unfamiliar conspecifics showing positive facial expressions compared to images showing 1655 negative horse faces (Wathan et al., 2016). Boggiani et al. (2018) used an approach-avoidance 1656 paradigm to investigate modulating effects of social aggression on attention in tufted 1657 capuchin monkeys (Sapajus sp.). A food reward (peanut) placed below the image of a 1658 putatively negative stimulus (*i.e.* the image of a laboratory technician wearing a capture 1659 glove) was approached faster and more frequently than a peanut placed below a neutral 1660 image (*i.e.* the image of a familiar person) by monkeys, after receiving social threats from 1661 dominant conspecifics whilst receiving food rewards (Boggiani et al., 2018). The authors 1662 interpreted these observations as evidence of socially induced AB as the test monkey did not 1663 show any differences in latencies to take the reward before receiving aggression from 1664 dominant bystanders, but this does not explain why the monkeys did not seem to be 1665 frightened by the negative stimulus (akin to a predator) as they approached rather than avoid 1666 it. Testing sheep in the approach-avoidance paradigm in which animals were presented with 1667 positive (conspecifics ruminating) and negative (dog) videos resulted in inconclusive findings 1668 as animals' behavioural responses did not vary consistently according to the presumed

valence of the stimuli, suggesting that videos might not be the ideal way to present valencedcues for sheep (Raoult & Gygax, 2018).

1671 Another way to infer attention from spontaneous animal behaviour is to record the 1672 disruption of non-acquired ongoing behaviours in response to a threat stimulus. For 1673 instance, Brilot & Bateson (2012) exposed starlings to conspecific alarm calls. Attention to the 1674 threat was estimated based on latency to begin feeding, expression of vigilance behaviours 1675 (*i.e.* head-up/head-down bouts), and the rate of head-up bouts during the first feeding bout 1676 following the alarm call. Birds with access to a water bath, important for feather health and 1677 maintaining the ability to escape predation (Brilot, Asher, et al., 2009), had shorter latencies 1678 to feed and decreased head-up bouts compared to birds without baths. This suggests the 1679 latter exhibited heightened vigilance (in absence of a visible threat, or possibly it was due to 1680 an increased perception of vulnerability to predation due to poorer feather condition (Brilot 1681 & Bateson, 2012)). Studies in domestic chickens used a similar test set-up (i.e. playback of 1682 conspecific alarm call in presence of food), however latency to resume feeding following the 1683 threat stimulus was deemed as a poor measure of AB when testing highly fearful hens 1684 (compared to low fearful group of hens) as these birds failed to eat at all during testing 1685 (Campbell et al., 2019a). These studies demonstrate that, besides visual stimuli, AB to threat 1686 can be induced by stimuli of other modalities, *e.g.* auditory cues. However, using auditory stimuli could make it even more challenging to measure AB, due to the difficulty in measuring 1687 1688 the attentional response to auditory stimuli.

1689 Lastly, **distractibility tests** measuring attentional shift away from task-relevant information to 1690 task-irrelevant cues (distractors) have been used as a proxy of attention. For example, 1691 Trevarthen et al. (2019) tested mice's attentional response to unexpected, valenced, task-1692 irrelevant distractor cues (*i.e.* positive – food; negative - light flash) appearing within the 1693 runway the animals were previously retrained to traverse to return to their home cage. 1694 Overall, mice were slower to approach the light and spent more time with the food suggesting that the task-irrelevant cues had the predicted distracting effects. However, negative affect 1695 1696 manipulation (*i.e.* tail handling as opposed to tunnel handling) reduced mice's latencies to 1697 approach both test stimuli against the authors' predictions, possibly because of modulating 1698 effects of emotional arousal rather than affective valence on mice's running speed to return 1699 to the home cage (Trevarthen et al., 2019).
1700 Table 2.2 Overview of commonly applied ABT in animals.

Experimental design	Schema	Stimuli	AB measure	Targeted attention stage
Dot-probe task	Congruent trial (-) (0) time (0) (0) (0) (0) (0) (0) (0) (0)	<i>E.g.</i> negative (-) vs. neutral (0) conspecific faces; Conditioned probe (blue dot) as target stimulus	AB to threat: RT to locate probe on congruent trials < RT on incongruent trials AB away from threat: RT to locate probe on congruent trials > RT on incongruent trials	Attention capture & orientation
Emotional-Stroop task	$\begin{array}{c c} \hline & \hline & \hline \\ \hline & \hline \\ \hline \\ \hline \\ \hline \\ \hline \\ \hline \\$	Conditioned stimulus = blue dot/frame; Test stimuli = negative (-) vs. neutral (0) conspecific faces; or non-social stimuli (<i>e.g.</i> images of food, capture gloves); or familiar humans caretaker, vet, stranger)	AB to threat: RT to touch frame with (-) > RT to touch frame with (0) or baseline (dot)	Later stages of attention and cognitive processing

Experimental design	Schema	Stimuli	AB measure	Targeted attention stage
Response slowing task	Training trial Target = grey square Image: Image in the second s	Conditioned stimulus = grey square; Test stimuli = negative (-) conspecific faces with direct gaze; Control stimuli = (0) conspecific faces with averted gaze	AB to threat: RT to touch square with (-) face images > RT to touch (0) face images	Later stages of attention and cognitive processing
Visual search task	A) Image: B) <	 A) Fear-relevant target (<i>e.g.</i> snake image) placed among distractors images (e.g. flowers) B) Fear-irrelevant target (<i>e.g.</i> flower) placed among negative distractor images 	AB to threat: RT to detect fear- relevant stimulus among distractors < RT to detect fear- irrelevant images among negative distractors	Attention capture; distraction

Experimental design Schema Stimuli Targeted AB measure attention stage Single presentation Threat stimulus AB to threat: Attentional test (*e.q.* dog) measured as looking engagement Food presented in and vigilance time to threat outside area with location, vigilance food behaviours, time to start feeding **Dual presentation** Negative (-) and AB to threat: Attention **RIGHT SCREEN** LEFT SCREEN neutral (0) attention duration test capture; conspecific face to (-) > Attention attention (preferential looking pairs shown duration to (0); disengagement Trial time test) simultaneously Latency initial gaze and avoidance (0) **10s** (-) to (-) < latency initial gaze to (0) AB away from threat: time Attention duration to (-) < Attention duration to (0)

Chapter 2 | Attention to emotional stimuli as an indicator of affect – a literature review

1702 **2.9.3** Influence of experimental mediators on attention and AB

From the above-mentioned studies, it is clear that AB in animals has been investigated using
diverse testing procedures. In the following, I will discuss how experimental dissimilarities,
particular those concerning the test stimuli, might influence animals' responses in ABTs.
Given the small number of auditory ABTs, I will only focus on studies using visual stimuli.

1707 **2.9.3.1 Biological and acquired test stimuli**

1708 From an evolutionary perspective, it is unsurprising that biologically relevant (innate) signals 1709 (e.g. predator or prey cues, social cues) are more likely attended to than non-biological (i.e. 1710 learnt through association) information. Lacreuse et al. (2013) reported that humans, but not 1711 macaques, showed attentional avoidance of images associated with threat (*i.e.* syringes, guns, 1712 or arms), while both species exhibited AB toward negative social images. These findings 1713 indicate that negative AB may not be generalizable across threat stimuli. Moreover, 1714 experience (*i.e.* previous learning and memory) with associated threat cues might be an 1715 important modulator of attention. Testing anxiety-related AB in humans using disorder-1716 congruent threat stimuli (e.g. trauma-related cues in posttraumatic stress disorder) show 1717 larger effects than tests with disorder-incongruent threat stimuli (Mathews & MacLeod, 1994; 1718 Pergamin-Hight et al., 2015). In animals, it might be challenging to reliably test associations 1719 between non-biological stimuli. One common method is to observe approach/avoidance 1720 behaviours, but in Boggiani et al.'s (2018) study tufted capuchin monkeys approached food 1721 reward placed below the image of a person with a capture glove (presumably negatively 1722 associated as animals were routinely captured with the depict equipment) faster than 1723 rewards placed by the image or a familiar student. Hence, in this example, stimulus appraisal 1724 and behavioural response go against each other unless the monkeys interpreted the students' 1725 direct gaze as more threatening than the face of the vet whose face was covered by a 1726 protective hat.

1727 It should be noted that apart from emotional signals, faces also convey other social cues such 1728 as dominance, sex and familiarity which could mediate attention (Lewis et al., 2021). In 1729 bonobos and chimpanzees, social attention is guided by sex-based dominance, but AB to 1730 images of dominant sex faces is also dependent on familiarity (Lewis et al., 2021). Similarly, 1731 the age of the stimulus animal can influence social attention. Sato et al. (2012) reported a 1732 visual preference for infantile features in nonhuman primates equivalent to the human

preference for baby schema (Sato et al., 2012). Whether other species might show anattentional preference for younger than older faces needs further investigation.

1735 Testing sheep's vigilance and attention to a live dog vs. conspecific images, no significant 1736 difference between treatment groups (*i.e.* pharmacologically induced 'anxious', 'happy', or 1737 'calm' sheep) was found by Monk et al. (2019), possibly because older animals were tested 1738 which all had previous experience with dogs during routine farm procedures which might 1739 have influenced their attentional responses. However, given that sheep were exposed to a 1740 social and a non-social stimulus, the animals' behaviour could have also been modulated by 1741 individuals' motivation for social contact (assuming the images were indeed perceived as 1742 conspecific) since the test sheep were visually, but not socially, isolated from conspecifics.

When selecting test stimuli, whether intra- and/or inter-specific stimuli are used might need consideration. Dogs fixate on images of conspecific faces for longer than images of human faces, suggesting sensitivity differences for processing intra- and inter-species cues (Somppi et al., 2012). This idea is supported by later findings that conspecifics' faces evoked heightened attention in dogs but threatening human faces elicited an attentional avoidance response (Somppi et al., 2016).

Overall, the difference in AB responses to biological and acquired stimuli might not be easily differentiated, and individuals might vary in their experiences with acquired stimuli or exposure to biological cues. Nonetheless, the potential influence of the above-addressed factors on animals' attention should be considered.

1753 2.9.3.2 Stimulus intensity

1754 Besides stimulus valence, the intensity of emotional information (i.e. arousal) needs to be 1755 considered as both aspects influence stimulus saliency (Lundqvist et al., 2014). High-arousal 1756 stimuli enhance negative attention bias relative to low-arousal stimuli (Yuan et al., 2019). For 1757 instance, RTs in a human pictorial dot-probe task was found to be linearly slowed down with 1758 increasing stimulus intensity (*i.e.* high threat – mutilated face; mild threat – men with a knife; 1759 neutral – hair dryer; (Koster et al., 2004). Hence, it might be crucial to control for stimulus 1760 intensity when testing valence-related AB responses in animals. Practically speaking, this 1761 might be difficult to achieve since the verbal account of the perceived stimulus intensity is the 1762 most direct measure in humans. Maybe psychophysiological measures, such as changes in

1763 heart rate or skin temperature measured non-invasively via infrared thermography could be 1764 used to assess stimulus intensity (van Rooijen et al., 2017). In primates, a decrease in facial 1765 skin temperature (e.g. nose) indicates states of negative high arousal (Parr, 2001). Another 1766 option could be measuring pupillary diameter changes, e.g. in conjunction with eye tracking, 1767 as larger pupil diameter is associated with activation of the SNS reflecting greater emotional arousal (van Rooijen et al., 2017). Since these techniques allow non-invasive and remote 1768 1769 monitoring of physiological responses, it might even be useful to measure attentional and 1770 bodily responses simultaneously and in real time.

1771 **2.9.3.3 Live stimuli and artificial stimuli**

1772 Artificial (e.g. screen images, photos), as well as live stimuli, are used in animal AB studies. 1773 Both have considerable advantages, but also significant shortcomings. live stimuli (e.g. dog 1774 and food in Lee et al., 2016, 2018; Monk et al., 2018a,b, 2019) should possess all perceptual 1775 characteristics to evoke innate attentional responses, but their controlled presentation can 1776 be challenging (e.g. a dog might move or bark more/less across trials or animals, as was 1777 assessed in Monk et al., 2018b). Artificial representations of objects/individuals (e.g. computer images, photos) guarantee consistent stimulus presentation across trials or test 1778 1779 animals, but often only display a censored amount of sensory information (see Griffin et al. 1780 (2000) for how this affects training for avoiding predators). For instance, faces are presented 1781 on a blank background (e.g. Bellegarde et al., 2017a,b; Bethell et al., 2012; Howarth et al., 1782 2021; Wathan et al., 2016) without the attached body, but in real life, faces are always 1783 encountered in the presence of a body. Hence, using stimuli that include bodily expressions 1784 rather than isolated faces might add to the ecological validity of the task (van Rooijen et al., 1785 2017). However, standardization of all aspects of the visual scene as much as possible is also 1786 critical.

When using artificial stimuli, it is essential to know how the test animals perceive these stimuli since electronic devices (*e.g.* computer screens, projectors) are designed for human vision, hence may not produce the same visual information impressions in non-human species (Fleishman & Endler, 2000). Perceptual differences might also occur depending on the presentation medium. Image recognition in sheep has been reported in several studies using printed images, but Raoult & Gygax (2018) concluded that computer projections may not be suitable representations of valenced stimuli in sheep, as animals did not respond significantly

1794 differently to videos of dogs (negative) and conspecifics (positive). Instead, naturalistic cues 1795 (*i.e.* acoustic recordings of a dog barking, or sheep bleating) appeared to be somewhat more 1796 appropriate for sheep, as negative mood enhanced sheep's attention to the negative sounds 1797 (Raoult & Gygax, 2019) as would be predicted. Hence, for species where the perceptual 1798 abilities are not fully understood, it might be more sensible to rely on stimuli resembling 1799 naturalistic cues as much as possible. When test stimuli differ not only in terms of their 1800 emotional information but also in other stimulus aspects, *e.g.* real-life non-social cue (dog) 1801 vs. artificial social cue (conspecific image) in Monk et al. (2018b, 2019, 2020), it might be 1802 difficult to explain attentional response exclusively by the emotional content they putatively 1803 convey.

1804 Winter et al., (2015) proposed an intriguing solution to improve stimulus presentation in free-1805 looking time tasks. They suggest that future studies could use robotic animals as they have 1806 advantages over real or two-dimensional stimuli in that their appearance can be controlled 1807 and manipulated independently from their behaviour, in other words, they can move and 1808 'behave' in a standardized fashion (Winters et al., 2015). A range of animal robots are already 1809 used as social cues in the field of animal cognition research (Frohnwieser et al., 2016) so their 1810 use as stimuli in ABTs seems reasonable, as long as they copy all sensory characteristics (e.g. 1811 visual appearance, smell, acoustic properties, which would need scientific validation) of their 1812 real-life counterparts.

1813 2.9.4 Subject-related mediating factors

1814 Several individual characteristics, as well as underlying emotional states, might mediate AB 1815 responses in animals. Mandalaywala et al. (2014) reported that early life experience (i.e. 1816 protectiveness and social rank of the mother) modulates infant rhesus monkeys' attention to 1817 threatening social stimuli. Nine-months old animals whose mothers were high-ranking and 1818 more protective displayed a greater vigilance for threat than infants of lower-ranking and less 1819 protective mothers (Mandalaywala et al., 2014), indicating that social attention might be 1820 shaped directly (*i.e. via* protectiveness) and indirectly (via mother's social rank). In pigs (Sus 1821 scrofa domesticus), however, rather than early life conditions (barren vs. enriched housing), 1822 personality (*i.e.* reactive or proactive coping style) appears to mediate animals' attention to 1823 threat (*i.e.* loud noise, movement of a door being raised up and down, and flashing light, Luo 1824 et al., 2019). Personality has also been investigated as a mediating factor in chimpanzees'

responses in an emotional Stroop-like task, but personality traits (*i.e.* anxiety/aggression) as rated by animal caretakers and research assistants were not significantly associated with animals' task performance (Allritz et al., 2016). In addition to the above, general characteristics such as the age and sex of the test animal might influence attentional responses to the images of conspecifics (Lewis et al., 2021; Parr et al., 2013).

1830 **2.10 Conclusion**

1831 Based on the reports of attentional prioritisation in animals in response to emotional stimuli, 1832 as well as evidence of affect-mediated shifts in attention, I conclude that testing AB has the 1833 potential to further our understanding of how animals perceive their external environments 1834 and ultimately animals' affective state. However, several important experimental limitations 1835 currently hamper the validity of AB as a proxy for affect. Here, I identified several 1836 experimental discrepancies, including differences in test paradigms, outcome measures and 1837 the type of stimuli used (e.g. innate or associated, social or non-social, real-life or artificial 1838 cues). These discrepancies are likely to contribute to the heterogeneity of existing AB findings. 1839 Hence, by considering these problems and developing more refined species-appropriate test 1840 protocols, and being able to predict emotion-dependent AB direction, the evaluation of 1841 animals' attentional response to emotional stimuli could be an insightful method to assess 1842 animal emotions in future studies. With this in mind, I developed a visual AB test in horses 1843 described in Chapter 4 using competing biologically relevant stimuli (conspecific face images).

Chapter 3 | General methods 1844

1845 Summary

1846 This chapter gives details about the general methods applied throughout my thesis, including 1847 a description of the study population and sites, explanations of pilot tests and experimental 1848 protocol development, and the assessment of horses' welfare (condensed into an individual 1849 score). Final methods and experimental set-ups are detailed in the experimental chapters.

1850 3.1 **Ethical approval**

1851 All experiments were approved by the Animal Ethical Review Committee of the University of 1852 Plymouth (ETHICS-41-2020; ETHICS-42-2020). All experimental procedures were below the 1853 threshold for regulation by the UK Animals (Scientific Procedures) Act 1986 (ASPA). Animal 1854 owners consented to the use of their animals and were responsible for the care and health of 1855 their horses. Data collected consisted of non-invasive measurements, *i.e.* behaviour 1856 observations, and recording of heart rate and surface temperatures. All horses were 1857 habituated to the experimental equipment before testing. In Chapter 5, positive 1858 reinforcement (i.e. food) was used for training/testing the animals. In Chapter 6, tests 1859 comprised trials with immediate or withheld food rewards. The number of withheld reward 1860 trials was kept at a minimum (*i.e.* 2x10 trials in two sessions) determined through pilot studies 1861 (see section 3.2). After testing, all horses received a food reward (carrots) to ensure a positive 1862 association with the experimenter.

1863 Justification of animal numbers (also see section 3.3), reasons for animal exclusions, blinding 1864 and randomisation are specified in each experimental chapter. All horses remained at their 1865 home facility at the end of the experiments or were relocated for reasons not associated with 1866 my studies.

Pilot studies and conceptualising of experiments 1867 3.2

1868 UK riding school horses were chosen as the study population as this cohort is likely to 1869 experience varying degrees of welfare states associated with routine housing and 1870 management procedures (Hotchkiss et al. 2007; Hockenhull and Creighton 2015; Horseman et al. 2016, see 1.3.4.2 for details). Pilot studies were conducted at the Mare and Foal 1871 1872 sanctuary in Totnes (Devon, UK) between January and December 2019 using rescue horses

1873 (N=15, various breeds, 9.78±4.42 years) with a history of poor welfare states of varying 1874 degrees. However, only animals deemed rehabilitated (*i.e.* ready to be rehomed) were 1875 enrolled in my pilot studies and their management corresponded to common housing 1876 practices of UK riding school horses (*i.e.* kept in single stalls allowing varying degrees of 1877 physical contact to conspecifics, limited pasture access, restricted roughage intake).

1878 My first pilot study explored applying auditory social cues as emotional stimuli in an attention 1879 bias test. Horses are sensitive to vocalisations of conspecifics and show significantly different 1880 behavioural and physiological reactions to positively/negatively-valenced conspecific calls 1881 (Briefer et al. 2015). I exposed the horses to the vocalisations of unfamiliar conspecifics whilst 1882 feeding with their heads down in their stalls. The recordings were obtained from six horses 1883 housed at a different yard, Duchy College, Cornwall, during food anticipation (*i.e.* positive 1884 vocalisations (whinnies, nickering)) and social encounters (*i.e.* negative vocalisations (squeals, 1885 grunts)). All test stimuli appeared to induce similar behavioural responses (*i.e.* head raise and 1886 turn towards the source of the sound with ears pointing forward) with no apparent 1887 behavioural differences between test cues, which made it difficult to judge whether horses 1888 truly perceived the sounds as differently valenced. As background noises (e.g. people talking 1889 in the distance, door banging, bird sounds) were occurring during testing it was not clear 1890 whether the horses' responses were elicited by the test stimuli or other farm sounds. I 1891 hypothesised that valenced calls of unfamiliar conspecifics would capture horses' attention 1892 more than the other farm noises assuming that horses were habituated to the latter and had 1893 no associations with certain background noises I was unable to control during testing (e.g. 1894 bird sounds).

1895 To test this, I investigated whether the test stimuli would cause changes in horses' behaviour 1896 whilst performing a cognitive task, similarly to Rochais et al. (2017a), who tested horses' 1897 distractibility with novel sounds in relation to their performance in a work task (i.e. lunge 1898 training). Here, the horses were trained to perform a visual discrimination task comprising a 1899 flashing light cueing one of two possible reward locations (i.e. left or right hole within a 1900 wooden wall erected in a familiar pen). The aim was to play the test stimuli whilst the horse 1901 was approaching the mid-point between the start line (opposite end of the pen at 10 m 1902 distance) and the reward locations. I measured the effects of the test cues on latency and 1903 response accuracy, to infer horses' attention to the test sounds (*i.e.* distractibility from the

1904 ongoing task). However, none of the horses had learnt the discrimination task after several 1905 weeks of training. Most horses returned to where they received a reward last, and hence 1906 seemed to rely on spatial (location of reward on the previous trial) rather than visual (location 1907 of flashing light) cues in this task. Furthermore, it was challenging to control background 1908 noises occurring on the farm.

1909 For these reasons, I decided to develop a visual AB test requiring no learning. Initially, I 1910 planned to use pairs test images comprising either social (*i.e.* facial expression of conspecifics 1911 (Wathan et al. 2016, Bellegarde et al. 2017a,b) or non-social cues (e.g. familiar food bucket 1912 (McBride & Morton, 2018)) and present side-by-side images putatively differing in emotional 1913 valence as previously done in other animals (e.g. macaques (Bethell et al., 2012)). Matching 1914 object images to ensure consistency in luminescence and pixelated areas to create stimulus 1915 pairs only differing in the emotional content was not possible. Therefore, I only used image 1916 pairs of unfamiliar conspecifics displaying facial expressions associated with different 1917 emotional situations in the final test set-up described in Chapter 4.

1918 After piloting the AB test, I wondered to what extent horses indeed recognised the on-screen 1919 images as representations of conspecifics. Several horse studies have reported that horses 1920 possess image recognition ability (Hanggi, 2001; Lansade et al., 2020a; Lansade et al., 2020b; 1921 Smith et al., 2016; Trösch et al., 2020), but the scientific evidence might be questioned in light 1922 of several experimental limitations (Amici, 2019; see Chapter 5 for an in-depth discussion). 1923 Moreover, my pilot horses attended to the images, but did not respond to them as if meeting 1924 an unfamiliar conspecific (*i.e.* approaching, vocalising (whining)), which would have indicated 1925 that they perceived the images as other horses. To test horses' spontaneous response to on-1926 screen images in a more controlled manner, I piloted the object recognition test described in 1927 Chapter 5 in the same yard with the same pilot horses.

The results derived from the cognitive tests motivated me to explore physiological parameters when measuring horses' emotional states during testing. I collaborated with Dr Marco Ramirez Montes De Oca (University of Bristol), who has expertise in Infrared Thermography (IRT) and measuring emotion-related facial surface temperature changes in cattle, to test short-term changes in emotions during testing (as opposed to longer-lasting mood states measured through welfare scores that I addressed with measuring AB). To induce changes in horses' emotional states, horses received a food reward supposedly inducing a

more positive emotional state, or reward was withheld whilst horses performed the conditioned response behaviour they had learnt in the object recognition test (Chapter 5). To ensure that horses would still perform the response following a period of withheld reward trials (for controlling order and time effect when collecting physiological data), we conducted pilot trials with 8 horses. We found that animals stopped performing the conditioned response following an average of 10 withheld reward trials, which is the number of trials we tested horses in the test sessions described in Chapter 6.

1942 **3.3 Calculation of sample sizes**

1943 The sample size for the AB experiment (Chapter 4) was estimated through power analysis (80 1944 % power, 2-tailed, $p \le 0.05$) using the mean durations horses attended to the test images 1945 during pilot testing. The estimated required sample size was 113. I aimed to recruit horses 1946 from six yards with approx. 20 animals from each. However, due to COVID-19 restrictions, I 1947 had to reduce yard numbers and tested fewer animals. Animal number in Chapter 5 (image 1948 recognition test) was dependent on the number of horses reaching learning criterion (*i.e.* not 1949 based on power calculations, but animal number was significantly bigger than previous horses 1950 cognition studies (e.g. N=3 in Hanggi (2001), 2001; N=11 in Lansade et al. (2020b)). Horse 1951 numbers for testing physiological changes presented in Chapter 6 was determined through 1952 power analysis using data collected in an IRT study (Ramirez Montes De Oca, 2021) on calf 1953 disbudding and social separation (G-power, F-tests, f^2 =0.3, α =0.05, 0.8 power, number of 1954 predictors 1) suggesting an estimated number of 25 animals.

1955 **3.4 Study population recruitment and yard descriptions**

To assess the influence of welfare conditions on cognitive processing, I aimed to recruit yards varying in their housing/management procedures. I searched for facilities keeping their horses in conditions relative close to the *naturalistic* lifestyle of this species (*i.e.* in stable groups, with free access to roughage and outdoor space (Waring, 2003), or comparatively more *restrictive* meaning that horses' access to social, food or environmental resources was limited.

1962 I identified potential study sites within a 20-mile radius around Bristol *via* web search and
1963 word of mouth (*i.e.* asking horse-owning colleagues and yard owners for recommendations).
1964 Following email contact, I initially visited the sites to evaluate their suitability (*i.e.* whether an

indoor arena/weather-protected space existed for AB testing) and feasibility in terms of
animal availability (yards with >10 animals). For 11 months (June 2020 to May 2021), I
conducted experiments at three different facilities located in North Somerset and Bristol.

1968 Data were collected at the first facility (yard 1) between June and August 2020 using 11 1969 privately owned horses (14.72±6.06 years, various breeds) which were kept together in one 1970 group outdoors all year round on a pasture/track system in one stable herd (Figure 3.1 A). 1971 Grass and hay were available ad libitum. Each animal received supplement feed (brands 1972 varied based on owners' choices) every morning. Water was always freely available via 1973 automatic drinkers. Animal health was checked regularly by the owners. Horses were ridden 1974 regularly (approx. 2-7h a week) but hours and type of work (*i.e.* 'hacking', trail riding, or 1975 pleasure riding) varied between animals and between the weeks of data collection. Overall, 1976 these horses were housed in conditions relatively close to equines' ethological needs (i.e. 1977 *naturalistic*) compared to animals from the other study sites.

1978 I collected data in yard 2 in July – August 2020. All horses (N=17, 10.76±2.63 years, various 1979 breeds) were riding school horses and were kept in single stalls (9.87±2.09 m², wood-shaving 1980 bedding) with limited social contact (i.e. animals had visual but no physical contact with 1981 neighbouring conspecifics, Figure 3.1 B, apart from three horses that also had head-to-head 1982 contact with their neighbour). Access to pasture was provided within the same group of 1983 conspecifics (3 groups of >10 animals; some horses not enrolled in my studies) during parts 1984 of the day and/or at night (depending on weather conditions and riding lessons, the duration 1985 varying between 2-16 h/day). Roughage amount was restricted as horses were fed loose hay 1986 2-3 per day (amount roughly adjusted based on horses' weight). Water was freely accessible 1987 through automatic troughs. No horse received additional hard feed. All animals were on a 1988 'diet' for three months before my data collection due to COVID-19-related reduction in 1989 workload, hence horses were not ridden as often as before the pandemic, but exercised for 1990 at least 30min during weekdays with some horses additionally being used for riding lessons. 1991 The maximum workload was 10h per week.

At yard 3, animals (N=19, 16.83±6.09 years, various breeds) were riding school horses kept in three different barns. Two barns housed horses in single stalls (approx. 4.7x2.75 m², straw bedding, Figure 3.1 C) which allowed visual contact but no physical contact among horses (except for one tall horse that could lift its head over the stall wall and have head-to-head

1996 contact with its neighbour). The third barn contained tie-stalls (approx. 2x2 m², rubber mats 1997 covered with thin layer shavings/straw, Figure 3.1 D) allowing the ponies to have head-to-1998 head contact with neighbours (if present) and to lie down, but movements were otherwise 1999 limited by a rope (approx. 1.5 m long). All horses were provided with hay nets 2-3x per day 2000 and received an adjusted diet (Thunderbrook Equestrian) once per day. Water was provided 2001 via buckets. Pasture access was limited, and varied across days and horses (i.e. kept on 2002 pasture during the night, weather permitting, and/or parts of the day when not used for riding 2003 lessons). Data were collected here from September 2020 – April 2021, but paused between 2004 November 2020 – April 2021 as all horses were moved to pasture (full-time) during these months due to COVID-19-enforced yard closure during the 3rd UK lockdown. More 2005 2006 information to each test horse is provided in Table 3.5 (at the end of this chapter).



2007

2008 Figure 3.1 Example images of housing conditions in (A) yard 1, (B) yard 2 and (C+D) yard 3.

2010

3.5 Horse welfare assessment

2011 As explained in the introduction, I follow the view of others that affective states, and 2012 therefore subjective states, are central to animal welfare (e.g. Duncan 2005; Browning 2022a, 2013 b). As emotions are elicited by stimuli of varying characteristics (external or internal, social or 2014 non-social, animate or inanimate), I decided to examine horses' living conditions as 2015 comprehensively as possible. Several environmental, management, health and social factors 2016 have been found to compromise equine welfare (Hockenhull & Creighton, 2015; Horseman 2017 et al., 2016; Hotchkiss et al., 2007). Animal welfare is a multifaceted concept linked to many 2018 different aspects encompassing an animal's life, but there is currently no scientific approach 2019 to measuring 'welfare state' that allows balancing all contributing relevant factors (Appleby 2020 et al., 2018, p. 18). Therefore, a combined evaluation of all these different factors might be 2021 the most reliable approximation of individuals' welfare state currently possible, while the 2022 direct measurement of welfare is still unattainable (Browning, 2022a,b). I developed a horse 2023 welfare assessment protocol to synthesise a single score, putatively reflecting individuals' 2024 potential to experience positive (or less positive) emotional states concomitant to their 2025 welfare state. This score was used to investigate the relationship between welfare, and 2026 cognitive processing and emotions in Chapter 4 (AB test) and 5 (image recognition test), and 2027 to explore individual effects on physiological changes in Chapter 6.

2028 **3.5.1 Welfare factors**

2029 Relevant welfare factors were derived from the AWIN welfare assessment protocol for horses 2030 (AWIN, 2015), Dalla Costa et al. (2014), and other equine welfare studies (see references in 2031 Table 3.1). The assessment of certain welfare aspects recommended in AWIN (2015) was 2032 modified (*i.e.* human interaction test, behavioural assessment) or excluded (*i.e.* the horse 2033 grimace scale, tactile sensitivity testing) for reasons stated in sections 3.6.1 and 3.6.2. I 2034 applied a multi-component approach comprising environmental, social and animal-based 2035 assessment factors (Table 3.1). For each factor, a score of 0–1 was assigned depending on the 2036 presence/absence of the resource. For factors with more than two possible outcomes, fractal 2037 values were assigned to each possible condition (e.g. social contact score in Table 3.1). All 2038 factors were added to create a single welfare score with a maximum value of 20 with higher 2039 scores reflecting a greater likelihood to experience positive welfare, and therefore better 2040 chances of positive emotional states.

Table 3.1 Description of welfare factors, which all equally weighted into the final welfare score.

2042 References to relevant literature are given in brackets. * indicates all animal-based indicators derived 2043 from the AWIN welfare assessment protocol for horses.

#	Welfare indicator Answer Description [References] option		Description of answer option	Score
Envi	ronmental factors			
1	Confinement [McGreevy 1995, Heleski et al.	Always	The horse has no outdoor access or is outside for less than 12h per week (excluding work, husbandry procedures).	0
	2002, Henry et al. 2017, Loeckener	Part-time	The horse has outdoor access for at least 12h per week (excluding work, husbandry procedures).	0.5
et al., 2015]		Never	The horse is always kept outdoors (<i>e.g.</i> on paddock or on pasture) (excluding work, husbandry procedures).	1
2	Stall space sufficiency*	No	Stall dimensions are smaller than required according to the British Horse Society (BHS) guidelines.	0
		Yes	Stall dimensions are within or bigger than required according to the BHS guidelines. Horses always kept outdoors were given a score of 1.	1
3	Water availability*	No	Water is not always freely accessible or manually filled (<i>e.g.</i> buckets).	0
		Yes	Water is always freely accessible (<i>e.g. via</i> an automatic system).	1
4	Ventilation	No	No ventilation is provided.	0
	Holocombe et al., 2001]	Partially	Ventilation is provided during parts of the day but not always (<i>e.g.</i> barn doors are shut at night preventing constant airflow).	0.5
		Always	Ventilation is always provided (<i>e.g. via</i> an air- conditioning system or horses are kept outside).	1
5	Forage opportunity	Never	The horse has no opportunity to forage.	0
	[Sharpe et a., 2019, Ninomiya et al., 2004]	Restricted	The horse has restricted (time/amount) forage opportunity through the provision of forage at least twice per day.	0.5
		Always	The horse can always forage.	1

#	Welfare indicator [References]	Answer option	Description of answer option	Score
6	Forage diversity	rage diversity No No forage diversity is available.		0
	[Goodwin et al., 2002, Ninomiya et al 2004]	Yes, part- time	The horse can choose between forage types for parts of the day (<i>e.g.</i> grass and hay available on pasture) but not full-time (<i>e.g.</i> , indoors only hay was provided).	0.5
		Yes	The horse has a choice in forages through the provision of at least two forage options (<i>e.g.</i> hay available on grass pasture).	1
7	Bedding type [Nazarenko et al., 2018, Wehrhahn et al., 2010,	No bedding/ Rubber mat only	No bedding is provided.	0
	Pederson et al., 2004, McGreevy	(Wood) shavings	Wood shavings are provided as bedding.	0
	1995]	Straw	Straw is provided as bedding.	1
8	Visual scope	Never	The horse cannot broaden its visual scope (<i>e.g.</i> no window, closed stable doors)	0
	2008, Ruet et al., 2009]	Part-time	The horse can broaden its visual scope (<i>e.g. via</i> a half-closed stable door, windows, access to pasture) for parts of the day.	0.5
		Always	The horse always can broaden its visual scope (<i>e.g. via</i> windows, half-open stable door, full-time on pasture).	1
Socia	ıl indicators			
9	Social contact (daily duration and level of	<u>Full-time</u> no social contact	The horse never has social contact.	0
	physical contact) [Lansade et al., 2007, Yarnell et al., 2015, Lesimple et al., 2019]	Part-time no social contact and part- time restricted	The horse has no social contact at times (<i>e.g.</i> in the stable) but has restricted (<i>e.g.</i> only visual) contact with other horses during parts of the day (<i>e.g.</i> top half of the stable door is opened during the day allowing it to see other horses).	0.25
		<u>Full-time</u> restricted	The horse can sniff and/or view other horses <i>e.g.</i> through stable bars but cannot physically interact further than sniffing with other horses.	0.5

#	Welfare indicator [References]	Answer option	Description of answer option	Score
9	Social contact (continued)	Part-time restricted and part- time full	The horse is restricted in interacting with other horses (<i>e.g.</i> in the stable) but full social contact (full-body) with others for parts of the day (<i>e.g.</i> on pasture).	0.75
		<u>Full-time</u> social	The horse always has full social contact (full-body) with other horses.	1
10	Social stability (opportunity to maintain social relationships)	Never	The horse has no social contact or social partners are changing (<i>e.g.</i> horses are allocated to neighbouring stables at random, mixing neighbours frequently).	0
	[Fureix et al., 2013, Waring, 2003]	Partially	The horse has the same social partner(s) inside the stable but social partner(s) outside vary frequently (or <i>vice versa</i>).	0.5
		Full-time	The horse is kept with the same social partner(s) continuously.	1
Anin	nal-based indicators*			
11	Signs of practice- induced injuries	Yes	Injuries and/or healed scars (<i>e.g.</i> fresh skin lesions, white hairs, and bald skin patches) likely caused by management procedures or ill-fitting tack (<i>e.g.</i> lesions mouth corners, white hair on withers) are present.	0
		No	The horse has no signs of injuries linked to management procedures.	1
12	Signs of other injuries	Yes	Injuries and/or healed scars (<i>e.g.</i> fresh skin lesions, white hairs, bald skin patches) bigger than 1x2cm ² or longer than 4 cm are not likely to be caused by management procedure (<i>e.g.</i> bite marks on the neck) are present.	0
		No	No injuries are present.	1
13	Nasal discharge	Yes	Flow of discharge (watery, transparent or yellow/green) is present from one or both nostrils.	0
		No	Nostrils have no discharge.	1
14	Lameness	Present	The horse's gait is altered with uneven weight bearing (limp) whilst standing or walking. If the horse received treatment for lameness during the time of the study it was scored as lame regardless of whether it presented as lame.	0

#	Welfare indicator [References]	Answer option	Description of answer option	Score
14	Lameness (continued)	Not present	The horse is walking/standing with even weight bearing on all four feet and was not treated for lameness during the study.	1
15	Swollen joints	Present	Swelling of joints is evident in increased diameter of one or more limbs.	0
		Not present	No swelling of joints present.	1
16	Manure consistency	Abnormal	Manure consistency is different from the normal (shaped) texture (<i>e.g.</i> too watery (cow dung-like with no shape or too dry, <i>i.e.</i> manure very hard).	0
		Normal	The horse's manure has normal consistency.	1
17	Abnormal hoof condition/ shape	Present	One or more hooves are overgrown, cracked or misshaped.	0
		Absent	No signs of hoof abnormalities are present.	1
Othe	er indicators			
18	Abnormal behaviour (assessed during behaviour	Present	The horse displayed a form of abnormal behaviour (<i>e.g.</i> teeth grinding, cribbing, head tossing, abnormal posture during resting, <i>i.e.</i> withdrawn state) during observation periods.	0
	observation during 5x1h sessions) [Dalla Costa et al., 2016]	Absent	The horse was not observed displaying abnormal behaviours during observation periods.	1
19	Owner information on horse health [†]	Acute or chronic condition present	The horse had an acute or chronic condition possibly compromising welfare at the time of the study.	0
		Reported by the owner as healthy	The owner stated that the horses had no known underlying health conditions.	1
20	Human interaction (voluntary (VAT) an approach test (FAT) combined) [Fureix o Dalla Costa et al., 2	score* Id forced) results et al., 2010; 016]	See section 3.5.2.2. for the description.	

2044 + We asked owners to provide information on horses' health history as we were unable to schedule a veterinarian to perform health checks due to monetary constraints.

2047

2048 **3.5.2.1** Scan sampling of abnormal horse behaviour

3.5.2 Welfare factors modified from the AWIN protocol

2049 Similar to other captive non-domesticated species, atypical, stereotypic (ST) and abnormal 2050 repetitive behaviours (ARB) are commonly observed in horses (Hothersall & Casey, 2012; 2051 McBride & Hemmings, 2009; McGreevy et al., 1995a; Roberts et al., 2017). The development 2052 of STs and ARB has been related to animals' past or present experiences of suboptimal 2053 environments (horse reviews: Hothersall and Casey 2012; Roberts et al. 2017; Krueger et al. 2054 2021) resulting in frustration-like stress (Mason et al., 2001; Mason & Latham, 2004) when 2055 animals are prevented from performing highly motivated behaviours (e.g. feeding, mating, 2056 migration; Mason and Rushen 2006). However, STs might also be linked to malfunction of 2057 different brain regions (e.g. alternations of dopamine receptors in the mesoaccumbens 2058 dopamine pathway (McBride & Hemmings, 2005; McBride & Hemmings, 2009)), although it 2059 is still unknown whether the neuroanatomical changes are an effect rather than a cause of 2060 STs (Tatemoto et al., 2022). Interestingly, horse breeds such as thoroughbreds (Bachmann et 2061 al., 2003; Ninomiya et al., 2007) and warmbloods (Wickens & Heleski, 2010) seem to be more 2062 susceptible to the development of STs than others. On the other hand, these breeds are used 2063 more in performance disciplines (e.g. dressage, racing) associated with intense management 2064 regimes, which could also explain the increased prevalence of STs in these breeds (Roberts et 2065 al., 2017). Manifestations of STs and ARB can have secondary welfare implications, *e.g.* by 2066 causing physiological pathologies, such as pain and colics (Archer et al., 2008), excessive wear 2067 of incisors (McBride & Hemmings, 2009), or muscle fatigue (Ninomiya et al., 2007). In 2068 addition, compromised equine welfare can be detectable as abnormal body postures (Fureix 2069 et al., 2010, 2011; Fureix et al., 2012; Lesimple et al., 2010). For instance, greater time spent 2070 displaying a so-called 'withdrawn' posture, characterized by standing motionless with eyes 2071 open with a stretched neck and a similar height between neck and back, might reflect a 2072 depression-like state in horses (e.g. predicting reduced sucrose consumption, a proxy for 2073 anhedonia; Fureix et al. 2012, 2015).

I evaluated whether animals performed the above-mentioned abnormal behaviours by scan
sampling their activity every 5-min in 3x1h sessions/day (between 8-10 am, 11-1 pm and 2-4
pm) over 5 days, with the initial intention to create a time budget (Auer et al. 2021; ethogram
shown in Appendix B Table B1). Comparing time budgets across animals and/or yards,

however, appeared difficult to interpret. Indeed, resource availability (*i.e.* access to hay and pasture, which would directly influence *e.g.* amount of time spent feeding) varied both between yards and daily within the yard. Therefore, horses were instead scored in the welfare assessment protocol as displaying any forms of abnormal behaviours (*i.e.* STs, ARB, withdrawn postures) or not (#18 in Table 3.1).

2083 3.5.2.2 Human interaction score

2084 Since chronic back pain is associated with increased horses' aggressiveness toward humans 2085 (Fureix et al., 2010), I conducted two types of human-horse interaction tests and combined 2086 these results into a human interaction score (HIS, #20 in Table 3.1). The test protocols were 2087 adapted from the voluntary approach test (VAT) and forced approach test (FAT) described in 2088 Fureix et al., (2010). For the VAT, the experimenter entered the horse's stall but remained at 2089 the door motionless (i.e. approx. 2-3m distance to the horse), thereby looking at the horse's 2090 shoulder (whichever side was closest to the human) until the horse approached the 2091 experimenter or 5min elapsed. The FAT was conducted similarly, but the experimenter 2092 approached the horse's shoulder (whichever side was closest) by walking slowly towards the 2093 horse with her right arm stretched out in front of her and briefly touching the top of the 2094 horse's shoulder/withers. Horse behaviour was scored from videos by a single observer 2095 according following the ethogram summarised in Table 3.2.

Table 3.2 Scoring system of the voluntary and forced approach test. Descriptions apply to both testsunless stated otherwise.

Behaviour	Description	Score
Positive response	The horse stops feeding, lifts its head up and approaches the experimenter with forward-pointing ears, possibly sniffing the experimenter.	1
Attentive	The horse is continuing to feed but moves its ears forwards when the experimenter appear s, the head is lifted or/and its head is rotated towards the experimenter.	0.5
Feeding — inattentive (to the experimenter)	The horse is feeding (with its head down), no obvious change in behaviour when the experimenter appears. Ears remain in the same position as before the experimenter entered.	0

Behaviour	Description	Score
Negative - threat	The horse rotates its ears backwards , head possibly lifting or moving towards the experimenter or remains in the same position as before the experimenter entered.	-0.5
Negative -avoidant	The horse lifts its head, possibly startles upon entry of the experimenter, ears backwards and the horse actively moves away from the experimenter.	-0.5
Negative - aggressive	The horse lifts its head and approaches the experimenter (walking or with the head moving towards the experimenter) while displaying a threatening attitude (ears back and teeth showing).	-1

2098

2099 3.6 Excluded welfare factors

The AWIN (2015) protocol recommends several welfare factors, which I excluded from my assessment protocol either for practical reasons (*i.e.* eye discharge, since horses at yard 1 worn fly/eye masks during the assessment) or for reasons explained below.

2103 3.6.1.1 Horse grimace scale

2104 Dalla Costa et al. (2014) developed a horse grimace scale (HGS) to standardise the measurement of pain using facial expressions, previously validated for several other animal 2105 2106 species (e.g. mice, pigs, and sheep, reviewed in Evangelista et al. 2022). The AWIN protocol 2107 recommends using the HGS to evaluate the presence of pain in assessed horses. For intra-2108 and inter-observer comparison, I photographed horses' faces (but in yard 1 horses since they 2109 wore fly masks) when horses were resting in their stalls. All images were taken from a lateral 2110 angle perpendicular to the horse's eye (see image in Table 3.4) of the left/right head side. I 2111 quietly approached the horse's stall and took images at approx. 2–4 m distance to the horse 2112 from outside the stall without visibly disturbing the animal. When possible, images of both sides of the head were taken, but some animals always rested in the same 2113 2114 position/orientation towards the stall door, meaning that the image sides/numbers were not 2115 balanced within and across horses. Images were taken on three different days (*i.e.* max. six 2116 images/horse) resulting in 98 images in total (yard 2 N=67; yard 3 N=31). These were 2117 independently scored by three raters (via an online questionnaire; one not blinded; two

blinded). For intra-observer reliability testing, 30 images (30%) were randomly selected,
which were repeated within set of the original test images. Each rater scored six facial action
units (FAUs, Table 3.3.) according to the HGS scoring system by Dalla Costa et al. (2014).

Intra-observer reliability was high (rater 1 0.81, rater 2 0.89, rater 3 0.83). However, interobserver reliability was overall moderate (Fleiss kappa 0.49) and varied largely within different FAUs (see Table 3.3). Inter-observer reliability was not improved *via* training (*i.e.* all three raters discussed/practised together FAUs scoring using five test images). Further discussion with Dalla Costa to clarify subtle difficulties in scoring 'tension above the eyes', 'strained mouth' and 'strained nostrils' did not improve inter-observer reliability, which is why I excluded the HGS results from my welfare assessment protocol.

2128

Table 3.3 Results of inter-observer reliability testing (weighted kappa) between the three independentraters.

FAU	Rater 1+2	Rater 1+3	Rater 2+3
Stiff backwards ears	0.31	0.604	0.549
Tension above the eyes	0.218	0.456	0.304
Eyes closure	0.534	0.824	0.575
Strained chewing muscles	0.424	0.512	0.44
Strained mouth	0.183	0.07	0.06
Strained nostrils	0.496	0.442	0.292

2131

2132

Table 3.4 Overview of facial action units (FAUs) evaluated for horse grimace scale scoring after Dalla Costa et al. (2014). For each FAU, the raters scored whether the description was obviously present, moderately present or obviously absent, following the exact wording of Dalla Costa et al. (2014) as shown in the table. The example image shows on of my test horses. Blue markings indicate the evaluated FAUs.



2139

2140

2142 **3.6.1.2** Mechanical nociceptive threshold

The majority of my test animals were riding school horses kept in restrictive living conditions 2143 2144 and musculoskeletal disorders can be highly prevalent in these animals (Fureix et al., 2010; 2145 Fureix et al., 2012; Lesimple et al., 2010, 2012, 2013). For instance, 73% (N=43) of riding school 2146 horses housed under restrictive conditions were severely affected by vertebral problems. 2147 Riding school horses might also be particularly prone to vertebral problems due to the type 2148 of riding style (i.e. high horse neck position positively correlated with vertebral areas affected, 2149 (Lesimple et al., 2010). Back pain in horses is however difficult to evaluate, even for professionals (e.g. manual palpation by practitioners, Lesimple et al. 2013). Therefore, I 2150 2151 decided to assess horses' dorsal thoracolumbar area (i.e. the back) utilising pressure 2152 algometry (PA), as this tool allows quantifying pressure sensitivity in a standardised manner 2153 through mechanical nociceptive threshold (MNT) testing (De Heus et al., 2010; Haussler, 2154 2020; Haussler & Erb, 2006). PA has been used in horses to quantify acute pain sites (e.g. 2155 surgical sites) from non-pain sites (Haussler & Erb, 2006), musculoskeletal pain due to 2156 orthopaedic injuries (Varcoe-Cocks et al., 2006) and clinical signs of back pain (Long et al., 2157 2020).

2158 The testing procedure involved a handler standing next to the horse's head with his back to 2159 the experimenter who applied a handheld pressure algometer (ProD-Plus; Topcat Metrology) 2160 with a flat-head tip (Ø 2mm, *i.e.* the device part contacting the horses' skin, see Figure 3.2 A) 2161 perpendicular to five marked areas on the horse's back (see Figure 3.2 B, De Heus et al. 2010). 2162 The pressure application rate was kept constant at 2 N/s ensured by warning lights turning on 2163 if the rate changed by 0.5 N/s. The moment at which the horse showed a reaction (e.g. muscle 2164 switch, tail flick, head shake, signalled by the handler and/or experimenter by saying "pickle") 2165 was defined as the endpoint of the testing. The maximum force applied at this moment was 2166 noted as the MNT value. In total, three measurements of each side (*i.e.* left/right) of each test 2167 area were taken in one single session. Each measurement followed a 4-6s break in which the 2168 handler took note of the MNTs and the experimenter adjusted her position to assess a 2169 different test area. To avoid confounding effect of de-/sensitisation of the tested areas, the 2170 order and side of measured areas were randomised within the test session. Following a 15-2171 min break, in which the horse was left alone its stall, the test procedure was repeated. I 2172 hypothesised that horses with significantly lower thresholds compared to MNT described for

2173 healthy horses (De Heus et al., 2010; Haussler & Erb, 2006; Menke et al., 2016; Taylor et al., 2174 2016) were more likely at risk to experience musculoskeletal problems (and associated pain) 2175 than horses with thresholds around the average MNTs of healthy horses. In addition, I 2176 intended to analyse the relationship between horses' MNT and HGS results. However, during testing, it became clear that it was difficult to achieve agreement between the handler and 2177 experimenter when detecting endpoints, hence determining a clear stop point was not 2178 2179 straightforward, as has previously been pointed out as a criticism by others (e.g. Love et al. 2180 2011). For instance, sometimes the handler signalled an endpoint before the experimenter started applying pressure. Two horses actively moved their necks away from the 2181 2182 experimenter when trying to apply the device, maybe because of similarities in the 2183 experimenter's movements during testing and the administration of neck injections. The 2184 animals might have avoided the experimenter due to negative associations with painful 2185 injection events.

I collected data from 16 horses at yard 2. However, given the described limitations in determining endpoints and time constraints to refine the experimental protocol, I did not conduct this type of data collection at yard 3 and also did not consider analysing further the obtained data.

2190 2191



2192

Figure 3.2 Example images of A) application of the pressure algometry device and B) test areas on the horse's back. Five locations were tested approx. 10 cm lateral to the dorsal midline on both sides of the horse: 1) middle of the splenius muscle at the third vertebrate, 2) trapezius muscle below the withers, 3) thoracic longissimus muscle at the middle of the saddle area (13th thoracic vertebral level), 4) thoracic longissimus muscle at the end of saddle area (approx. 18th thoracic vertebral level), and 5) mid-portion of the middle gluteal muscle at the third lumbar vertebral level.

Hors	e ID	Yard ID	Sex	Age	Breed	Welfare score	Used in AB test (Chapter 4)	Used in image test (Chapter 5)	Used to test physiological indicators Chapter 6
1	_	1	ď	7	Cob	17.42	yes	no	no
2	2	1	ď	19	Welsh Pony	17.00	yes	no	no
3	3	1	ď	24	New Forest Pony	16.75	no	no	no
4	Ļ	1	ď	13	Thoroughbred	19.00	yes	no	no
5	5	1	ď	7	Irish draft x	17.58	yes	no	no
6	5	1	ď	17	Thoroughbred	15.92	yes	no	no
7	,	1	ę	10	Thoroughbred	16.42	yes	no	no
8	3	1	ď	20	Thoroughbred	15.58	yes	no	no
9)	1	ç	9	Appaloosa	17.42	yes	no	no
1	0	1	ď	14	Connemara Pony x Camargue Pony	16.75	yes	no	no
1	1	1	ç	22	Connemara Pony	16.75	yes	no	no
1	2	2	ď	13	Cob X	13.17	no	yes	yes
1	3	2	ď	7	Cob	15.92	yes	yes	yes
14	4	2	ď	14	Cob	14.33	no	yes	yes
1	5	2	ç	7	Fell	15.42	yes	yes	yes
1	6	2	ď	15	Irish Cob	14.33	yes	yes	yes

2200 Table 3.5 Information to horses, individual welfare scores and their use in experimental Chapters 4-6.

Horse ID	Yard ID	Sex	Age	Breed	Welfare score	Used in AB test (Chapter 4)	Used in image test (Chapter 5)	Used to test physiological indicators Chapter 6
17	2	ę	10	Cob	15.25	yes	yes	yes
18	2	ď	9	Welsh Cob X	12.92	yes	yes	yes
19	2	ę	13	ISH	11.33	yes	yes	yes
20	2	ď	7	ISH	12.42	yes	yes	yes
21	2	ď	9	ISH	15.50	yes	yes	yes
22	2	ď	14	Cob	14.17	yes	yes	yes
23	2	ę	11	Cob	15.25	yes	yes	yes
24	2	ď	12	Welsh	13.67	yes	yes	yes
25	2	ď	11	Cob	12.92	yes	yes	yes
26	2	ę	9	Cob	14.25	yes	yes	yes
27	2	ę	9	Cob	15.58	yes	yes	yes
28	2	ď	13	ISH	13.25	yes	yes	yes
29	3	ď	14	Welsh x Anglo Arab	17.25	yes	yes	yes
30	3	ę	10	Cob x Thoroughbred	12.75	yes	yes	yes
31	3	ę	14	Percheron Cross	15.75	yes	yes	yes
32	3	ď	15	Welsh pony	16.75	yes	yes	yes
33	3	ď	17	Connemara Grey	15.42	yes	yes	yes

						Used in AB test	Used in image test	Used to test physiological indicators
Horse ID	Yard ID	Sex	Age	Breed	Welfare score	(Chapter 4)	(Chapter 5)	Chapter 6
34	3	ď	9	Cob	16.75	yes	yes	yes
35	3	ď	18	Shire Cross Irish Pony	16.25	yes	yes	yes
36	3	ę	4	Shire x Thoroughbred	15.75	yes	yes	yes
37	3	ď	NA	NA	NA	yes	yes	no
38	3	ę	15	Welsh pony	16.92	yes	yes	yes
39	3	ę	18	Welsh pony	13.92	yes	yes	no
40	3	ď	9	New Forest Pony	17.75	yes	yes	yes
41	3	ď	21	Dartmoor Pony	12.75	yes	yes	yes
42	3	ď	26	Welsh pony	13.25	yes	yes	yes
43	3	ď	19	New Forest Pony	12.75	yes	yes	no
44	3	ď	26	Welsh Pony	15.25	yes	yes	yes
45	3	Ŷ	16	Exmoor Pony	14.75	yes	yes	no
46	3	Ŷ	24	New Forest Pony	14.25	yes	yes	yes
47	3	ď	28	New Forest Pony	14.75	yes	yes	yes

2204 Summary

2205 As described in Chapter 2, negative affect mediates animals' attention to competing 2206 emotional stimuli, e.g. images of threatening vs. non-threatening faces of conspecifics; similar 2207 to anxiety-related enhanced attention to threat reported in humans. Here, I present the 2208 development of a visual attention bias (AB) test in horses. If horses show a preferential 2209 attention allocation towards images of unfamiliar conspecifics with negative, comparatively 2210 more neutral and positive facial expressions, it might be possible to infer emotional valence 2211 from the direction of AB as previously suggested in other animal studies (see Chapter 2). 2212 Hence, I assess the relationship between individuals' welfare state (proxy of affective state) 2213 and attention. As in humans, I predicted that horses exhibit AB to negative stimuli compared 2214 with comparatively more neutral or positive stimuli and that low welfare state (*i.e.* negative 2215 affect) would pronounce AB to threat and increase overall attention to the stimuli as a sign of 2216 increased vigilance. My exploration of horses' attention to positive relative to neutral stimuli, 2217 is a comparison rarely considered in other animal AB studies, and therefore there are no 2218 specific predictions. I found that horses attended to the three stimulus types for similar 2219 lengths of time and lower welfare states predicted shorter attention to the negative stimulus 2220 contrarily to my predictions. Moreover, horses from the three facilities significantly differed 2221 in their overall attention to the stimuli likely due to variations in the surrounding testing 2222 conditions. My findings highlight that external influences on attention require consideration 2223 for conducting cognitive tests at different facilities. As I will discuss, a further methodological 2224 investigation is needed in order to evaluate the potential of AB as an indicator of affective 2225 valence in horses.

2226 4.1 Introduction

2227 Measuring changes in cognitive processing (e.q. biases in attention, memory, and decision-2228 making) is an important approach to studying non-human animal emotions, since affective 2229 valence (*i.e.* whether the emotional state is pleasant or unpleasant) can be inferred from the 2230 direction of cognitive biases (Mendl et al., 2009, 2010; Paul et al., 2005). Negative emotional 2231 states are associated with a greater expectation of negative outcomes under ambiguity 2232 ('pessimism'), attention allocation toward negative information, and a greater likelihood of 2233 negative memory recall (Mendl et al., 2010). Cognitive biases, primarily in decision-making 2234 when facing ambiguity, have now been demonstrated in various mammal, bird, and insect 2235 species (reviewed in Gygax, 2014; Lagisz et al., 2020). Applying judgement bias tasks ('JBTs') 2236 has the advantage that *a priori* predictions can be made about the influence of underlying 2237 mood (*i.e.* longer-lasting states arising from the accumulation of short-lasting emotions in 2238 response to specific stimuli) on decision-making under ambiguity (Mendl et al., 2009).

2239 However, JBTs have significant limitations (Bethell, 2015; Crump et al., 2018; Roelofs et al., 2240 2016). These can include; i) labour-intensive training phases (e.g. discrimination training of 2241 positive/negative cues before animals' responses to intermediate cues can be tested), ii) 2242 sampling biases towards 'learners' (and exclusion of non-learning animals, which possibly 2243 have more negative affect since severe stress can impair cognitive abilities (Mendl, 1999), iii) 2244 ambiguous cues could be perceived as novel rather than intermediate (mixed 2245 positive/negative) stimuli (Bateson & Nettle, 2015; Doyle et al., 2010; Roelofs et al., 2016), 2246 and/or iv) discrimination training could act as cognitive enrichment for animals kept in barren 2247 environments (hence influencing affect and responses during testing) (Roelofs et al., 2016). 2248 Developing new cognitive measures free from these caveats is therefore desirable (Roelofs et 2249 al., 2016; Crump et al., 2018).

Assessing affect-driven attentional biases (ABs), the emotion-induced differential allocation of attention to certain information compared to others has been promoted as an alternative, and potentially more practical, cognitive marker of animal affect (see Chapter 2; Crump et al., 2018). Unlike JBTs, attention bias tests (ABTs) rely on innate tendencies to cognitively process sensory information, such as biologically relevant visual stimuli (*e.g.* aggressive faces, predator cues). AB to these stimuli can be assessed, depending on the experimental paradigm, with no, or relatively little, animal training (*e.g.* habituation to test apparatus

(Lewis et al., 2021), training animals to sit by a target for testing (Howarth et al., 2021)).
Attentional state can be behaviourally quantified, *e.g. via* measuring eye gaze or head
movements toward visual cues, thereby measuring looking time duration or location of target
fixation (Winters et al., 2015).

2261 In humans, negative affect (e.g. trait and state anxiety) is associated with the modulation of 2262 attention towards threatening information (Bar-Haim et al., 2007; MacLeod et al., 1986; 2263 Yiend, 2010). For example, when simultaneously presented with images of angry and neutral 2264 faces, anxious subjects are more likely and faster to attend to angry than neutral faces, 2265 relative to control subjects (Mogg et al., 2000). Attention to threat also appears to be 2266 modulated by affect in non-human animals (reviewed Crump et al., 2018). For example, in 2267 sheep (Lee et al., 2016) and cattle (Lee et al., 2018), the pharmacological induction of anxiety 2268 results in longer looking duration and increased vigilance towards a threatening stimulus 2269 (dog)¹. In contrast, administering an anxiolytic lowers sheep's attention to the threat (Lee et 2270 al., 2016; Monk et al., 2018a). I aimed to develop a test to assess whether attentional biases 2271 to emotional stimuli can be used as a cognitive proxy of emotional state in horses. Several 2272 housing- and management-related issues have been identified to compromise equine well-2273 being (e.g. Hockenhull & Creighton, 2014a, 2015; Lesimple, 2020; Minero & Canali, 2009). Yet, 2274 in comparison to other farm animals, horse welfare studies are relatively scarce (Hockenhull 2275 & Whay, 2014), and developing more accurate measures of horse affect is therefore 2276 important.

To investigate whether horses exhibit AB towards emotionally-valenced visual social cues, I presented horses of varying welfare levels (quantified as welfare score as described in Chapter 3) with paired images showing facial expressions of conspecifics of different valence (*i.e.* negative, positive, or comparatively neutral faces, Waring, 2003; Wathan et al., 2016).

I predicted that horses would attend to negative faces for longer than neutral or positive faces
(because attention to threat is adaptive, Bar-Haim et al., 2007), and even more so in horses
with lower (*i.e.* poorer) welfare scores since negative affect increases attention to threat in
other ungulates (Lee et al., 2016, 2018; Monk et al., 2018a). I also explored whether welfare

¹ although see Monk et al., 2018b (in sheep) and Bethell et al., 2012 (in macaques) for alternative responses such as avoidance of the threat stimulus

2285 score would modulate how horses attend to the positive relative to the neutral face, i.e. a 2286 bias in positivity that could be relevant to assessing the quality of life since good welfare 2287 requires the presence of positive experiences (Boissy et al., 2007). ABs favouring positive 2288 information are understudied in both humans and animals (Crump et al., 2018), making the 2289 exact direction and strength of such a bias difficult to predict. Finally, I expected that horses 2290 with lower welfare scores would show longer overall attention to both stimuli since negative 2291 affect can increase vigilance (i.e. increased monitoring of the environment for potential 2292 threat) in other ungulates (Lee et al., 2016, 2018; Monk et al., 2018a, 2018b).

2293 **4.2 Methods**

2294 4.2.1 Ethical approval

This study was approved by the Animal Ethical Review Committee of the University of Plymouth (ETHICS-41-2020). The experimental procedure was below the threshold for regulation by the UK Animals (Scientific Procedures) Act 1986 (ASPA) and followed the Essential 10 ARRIVE guidelines 2.0 (Percie du Sert et al., 2020). The owners consented to the use of their animals and were responsible for horse care and health checking. The animals remained at their home facility at the end of the study.

2301 4.2.2 Animals and housing

2302 Forty-seven horses from three different yards were enrolled in this study. Three horses were 2303 excluded from data collection or analysis (one due to camera failure, one because of 2304 disruptions during testing, and one horse was reluctant to approach the test equipment), 2305 resulting in 44 animals used for data analysis (yard 1: N=11; yard 2: N=16; yard 3: N=17). The 2306 horses belonged to private owners (yard 1 horses) and riding schools (horses from yard 2 and 2307 3), were between 4 and 28 years old (13.84±5.92 years), of mixed breeds, and 17 (39%) were 2308 females. The housing conditions varied yards and animals as described in Chapter 3. Briefly, 2309 yard 1 offered relative 'naturalistic' housing conditions (*i.e.* horses lived in one stable group 2310 with free access to pasture and roughage). Yard 2 and 3 horses experience comparatively 2311 more 'restrictive' living conditions as their social contact, outdoor access and roughage intake 2312 were limited.

2313 4.2.3 Welfare assessment

As described in Chapter 3, I evaluated horses' individual welfare using a holistic approach comprising multiple assessment steps (including environmental, social, and animal-based factors) from which I calculated a single welfare score for each animal. Table 3.1 in Chapter 3 provides the full list of assessed factors. The maximum score was 20 with higher scores putatively reflecting a greater likelihood to experience positive welfare and positive affective states.

2320 4.2.4 AB testing

To assess AB, I presented horses of varying welfare levels with pairs of pictures of horse faces showing facial features associated with three emotional states previously described in Wathan et al. (2016). These states were *negative* (*i.e.* horse in agonistic interactions with conspecifics), *positive* (*i.e.* during food anticipation) or a comparatively intermediate state in comparison to the former two, hereafter termed *neutral* (*i.e.* whilst resting).

2326 4.2.4.1 Stimuli preparation (prior to testing)

2327 Four stimulus horses unfamiliar to test subjects were photographed during each of the three 2328 situations (negative, positive and neutral) using a NIKON D3200 camera (focal length 30.00 2329 mm). Images were edited so that only the horses' head on a white background was visible 2330 (see Table 4.2 for example images and Appendix C Table C1 for all test stimuli). Mean 2331 luminosity values were extracted for each image via the histogram function in Adobe 2332 Photoshop. To compare the mean luminosity across stimulus types (negative, positive, 2333 neutral) and horses, I used a two-way ANOVA with Tukey post hoc pairwise comparison tests. 2334 There was no effect of stimulus type (F=0.65, p=0.55), but there were significant differences 2335 between images of stimulus horses A-B, A-C, A-D and B-C (all p-values <0.05). However, the 2336 mean luminosity of images of stimulus horses B and D (p=0.11), or C and D (p=0.71) did not 2337 significantly differ. Therefore, I controlled for stimulus horse ID in the later analysis. To ensure that the face images I determined as 'negative', 'neutral' and 'positive' were accurately 2338 2339 categorised based on the facial patterns displayed by the stimulus horses, the final images 2340 were independently rated by 11 observers (members of Animal Behaviour and Welfare 2341 groups at Plymouth and Bristol University; group 1: N=6, α =0.55; group 2: N=5, α =0.49, 2342 Krippendorff's alpha) blinded to the situation during which the pictures were taken.

2343 Pictures were subsequently arranged so that during testing, each subject horse was exposed 2344 to one Microsoft Powerpoint presentation containing a total of six stimulus slides, each 2345 proceeded by a blank inter-trial slide. Each stimulus slide showed two side-by-side images of 2346 the same horse displaying differing facial expressions (*i.e.* negative-positive, positive-neutral, 2347 or negative-neutral). Two stimulus horses were randomly chosen from the four available to 2348 create the presentations. The order of stimulus slides was pseudo-randomised so that if the 2349 same type of stimulus pairing (e.g. positive vs. neutral) happened to be shown on two 2350 consecutive trials, the type of facial expression was counterbalanced across sides (e.g. 2351 positive (left) – neutral (right) followed by neutral (left) – positive (right)). Four presentations, each containing a different combination of stimulus horses, were created in total. To ensure 2352 2353 I remained blind to the content of the presentations, another person not involved in testing 2354 pseudo-randomised (i.e. the four presentations formed a 'set'; order within sets was 2355 randomised and repeated 12 times to attain a list of 48 blinded presentations with a balanced 2356 number of the same presentation in the list) and renamed the presentations with numbers 2357 (1-48), which were then selected by me during test-setup in ascending order.

2358 4.2.4.2 Experimental set-up and procedure.

2359 The test area was a familiar arena equipped with two back projection screens (HOIN, 140 cm 2360 H x 212 cm W, 254 cm in diameter, placed 1m above ground), and two LCD projectors 2361 (HITACHI CP-X303WN) connected to a computer (Lenovo Thinkpad 13) and positioned at 50 2362 cm height in 2.5 m distance behind the screens Figure 4.1. At yard 1, the projectors were 2363 always placed in front of the screen for practical reasons of space limitation. At yard 2, back 2364 projection seemed unsuitable (*i.e.* images barely visible with human eyes) for eight horses 2365 and therefore images were front projected for these horses. Two cameras (GoPro Hero 3, 3+) 2366 were used for recording the horse's behaviour from the front (i.e. horse is visible from the 2367 front, stimuli not visible) and back views (*i.e.* horse visible from behind, both screens in full 2368 view). The back view videos were subsequently edited (VSDC free video editor software) for 2369 blinding purposes, by placing white squares over the screens so that only the stimulus horses' 2370 muzzle was visible. This allowed me to record the start and the end of a trial when extracting 2371 the tested horse's behaviour from the footage, whilst remaining blind to the type of stimuli 2372 shown on screen.


2373

2374 Figure 4.1 Diagram of the AB test set-up. The test stimuli were projected onto two screens (1a, b) via 2375 two LCD projectors (2a, b) receiving input from a laptop (3). The midpoint of two ground poles (3m 2376 long, Ø 98mm) placed in front of the screens (4a, b) marked the screens' centre point. A ground pole 2377 (5) ensured consistent distance between animals and the screens. Horses were positioned with their 2378 front feet as close as possible to the midpoint of the pole. Perpendicular to this location, another pole 2379 (6) was placed on the ground, which marked the midline between both screens as a visual reference 2380 for horses' head position during testing and post hoc behaviour analysis. Video cameras were 2381 positioned so that footage was recorded from the front (7a) and back (7b).

2382

2383 The horse was habituated to the test equipment (screen turned off), first from an approx. 2384 10m distance, then from the test position with head and body aligned with pole 6 in Figure 2385 4.1 (reference pole hereafter) and standing as close as possible to the horizontal pole (5 in 2386 Figure 4.1). Once the horse stood quietly in front of the screens with its head in a relaxed 2387 position (*i.e.* with the muzzle at approx. chest height), the experimenter (E hereafter) stood 2388 next to the horse's shoulder with her head directed to the horse's withers (*i.e.* E standing at 2389 a 90° angle to the horses' shoulder). Whilst glancing towards the horse's neck to judge its 2390 alignment with the reference pole, E turned on both screens using a remote control (Kensington wireless presenter) to show the first stimulus slide (*i.e.* started the first trial) at 2391

2392 the moment the horses' head was aligned with the reference pole. The images were shown 2393 for at least 20s regardless of whether the horse looked at the screens or not (see ethogram 2394 Table 4.1). If the horse was still looking at the screens after 20 s, the images were shown until 2395 the horse no longer appeared to look at the stimuli². When done, E showed a blank slide on 2396 the screen, switched her position to stand at the horse's opposite shoulder, corrected the 2397 horse's position if necessary, and started a new trial as soon as the horse's head aligned with 2398 the reference pole. All six trials were conducted this way. Disturbances (e.g. noise, a person 2399 unexpectedly entering the test area) were signalled by E by directly looking into the back 2400 camera and mouthing the word "disturbance" to identify trials to be excluded during video 2401 analysis. Otherwise, E did not look at the screens or cameras and fixed her gaze on the horse 2402 to remain blinded to the image contents throughout testing.

2403 4.2.4.3 Behavioural analysis

2404 The horse behaviour was scored from video footage by me (intra-observer reliability assessed 2405 with the icc function in the psych R package: 0.932) using Noldus The Observer XT 14. A 2406 second observer not involved in the study independently scored videos of 10 horses (22.7 % 2407 of video samples randomly assigned) and inter-observer reliability was 0.868. The ethogram 2408 is summarised in Table 4.1. Given the lateral eye position and their limited movement within 2409 the eye socket, it is difficult to judge what horses are fixating on by tracking the eyes 2410 themselves. I thus used proxies of overt visual attention as previously described in horses 2411 (Waring, 2003; Wathan et al., 2016), *i.e.* head and ear positions. Head and ear positions were 2412 scored continuously whilst images were shown. A horse was assessed as attending to a 2413 stimulus if it turned its head either to the right or the left with both or one ear forward 2414 suggesting engagement of the visual (binocular) system to look at the screen ahead of it. If a 2415 horse kept its head turned towards one screen whilst moving both its ears from forward to a 2416 lateral or backward position for more than 1s, it was no longer considered attending the 2417 screen. Only time periods >1s where the horse was attending to either screen were included 2418 in the final data analysis as movement lasting <1s were difficult to score. Although horses 2419 possibly looked at the stimuli whilst their head pointed straight forwards (i.e. both eyes at 2420 equal distance to the stimuli, see Table 4.1), I did not include these time points in the final

² Based on pilot observations showing variability between horses, with some paying attention to the screen for very short periods of time (<5s) whilst others were looking at the screens for longer than 2min.

- 2421 analysis as it was impossible to determine which stimuli the horses were attending to without
- 2422 clear indication of the side they were looking at.

2423 Table 4.1 Ethogram of horse behaviours scored during the attention bias test. Behaviours indicating

that the horses were no longer looking at the stimuli once the first 20s of stimulus presentation

2425 elapsed were categorised as behavioural events terminating the trial.

Behaviour	Description				
Behaviours indicating attention towards the stimuli					
Head position (sig	de scored from the horse's perspective)				
Head turn right	Head is held at, or higher than, chest height, with the centre of the forehead rotated away from the midline towards the left corner of the front view video footage (head direction scored from the horse's perspective hence contralateral side when seen from front view). The left eye is more visible on front view footage than the right eye (which might be partially visible or not at all).				
Head turn left	Head is held at, or higher than, chest height, with the centre of the forehead rotated away from the midline towards the right corner of the video frame. The right eye is more visible on front view footage than the left eye (which might be partially visible or not at all).				
Ear position					
Ears both forward	Head is held at, or higher than, chest height, and both ears are rotated forwards so that the insides of the ears are fully visible from frontal view footage while the outer parts of the ears (pinnae), apart from the tips, are not.				
Ears both back	Head is held at, or higher than, chest height, the ears are rotated backwards so that the insides of the ears are not visible from frontal view footage, but the pinnae become more visible and the tip of the ears are pointing backwards, or are possibly not visible if ears are pressed flat against head.				
Ears lateral	Head is held at, or higher than, chest height, the ears are rotated outwards so that the inside of both ears are only partly visible, the pinnae become more visible at the base of the head, and tips of the ears point outwards.				
Right ear forward	The right ear is rotated forwards so that the inside of the right ear is fully visible while the left ear is in lateral or backward position.				
Left ear forward	The left ear is rotated forwards so that the inside of the left ear is fully visible while the right ear is in lateral or backward position.				

Behaviour	Description					
Behaviours indic	Behaviours indicating attention away from stimuli terminating a trial					
Head turn away right	The horse's neck is bend to its right-hand side with its head at least in profile view with the forehead no longer visible from frontal view footage.					
Head turn away left	The horse's neck is bend to its left-hand side with its head at least in profile view with the forehead no longer visible from frontal view footage.					
Horse interacting with the handler	Physical contact between any part of the horse's head and the handler (<i>e.g.</i> sniffing, head rubbing).					
Exploring the ground	The horse is sniffing the ground/poles.					

2426

2427 4.2.5 Data analysis

Data were analysed in R (version 4.0.3 (R Core team, 2021)). For each horse, four indices of 2428 2429 attention were calculated per trial. Firstly, all time periods (>1 s) a horse was observed looking 2430 at one of the stimuli (i.e. negative/positive/neutral) were combined to assess the duration of 2431 attention to the stimulus (i.e. negative faces considered as the stimulus in trials with negative-positive and negative-neutral pairs, positive faces considered as the stimulus in 2432 2433 positive-neutral stimulus pair trials). Likewise, the duration of attention to the opposite 2434 stimulus was assessed by combining all time points the horses was recorded looking at the 2435 opposite image (e.g. in negative stimulus pairs, the opposite stimulus was either positive or 2436 neutral). Total attention duration was assessed as the sum of durations attending the 2437 stimulus and opposite stimulus. Lastly, the proportion of time looking at the stimulus was 2438 calculated relative to the total duration of attending both stimuli. Assessing both absolute 2439 duration and proportional values when measuring attention biases might be important (Lewis 2440 et al., 2021). Absolute differences give a direct measure of the difference in attention duration 2441 to the stimuli, but variation in overall looking duration can influence this measure of AB 2442 because of differences in absolute looking times between individuals (Lewis et al., 2021; 2443 Lonsdorf et al., 2019). Hence, to control for the differences in absolute attention duration, 2444 proportional values were calculated, noting that this would amplify strongly biased looking

durations on trials of overall shorter durations (Lewis et al., 2021). Moreover, a value of 0.001
was added to all absolute duration values so that trials with zero attention durations for one
of the two stimuli were adjusted to 0.001 thereby enabling inclusion in the statistical analysis
of trials where horses looked at one stimulus only.

Trials during which the horse did not attend to either of the stimuli, as well as trials with external disturbances (e.g. loud banging, dog barking), were excluded. This yielded an unbalanced data set with both data points per stimulus pair not being available for all horses. Consequently, I only considered the first trial a horse attended to each of the three stimulus pairs to assess AB at the group level. Since duration data (attention to the stimulus, attention to the opposite stimulus and total attention duration) were not normally distributed (Shapiro-Wilk test, p<0.001), central tendency measures are reported as median and Q1-Q3 range. Wilcoxon signed-rank tests were used to compare attention duration and proportion of time horses looked at the images (i.e. stimulus and opposite stimulus) presented as pairs. The number of trials per yard used for this analysis is shown in Table 4.2.

- .

2/72	Table 4.2 Example image of stimuli and number of tri	als nor stimulus nair used for AB testing
24/Z	Table 4.2 Example image of sumuli and number of th	als per stillulus pair used for AB testillg

Stimulus pair	Yard	N trials
Negative-neutral	1	10
	2	15
	3	15
Negative-positive	1	10
	2	14
	3	17
Positive-neutral	1	7
	2	15
	3	17

2473

F V

To investigate the influence of the putative emotional state, my predictions of how horseswould attend to the emotional stimuli in relation to their welfare were as follow:

Attention bias to negativity is reflected in (a) longer attention duration and (b) greater proportion of time attending to the negative horse face in comparison to putatively more positive stimuli (*i.e.* a comparatively more neutral or a positive horse face). We expected that horses with lower welfare scores would attend to the negative stimuli for longer and for a greater proportion of time than horses with comparatively higher welfare scores.

Attention bias to positivity measured as (a) attention duration and (b) proportion of time attending the positive horse face relative to the comparatively more neutral stimulus is modulated by welfare score (exploratory analysis, no direction of ADAB predicted due to the lack of testing ADAB towards positive stimuli in current literature).

Enhanced attention (vigilance) is reflected in a relationship between total attention duration
 and welfare score, with horses scoring lower on welfare showing more attention to the stimuli
 compared to horses with comparatively higher welfare scores.

2488 To test my hypotheses, I fitted separate linear mixed-effect models (lmer function in the lme4 2489 package, Bates et al., 2015) with attention duration to the stimulus (*i.e.* negative or positive; 2490 square root transformed for better model fit), and total attention duration (log-transformed) 2491 as the dependent variables. Facility (yard1, 2, 3), horses' age, sex and welfare score, as well 2492 as opposite stimulus type (negative/neutral/positive), stimulus side (left, right) and stimulus 2493 horse ID were set as fixed factors. Collinearity of fixed factors was checked with the vif 2494 function (from the Ime4 package) as factors with vif >5 would have indicated a problematic 2495 amount of collinearity between predictors (James et al., 2014), but this was not the case. I 2496 included all trials in which a horse attended to at least one of the images in this analysis and 2497 therefore included horse ID as a random factor in the models. The drop1 function (base R) 2498 was used for all models to exclude all non-significant factors (except welfare score, which was 2499 included in all models) to reach the minimum adequate model. The anova function with type 2500 III sum of squares implemented from the 'car' package (Fox & Weisberg, 2019) was used to 2501 generate p-values for individual factors within each model. Results were reported as 2502 significant at a threshold of p≤0.05 and results between 0.05 and 0.1 were reported as 'trends' 2503 given that these might convey meaningful biological variation (Stoehr, 1999). For model 2504 validation, model residual diagnostics were done using the 'DHARMa' package (Hartig & 2505 Lohse, 2022). Details of statistical parameters from the final model outputs are presented in 2506 Tables 4.4 - 4.8 at the end of this chapter.

Proportional data were analysed in the same way except that the dependent variables (*i.e.* the proportion of time looking at negative or positive relative to the opposite) were modelled using the glmmTMB function (glmmTMB package, Brooks et al., 2017). This allowed the modelling of beta-distributed data, which is the most appropriate fit for data bounded between 0 and 1 (Douma & Weedon, 2019).

- 2512 For pairwise comparisons of welfare scores between the three yards, Kruskal Wallis test with
- post hoc Dunn test (dunnTest function from FSA package (Ogle et al., 2022), p-values adjusted
- with Holm method) was used.

2515 **4.3 Results**

2516 4.3.1 Horses' attention to the stimuli when first attending to the images

2517 Considering the absolute duration when horses attended to the stimulus pairs for the first 2518 time, I found no statistically significant evidence for AB as horses viewed all images for similar 2519 lengths of time (Figure 4.2A). Likewise, the proportion of time horses looked at the images 2520 for the first time did not significantly differ within the stimulus pairs (Figure 4.2B).



2521

Figure 4.2 Absolute duration (A) and proportion of time (B) recorded when horses attended to the stimulus pairs for the first time. Boxplots show data at group level and symbols show individual data points. Wilcoxon signed rank test, p>0.05 for all stimulus pairs.

2525

2527 4.3.2 Testing hypothesis 1: Is attention to negative stimuli influenced by welfare score?

2528 Welfare score significantly predicted the absolute duration horses looked at the negative 2529 stimuli (X²₁=4.56, p=0.03, see Table 4.3), but in contrast to my prediction, a higher welfare 2530 score (*i.e.* greater likelihood for positive emotional states) predicted longer absolute attention 2531 duration to the negative stimuli (Figure 4.23). Moreover, facility had a significant effect on 2532 attention duration (X²₂=23.15, p<0.001, see Table 4.3). Yard 2 horses looked at the negative 2533 stimulus for significantly longer than horses from yard 1 (z=4.68, p<0.001) and yard 3 (z=-3.28, 2534 p=0.002), and horses from yard 3 attended to the negative images for significantly longer than 2535 horses from yard 1 (z=2.67, p=0.01). Mean welfare scores significantly differed between yards 2536 (mean±SD, yard 1: 16.95±0.9; yard 2: 14.18±1.31; yard 3: 14.95±1.35; all p<0.001), but there 2537 was no significant interaction between yard and welfare score on attention measures $(X^{2}_{2}=0.17, p=0.92)$. Moreover, the type of opposite stimulus significantly influenced the 2538 2539 duration horses attended to the negative stimulus ($X_{1}^{2}=4.04$, p=0.04, Table 4.3). When paired with the positive stimulus, horses looked at the negative stimulus for significantly longer than 2540 2541 when it was presented with the neutral stimulus (z=2.01, p=0.04). Considering the proportion 2542 of time horses attended to the negative stimulus, I found no significant effect of welfare score 2543 $(X^{2}_{1}=2.24, p=0.13)$, or any of the other factors I investigated (p-values from 0.17 to 0.99, Table 2544 4.4).



2545

Figure 4.3 Attention duration (raw data presented) to the negative stimulus predicted by welfare score. Different shapes and colours indicate data points for each yard.

2549 **4.3.3** Testing hypothesis 2: Is attention to positive stimuli influenced by welfare score?

Neither the absolute duration of attention to the positive stimulus (X^{2}_{1} =0.79, p=0.37, Table 4.5) nor the proportion of time horses looked at the positive stimulus (X^{2}_{1} =0.78, p=0.37, Table 4.6) was significantly predicted by welfare score or any of the other factors I explored (duration: p-values from 0.26 to 0.74, the proportion of time horses looked: p-values from 0.16 to 0.94).

2555 4.3.4 Testing hypothesis 3: Is enhanced attention to all stimuli predicted by welfare score? 2556 The total duration, assessed as horses attending to both images, was not significantly 2557 predicted by welfare score ($X_{1}^{2}=0.87$, p=0.34, Table 4.7). However, facility had a significant 2558 effect on attention duration (X_{2}^{2} =49.41, p<0.001). Horses from yard 2 attended to both stimuli 2559 for significantly longer than horses from yard 1 (z=6.99, p<0.001) and yard 3 (z=-2.56, p=0.03), 2560 and horses from yard 3 spent significantly more time looking at the images than horses from 2561 yard 1 (z=5.95, p=0.02). Stimulus horse had a significant effect on horses' overall attention to 2562 the images (X²₁=8.15, p=0.04, Table 4.7) with images from stimulus horse B being attended to 2563 for significantly longer than images from stimulus horse D (z=-2.7, p=0.03), but there were no 2564 significant differences in horses' total attention to the other stimulus horse combinations (p-2565 values from 0.15-0.96). Figure 4.4 show the estimated marginal means for facility and 2566 stimulus horse ID.





2569

2570 **4.4 Discussion**

2571 Previous studies in other species (*e.g.* macaques, sheep, cattle) reported that visual attention 2572 to emotional stimuli can be driven by the subject's emotional state (Crump et al., 2018), but 2573 whether this is true for horses has not yet been validated. Here, I tested if horses exhibit AB 2574 when presented with a combination of two digital images of unfamiliar conspecifics' faces 2575 showing either negative, positive or comparatively neutral facial expressions. I expected that 2576 negative stimuli would be attended for longer when presented alongside comparatively more 2577 positive or neutral stimuli, reflecting adaptive AB to threatening information (Bar Haim et al., 2578 2007), and that this bias would be more pronounced in horses with poorer welfare scores 2579 (proxy for poorer underlying affective state) as a result of negative affect modulation on 2580 attention to threat as reported in humans (Bar Haim et al., 2007) and other animals (Crump 2581 et al., 2018).

2582 In contrast to my predictions, I found no significant evidence of AB to specific types of 2583 valenced images as horses attended to all stimuli for similar lengths of time. However, I did 2584 observe a moderating effect of welfare score on AB but in the opposite direction to my 2585 predictions; horses with higher welfare scores spent more time attending to negative stimuli. 2586 More, attention appeared to be modulated by several experimental factors. AB to negative 2587 stimuli was significantly longer when paired with positive than neutral stimuli and stimulus 2588 horse ID had a significant effect on overall attention. Furthermore, horses from the three 2589 facilities differed significantly in their absolute duration of attending to the negative stimuli 2590 suggesting that variations in test surroundings might have influenced horses' attention 2591 responses. The same significant yard effect was found considering horses' overall attention 2592 to the stimulus pairs. Lastly, horses exhibited no AB when presented with positive-neutral 2593 stimulus images.

2594 While innate preferential attentional processing of emotional faces has been well described 2595 in humans, knowledge about what type of information other (non-primate) animals extract 2596 from faces is still limited (Leopold & Rhodes, 2010). Wathan et al. (2016) found that horses 2597 were more likely to approach the printed images of unknown conspecifics' positive faces than 2598 negative faces (Wathan et al., 2016), which is an appropriate response to avoid social conflict, 2599 possibly reflecting an adaptive behaviour in this social-living species. However, 2600 approach/avoidance response behaviours are controlled by stages of cognitive processing

that follow initial attentional processing, and stimulus appraisal (*e.g.* stimulus relevance to the individual) before an appropriate response is determined (Scherer, 1999). Hence, my findings are not necessarily comparable with the previous horse study, as I tested responses putatively linked to different stages of stimulus processing (*i.e.* looking here *versus* approaching/avoiding in Wathan et al 2016).

2606 Familiarity of the horses used as stimuli might also have influenced my results. In goats, 2607 images of negative faces of familiar conspecifics were attended to for longer than positive 2608 faces (Bellegarde et al., 2017a) and videos of familiar sheep in agonistic interactions were 2609 attended to more (*i.e.* longer time spent head turned with ears forward to the screen) than 2610 videos of sheep whilst ruminating (Vögeli et al., 2015). In contrast, my horses were presented 2611 with images of *unfamiliar* individuals, which likely present important information; horses 2612 naturally live in stable groups and the introduction of new group members increases 2613 aggression (Fureix et al., 2012a). It is therefore possible that subjects attended to all images equally because the pictures of unknown individuals conveyed important non-emotional 2614 2615 information, *e.g.* face-based signals such as identity, sex and age (Burke & Sulikowski, 2013). 2616 It would thus be interesting to see whether images of familiar horse faces would lead to 2617 different results, considering e.g. that chimpanzees and bonobos show preferential attention 2618 to familiar over unfamiliar conspecifics, although only when viewing individuals of the dominant sex suggesting that socioecological factors modulate social attention in these 2619 2620 species (Lewis et al., 2021). If so, this would suggest that emotional cues might be less salient 2621 than familiarity cues, which is important to consider when choosing visual stimuli in similar 2622 test paradigms in future.

2623 In addition to familiarity, other facial characteristics (e.g. coat colour, head shape, facial 2624 markings) might contribute to preferential attention unrelated to emotional state. I used four 2625 different stimulus horses and found that stimulus horse ID had significant effects on total 2626 attention duration (*i.e.* considering both stimuli combined). Horses looked longer at stimulus 2627 horse B (dark brown coloured horse face with two small white areas on its forehead and 2628 muzzle) than stimulus horse D (chestnut-coloured horse face with a wide white vertical stripe 2629 covering (blaze) most of its head). Mean luminosity values did not significantly differ between 2630 the images of these horses. It is possible that horse D was less salient on a white background 2631 (*i.e.* more difficult to perceive) compared to horse Bellegarde et al. (2017a) similarly observed

that images of one goat were more attended than those of the other three stimulus goats (regardless of emotional information the faces conveyed), although instead of facial characteristics, social rank might have influenced goats' attention in this study as the stimulus goats were familiar conspecifics (Bellegarde et al., 2017a).

2636 It is also possible that the horses did not recognise the images as conspecific faces. My test 2637 stimuli were computer-generated projections, *i.e.* artificial images made for the human eye 2638 while equine vision differs from ours (e.g. colour perception and scotopic vision, (Rørvang et 2639 al., 2020). Despite these differences in sensory processing, several studies have reported that 2640 horses can successfully discriminate between artificial images such as printed photos of 2641 objects (Hanggi, 2001), horses (Wathan et al. 2016) and humans (Proops et al., 2018; Smith 2642 et al., 2016a), screen images of humans (Lansade et al., 2020a,b), and videos of human-horse 2643 interactions (Trösch et al., 2020). Accordingly, I considered screen images as suitable visual 2644 stimuli for horses. Existing evidence of image recognition in horses should nonetheless be 2645 approached with caution because previous findings might have alternative explanations (e.g. 2646 rapid learning and generalisation in Hanggi (2001), lack of control conditions in Proops et al., 2647 2018, Wathan et al., 2016, and statistical weaknesses, e.g. Amici (2019) discussed in detail in 2648 Chapter 5). For this reason, I conducted a follow-up study (testing 27 horses from yard 2 and 2649 3), which is presented in Chapter 5 showing that after learning to discriminate between two 2650 real-life objects of different colours and shapes, only one horse spontaneously approached 2651 the rewarded object when presented with on-screen images of these objects. These 2652 observations suggest that horses' ability to recognise objects from images might be highly 2653 variable and cannot be generalised. However, that the horses were not able to reliably 2654 transfer from inanimate objects to images in my follow-up study does not eliminate the 2655 possibility that horses might recognise biologically more important cues (such as conspecifics 2656 faces) from images.

The finding that absolute attention duration to the negative stimulus was modulated by welfare score might support this argument. Contrary to my predictions, horses with lower welfare scores attended to negative face images for shorter durations compared to horses with higher welfare scores. Avoidance of threat cues has been described in other animals. Macaques direct attention away from threat following a stressful handling procedure although they show sustained attention to aggressive faces of conspecifics during enriched

housing conditions (Bethell et al., 2012). In sheep, the pharmacological induction of anxietylike states increased attention to threat (live dog) compared to controls (Lee et al., 2016), although Monk et al. (2018b) found the opposite, and Monk et al. (2019, 2020) noted no significant effects. In humans affected by social phobias, avoidance of anxiety-eliciting stimuli (*e.g.* social situations) is a form of coping (Chen et al., 2020). Hence, horses with putatively poorer welfare might have disengaged from the negative faces faster to avert further negative experiences.

2670 I also assessed whether horses showed AB to the positive stimulus in relation to the neutral 2671 stimulus, as prioritised attending to positive information over neutral information might also 2672 be of biological significance (Gupta, 2019; Pool et al., 2016). However, horses attended to 2673 both types of stimuli equally and welfare score did not modulate attention to positive images. 2674 Nevertheless, when paired with the positive, stimulus horses' attention to the negative 2675 stimulus was significantly longer than when paired with the neutral stimulus. This observation might suggest that attentional processing for potential threat could be influenced by the type 2676 2677 of competing emotional information. Maybe attentional resources were less in demand when 2678 viewing the positive stimuli (*i.e.* horses rapidly recognised the emotional valence of both 2679 stimuli and attended to the negative stimulus for longer) compared to when viewing the 2680 neutral stimuli. The latter might be more ambiguous in terms of the emotional information 2681 conveyed and hence may have required longer attentional processing leaving less time to 2682 attend to the paired negative stimulus. I am not aware of any other animal studies testing the 2683 varying effects of positive and comparatively more neutral stimuli on animals' attention to 2684 threat. However, Bellegarde et al., (2017a) tested goats with morphed images (merging 2685 negative and positive goat faces at different gradations, *i.e.* 25%, 50%, or 75% positive) 2686 putatively presenting intermediate cues to positive and negative goat faces and found that 2687 the merged stimuli did not significantly differ in the time they were attended compared to 2688 the original faces.

2689 My inference of horses' affective states was indirect since the direct assessment of animal 2690 welfare is currently impossible. Therefore, I should acknowledge that inferring from the 2691 welfare score was only an approximation of the likelihood that subjects were experiencing 2692 negative/positive affective states. Since a scientific concept that allows balancing different 2693 aspects contribution to welfare is not available (Appleby et al., 2018), I aimed to evaluate all

2694 factors influencing horses' well-being as comprehensively as possible without prioritising one 2695 over the other. Moreover, the presence (or absence) of certain resources may not determine 2696 animals' experience of their environment to the same extent because variation in coping 2697 styles/abilities may influence individuals' perception of their environmental conditions, and 2698 ultimately their emotional state. I planned to use manipulations likely to induce specific 2699 (short-term) emotional states as done in other AB studies, but my experimental plans were 2700 significantly affected by Covid-19 restrictions preventing me from achieving my estimated 2701 sample size or imposing experimental treatments. In addition to this, would be interesting to 2702 test whether horses' immediate emotional response evoked by the test stimuli has an effect 2703 on their attention. Viewing negative stimuli might induced a more negatively-valenced 2704 emotional state in the observing horse which could have modulating effects on attention. One 2705 study reported that watching videos of positive or negative human-horse interactions induces 2706 congruent emotions in the observing horses proposing evidence of emotional contagion 2707 (Trösch et al., 2020). In order to detect shifts in emotional state in response to the valenced 2708 stimuli, I recorded horses' cardiac activity via ECG during AB testing, but I could not use these 2709 recordings for further data analysis due to discrepancies in the temporal requirements 2710 needed for this type of data (minimum of 5min, von Borell et al. 2007) and much shorter 2711 attentional responses observed in the horses. Furthermore, facility, but not welfare score, 2712 had a significant effect on how long both images were attended. Although welfare scores 2713 differed significantly between facilities, no interaction between both factors was found. This 2714 implies that non-welfare-related factors varying between facilities, such as environmental 2715 conditions during testing (e.g. lighting, noise) and disparities in the test setup may have 2716 induced variations in horses' attention to the stimuli (see e.g. Rosenberger et al. (2021) 2717 showing that goats' cognitive performance can be influenced by the testing site). Certain 2718 cognitive tests may be difficult to apply in farm environments where not all external 2719 conditions can be controlled. Yet, refining the presentation of visual stimuli (e.g. using a horse 2720 box/trailer as test area in which lighting can be controlled more easily) or using validated 2721 visual stimuli which can be perceived by horses under fluctuating light conditions. In order to 2722 establish AB as a measure of affect in horses, it is therefore necessary to gain more knowledge 2723 about factors (besides emotional state) that might modulate animals' attention to emotive 2724 stimuli such as facial expressions. Since cognition and welfare are interconnected,

2725 understanding how animals perceive and process emotive stimuli in their environment is 2726 necessary for the successful implementation of improved housing and management 2727 conditions.

2728 **4.5 Conclusion**

2729 The current study is the first to investigate affect-related AB in horses. Overall, I found no 2730 evidence of attentional biases when horses viewed image pairs of unfamiliar conspecifics' 2731 showing negative, positive or comparatively more neutral facial expressions. However, 2732 attention to the negative emotional stimuli was modulated by welfare score, with horses 2733 assessed as having putatively better welfare attending to the negative stimuli significantly 2734 longer than horses with putatively poorer welfare. In addition, emotion-unrelated stimulus 2735 characteristics as well as experimental circumstances associated with testing location may 2736 have influenced their response to visual stimuli. These findings emphasis the need for further 2737 refinements to experimental designs in order to test attention bias in horses, including a 2738 continuing need to expand our understand of horse perceptual abilities and the development, 2739 or identification of better species-appropriate test stimuli.

- 2741 Table 4.3 Model outputs from the LMER analysis with attention
- 2742 to the negative stimulus as dependent variable.

- 2744 Table 4.4 Model outputs from the glmmTMB analysis with proportion
- 2745 of time attending to the negative stimulus as dependent variable.

	Duration Stimulus (negative)		
Predictors	Estimates	CI	p
(Intercept)	-4.37	-	0.109
		9.72 – 0.98	
OppositeStimulus [positive]	0.81	0.02 - 1.60	0.044
Yard [yard 2]	3.19	1.85 – 4.52	<0.001
Yard [yard 3]	1.62	0.43 – 2.82	0.007
WelfareScore	0.34	0.03 – 0.65	0.033
Random Effects			
σ^2	5.46		
T 00 ID	0.00		
N ID	44		
Observations	136		
Marginal R ² / Conditional R ²	0.166 / N	A	

	Proportion Stimulus		
Predictors	Estimates	CI	р
(Intercept)	1.03	0.40 - 2.65	0.945
OppositeStimulus [positive]	1.22	0.76 – 1.93	0.410
SideStimulus [Right]	0.83	0.52 – 1.32	0.438
Sex [m]	0.70	0.42 - 1.17	0.174
Yard [yard 2]	1.49	0.68 - 3.28	0.324
Yard [yard 3]	0.90	0.45 – 1.80	0.757
StimID [B]	1.09	0.53 – 2.24	0.820
StimID [C]	0.98	0.49 – 1.97	0.962
StimID [D]	1.01	0.51 - 1.98	0.981
WelfareScore	1.26	0.93 – 1.70	0.135
Random Effects			
σ^2	0.04		

0.00

44

136

0.747 / NA

2743

124

2746

τ_{00 ID}

N ID

Observations

Marginal R² / Conditional R²

-

2747 Table 4.5 Model outputs from the LMER analysis with attention to the 2751 Table 4.6 Model outputs from the glmmTMB analysis with proportion of 2748 positive stimulus as dependent variable.

2752 time attending to the positive stimulus as dependent variable.

	Duratio	on Stimulus (po	ositive)		Proportio	on Stimulus (_l	
redictors	Estimates	CI	р	Predictors	Estimates	CI	
(Intercept)	3.62	1.66 – 5.58	<0.001	(Intercept)	2.32	0.71 – 7.57	
Sex [m]	-0.55	-1.71 – 0.62	0.359	Sex [m]	0.63	0.31 - 1.26	
SideStimulus [Right]	-0.69	-1.90 – 0.52	0.265	Yard [yard 2]	0.81	0.23 – 2.91	
Yard [yard 2]	0.71	-1.45 – 2.88	0.518	Yard [yard 3]	0.89	0.30 – 2.70	
Yard [yard 3]	0.75	-1.15 – 2.65	0.439	StimID [B]	1.27	0.48 – 3.37	
StimID [B]	0.52	-1.12 – 2.17	0.533	StimID [C]	0.85	0.34 – 2.15	
StimID [C]	0.10	-1.46 – 1.67	0.897	StimID [D]	0.59	0.22 - 1.62	
StimID [D]	-1.00	-2.68 - 0.68	0.242	SideStimulus [Right]	0.87	0.42 – 1.77	
WelfareScore	-0.33	-1.07 – 0.40	0.372	WelfareScore	0.82	0.52 – 1.28	
Random Effects				Random Effects			
σ^2	5.21			σ^2	-0.16		
τ _{00 ID}	0.00			τ ₀₀ ιd	0.00		
N _{ID}	40			N _{ID}	40		
Observations	65			Observations	65		
Marginal R ² / Conditional R ²	0.095 / N	IA		Marginal R ² / Condition	onal R ² -4.237 / I	NA	
				2753			

2750

- 2755 Table 4.7 Model outputs from the LMER analysis with total attention 2758
- 2756 duration as dependent variable

_

	Total attention duration		
Predictors	Estimates	CI	p
(Intercept)	2.21	1.81 - 2.61	<0.001
StimulusPair [positive_negative]	0.17	-0.08 - 0.42	0.178
StimulusPair [positive_neutral]	0.12	-0.13 - 0.38	0.342
Sex [m]	-0.03	-0.25 - 0.18	0.771
Yard [yard 2]	1.28	0.92 – 1.65	<0.001
Yard [yard 3]	0.97	0.64 – 1.29	<0.001
StimID [B]	0.20	-0.11 - 0.51	0.207
StimID [C]	-0.12	-0.43 - 0.18	0.417
StimID [D]	-0.19	-0.49 - 0.10	0.203
WelfareScore	0.06	-0.07 – 0.20	0.349
Random Effects			
σ²	0.54		
τ _{00 ID}	0.00		
N ID	44		
Observations	201		
Marginal R ² / Conditional R ²	0.283 / N	A	

2761 Summary

2762 In Chapter 4, I reported that horses looked at the images of conspecifics, putatively showing 2763 differently valenced facial expressions, for similar lengths of time. One explanation for not 2764 finding any significant evidence of AB might be that the horses did not recognise the image 2765 as representations of conspecifics. Previous studies in horses have reported that this species 2766 possess the ability to recognise objects and individuals from 2-D representations (e.g. printed 2767 images, computer screens) and vice versa. However, these findings underly several experimental limitations. Hence, in this part of the chapter, I investigated image recognition 2768 2769 in the same horses tested in Chapter 5. I developed a relative simplistic test protocol in which 2770 horses were first trained to discriminate between two real-life objects, before their responses 2771 to the images of the object displayed as on-screen images were tested. I used objects rather 2772 than images of other conspecifics as stimuli for practical reasons, including the more 2773 straightforward approach to train horses to exhibit distinguishable responses to objects which 2774 then allowed testing their spontaneous response to object images. I expected that horses 2775 would show the same conditioned response (*i.e.* muzzle contact object for food) to digital images of objects they had previously learnt to discriminate as an indication of image 2776 2777 recognition. At first image presentation, all but two horses spontaneously responded to the 2778 images with the learnt behaviour by contacting one of the two images with their muzzle, but 2779 the number of horses touching the correct image was not different from chance and only one 2780 horse touched the correct image above chance level when tested over 10 image trials. My 2781 findings thus question whether all horses recognise real-life objects from digital images as this 2782 ability might not be generalisable across individuals. I also discuss how methodological factors 2783 and individual differences (i.e. age, welfare state, learning ability) might have influenced 2784 horses' responses to the images, and highlight the importance of validating the suitability of 2785 this type of stimuli for cognitive studies in horses.

2786

2787

2788 5.1 Introduction

2789 Visual 2-dimensional representations (*e.g.* printed photos, digital images, silhouettes, videos) 2790 are used as substitutes for real-life objects, or individuals, in cognition studies of non-human 2791 animals, including horses. Screen-displayed visuals are of advantage in research as stimulus 2792 timing and presentation of identical stimuli can be repeatedly presented to the same or 2793 different subject animals (D'Eath, 1998). However, scientific evidence of object-image 2794 recognition in animals is not always consistent (reviewed in Bovet & Vauclair, 2000; Fagot, 2795 2000; Weisman & Spetch, 2010). This might be because pictures designed for the human eye 2796 may not result in the same sensory experiences in other species with different functional 2797 visual systems (Fagot & Parron, 2010; Weisman & Spetch, 2010). Moreover, how images are 2798 perceived and cognitively processed is not fully understood for most animal species (Fagot, 2799 2000; Fagot et al., 2010). For instance, Fagot et al. (2010) proposed that animals could 'read' 2800 images using different processing modes. In a mode of confusion, images and their real-life 2801 exemplars are perceived and treated as functionally and physically the same thing. 2802 Conversely, in a mode of *independence*, images could be perceived as different from their 2803 referents without making an association between objects and their images. In a processing 2804 mode of *equivalence*, images are understood as representations of their referents (*i.e.* images 2805 are used as referential cues for real-life objects, Fagot 2000; Fagot et al. 2010).

A variety of factors, including cognitive limitations or experience with images, could influence which processing mode is deployed by animals and ultimately lead to differences in how images are treated by humans and other animal species (Fagot & Parron, 2010). Therefore, the suitability of artificial representations (*e.g.* digital images, videos) for animal studies is likely to depend on the purpose of the stimuli. For instance, if images are used to imitate real stimuli in behavioural experiments, animals need to respond to images in a comparable way to how they respond to real stimuli (D'Eath, 1998).

Investigating image recognition is challenging because pictures can never be identical to their
3D referents given the lack of dimensionality, depth cues and olfactory characteristics, which
results in substantial sensory differences between objects and their 2D imitations (Aust &
Huber, 2006; Bovet & Vauclair, 2000). Prior to image processing, the perceptual abilities of
the viewer also need to be considered, for instance, whether an animal can identify an object

from an image despite the lack of depth cues or additional cues (*e.g.* reflectance of
photographic surface; Fagot & Parron, 2010).

2820 Unlike in humans, the visual field of horses is mainly monocular (*i.e.* visual input is received 2821 from just one eye; Waring, 2003). Binocular vision allowing depth perception is only possible 2822 within a relatively small area in front of the horse's head (55°- 65°; Hughes, 1977) extending 2823 downwards along the midsagittal plane (the vertical axis dividing the head in left/right) at 2824 approximately 75°, enabling horses to view the ground in front of them with both eyes (Duke-2825 Elder, 1958). A blind spot interrupts the almost panoramic visual field in front of the horses' 2826 forehead (Waring, 2003). In addition, visual acuity is much poorer in horses compared to most 2827 other terrestrial mammals (Rørvang et al., 2020). Horses have dichromatic vision resulting in 2828 similar colour perception to humans affected by red-green blindness (Hanggi et al., 2007). 2829 However, equine vision is highly adapted to low-light conditions with a high ratio of rods to 2830 cones and a reflecting tapetum lucidum enabling scotopic vision (*i.e.* ability to see under low-2831 light conditions) superior to that of humans (Hanggi & Ingersoll, 2009a). Given these visual 2832 differences, it appears that humans and horses see the world differently (Saslow, 2002). This 2833 raises the question of whether artificial stimuli such as digital images generated through 2834 computer projections are suitable representations of real-life objects for horses and other 2835 ungulate species sharing these traits (e.g. cattle, goats, sheep; Jacobs et al., 1998). Hence, 2836 further validation of whether horses recognise the content of digital stimuli is necessary.

2837 Generally, two different experimental approaches are applied to test image recognition in animals (reviewed in Bovet & Vauclair, 2000; Weisman & Spetch, 2010). For one, animals' 2838 2839 spontaneous responses to artificial representations of biologically relevant stimuli (e.g. 2840 photos of food, prey, predator or conspecifics) are tested as an indication of direct transfer 2841 (*i.e.* images are treated as the same as objects). In this case, the same adaptive behaviour is 2842 provoked by the artificial representations as if the real referent was present (Bovet & Vauclair, 2843 2000; Weisman & Spetch, 2010). A study in sheep, another ungulate species, found that 2844 animals respond to the image of a sheep with species-specific social behaviour (e.g. sniffing 2845 of the anogenital region and the head) and the sheep image appears to have fear-reducing 2846 effects on socially isolated sheep comparable to the presence of real conspecifics 2847 (Vandenheede & Bouissou, 1994). Interestingly, a human image did not result in the same 2848 fear response as elicited by a real human, suggesting that different stimuli types may be

2849 processed differently by sheep (*i.e.* sheep image possibly confused with a real sheep whereas 2850 the human image was not treated as a substitute; Vandenheede & Bouissou, 1994). Horses 2851 also respond to 2D and 3D horse imitations (photograph, life-size model) with sniffing 2852 behaviour near the head and flank areas corresponding to their natural approach of 2853 conspecifics, while an incomplete horse drawing and a dog image were not approached (Grzimek, 1943). These observations might suggest that horses can recognise conspecifics 2854 2855 based on specific cues, such as social cues conveyed by a near-realistic 3D model and 2856 photograph but not a drawing. However, approach and sniffing behaviours are also 2857 associated with exploration meaning that using explorative responses as outcome measures 2858 is not specific to image recognition alone and could result from other motivations, such as 2859 gathering novel information. Similar reasoning may apply to other studies that use 2860 spontaneous approach behaviours to indicate image recognition in horses (e.g. Smith et al., 2861 2016; Wathan et al., 2016). Physiological changes (mean heart rate) measured alongside 2862 horse behaviour were interpreted by the authors as support for horses' ability to differentiate 2863 between emotional stimuli, although cross-validation through multiple physiological 2864 measures (e.g. HRV indices to infer autonomic response; von Borell et al., 2007) could have 2865 strengthened these findings even more.

2866 An alternative to the above-described adaptive behaviour responses is studying animals' 2867 ability to transfer acquired (operant) responses associated with real-life objects to their 2868 pictorial representations (Bovet & Vauclair, 2000). For example, Cabe (1976) trained pigeons 2869 to discriminate between two solid objects (one rectangular block and a cross) by pecking the 2870 rewarded stimulus. The birds spontaneously transferred the learnt discrimination rule when 2871 the objects were replaced by pictorial representations (e.g. black-and-white photographs, 2872 white-on-black silhouettes) demonstrating that pigeons can recognise objects from images 2873 (Cabe, 1976). Using a similar approach, Hanggi (2001) reported that, after multiple 2874 presentations, horses (N=2) were able to transfer a learnt behaviour (contact object with their 2875 nose for food) from real objects (various toys varying in colour, shape and size) to their 2876 pictures, indicating image recognition. However, the ability to categorise images does not 2877 automatically provide evidence of representational insight (*i.e.* the subject understands what 2878 the image stands for; Aust & Huber, 2006). The horses might have learnt to discriminate 2879 between the images during repeated testing, e.g. based on invariant features between

images (*e.g.* colour, shapes, or distribution of light/dark patterns) unrelated to the real objects. According to the author, this explanation seems unlikely given the large number and diversity of objects tested (Hanggi, 2001). However, the same two horses were previously reported to understand shared characteristics between stimuli (i.e. pattern rules; Hanggi, 1999), indicating their ability of categorisation learning, which one animal was reported to still remember several years later (Hanggi & Ingersoll, 2009b).

2886 Experimental biases and ambiguity of outcome measures can further hamper the validity of 2887 image recognition evidence. For instance, it has been reported that horses can recognise 2888 humans from images because they were not only able to differentiate between happy and 2889 angry human faces, but also appear to possess emotional memory (Proops et al., 2018). 2890 Horses were described as reacting 'appropriately' following the theory of emotional 2891 lateralisation (*i.e.* left-eye bias for humans with angry faces and more time engaging in stress-2892 related displacement behaviours) when encountering the real human hours after they had 2893 seen a photo of the same person displaying an angry face. However, due to experimental 2894 limitations (e.g. horses kept in different conditions between tests, non-specificity of response 2895 behaviours (e.g. scratching, floor sniffing; these activities are also expressed in other contexts; 2896 Waring, 2003) and statistical weakness (e.g. no control conditions), the robustness of these 2897 findings has been questioned (Amici, 2019). Moreover, inferring evidence of recognition from 2898 emotional responses might not be straightforward in absence of control (*i.e.* non-emotional) 2899 comparisons. Hence, it is possible that horses' responses were driven by image-inherited cues 2900 unrelated to the emotional image content (e.g. image colours, brightness or contrast). The 2901 study by Lansade et al. (2020b) reduced experimental biases by training horses first to reliably 2902 select a screen image showing one of four human faces instead of images of objects (novel 2903 objects differing on each trial), thereby priming horses to respond to content-specific 2904 information. The horses significantly discriminated between familiar faces and a novel face. 2905 When a photo of the horses' keeper replaced the training faces, the animals again selected 2906 the keeper image at the above chance level suggesting that the keepers' faces were also 2907 identified as familiar. Alternatively, the keeper images might have been more similar to each 2908 of the training images than the novel images. In a follow-up study using on-screen images, 2909 Lansade et al. (2020a) controlled for this and found that horses can reliably select familiar

2910 faces paired against unfamiliar faces, despite removing photo colour, external cues2911 (hairstyle), or facial features (eyes).

2912 Overall, given a variety of experimental difficulties in this area, there is still a need for further 2913 evidence of the ability of horses to recognise the content of screen images and their 2914 relationship with real-life objects. The motivation of this study was therefore to test if horses 2915 spontaneously respond to digital images of two real-life objects, which they had previously 2916 learnt to discriminate. I predicted that horses would touch the images of the correct 2917 (rewarded) object at a level above chance if they recognised the images as real objects or 2918 representations of such. I only tested horses' transfer ability from real-life objects to on-2919 screen images, and not the reverse (*i.e.* training horses with images to test discrimination with 2920 their real-life counterparts), to gain evidence that digital images are suitable stimuli for 2921 cognitive tests in this species. For this, I developed a relatively simple and practical approach 2922 and also only used two real-life objects.

2923 Animals' performance in cognitive tests can be influenced by individual characteristics, 2924 including personality (Carere & Locurto, 2011; L. R. Dougherty & Guillette, 2018), learning 2925 speed, and motivation to engage in the task (reviewed in (Rowe & Healy, 2014). In horses, 2926 age (Krueger et al., 2014), sex (Murphy et al., 2004), but also emotional state (Christensen et 2927 al., 2012; Valenchon et al., 2013), and welfare status (reviewed in Hausberger et al., 2019) 2928 have been identified as sources of individual variation in cognitive performance. Therefore, I 2929 tested each horse in a total of 10 trials and assessed the effects of intrinsic (*i.e.* age, welfare 2930 score) and experimental factors (e.g. type of target, trial order, facility) on horses' 2931 performance.

2932 **5.2 Methods**

2933 5.2.1 Ethical statement

This study was approved by the Animal Welfare and Ethical Review Body of the University of Plymouth (ETHICS-41-2020). The experimental procedure was below the threshold for regulation by the UK Animals (Scientific Procedures) Act 1986 (ASPA) and followed the Essential 10 ARRIVE guidelines 2.0 (Percie du Sert et al., 2020). The horses belonged to two UK riding schools that consented to the use of their animals. Housing, care and health check were provided by the riding schools. The animals remained at their home facility at the end

of the study, except for one horse that was relocated during our data collection for reasons not related to this study. Horses that did not learn object discrimination in stage 1 were excluded from the object recognition test in stage 2.

2943 5.2.2 Animals and housing

2944 In total, 36 horses of mixed breeds from two UK riding schools (yard A: N=17, mean±SD age 2945 10.6±2.5 years; yard B: N=19, 16.6±6.5 years, of which three animals did not complete training 2946 at this yard as one was relocated and two became aggressive towards nearby conspecifics 2947 during training) were trained in an object discrimination test (ODT, stage 1). All horses that 2948 completed stage 1 (i.e. discrimination between the real objects; N=28) were tested in the on-2949 screen object recognition test (ORT, stage 2). However, one horse was scared of the test setup 2950 and was therefore excluded from testing, resulting in a total of 27 horses (16 from yard A of 2951 which 6 were females, 11 from yard B of which 4 were females) used in the ORT. In both 2952 facilities, horses were kept in single stalls, or tie-stalls, with full, or limited visual/physical 2953 contact with conspecifics during daytime (further details in Chapter 3). All horses had pasture 2954 access (in stable groups) at night and/or during parts of the day. Hay provision was restricted 2955 (*i.e.* facilities adjusted hay allowance based on body weight), and horses additionally received 2956 an adjusted diet (at yard B, brand Thunderbrook Equestrian), or not (at yard A where horses 2957 were 'on a diet' due to the lowered workload associated with COVID-19 restrictions). Water 2958 was freely accessible through automatic troughs in yard A and provided with water buckets 2959 in yard B.

2960 5.2.3 Experimental design

The experimental design consisted of two stages summarised in Figure 5.1. In stage 1, the horses were trained to discriminate between two real objects by touching the rewarded (target) object with their muzzle to receive a food reward before their spontaneous response to on-screen images was tested in stage 2.

2965 **5.2.3.1 Object discrimination – stage 1**

All horses were first trained inside their stall by a single familiar person (experimenter SK) to respond to real objects and discriminate between the target (rewarded) and an unrewarded object. The horses were able to move around freely (although six horses at yard B were tethered as they were kept in tie-stalls). Two objects ($kong^{TM}$: red dog toy, Ø 10 cm, length 16 cm; *ring*: doughnut-shaped dog toy, Ø 20 cm, depth 4 cm, with dark and light blue stripes, see

Figure 5.1) used as target objects were mounted onto a 50 cm wooden stick to facilitate the presentation of the objects in different positions and at distance to the experimenter. Which object a horse received as *target* (rewarded object) was pseudo-randomly allocated, ensuring that the number of horses trained with the same target was evenly distributed across yards. As only horses that completed ODT and learnt the discrimination within 200 trials (20 trials blocks) were used in ORT, the final number of horses tested in ORT with the ring and kong[™] as target object was 11 and 16, respectively.

2978 The first training step consisted of shaping horses' responses to the target object using 2979 instrumental conditioning. The experimenter moved toward the horse's shoulder (whichever 2980 side was most accessible) hiding the target behind her back. Standing at shoulder height, she 2981 then slowly moved the object into view for the horse and held the target at approx. 20-30 cm 2982 from the horse's muzzle (approx. 1.0-1.2 m above the ground depending on the horse's 2983 height). The horse could voluntarily move towards the object and contact with the object was 2984 never forced. Upon the first voluntary contact, the horse was instantly rewarded with a piece 2985 of carrot retrieved from a treat bag attached to the experimenter's waist at her back. At the 2986 same time, the target was moved behind the experimenter's back. Within 5 s of rewarding 2987 the horse, the same motion of moving the target near the horse's muzzle was repeated and 2988 the horse was instantly rewarded upon voluntary contact. All contacts with the object only 2989 (regardless of where on the object and with which part of the muzzle) were rewarded. The 2990 target training was repeated for 10 consecutive trials. The experimenter then left the stall to 2991 refill the treat bag again with 10 pieces of carrots and repeated this training step so that each 2992 horse received a total of 20 single-target trials.

2993 After a 2-min break, 10 single target trials were conducted again to remind the horses of the 2994 correct (familiar) target before a second unfamiliar object was introduced. The experimenter 2995 followed the same procedure as before to present the objects, except that now two objects 2996 were shown to the horses simultaneously for object discrimination training (ODT, see Figure 5.1). For this, the experimenter moved both objects simultaneously from behind her back to 2997 2998 in front of the horses' head holding each object by its handle in one hand at approx. 1.0-1.2 2999 m above the ground and with objects separated approx. 0.4-0.6 m. If horses touched the 3000 unrewarded object, the objects were shortly moved behind the experimenter's back for 5 s 3001 time-out before starting a new trial. If the unrewarded object was consecutively touched over

3002 three trials, the experimenter only presented the target to the horse (to remind it of the 3003 target, and guarantee that the horse received a reward and maintained motivation). The 3004 number of these *forced* trials was not recorded as this occurred rarely. If a horse did not touch 3005 any objects within 30 s, this response was regarded as incorrect, and a new trial was started. 3006 On each trial, the experimenter slightly altered her position relative to the horse, in which 3007 location and side, from the horse's perspective, the objects were shown, and alternated the 3008 hand used to reward the horse. These changes were done to avoid the horses developing side 3009 biases, or learning by association which object to contact relative to the handler (*e.g.* always 3010 chose object in experimenter's left hand). In addition, the side of object presentation was 3011 pseudo-randomly selected by the experimenter with the same object never being presented 3012 on the same side more than twice consecutively. Depending on horse availability, each horse received a maximum of 80 ODT training trials per day, which were conducted in eight trial 3013 3014 blocks (10 discrimination trials per block) with 2-min breaks between each block. Horses were 3015 trained over a maximum of 20 trial blocks (equal to 200 discrimination trials in total), and with 3016 a maximum of three days between training sessions. Training of three horses at yard 2 was 3017 interrupted due to COVID-19 restrictions and resumed 6 months later starting from ODT. For 3018 these horses, only trials conducted after the break were included in the data analysis. 3019 Learning criterion (LC) required to move to stage 2 (testing) was defined as performing eight 3020 or more correct responses per trial block over three consecutive trial blocks. The eight horses 3021 that did not reach LC within 20 trial blocks were not tested in stage 2.

3022 5.2.3.2 Object recognition test - stage 2

3023 Stage 2 consisted of the on-screen object recognition test (ORT) and was divided into three 3024 steps (see Figure 5.1). Pre-tests conducted in the horses' stall and the test arena using the 3025 real objects serving as verification of reliable discrimination performance before the horses 3026 were tested with images in the screen test.

3027 5.2.3.3 Pre-test in stall

The horses first received 10 single target trials conducted by the experimenter in the horses' home stalls. A second unfamiliar handler then entered the horses' stall alongside the experimenter to take hold of the lead rope, hence mimicking the handler's presence later in the test stage. The handler stood next to the horse's left shoulder, with his back turned to the horse and wearing noise-cancelling headphones to remain blinded to which of the two objects

3033 was the target. The experimenter presented the two objects for 10 trials as done in the ODT, 3034 except that the objects were now always presented in front of the horse's head at approx. 1-3035 1.5 m height, *i.e.* at a similar position as to where the images replacing the real object would 3036 later occur in the screen test. The handler's role was to reward the horse as indicated by the 3037 experimenter (saying her name to indicate an incorrect response, or the handlers' name to 3038 indicate a correct response) whilst remaining blind to the correct target to avoid any 3039 conscious or unconscious signalling from the handler (*i.e.* 'clever Hans effects', Pfungst & 3040 Rahn, 1911) during later stages of testing. If the horse performed ≥8 correct responses out of 3041 10 in the pre-test, it was immediately taken to the test area for the screen test. Horses that 3042 did not perform as such were re-tested in the same manner after a break (of varied duration 3043 for practical reasons, e.g. horse availability).

3044 5.2.3.4 Pre-screen test (PST)

3045 The horse was led into the test area (familiar indoor riding arena) where a back-projection 3046 polyvinyl chloride screen (1.6 m W x 2.5 m H) was set up. A multi-coloured pole (normally 3047 used as a training item and familiar to the horse) serving as a visual marker was placed on the 3048 ground directly in front of the screen at approx. 50 cm distance to indicate the position of the 3049 horse during testing. The horse was habituated to the screen (first turned off, then turned on 3050 not showing any images) and test equipment until it stood calmly in front of the screen. The 3051 screen was then turned off again and the handler positioned himself approx. 1 m away from 3052 the ground pole by the horse's left shoulder, turning his back towards the screen (position 3053 allowing him to stay blind to the images to be shown in the next phase). The experimenter 3054 stood in front of the horse (between the ground pole and screen) towards the right side of its 3055 head. She retrieved the real objects from a bucket and conducted 10 ODT trials following the 3056 same procedure as during the pre-test in the stall (*i.e.* the experimenter presented the objects 3057 and indicated to the blinded handler when to give deliver the reward). This was done to test 3058 if the horse still discriminated between the real objects in this different context (arena rather 3059 than stall). After five trials, the experimenter briefly moved behind the screen (out of view

from the horse³) to habituate the horse to her movement and absence. After five seconds,
she returned to her original position in front of the screen and conducted five more trials.

3062 If the horse performed ≥8 correct responses out of 10, the experimenter stepped behind the 3063 screen to start the screen test. If the horse performed below this level, it was led around the 3064 arena for approx. 2 min and the pre-screen test was repeated. In total, horses received a 3065 maximum of six pre-screen tests, with a maximum of three daily. All horses performed at the 3066 required criterion within six pre-screen tests. Table D1 in Appendix D shows the total number 3067 of pre-screen sessions conducted per horse before individuals moved on the screen test.

3068 **5.2.3.5 Screen test**

3069 In preparation for the screen test, each object was photographed three times using a Fujifilm 3070 X-T100 digital camera (focal lens 23 mm). Images were edited to remove the background so 3071 that only the object and wooden handle were visible in the final images (see Figure 5.1). Three 3072 versions of computer presentations (Microsoft PowerPoint) were created, each consisting of 3073 10 stimulus slides. Each slide contained one image of each object side-by-side on white 3074 background. Within the three presentations, the location of target images was balanced (50% 3075 left) and pseudo-randomised so that the target object was shown no more than twice in a 3076 row on the same side. The order and side of images varied between the three presentations 3077 to control for order effects. Additionally, the images were randomly rotated around their 3078 horizontal plane to change the position of the wooden handle. Later on screen, the images 3079 were shown approx. 1.1-1.2 m above the ground and at 0.5-0.7 m distance from each other.

Each stimulus slide was preceded by a white blank slide, except for the slides prior to stimulus slides 4, 7 and 9, which were black, indicating the points in the test at which real object trials were to be conducted (later described). Each horse was tested with only one out of the three presentations (equally spread across tested horses). Which presentation was projected was unknown to the experimenter at the time of testing, ensuring blindness to the target location (since the only slide she saw when starting playing the presentation was a blank slide).

3086 The screen test started immediately following the pre-screen test. The images were 3087 broadcasted from a laptop (Lenovo ThinkPad 13) *via* an LCD projector (HITACHI CP-

¹ but the experimenter could observe the horse *via* a web cam connected to a computer (Lenovo ideapad) serving as monitor

3088 WX3030WN) placed at approx. 2.5 m distance behind the screen. Standing next to the laptop, 3089 the experimenter used a remote control to start the slide show and advance the slides 3090 (thereby moving as little as possible to avoid any distracting noise). The first slide was blank 3091 but the experimenter advanced to the first stimulus slide as soon as the horse's head was 3092 straight in front of the screen (monitored *via* the webcam allowing to see the horse and the 3093 screen content). As soon as the horse contacted one of the images, the stimulus slide was 3094 immediately advanced to the next blank slide. At the same time, the experimenter indicated 3095 to the blind handler whether a reward should be delivered. A trial commenced as soon as the 3096 horses' head was straight in front of the screen again resulting in variable inter-trial intervals. 3097 The stimulus slides advanced automatically to the next blank slide after 20 s if no contact was 3098 made. In case the horse moved away from the screen immediately after trial onset (approx. 3099 within <2 s after stimulus onset), the presentation was moved to the previous blank slide and 3100 the trial was repeated as soon as the horse's head was back in a straight position in front of 3101 the screen.

3102 In total, 10 trials with images were conducted, interspersed with real object trials (where the 3103 experimenter returned to her position by the horse). Two object trials were conducted after 3104 image trials 3 and 6, and one object trial was conducted after image trial 8 (*i.e.* five object 3105 trials in total were conducted during the screen test). The real object trials were conducted 3106 as per the pre-screen test procedure, to remind the horses of the properties of the real 3107 objects, and to test whether they were still motivated to touch the objects, even if the images 3108 were not touched. To avoid horses learning to respond to the images when contacting the 3109 correct picture, a partial rewarding schedule was applied during the screen test (first and 3110 every third correct contact with the target image rewarded). Horses were always rewarded if 3111 they contacted the correct object on real object trials. Following the last stimulus trial (trial 3112 15), all horses received one last target trial (single object, not included in results) to ensure 3113 that all animals ended the testing with a positive experience. Horse behaviour was recorded 3114 throughout with three GoPro cameras (Hero 3+), and the number of correct responses was 3115 later extracted from the videos. Horses' responses to the objects/images in the screen test 3116 were extracted from footage and coded as 'correct' if the horses touched the rewarded 3117 object/image, and 'wrong' if the unrewarded object/image or if neither object/image was 3118 touched. Furthermore, the location (left/right) of the target image was recorded to assess

side effects. A second naïve coder analysed 30 % of the screen test videos, which were
selected at random (using the Excel random number generator and choosing the first 8 videos
after sorting in ascending order). Inter-observer reliability (Cohen's kappa) for coding the
response behaviours was very high (0.94).

3123



3124

3125 Figure 5.1 Experimental design comprising the objects discrimination training (ODT) and object 3126 recognition test (ORT). In stage 1 (A), horses learnt to contact a single rewarded object (target) with 3127 their muzzle to receive food during ODT. A second (unrewarded) object was subsequently added and 3128 horses trained to discriminate between both until it touched the correct object on ≥8 trials/10 over 3 3129 consecutive trials blocks. In stage 2 (B), horses were tested in three consecutive steps. A pre-screen 3130 test was first conducted in the horse's stall. When ≥ 8 correct responses were performed, the horse 3131 was moved to the test arena (illustrated as rectangle with dashed lines) and re-tested in a pre-screen 3132 test to ensure it performed reliably in the new environment. When ≥ 8 correct responses were 3133 performed, the horse was immediately tested with images on the screen (indicated by rectangle with 3134 solid black lines). During the screen test, the horse was presented with the real objects on five trials 3135 interspersed between the 10 image trials to test whether it was still motivated to touch the objects, 3136 even if the images were not touched.

3137

3138 5.2.4 Welfare assessment

Previous studies have suggested that welfare status can cause great individual variation in cognitive performance (reviewed in Hausberger et al., 2019). Therefore, I tested the effect of

3141 welfare state, *i.e.* the level of provided environmental resources (e.g. stall space, pasture 3142 access), social factors (e.g. ability and stability of social contact) and animal-based 3143 measurements (including health indicators, workload, abnormal behaviour), putatively 3144 contributing to good horse welfare on learning ability and test performance. The welfare 3145 score used for this analysis is described in Chapter 3. Briefly, for each welfare factor I 3146 evaluated, non-weighted numerical scores were given (0-1 indicating absence/presence of 3147 resource) and I combined all scores to calculate an overall welfare score (maximum score was 3148 20 with higher scores reflecting better welfare conditions).

3149 5.2.5 Statistical analysis

3150 Data were analysed in R (version 4.0.3 (R Core team, 2021)). Age and the welfare scores of 3151 horses between the yards were compared using Wilcoxon rank sum tests. The number of trial 3152 blocks needed to reach learning criterion in ODT was assessed as a measure of learning ability 3153 and followed a normal distribution (Shapiro-Wilk's test, p=0.09). Thus the effect of fixed 3154 factors (*i.e.* yard, target, sex) and covariates (*i.e.* age, welfare score) on learning ability was 3155 assessed by fitting generalised linear models (glm() function with Gaussian distribution in 3156 Ime4 package, Bates et al., 2015). Predictor covariance was checked with the vif() function 3157 from the car package (Fox & Weisberg, 2019), which indicated that age co-varied with the 3158 other fixed factors (vif=7.08). The effect of age on learning ability was therefore separately 3159 analysed using Pearson correlation test.

3160 Indication of recognition ability at group level was assessed by measuring whether the 3161 number of horses responding correctly and incorrectly on trial 1 of the screen test was 3162 significantly different from random using a Chi-square test. To test if the proportion of correct 3163 responses performed at group level in each of the ORT tests (*i.e.* pre-test, pre-screen test and 3164 screen test) was better than chance, one-sample Wilcoxon tests were used. Whether 3165 proportions of correct responses differed between trials following real object trials and trials 3166 following image trials was tested with a Chi-square test. Likewise, I tested the effect of reward 3167 delivery (i.e. received or withheld upon correct image contact) on subsequent trial 3168 performance using a Chi-square test.

Individual performance (correct/wrong response) during the 10 image trials was modelledusing generalised linear mixed models (GLMMs; glmer() function in lme4 package, binomial

- family) with target type (kong[™]/ring), target side (left/right), and trial order (after object/not
- after object) as categorical fixed factors, age and welfare score as covariates, and horse ID as
- 3173 random factor. P-values were exacted *via* the anova() function from the car package and
- 3174 reported as significant for $p \le 0.05$ and as trends for p < 0.1.

3175 **5.3 Results**

3176 **5.3.1** Learning ability during object discrimination training

3177 In total, 27 horses (16 out of 17 at yard A, 11 out of 16 at yard B) learnt to discriminate 3178 between the two objects. Overall, horses needed 11 trial blocks (median, Q1-Q3=7-15) to 3179 reach learning criterion (see individual LC results in Appendix D Table D1). Learning ability was predicted by target and yard, with horses from yard B (vs. yard A) and those trained with the 3180 ring (vs. with the kong[™]) needing more trials, but not by welfare level (see Table 5.1 for model 3181 3182 estimates). Pearson correlation test indicated a significant positive correlation between 3183 learning ability and age (t_{25} =4.09, r^2 =0.63, p=0.0003). Horses from yard A were significantly 3184 younger (mean±SD, 10.6±2.51; W=3950, p<0.0001) and had significantly lower welfare scores 3185 (14.1±1.30, W=4400, p<0.0001) than horses from yard B (age: 14.8±5.7, welfare score: 3186 15.5±1.73).

Table 5.1 Estimated regression parameters from the GLMM model. Learning ability (dependent
variable) predicted by welfare score, yard, sex and target type with information to the comparator
category in square brackets. P-values ≤0.05 are highlighted in bold.

Predictors	Estimates	Confidence Intervals	P-values
(Intercept)	12.80	-1.22 – 26.83	0.087
Welfare score	-0.45	-1.43 - 0.52	0.374
Yard [B]	4.67	1.46 – 7.89	0.009
Target [ring]	4.70	1.75 – 7.64	0.005
Sex [m]	1.27	-1.71 - 4.24	0.413
Observations	27		
R ²	0.471		

3190

3192 5.3.2 Objects Recognition Test

3193 5.3.2.1 Screen test

When the horses were first presented with the images, 92.6 % of the horses (25/27) spontaneously reacted to the images as trained, *i.e.* by contacting one of the two objects' images with their nose. However, the number of horses responding correctly by touching the target image (n=14) was not significantly different from the number of horses responding incorrectly (combining the 11 horses that contacted the image of the unrewarded object, and the 2 horses that did not contact the screen at all; X²₁=0.03, p=0.8).

3200 **5.3.2.2** Performance during the different stages of the ORT

3201 Figure 5.2 shows the proportion of correct responses during the pre-test (PT) and pre-screen 3202 test (PST) leading up to the screen test. Since all horses needed to perform at least 8 out of 3203 10 responses in the PT to move on to the PST, and in the PST to be tested with the images on 3204 the screen (which all horses did, although some animals were re-tested in PST, see Table D2 Appendix D), the effect of fixed factors (i.e. target, age, welfare score, sex) on individual 3205 3206 performance in the PT and PST tests was not further analysed. At group level, horses 3207 performed significantly better than chance (50%, V=36585, p<0.0001, Figure 5.2) in PT, PST 3208 (as required) and on object trials, but significantly below this threshold during image trials 3209 (V=7340, p<0.0001).



3210

3211Figure 5.2 Proportion of correct responses during each step of the object recognition test (ORT). The3212results of the screen test are shown separately for the 10 images trials ('screen test (images)') and 5

3213 real objects trials interspersed between image trials ('screen test (object)'). Dashed line indicates 50

3214 % correct (chance level performance) against which group level performance was tested (one-sample 3215 Wilcoxon test, **** p<0.0001 (note that performance above chance level during pre-test and pre-3216 screen test was required for the horses to move the screen test). Lines connecting boxplots show 3217 individual performances throughout the stages of the ORT. One horse touched the correct images 3218 significantly above chance level during screen test (images); data for this individual is indicated as bold 3219 line.

Considering individual performance over the 10 image trials, one horse performed above chance level by selecting the correct target images 9 times (p=0.021). Other horses (N=3) always contacted the correct image when making contact with the screen, but failed to touch the images on other trials (two horses did not touch the screen on four trials, one on two trials), and therefore were not considered to perform better than chance (6/10 and 8/10 correct, both p>0.1). An overview of individuals' performance when omitting trials where horses did not make any image contacts is provided in Appendix D Table D2.

3227 **5.3.2.3 Factors influencing response to the images**

Horses' response to the images (*i.e.* correct/wrong) was predicted by the type of preceding trial (p<0.001, model estimates shown in Table 5.2). Horses were more likely to respond correctly in trials following real object trials than in trials following images trials (X^{2}_{1} =8.45, p=0.004), although the proportion of horses touching the correct image was only 51.8% (Figure 5.3). Overall, horses did not make any image contacts on 144 trials (53.3%), whereas horses always approached the real-objects.

3234 Whether horses received a reward upon correct image contact or reward was (unexpectedly) 3235 withheld during a preceding image trial had no significant effect on horses' performance 3236 $(X^2_2=0.268, p=0.874)$. However, images presented on the right side were more like to result 3237 in correct responses than when the target was shown on the left (p=0.05, model estimates in 3238 Table 5.2). None of the other factors I explored (welfare score, yard, age, sex, target type, LA) 3239 had a significant effect on horses responses (see Table 5.2).


3240

Figure 5.3 Proportion of horses out of the 27 horses responding correctly or incorrectly depending on whether the preceding trial was refreshed with objects (yes) or not (no). More horses performed

whether the preceding trial was refreshed with objects (yes) or not (no). More horses performed correctly than incorrectly when the preceding trial was refreshed with objects (p=0.004).

3244

3245	Table 5.2 Model estimates of GLMM with 'response' as binary dependent variable (correct/wrong)
3246	and predictors with comparator information in square brackets. P-values ≤0.05 are highlighted in bold

Predictors	Odds Ratios	Confidence Interval	P-Value
(Intercept)	4.32	0.02 – 759.86	0.579
Yard [B]	0.53	0.15 – 1.85	0.317
Age	0.89	0.76 - 1.04	0.153
Welfare score	0.93	0.67 – 1.29	0.671
Side [R]	0.56	0.31 - 1.00	0.049
Target [ring]	1.87	0.60 - 5.87	0.282
AfterObject [yes]	2.98	1.65 - 5.39	<0.001
Learning ability	0.99	0.85 - 1.16	0.925
Sex [m]	1.63	0.61 - 4.36	0.334
Marginal R ² / Conditional R ²	0.183 / 0.344		

3247

3249 **5.4 Discussion**

3250 In this chapter, I investigated if horses can recognise real-life objects from on-screen images. 3251 The majority of horses initially reacted to images with the conditioned response (*i.e.* touching 3252 the target with their muzzle for food), but the number of horses touching the correct image 3253 was not significantly different from the number of horses contacting the wrong image. 3254 Therefore, performance at group level did not suggest that the horses recognised the real 3255 objects from their 2D representations shown on-screen. However, I found that more correct 3256 responses were being performed on image trials following real object trials, suggesting that 3257 horses' reactions to the images were not completely random. In fact, one horse selected the 3258 correct images at a level significantly above chance when tested repeatedly over 10 images 3259 trials, suggesting that this individual recognised the images either as the real object (confusion 3260 mode) or as a representation of it (equivalence mode; Fagot et al., 2010).

3261 Previous studies have reported that horses can recognise other individuals from photographs 3262 (Lansade et al., 2020a,b; Proops et al., 2018; Smith et al., 2016; Wathan et al., 2016). However, 3263 as presented in the introduction, the validity of existing evidence might be hampered by experimental limitations (see Amici (2019) for a discussion of Proops et al., 2018). Moreover, 3264 3265 discrimination ability is not automatic proof of recognition (Aust & Huber, 2006), and 3266 alternative mechanisms such as learning, categorisation (*i.e.* of biologically relevant objects 3267 such as food), or habituation might also influence animals' responses to repeated 3268 presentation with images (reviewed in Bovet & Vauclair, 2000). Here, I tested if horses would 3269 spontaneously respond to on-screen images with the same learnt response that they were 3270 trained to make to real objects, using a relatively low number of test trials and partial reward 3271 delivery to avoid learning. In contrast to previous reports, our horses failed to recognise the 3272 objects from images, except for one individual. Several aspects need to be considered to put 3273 our findings in context with previous findings.

When exposed to the images for the first time, all but two horses spontaneously responded to the images with the conditioned response, suggesting the horses made some association between images and objects since the stimuli provoked the learnt behaviour. I trained the horses to express their choice by contacting the target with their muzzle, because this conditioned behaviour is commonly used in horses tested in two-choice discrimination tests (*e.g.* Flannery, 1997; Hanggi, 2001, 2003; Lansade et al., 2020a,b). In retrospect, I question

3280 the suitability of this behaviour as a conditioned response. Horses naturally use their nose to 3281 explore unfamiliar items to gather olfactory/tactile information whilst inspecting novel 3282 objects (De Boyer Des Roches et al., 2008). Therefore, the horses might have contacted the 3283 images to explore the items rather than performing a conditioned behaviour. This might 3284 explain why I found no significant preference for either image at first presentation (trial 1). Utilising stimulus-specific adaptive responses as done in studies in other species (e.g. grasping 3285 3286 behaviour in marmosets (Oh et al., 2019), eating attempts of banana images in gorillas (Parron 3287 et al., 2008), or shaping behaviours distinctively different from normal horse behaviour (e.g. 3288 level pressing; (Dougherty & Lewis, 1991) could avoid this problem of ambiguity.

3289 Intriguingly, horses were nevertheless more likely to make correct responses to the images 3290 following real object trials than following image trials. Maybe responding to the real objects 3291 before seeing the images somehow facilitated horses' ability to transfer between the stimuli, 3292 despite perceptual differences (e.g. lack of depth cues), for instance by matching them based 3293 on relational sameness (e.g. shape). In fact, Flannery (1997) observed that horses have the 3294 capacity to learn higher-order discriminations based on relation between stimuli, such as 3295 geometric shapes. It could be that horses initially confused objects and images (*i.e.* seeing 3296 both as the same), but once they made physical contact with the images, the mismatch in 3297 sensory feedback (e.g. olfactory/tactile feedback) between the familiar object and images 3298 resulted in independent processing of both as completely different items. Moreover, cross-3299 modal differences (i.e. looks like target but does not smell/feel like target), might have 3300 stopped the horses from touching the images. Horses use cross-modal (visual/olfactory and 3301 auditory information) sensory input to recognise individuals (*e.g.* horses - Proops et al., 2009; 3302 humans - Lampe & Andre, 2012; Proops et al., 2013), but whether this is also true for 3303 identifying (familiar) objects has not been tested yet.

In addition, other experimental limitations might have influenced our findings. Work by one other group used digital stimuli (computer screens (Lansade et al. 2020a, b) and projections (Trösch et al., 2019, 2020), which is why I anticipated that this type of visual information would be suitable for this study. However, image quality and differences in colour perception of the images resulting from the use of the LCD projector (images generated from a light signal comprised of red, blue and green components but horses cannot perceive red/green colours) may have contributed to sensory image impressions different in horses to those generated

3311 by the real object, and to what humans see in digital images. Besides, the equine eye is 3312 adapted to dim light conditions and scattered light (e.g. from a bright light source such as a 3313 projector) can lead to loss of resolution (Hebel, 1976). One may wonder whether the close 3314 distance to the screen might have hindered our horses' ability to clearly see the items in front 3315 of them given the blind spot directly in front of their forehead and limited visual acuity in 3316 close proximity (Hebel, 1976; Rørvang et al., 2020; Timney & Macuda, 2001). Our setup seems 3317 appropriate, however, since others reported that horses successfully learn to discriminate 3318 between symbols of different shapes and sizes, and photographs, when standing directly (≤50 3319 cm) in front of a screen and contacting the stimuli with their muzzle (Gabor & Gerken, 2012; 3320 Tomonaga et al., 2015). However, varying the blinded handler position (always positioned on 3321 the left-hand side for practical reasons) should be considered for future work, since I found 3322 that targets presented on the right side were more likely to result in correct responses. The 3323 spatial relationship between cue, reward and response can influence discrimination learning 3324 (Hothersall et al., 2010; Miller & Murphy, 1964), which might explain why target location 3325 tended to affect performance.

3326 Maybe my results do not support previous reports of image recognition in horses because of 3327 the type of stimuli I used. From an adaptive perspective, processing visual cues of biological 3328 relevance are highly important, and images representing biologically relevant stimuli (e.g. 3329 prey, conspecifics, predators) are instrumental in studies of animal picture recognition where 3330 animals' spontaneous (initial) response to pictorial cues is tested (Bovet & Vauclair 2000). For 3331 instance, Kendrick et al. (1996) observed that sheep were much faster in learning to 3332 discriminate between images of conspecifics (familiar or unfamiliar) than between 3333 geometrical shape discrimination cues, possibly because sheep seem to cognitively process 3334 information associated with social familiarity (i.e. facial features of conspecifics) more 3335 efficiently than non-social cues. It seems probable that specialised sensory processing of 3336 social cues is also relevant to horses, since they show a range of postural and facial 3337 expressions for social communication (Waring, 2003; Wathan et al. 2016), and understand 3338 visual cues from humans (Proops & McComb, 2010). It seems therefore possible that equine 3339 studies using images of conspecifics (Wathan et al., 2016) or humans (Smith et al., 2016; 3340 Proops et al., 2018; Lansade et al., 2020b, a) tap into different sensory processing levels than 3341 when objects images are used. I chose real-life objects instead of images of conspecifics as

this allowed me to train and test horses' responses more easily under controlled conditions (*i.e.* excluding variation within the test stimuli). I also excluded food cues since disentangling animals' motivation to respond to food cues when food rewards are provided during repeated testing might be difficult. Nevertheless, I expected that the horses would pay attention to the on-screen stimuli if they perceive the images as equal to the real objects given that they had learnt to associate these with food (*i.e.* a biologically relevant resource).

3348 Digital images are increasingly applied in the study of horse cognition, but evidence that this 3349 species can recognise the content of digital images is still sparse. Hence, I investigated how 3350 horses spontaneously respond to on-screen images of known objects and did not consider 3351 testing the reverse (*i.e.* whether horses recognise real-life objects from images). Testing 3352 horses' transfer ability in the reverse direction would not answer the research question 3353 addressed here. It would require exposing the animals to screen images over multiple 3354 sessions to train the conditioned response behaviour, whereby horses might learn to respond 3355 to particular features of the images and therefore this would not allow investigation of 3356 spontaneous image recognition when applying the opposite transfer test. I do, however, 3357 encourage future research to study this further (considering confounding factors discussed in 3358 the introduction (section 5.1) regarding Hanggi et al., 2001), for instance, to understand what 3359 stimulus characteristics (e.g. colour - Hanggi et al., 2007; shape, or size - Hanggi, 2003; 3360 Tomonaga et al., 2015) drive recognition as these could be easily manipulated in digital images. Here, I only used two real-life objects distinctively differing in colour, shape and size 3361 3362 (stimulus features horses can generally discriminate) as using more items could have 3363 introduced more variability in individuals' responses making the interpretation of findings 3364 more difficult. Therefore, I believe my findings that horses overall did not perform reliably 3365 enough to suggest image recognition using two objects are of significance. However, I must acknowledge that my observations may not be generalizable as the use of different objects 3366 could have led to different findings. 3367

Only one out of 27 horses responded to the stimuli on screen above chance level suggesting that this individual recognised the images as objects or representations of such. Rapid learning seems unlikely given the experimental precautions I undertook. For example, I used partial reinforcement in the screen test to reduce the possibility that horses would respond to image-related cues, *i.e.* exhibit the muzzle contact as new behaviour specific to the images

rather than touching them because they were recognised as a replacement for the objects. Indeed, horses were not more likely to respond correctly to the images following images trials a reward was delivered upon correct response than when the reward was (unexpectedly) withheld. Nevertheless, learning cannot be completely excluded as partial reinforcement reduces, but does not exclude, the acquisition of a conditioned response compared to continuous reinforcement (Anselme, 2015; Gottlieb, 2005).

3379 My study design does not allow me to draw conclusions as to how this horse recognised the 3380 images (i.e. whether images and objects were seen as the same item (i.e. confused), the 3381 images seen as functional representations (equivalent) of the target or both processed 3382 independently). Nevertheless, this finding is interesting as it highlights the importance of 3383 considering individual variation in cognitive tests. In correspondence with other findings 3384 showing that older horses learn more slowly in a social learning task (Krueger et al., 2014), I 3385 found that older horses needed more trial blocks to learn the discrimination task. However, I 3386 found no association between welfare level, learning ability and test performance. Further 3387 study could investigate further inter-individual differences such as variations in personality or 3388 perceptual abilities on performance.

3389 **5.5 Conclusion**

3390 Only one of 27 horses responded to the images suggesting it recognised the images as objects 3391 or representations of such, while all other horses apparently failed to do so. As a species, 3392 horses may possess the basic capability to perceive the content of artificial visual stimuli such 3393 as digital images, but my findings indicate that image recognition might not be an ability that 3394 can be generalised across horses and testing situations. Regarding my results in the AB test 3395 (Chapter 4), the findings presented here further suggest that horses might have looked at the 3396 images of conspecifics for similar lengths of time as these were possibly not recognised as the 3397 faces of other horses. Further research is warranted to understand how horses perceive (at 3398 the sensory level) and interpret (at the cognitive level) images made for the human eye, 3399 especially if they are to be utilised as representations of real-life objects, as well as inter-3400 individual variations in such abilities. Until then, we do not know if humans and horses see 3401 eye to eye when viewing this type of artificial stimuli.

Chapter 6 I Changes in heart rate variability and facial surface temperatures in response to immediate and withheld reward

3404 Summary

3405 This chapter presents my investigation of changes in physiological parameters associated with 3406 emotional arousal and valence when horses experience a putative positive (immediate food 3407 reward, IR) and a less positive (withheld food reward, WR) emotional situation. Recording 3408 cardiac activity with an electrocardiogram (ECG) and detecting surface temperature changes 3409 via infrared thermography (IRT) allows the assessment of autonomic nervous system (ANS) 3410 activity, which plays an important role in emotional control. ANS activity has mainly been 3411 studied in animals experiencing aversive conditions, whereas physiological changes induced 3412 by positive emotions are less well understood. Hence, I aimed to apply IR and WR whilst 3413 horses' performed a conditioned behaviour (muzzle contact with a target object) to 3414 investigate physiological profiles in response to shifts in emotional states.

3415 There was a significant increase in mean HR from baseline during IR and WR, whereas 3416 maximum eye temperature decreased significantly during both reward phases. The observed 3417 changes in these arousal indices suggested that the reward treatments, whether positive or 3418 negative, induced states of elevated physiological activity. However, valence indicators (heart rate variability (HRV) indices, L-R maximum eye temperature) did not differ significantly 3419 3420 between reward phases. Maybe IR and WR did not induce the anticipated shifts in emotional state, or HRV and temperature measures are not suitable markers of valence as I will discuss 3421 3422 in this chapter.

3423

3424 6.1 Introduction

3425 Good animal welfare constitutes not only the absence of negative emotional states but also 3426 the presence of positive emotional experiences (Boissy et al., 2007). To identify conditions 3427 that are positive (*i.e.* rewarding) to animals, the validation of reliable markers of positive 3428 emotions is required. When considering the physiological component of animal emotions, 3429 research has mainly focussed on measuring physiological expressions/changes in situations 3430 likely to induce negative emotions (e.g. fear, stress (Veissier & Boissy, 2007; von Borell et al., 3431 2007)) and reliable physiological markers of positive emotional states are still lacking (Boissy 3432 et al., 2007). Moreover, many emotion-associated physiological changes are often indicative 3433 of emotional arousal rather than valence (Paul et al., 2005). Nonetheless, recording changes 3434 in physiological profiles can contribute to better insights into animals' emotional states, for 3435 instance, as additional measures when observing 'general' animal behaviour does not allow 3436 distinguishing between positive and negative emotions (e.g. locomotion, body posture, 3437 vocalisation (Desire et al., 2002)).

3438 In mammals, the ANS plays a crucial role in emotion regulation (Kreibig, 2010; Levenson, 3439 2014; Thayer & Brosschot, 2005) in combination with other physiological functions associated 3440 with bodily homeostasis (e.g. digestion, cardiac activity, and thermoregulation (Shields, 3441 1993)). More specifically, internal conditions are constantly monitored and physiological 3442 short-term adjustments are made by the ANS to enable the body to cope with changes in 3443 internal and external demands (Levenson, 2014). These autonomic adjustments control the 3444 bodily functions that ultimately determine health, illness, and longevity (Levenson, 2014). The 3445 ANS consists of two subsystems, the parasympathetic (PNS), which is responsible for 3446 physiological functions of energy conservation and restoration (e.g. rest and digestive 3447 functions), and keeping the body at an efficient minimum energy expenditure (i.e. 3448 homeostasis), whereas the sympathetic nervous system (SNS) prepares the body for stressful 3449 or emergencies (e.g. fight or flight response). Autonomic imbalance, characterized by 3450 decreased PNS activity and/or hyperactivity of the SNS is associated with several somatic and psychological conditions, which can become pathogenic when sustained over longer periods 3451 3452 (Thayer & Brosschot, 2005). Similarly, shorter-lasting negative emotional states, e.g. acute 3453 stress in response to environmental stimuli perceived as threat, can be accompanied by 3454 physiological profiles signalling adaptive shifts towards SNS dominance in humans (Kreibig,

3455 2010) and animals (von Borell et al., 2007). Stress is generally defined as a biological state of 3456 disrupted homeostasis induced by specific stimuli (*i.e.* stressors (Selye, 1976)). If a stimulus is 3457 perceived as stressor, this will result in the activation of the sympatho-adrenomedullary axis 3458 (SAM) as the first part of the stress response system (Godoy et al., 2018). This induces the 3459 secretion of norepinephrine, which ultimately initiates energy mobilisation, metabolic changes and other bodily and behavioural reactions to adapt to the stressor (reviewed in e.g. 3460 3461 Gaidica & Dantzer, 2020; Godoy et al., 2018). In contrast, high vagal tone (reflecting PNS 3462 activity) is associated with autonomic balance dominated by PNS regulation, which promotes 3463 good health and decreased susceptibility/increased ability to respond to stress (von Borrell et 3464 al., 2007).

3465 It is possible to measure shifts in the net interaction of the PNS and SNS (*i.e.* sympathovagal 3466 balance) in target organs receiving dual innervation, such as the heart. The activity of the 3467 sinoatrial node (SN), the 'pacemaker' of the heart, is controlled by the PNS (mainly supplied 3468 by the vagus nerve), which has dampening effect on heart rate (HR). Without vagal 3469 stimulation, the intrinsic (SN-generated) resting heart rate in humans would be 100-120 bpm 3470 (instead of 50-90 bpm (Nanchen, 2018)), henceforth much higher than under vagal influence 3471 (von Borrell et al., 2007). In contrast, the SNS has a stimulating effect on HR. In addition, by 3472 measuring cardiac activity non-invasively, e.g. via an electrocardiogram (ECG), indices of the 3473 irregularity of time intervals between two contractions of the cardiac ventricles, also called 3474 heart rate variability (HRV), can be derived as quantitative markers of autonomic activity 3475 (Taskforce, 1996). Fluctuations in inter-beat time intervals (IBIs) arise from the dynamic, but 3476 non-additive, interplay between PNS and SNS stimulation on the heart, a phenomenon 3477 enabling the flexibility of biological systems to adjust to different internal/external conditions, 3478 including emotionally-valenced situations (He, 2020; von Borell et al., 2007). Measuring HRV 3479 has the further advantage of detecting physiological changes in real-time, *i.e.* shortly 3480 following exposure to emotional stimuli, as vagal-induced changes in HR occur rapidly (<5s) 3481 and SNS-controlled changes in HRV occur with 20-30s delay (due to slower releases of 3482 norepinephrine from SNS nerve cells as compared to PNS acetylcholine; von Borrell et al., 3483 2007). Importantly, the separate effects of PNS (vagus nerve) and SNS cannot be estimated by simple addition or subtraction of the relative components of the two systems, as vagal and 3484 3485 SNS modulation on cardiac activity do not necessarily have antagonistic effects. That is, a

reduction in one does not warrant an increase in the other (von Borrell et al., 2007). An increase in HR, as opposed to HRV, can be induced by elevated sympathetic activity, a reduction in vagal tone, or both (von Borrell et al., 2007). Hence, HR is not a suitable measure of emotional valence, but increased HR is indicative of heightened emotional arousal (von Borell et al., 2007; Wascher, 2021).

3491 The assessment of HRV has predominantly been used to measure stress (chronic and acute) 3492 in farm animals (reviewed in von Borrell et al., 2007; Gaidica & Dantzer, 2020). In horses, stress-related low HRV caused by a shift towards SNS dominance (i.e. adaptive switch to 3493 3494 sympathetic activation; Gaidica & Dantzer, 2020) has been reported in response to an acute 3495 psychological stressor, *i.e.* waiting for food (Bachmann et al., 2003), exposure to novel objects 3496 (Visser et al., 2002), or when being forced to move backwards continuously for 3 min 3497 (Rietmann et al., 2004). There is some evidence that positive emotional states alter HRV in 3498 dogs (e.g. Katayama et al., 2016; Zupan et al., 2016) and cattle (e.g. Lange et al., 2020, 2021). 3499 However, only a limited number of studies have measured HRV in horses subjected to 3500 situations presumably eliciting positive emotional states. Wiśniewska et al. (2019) tested the 3501 effect of music as a relaxation tool in geriatric horses. Horses exposed to 'new age' music for 3502 several hours per day over 28 days had significantly lower HR and higher HRV (vagal tone) 3503 activity, but only during the initial week of exposure while indices were not significantly 3504 different compared to control horses (not exposed to music) in the subsequent 2-3 weeks of 3505 music exposure, possibly because of habituation effects (Wiśniewska et al., 2019). Scopa et 3506 al. (2020) reported a significant decrease in HR and a significant increase in vagal tone when 3507 horses were brushed, suggesting grooming-induced relaxation effects, which was modulated 3508 by handler familiarity, as an increment in vagal tone was significantly higher when the person 3509 brushing the horses was a familiar rather than an unfamiliar person.

Another approach to quantify emotion-related ANS activity non-invasively is to measure changes in cutaneous temperatures of bodily or facial regions via infrared thermography (IRT, reviewed in Travain & Valsecchi, 2021). IRT allows the detection of infrared radiation emitted from surfaces, including body areas, with the advantage of detecting temperature changes in real-time and it is contact-free (*i.e.* without animal handling (Ioannou et al., 2014; Travain & Valsecchi, 2021)). Core (*i.e.* internal body temperature) and peripheral temperature (recorded at body surface level, *i.e.* cutaneous temperature) are controlled by the ANS via

3517 SNS-regulated changes in resistance in the vascular system (e.g. arteries, veins, or 3518 arteriovenous anastomoses, which are direct connections between arteries and veins) 3519 thereby altering blood pressure and blood flow, and ultimately modulating heat dissipation 3520 (Ioannou et al., 2014; Ziemssen & Siepmann, 2019). Both physical and psychological stress 3521 can evoke physiological changes leading to changes in core and peripheral temperatures, 3522 which can be used as a marker of emotional states in animals (Travain & Valsecchi, 2021). For 3523 instance, stress-induced hyperthermia (SIH) is a phenomenon referring to the increase in core 3524 temperature in response to an acute or repeated physical or psychological stressor 3525 (Bouwknecht et al., 2007; Nakamura, 2015) as a result of blood flow to internal organs (e.g. 3526 the brain, muscles) being prioritised. Cutaneous vasoconstriction further contributes to 3527 hyperthermia by reducing blood outflow to skin arteries and consequently lowering 3528 peripheral skin temperature (Vianna & Carrive, 2005).

3529 SIH is conserved across various endothermic species (Bouwknecht et al., 2007). For instance, 3530 placing rats in a chamber previously used as a shock box for fear-conditioning caused a 3531 significant drop in tail temperature and paw temperatures, and a significant increase in core 3532 temperature, indicating that psychological stress can induce SIH in these animals (Vianna & 3533 Carrive, 2005). In cattle, both physical stress (e.g. pain associated with castration (Stewart et 3534 al., 2010a) and disbudding (Stewart et al., 2008a)), and psychological stress (fear related to 3535 handling procedures, (Stewart et al., 2008b) induced eye temperature changes reflecting 3536 peripheral temperature changes. Stewart et al. (2008a) observed a rapid drop in the 3537 temperature of the area of the medial posterior palpebral border of the lower eyelid and the 3538 lacrimal caruncle (i.e. small pink gland in the inner corner of the eye) in calves following 3539 disbudding without local anaesthetic. The rapid decrease in eye temperature was thought to 3540 be the result of SNS-mediated changes in blood flow in the capillary beds in the lacrimal 3541 caruncle and medial canthus (inner corner of the eye) area causing a reduction in surface 3542 temperature in this area. The authors confirmed this hypothesis in a later experiment in which 3543 epinephrine infusion simulating increased SNS activity also significantly decreased eye 3544 temperature (Stewart et al., 2010b). Eye temperature increased and was significantly higher 3545 than baseline following the initial drop (lasting 5 min) when disbudding with and without local 3546 anaesthetic, possible due to the change in sympathovagal balance and switch from 3547 vasoconstriction to vasodilation resulting in increased blood flow to dissipate heat

accumulated during SIH (Stewart et al., 2008a). These findings demonstrate that changes in
ANS activity caused by negative situations associated with physical or mental stress can be
detected in peripheral temperature changes.

3551 It is possible to record thermal changes in other regions of interest (ROIs), including animals' 3552 ears (in e.q. dogs - Riemer et al., 2016; rabbits (Oryctolagus cuniculus) - Ludwig et al., 2007), 3553 the nose (in e.g. macaques - Nakayama et al., 2005; cattle - Proctor & Carder, 2015) or other 3554 cutaneous facial areas important for thermoregulation (*e.g.* the comb and wattle in chickens 3555 (Edgar et al., 2013; Herborn et al., 2015, 2018); forehead in cattle (Salles et al., 2016), or sheep 3556 (Joy et al., 2022)). In dogs, a brief separation from their owner decreased ear temperature, 3557 but the opposite effect was found when a person (either owner or stranger) returned to the 3558 test room (Riemer et al., 2016). As with HRV, the focus of IRT studies has been on negative 3559 emotional states (reviewed in Travain & Valsecchi, 2021), while fewer studies report 3560 thermographic changes related to positive emotional states (e.g. in chickens – Moe et al., 3561 2012; dogs – Travain et al., 2016). A study in dogs investigated the physiological response 3562 profiles to feeding palatable food treats (a putative positive event) using both HRV and IRT. 3563 Dogs' eye temperature and mean HR increased while receiving treats, but no clear changes 3564 in HRV emerged (Travain et al., 2016).

3565 Overall IRT-detected surface temperature changes are considered a valuable indicator of 3566 arousal (Travain & Valsecchi, 2021). Recently, a novel theory proposed that asymmetric 3567 temperature profiles in bilateral facial areas (e.g. eyes, ears, nostrils) could be linked to 3568 valence (Ramirez Montes De Oca, 2021; Telkanranta et al., 2019) in line with the emotional 3569 valence lateralisation hypothesis. According to this hypothesis, negative emotions are 3570 associated with right hemispheric dominance, while positive emotions are linked to greater 3571 activity in the left hemisphere (Davidson, 1995; Davidson et al., 1990; Leliveld et al., 2013). 3572 There are different views on how asymmetric hemispheric activity could affect surface 3573 temperatures of bilateral facial regions (illustrated in Figure 6.1). Firstly, an increase in ear or 3574 eye temperatures might result from increases in blood flow in areas with neural connections 3575 to enhance sensory efficiency (Ramirez Montes De Oca, 2021). In animals with bilaterally 3576 placed eyes (e.g. horses, cows, chickens) neural connections from the eyes predominantly 3577 project to the contralateral side of the brain (*i.e.* >80% of optic nerves crossover to the 3578 contralateral side of the brain (Herron et al., 1978)). Hence, an increase in right hemispheric

3579 activity could cause an increase in left eye temperature and vice versa (i.e. left hemispheric 3580 activation results in increased right eye temperature; Ramirez Montes De Oca, 2021). 3581 Alternatively, activation of a hemisphere could cause warming of adjacent facial areas due to 3582 heat conduction. This would result in temperature increases in the ipsilateral eyes or ears (Ramirez Montes De Oca, 2021; Telkanranta et al., 2019). Moreover, temperatures of areas 3583 3584 around the eyes or ears might increase due to increased muscular activity and/or blood flow, 3585 inducing regional (ipsilateral) temperature changes (Ioannou et al., 2015). Chickens exposed 3586 to a familiar room have higher right vs. left nostril temperatures, and opposite temperature 3587 asymmetries were observed when chickens were exposed to an unfamiliar room (Telkanranta 3588 et al., 2019). Higher right vs. left eye temperature was measured in cows whilst feeding when 3589 animals received an adverse handling experience before feeding. In contrast, higher left vs. 3590 right eye temperature was recorded in cows that did not receive a negative handling 3591 experience before feeding (Telkanranta et al., 2019). These findings suggest that temperature 3592 asymmetries measured in the eyes might result from heat convection of the hemispheres 3593 with higher temperature in areas close to the dominating brain side (Figure 6.1 B). However, 3594 disbudded dairy calves (*i.e.* animals with previous experience of a painful procedure) showed 3595 higher temperatures in the left relative to the right eye (measured in the medial canthus area) 3596 whilst observing a conspecific being disbudded (Ramirez Montes De Oca, 2021). The eye 3597 temperature profiles observed here seem to support the theory that increased right 3598 hemispheric activity during negative emotion processing induces higher temperatures in 3599 associated sensory organs of the contralateral side of the body (Figure 6.1 A). In line with this, 3600 calves approaching an observer (presumably perceiving a person as more positive than calves 3601 moving away or not approaching) also showed higher right than left eye temperatures 3602 (Ramirez Montes De Oca, 2021), further supporting the idea that lateralised temperature 3603 profiles could be related to emotional valence.



3604

Figure 6.1 Illustration of the three main theories of how lateralised hemispheric activity during 3605 3606 negative emotion processing might induce asymmetric eye temperatures. Negative emotions 3607 induced higher activity (white arrow) in the right hemisphere (RH) than in the left hemisphere 3608 (LH). This might result in warmer (red thermometer) left than right eye temperatures 3609 according to the sensory lateralisation theory (A). Alternatively, (B) heat convection from the 3610 RH might induce warmer right then left eye temperatures, or (C) the muscles and vascular 3611 systems around the right eye are more activated to increase sensory processing on that side, due to increased RH activation during negative emotion processing 3612

3613

3614 I investigated changes in cardiac activity and thermal profiles of facial areas (eyes, ears, hair 3615 whorl - centre of forehead where a small area of bare skin is visible due to parting in hair) as 3616 novel indicators of arousal (Ramirez Montes De Oca, 2021) to measure emotional state in 3617 horses. Receiving a palatable food reward is considered to induce positive emotional states 3618 in animals, including in horses (Fortin et al., 2018; Hintze et al., 2016; Sankey et al., 2010). I applied two different reward schedules to induce i) a putatively more positive emotional state 3619 3620 compared to rest (baseline) by delivering an immediate food reward (IR) and ii) a less positive 3621 emotional state by withholding food reward (WR), in each case following horses' performance of a conditioned response behaviour (contact target with their muzzle as described in Chapter 3622 5). I hypothesised that the physiological profiles assessed as mean heart rate and maximum 3623 3624 temperatures of the eyes (EyeMax) would significantly change from baseline as a reflection 3625 of increased arousal. As shown in Table 6.1 (see method section), I hypothesised that indices (potentially) linked to emotional valence (i.e. HRV, L-R eye difference in maximum 3626 3627 temperature) would significantly differ during IR and WR as a result of valence-induced shifts 3628 in these physiological markers.

3629 Apart from the emotional state, physiological processes involving the ANS can be influenced 3630 by several intrinsic factors, e.g. age can affect thermal radiation, and older horses have lower 3631 eye temperatures than younger horses (Negro et al., 2018). Furthermore, sex can influence 3632 IRT results (Knoch et al., 2022). I also considered experience with the target object as an 3633 important factor potentially modulating emotional response during testing. Horses needing longer to learn the conditioned response during the training originally conducted for the 3634 3635 object discrimination test presented in Chapter 5 had more exposure to the target object than 3636 horses needing fewer trials. Moreover, mental effort during discrimination learning and no 3637 reward delivery on incorrect trials could have resulted in individual differences in stress 3638 perception, possibly triggering longer-lasting emotional effects (Langbein et al., 2004). 3639 Therefore, I tested the effect of these factors, alongside welfare score (i.e. a proxy of 3640 emotional state), on horses' cardiac and thermal response during IR and WR.

3641 **6.2 Methods**

3642 6.2.1 Ethical statement

This study was approved by the Animal Welfare and Ethical Review Body of the University of Plymouth (ETHICS-42-2020), was below the threshold for regulation by the UK Animals (Scientific Procedures) Act 1986 (ASPA), and followed the Essential 10 ARRIVE guidelines 2.0 (Percie du Sert et al., 2020). The animals used in this study belonged to two UK riding schools that consented to the use of their horses for our experiment.

3648 6.2.2 Animals and housing

3649 Originally, 32 horses (12 females, age 13.12±5.48 years, various breeds) were tested. Six 3650 animals were excluded from the HRV data analysis, two due to measurement errors during 3651 the ECG recording (errors in 5min recordings resulted in >5% corrected beats (Ille et al., 3652 2014)), two were physically more active during breaks than the other animals (as they were 3653 kept in paddocks with more space to move than the other horses), and two horses due to cardiac arrhythmias (2nd AV blocks occurring >5 times during baseline recordings (Flethøj et 3654 3655 al., 2016; Mitchell & Schwarzwald, 2021). Therefore, the total number of animals included in 3656 the HRV data analysis was 26 (9 females, age 12.73±5.08 years). For the IRT data analysis, 3657 recordings from all 32 horses were used.

Horse husbandry is detailed in Chapter 3. Briefly, both facilities kept horses in single stalls, or tie-stalls, with full, or limited visual/physical contact with conspecifics. Animals had pasture access (in consistent groups) and hay. Horses from yard 2 additionally received an adjusted diet (brand Thunderbrook Equestrian). Animals from yard 1 did not receive any additional feed. All animals were used in riding lessons >10h per week. Water was freely accessible through automatic troughs in yard 1 and with water buckets in yard 2.

3664 6.2.3 Experimental design and testing procedure

Data were collected at yard 1 during July – August 2020 and at yard 2 between April-May 2021. One experimenter (E1), who handled the horses, and a second experimenter, who filmed the animal (E2), conducted the tests and the horses remained in their stalls during testing.

3669 All horses had previously learnt to touch one of two objects with their muzzle to receive a 3670 food reward as described in Chapter 5. Briefly, instrumental conditioning and positive 3671 reinforcement (pieces of carrot) were used to train horses' to touch one of two objects (two dog toys of similar size; pseudo-randomly counterbalanced across animals). During training, 3672 3673 animals received a maximum of 20 trial blocks (10 trials per block, no more than 8 trial blocks 3674 (80 trials)per day), or until learning criterion (LC, >80 % correct within three consecutive trial 3675 blocks) was reached. Hence, depending on learning ability (LA; measured as the number of 3676 trial blocks needed to reach LC), the amount of time/trials horses were previously exposed to 3677 the target object differed among horses (see Appendix D Table D1 for LA results). However, 3678 horses were presented with only the target object in this experiment and no discrimination 3679 was required.

3680 6.2.4 Reward schedule

3681 To control for time effects on physiological response, horses were pseudo-randomly allocated 3682 to one of two groups (each horse received a random number via the RAND function in Excel; 3683 horses with even numbers were assigned to group A; animal numbers were balanced across 3684 groups by moving the last assigned animals to the group with a lower number of animals). Horses from group A (N=16, for HRV data N=12) received immediate reward (IR) in phase 1, 3685 3686 comprising four consecutive trial blocks (T1-4) and 10 object trials per trial block (mean±SD 3687 duration 71.2±11s). Horses from group B (N=16, for HRV data N=14) received IR during the 3688 first two trial blocks (T1-2) of phase 1, but reward was withheld during T3-4. In phase 2, horses

3689 group A received IR in T5-6, but the reward was withheld in T7-8, whereas the other group

3690 received IR throughout all trial blocks. Trial blocks were interspersed with 2-min breaks during

3691 which the experimenters moved out of sight while the horses remained in their stalls. A 10-

3692 min break followed phase 1 before resuming phase 2. A summary of the reward schedule and

3693 details of data extraction is provided in Figure 6.2.



3694

Figure 6.2 Reward schedule and data collection timeline. During each test phase (T), horses were 3695 3696 prompted to contact the target in 10 trials, but the reward was only delivered upon contact during 3697 immediate reward (IR) phases and not during withheld reward (WR) phases. Black arrows indicate the 3698 number of IRT data points extracted per test phase. For HRV data extraction, two consecutive test phases of the same reward schedule, including the 2min-break between test phases, were combined 3699 3700 resulting in 5min blocks for the four test sessions (S1-4) containing two trial sessions per sessions. 3701 Horses were left alone in their stalls during a 10-min break following S2. Before testing, 5min baseline 3702 HRV data were recorded in B0 (S0) followed by 2min baseline (B1) recordings of IRT data during which 3703 the handler stood next to the horse.

3704 6.2.5 Testing procedure

3705 Horses were kept in their stalls for a minimum of 1 hour before the experiment, hence animals 3706 were not ridden, or on pasture, immediately before testing. E1 entered the stall, put a head collar around the horse's head, and tucked the forelock (part of the mane between the ears 3707 3708 that falls forward onto the forehead) underneath the top of the collar behind the ears. E1 3709 moved the horse to a position where it was standing with its head toward the open stall door 3710 so that E2 could position the IRT camera at 2 m distance directly in front of the horses' head. E1 stood either left or right (side counterbalanced across trial blocks) at the horses' shoulder 3711 3712 height. To start the test, E1 moved the target (hidden behind their back until this point) 3713 directly in front of the horse's muzzle at approx. 50 cm distance from the animal. During IR, 3714 horses received a piece of carrot within <5 s after touching the object; E1 moved the target 3715 behind her back whilst simultaneously retrieving a piece of carrot from a small bag attached

to her waist (using the opposite hand to the one holding the object). A total of 10 target trials
within each trial session (T1-8 in Figure 6.2) were conducted this way. During WR, the target
was presented to the horses in the same manner, held in position for a max. of 10 s, or until
the horses had touched it, before moving it behind their back for 3 s. Unlike in IR, no reward
was delivered and the next trial started by raising the object in position again.

3721 6.2.6 Measurements

3722 Horse behaviour was video-recorded (via two GoPro Hero 3 cameras positioned near the door 3723 facing into the stall and inside the stall facing the door) throughout the test and breaks. The 3724 number of trials horses made muzzle contact with the target was recorded live at the end of 3725 each trial block. During the WR phase, only one horse stopped responding to the target after not being rewarded on 8 trials during the 2nd reward phase (T4 in Figure 6.2). All other horses 3726 3727 performed object contacts during IR and WR phases regardless of whether they received a 3728 reward or not. Therefore, I did not consider the number of target contacts (putatively indicating motivation to perform that task) as a useful behavioural indicator and instead only 3729 3730 focused on physiological changes in my final analysis.

3731 The cardiac response was recorded with a III lead ECG device (Televet 100 software version 3732 5.1.2, Engel Engineering Services GmbH, Germany, 500Hz sampling rate) using a modified 3733 base-apex configuration (left arm electrode (LA), left leg electrode (LL), neutral electrode (N), 3734 right arm electrode (RA), shown in Figure 6.3). All ECG traces were visually checked in the 3735 Televet software using automated detection of irregular heartbeats (maximum deviation 20 3736 %, *i.e.* IBIs differing more than 20% from the previous IBIs were detected as outliers (Mitchell, 3737 2019), before exporting the IBI values into Kubios (Kubios HRV Standard 3.4.1, Kubios Oy, 3738 Finland) for extraction of HRV indices within 5min time bins (see Figure 6.2 for details).



3739

Figure 6.3 Illustration of the modified base-apex configuration used for ECG recording. The left arm electrode (LA, green) was placed over the apex beat area of the heart, on the thorax, caudal to the left elbow. The left leg electrode (LL, yellow) was positioned just over LA to create an additional lead. The neutral electrode (N, black) was placed behind the left shoulder blade, near the withers. The right arm electrode (RA, red) was placed opposite electrode, hence, not visible in this image.

3745 It is important to note that HRV indices describe the degree of autonomic modulation of 3746 cardiac activity rather than the absolute ANS tone (Taskforce, 1996). There are two main 3747 mathematical approaches to estimating HRV indices. The assessment of time-domain 3748 parameters (*i.e.* data assessed as a function of time) is the most straightforward method. 3749 These indices reflect the variability of cardiac activity (Wang & Huang, 2012). The second 3750 approach is using power spectral indices, which typically contain low-, and high-frequency 3751 values, as well as their ratio, which are derived from the spectral signal power, detected 3752 within certain frequency bands that are species-specific (Kuwahara et al., 1996; von Borell et 3753 al., 2007). More details on indices derived from time- and frequency-domain analysis are 3754 given in Table 6.1.

Table 6.1 Overview of time-domain and frequency-domain HRV indices and IRT measures and their
 predicted changes during IR and WR. (HR = heart rate; IBIs = inter-beat intervals)

Emotion index	Parameter	Description	Measure of	Prediction IR	Prediction WR							
	HRV indices											
Arousal	Mean HR	Mean of instantaneous HR values derived within a specific time period (5min)	PNS and SNS on HR; emotional arousal	1	ſ							

Emotion index	Parameter	Description	Measure of	Prediction IR	Prediction WR
Valence	SDNN	Standard deviation of all IBIs (within 5min period)	Estimate of overall HRV, reflecting PNS and SNS influence on HR	\downarrow	个
	RMSSD	Square root of the mean squared differences of successive IBIs	PNS (vagal tone) stimulation on HR; highly correlated with HF	\uparrow	\downarrow
	HF	Power in the high frequency range (0.07-0.6 Hz)†	Degree of PNS modulation (vagal tone) on cardiac activity; highly correlated with RMSSD	1	\checkmark
	LF/HF ratio	Ratio of low and high frequency spectral components	Sympathovagal balance modulation on HR	\checkmark	ſ
			IRT indices		
Arousal	EyeMax EarMax WhorlMax			\rightarrow -	\rightarrow \rightarrow -
Valence	L-R Eye Max			v NP	₩ NP

3757

3758 NP no directional prediction

3759

Thermal videos were taken with a FLIR T660[™] camera (FLIR SYSTEMS, Inc. USA; resolution 3760 3761 640x480 pixels, 60 FPS, sensitivity 0.02 °C, and accuracy ± 1%). Skin emissivity was fixed at 3762 0.98 (Tattersall, 2016; Westermann et al., 2013). Environmental factors (temperature and humidity) were taken (ORIA wireless thermo- and hygrometer) and updated on the camera 3763 3764 before the baseline recording and at the beginning of each session. Horses' heads were filmed 3765 from a frontal angle at approx. 90° to the horses' forehead. To ensure the appropriate camera 3766 angle, and both eyes being visible on the recordings, E2, holding the camera, slightly adjusted 3767 their position depending on the horses' movement. E1 always presented the target centrally 3768 between the horses' heads and the camera at approx. 30-40 cm distance to the horses' 3769 muzzle.

3770 From each video, three frames showing the horses' head with a sidewise variation of less than 3771 30° between the camera and nostrils (*i.e.* showing both eyes at an equal distance to the 3772 camera) were extracted by E2 using the FLIR Tools v.6.4.18039.1003 software (FlirSystems Inc, Oregon, USA). Thermal data were only extracted from frames where the maximum 3773

temperatures, indicated by the red triangles in Figure 6.4, were automatically detected within the medial canthus (inner corner of the eye, see Figure 6.4), and the hair whorl (measured as the warmest point on the forehead area) and the ear base (the area where the pinna (outer ear) merges into the skull) were not covered by the forelock (tucked back before the start of testing but occasionally hair fell back onto forehead during testing) to ensure the same areas were measured across all thermal images.



3780

Figure 6.4 Example of a thermal image used for extraction of maximum temperatures (marked by red
triangles) taken in the eyes (left (EI1) and right (EI2)), ear base (left (Li1 and right (Li2)) and hair whorl
(EI3).

3784 6.2.7 Statistical analysis

Data were analysed in R (version 4.0.3, R Core team, 2021). Each cardiac measure was analysed separately using a linear mixed model (Imer function from Ime4 package (Bates et al., 2015)). Test phase (baseline/IR/WR) and group (A/B), individual factors (*i.e.* age, sex, yard, target object, learning ability (LA)) were included in the models as fixed factors and welfare score as a covariate (collinearity of the factors was checked with vif function from car package, Fox & Weisberg, 2019). Horse ID was set as a random factor to control for repeated

3791 measurements. HF values and LF/HF ratio were log-transformed for a better fit of model3792 assumptions.

3793 For IRT measures, data of each region of interest (ROI, i.e. maximum temperature of each eye 3794 (EyeMax), ear base (EarMax), and the hair whorl (WhorlMax)) were analysed in the same 3795 manner as the cardiac data using linear mixed effect models. In addition to the factors 3796 mentioned above, ambient temperature and frame time (*i.e.* the time (s) the image was taken 3797 relative to the start of the test session in order to control for time effects in physiological 3798 responses), ROI side (left/right) and handler side (same/different (relative to ROI)) were 3799 included in the models. When controlling for yard and ambient temperature in the models 3800 using the full data set, problems with model convergence occurred, likely because the 3801 ambient temperature differed significantly between the two facilities with significantly 3802 warmer ambient temperatures measured in yard 1 (19.23°C ± 1.78, Q1-Q3: 14-25°C) than in 3803 yard 2 (mean 11.55°C±1.48, 8-14°C, two-sample t-test, p<0.0001). The surrounding 3804 temperature can have a significant influence on the heat exchange between an organism and 3805 its environment (Arduini et al., 2017; Collier et al., 2006). Therefore, I divided the data set by 3806 yard and carried out separate analyses for each facility. To assess asymmetries in eye 3807 temperatures (L-R EyeMax), the maximum values of the right ROI were subtracted from the 3808 maximum values recorded for the left ROI within each image.

3809 For all models, Type III tests (car package) with contrasts and the drop1 function (Ime4 3810 package) were used to remove non-significant terms from the models. Since test phase was 3811 the main factor of interest, it was included in all models even if not identified as a significant 3812 term in the full model. Model fit of the final model was compared against the null model (only 3813 containing ID as a random factor) with likelihood ratio tests (anova function in car package). 3814 The anova function with type III sum of squares was used to generate p-values for each fixed 3815 factor within the models. Tukey post hoc comparisons were used to test significant 3816 differences in factors between test phases. The significance level was set at P≤0.05 and p-3817 values between 0.05-0.1 were reported as trends. Model residuals were checked with the 3818 DHARMa package (Hartig & Lohse, 2022). All results are graphically presented as estimated 3819 marginal means (EMMs) ± SE unless stated otherwise.

3820 **6.3 Results**

- 3821 6.3.1 Changes in arousal markers during IR and WR
- 3822 As predicted, mean heart rate significantly increased from baseline during both IR and WR
- 3823 phases (Figure 6.5; IR: z=6.439, WR: z=4.098, both p<0.001) with no significant difference
- 3824 between reward phases (z=1.419, p=0.327).



3825

Figure 6.5 Mean HR recorded during test phases. Data presented as EMMs±SE, different lettersindicate significant differences at p<0.05.

3828 At Yard 1, EyeMax significantly decreased from baseline during IR (-0.3°C, z=-7.76, p<0.001)

and WR (-0.3°C, z=-6.52, p<0.001) and no significant difference between the reward phases (z=-0.74, p=0.73) was found as expected (Figure 6.6 A). In yard 2 (Figure 6.6 B), EyeMax dropped significantly from baseline during IR (-0.3°C, z=-4.88, p<0.0001) and but not WR (-0.2°C, z=-2.25, p=0.059) as eye temperatures remained significantly higher during WR than during IR (z= -3.27, p=0.003).

Unexpectedly, the handler side predicted EyeMax ($X_{1}^{2}=7.96$, p=0.004) in yard 1, with higher EyeMax when the eye side was the same as the side the handler was standing next to the horse (*e.g.* right eye side and right handler side) than when eye side and handler side were different, and the interaction between handler tended to be more pronounced when measured on the right side ($X_{1}^{2}=2.75$, p=0.09). The same significant relationship between eye side and handler side was observed in yard 2 ($X_{1}^{2}=7.55$, p=0.005), but without a significant interaction between eye and handler side ($X_{1}^{2}=0.24$, p=0.61).





Figure 6.6 Maximum temperatures of the medial canthus of the eye (EyeMax) recorded during test
phases at yard 1 (A) and yard 2 (B). Data presented as EMMs±SE, different letters indicate significant
differences at p<0.05.

3846



3847

Figure 6.7 Maximum ear base temperatures (EarMax) recorded during test phases at yard 1 (A) and yard 2 (B). Data presented as EMMs±SE, different letters indicate significant differences at p<0.05.

3850

Considering EarMax, in yard 1, temperatures dropped significantly during WR compared to baseline (z=-3.37, p=0.02) and IR (z=6.63, p<0.001), but there was no significant difference

between baseline and IR (z=0.8, p=0.69; Figure 6.7 A). The left ear tended to be warmer than the right (z=-1.76, p=0.07), but unlike with the eyes, handler side did not affect ear temperatures (X^{2}_{1} =0.25, p=0.62). No significant effects of reward (X^{2}_{1} =1.96, p=0.37) or handler side (X^{2}_{1} =0.001, p=0.96) on ear temperatures were found in yard 2 and ear temperatures did not significantly differ between test phases (Figure 6.7 B).

In yard 1, WhorlMax temperatures differed significantly decreased from baseline during reward phases (baseline – IR: z=-5.05, p<0.001, baseline – WR: z=-6.45, p<0.001) and was significantly lower during IR than WR (z= 3.07, p=0.005, Figure 6.8 A). In yard 2, test phase predicted WhorlMax temperature (X^2_2 =48.48, p<0.0001). However, the decrease in temperature was only significant between baseline and IR (z= -6.95, p<0.0001), and baseline and WR (z= -5.19, p<0.0001), but not between treatment phases (z= -1.58, p=0.246, see Figure 6.8 B).



3865

Figure 6.8 Maximum hair whorl temperatures (WhorlMax) recorded during test phases at yard 1 (A)
 and yard 2 (B). Data presented as EMMs±SE, different letters indicate significant differences at p<0.05.

3868

3869 6.3.2 Changes in valence markers during IR and WR

Against my predictions, I found no significant changes in RMSSD, predominantly reflecting PNS modulation on the heart, from baseline to IR (z=1.211, p=0.443), WR (z= -0.296, p=0.952), or between IR and WR (z=1.573, p=0.254, Figure 6.9 A). Likewise, there were no significant changes in HF (another marker of PNS (vagal) modulation) values from baseline during both

3874 test phases (X²₂=4.09, p=0.129, Figure 6.9 B). SDNN, an index of the combined activity of PNS 3875 and SNS stimulation on cardiac activity, significantly increased during IR (z=3.395, p=0.002) 3876 and there was no significant difference between baseline and WR (z=2.024, p=0.104), or IR 3877 and WR (z=0.916, p=0.626, Figure 6.9 C). I expected that LF/HF ratio, reflecting sympathovagal 3878 balance, would be significantly lower during IR than WR as an indication of increased PNS activity during the presumably more positive test phase, and significantly higher LF/HF values 3879 during WR due to increased SNS activity. Instead, LF/HF significantly increased from baselines 3880 3881 during IR (z= 2.559, p=0.027) and WR (z= 3.395, p=0.001) with no significant difference 3882 between reward phases (z=-1.599, p=0.242, Figure 6.9 D).



3883

Figure 6.9 Change of HRV indices linked to emotional valence measured as (A) RMSSD, (B) HF, (C)
 SDNN, and (D) LF/HF ratio (log-transformed). Data presented as EMMs±SE, different letters indicate
 significant differences at p<0.05.

Regarding thermal changes, asymmetries in EyeMax measured as L-R maximum eye temperature differences were not significantly different between baseline and reward phases in yard 1 (Figure 6.10 A). However, in yard 2, L-R EyeMax difference was significantly predicted by test phase (X²₂= 8.46, p=0.014) with significant changes in asymmetries between IR and WR (z=-2.73, p=0.016). Averaged L-R EyeMax were significantly lower during IR than WR,

3892 meaning that horses' right eye were warmer than their left eye when immediately rewarded

3893 than when reward was withheld. However, overall the eye temperature difference recorded

3894 in IR and WR were not significantly differed from baseline (Figure 6.10 B).



3895

Figure 6.10 Differences in asymmetric EyeMax assessed during baseline and reward phases in (A) yard1 and (B) yard 2.

3898

3899 6.3.3 Other factors modulating cardiac and thermal responses

3900 Age and sex had no significant effect on HRV indices (all p<0.05). However, higher number of 3901 trials blocks to reach learning criterion (*i.e.* lower learning ability (LA)) significantly predicted 3902 lower SDNN, RMSDD, and HF values compared to horses with fewer trials blocks. Moreover, 3903 lower welfare scores (i.e. lower likelihood to experience positive emotional states) significantly predicted lower SDNN (X²₁= 15.671, p<0.0001), RMSSD (X²₁= 15.258, p<0.0001), 3904 3905 and HF values (X_{1}^{2} = 21.22, p<0.0001). Figure 6.11 shows the relationship between SDNN, 3906 RMSSD and HF values and welfare score as predicted by each model. There was no effect of 3907 yard on any of the explored HRV indices (all p<0.05).

3910 Model estimates of all final models are provided at the end of this chapter in Tables 6.2-6.4.

³⁹⁰⁸ None of the factors I explored (age, group, welfare score, LA) had a significant effect on the3909 IRT markers.





Figure 6.11 Relationship between A) SDNN, B) RMSSD and C) HF values and welfare score as predictedby the linear mixed model.

3914

3911

3915 6.4 Discussion

3916 I investigated changes in HRV and IRT indices as markers of emotional state in horses. I 3917 expected that delivering a food reward immediately (IR) following horses' performance of a 3918 conditioned behaviour (contact target with the muzzle) would result in a shift towards a more 3919 positive emotional state, while withholding the reward (WR) would elicit a potentially more 3920 negative emotional state, and that both phases would increase arousal relative to baseline. 3921 Correspondingly, I predicted that physiological indices associated with arousal (i.e. mean HR, 3922 EyeMax, EarMax, WhorlMax) would significantly differ during the reward phases compared 3923 to baseline. Parameters reflecting valence (*i.e.* RMSSD, HF, SDNN, LF/HF ratio, L-R EyeMax) 3924 were predicted to be significantly different during IR and WR with significant changes from 3925 baseline.

3926 As expected, mean HR significantly increased from baseline during IR and WR without a 3927 significant difference between the two phases, suggesting that immediate and withheld

3928 reward induced similar heightened levels of physiological activity. The intensity of affective 3929 stimulation (*i.e.* arousal) can increase when animals experience both positive and negative 3930 situations (Paul et al., 2005; Mendl et al., 2010; Wascher, 2021). Frustration-induced high 3931 arousal reflected in increased HR was reported in horses denied access to a visible food 3932 resource causing acute stress (Bachmann et al., 2003). Learnt food arrival-related cues (e.g. 3933 animal caretaker first entering stable in the morning, the sound of concentrate pellets being 3934 collected from silo (Peters et al., 2012)), or reuniting a socially isolated horse with its group 3935 members (Briefer et al., 2015), also significantly elevated horses' HR confirming that positive 3936 situations heighten arousal measurable in changes of cardiac activity. Since HR is modulated 3937 by PNS and SNS activity, the change in mean HR could theoretically be the result of PNS 3938 withdrawal or increased SNS activity (Porges, 2007; von Borell et al., 2007). The finding that 3939 EyeMax significantly decreased during IR and WR, likely due to sympathetically mediated 3940 vasoconstriction in peripheral capillary beds in the medial canthus as reported in cattle 3941 experiencing stress (Stewart al., 2008a, b, 2010), might suggest that the change in mean HR 3942 was triggered by elevated SNS activity. In contrast to my findings, several equine IRT studies 3943 reported increased eye temperature in response to stress (e.g. novel object exposure - Dai et 3944 al., 2015; jumping competitions - Bartolomé et al., 2013; or trotting races - Negro et al., 3945 2018). Stewart et al. (2008a, b, 2010) noted the temporal effect in eye temperature changes. 3946 Temperatures in the medial canthus initially dropped in response to stress (-0.27 °C, 2-5min 3947 after treatment) before significantly increasing and overshooting baseline values (from 5min 3948 post-treatment; Stewart et al., 2008a). It is possible that the thermal changes in EyeMax found 3949 in this study do not correspond with previous observations in horses, because thermal 3950 responses were measured simultaneously with the animals exposure to IR and WR, hence the 3951 changes recorded here might indicate different stages of physiological responses to an 3952 emotional situation compared to other equine studies that recorded thermal response >5min 3953 post-treatment (Dai et al., 2013; Bartolomé et al., 2013; Negro et al., 2018). Moreover, 3954 physical activity level can influence IRT results as heat dissipating from skeletal muscles can 3955 cause peripheral temperature changes (Travain & Valsecchi, 2021). Unlike in the previous 3956 equine studies where thermal data was derived from moving horses (e.g. riding in Bartolomé 3957 et al., 2013; racing in Negro et al., 2018), the horses in this study moved very little (i.e. 3958 stretching neck to contact target with the muzzle).

3959 I investigated temperature changes in the ears (EarMax) and the hair whorl (WhorlMax) as 3960 additional indices of arousal-related changes in peripheral blood flow expecting that the 3961 direction of change would mirror eye temperature changes due to similar SNS-mediated 3962 changes in vasoconstriction in these areas. No clear patterns in EarMax recorded during IR 3963 and WR emerged. EarMax was derived from maximum temperatures recorded at the ear base 3964 (i.e. area between the skull and the pinna (outer part of the ear)), which was more difficult to 3965 identify consistently across horses/frames than the other ROIs. Also, this area is covered by 3966 hair and coat colour, or coat density, which can affect thermal readings (McCafferty et al., 3967 2015; Okada et al., 2013; Tattersall, 2016). Therefore, this ROI may not be an appropriate 3968 body area to investigate arousal-related temperature changes. WhorlMax, however, 3969 significantly decreased from baseline during IR and WR suggesting that this area might be a 3970 suitable indicator of arousal, for instance in situations where it is not possible to derive 3971 thermal data from horses' eyes (e.g. when eyes are obscured, e.g. by stable bars or blinkers). 3972 Ramirez Montes De Oca (2021) proposed the investigation of hair whorl temperature as an 3973 arousal index, noting a significant temperature increase in calves' hair whorl temperature 3974 following separation from their dam, which continued to increase during 60 min post-3975 separation (measured in 5-min bouts every 10min). The increase likely occurred due to 3976 modulatory physiological mechanisms (possibly heat dissipation following SIH) and increasing 3977 PNS activity following the initial higher activity of the SNS. Unlike in calves, horses' WhorlMax 3978 significantly decreased, which could be explained by the differences in recording times, as 3979 discussed for EyeMax.

3980 RMSSD and HF, which are HRV indices primarily reflecting parasympathetic modulation of 3981 cardiac activity (von Borell et al., 2007), did not significantly change during IR or WR implying 3982 the reward phases did not alter PNS activity. PNS-dominant sympathovagal balance prevails 3983 in resting horses (Kuwahara et al., 1996), but I would have expected a further increase in PNS 3984 activity during IR as an indication of a shift towards a more positive emotional state, unless 3985 feeding horses' with carrot pieces was not as rewarding as anticipated. In pigs trained to 3986 associate an acoustic stimulus with food, RMSSD significantly increases when hearing the 3987 conditioned cue (Zebunke et al., 2011). However, Peters et al. (2012) also found no significant 3988 changes in RMSSD and HF while horses anticipated the arrival of concentrate (food) during 3989 their normal feeding time, despite the significant increase in mean HR and performance of

3990 anticipatory behaviours (e.g. increased movement and investigation behaviours such as 3991 licking and sniffing). Horses shown a familiar food bucket, but not receiving food had 3992 significantly lower HF values during and following the presentation of the food stimulus 3993 (Bachmann et al., 2013) demonstrating that withholding food <u>can</u> induce stress and lower 3994 PNS activity, which is what I expected to observe during WR. Maybe the horses would have 3995 responded differently if a more palatable treat was used as a reward. Still, this does not 3996 explain why all horses, except one, continued to work for receiving carrots by contacting the 3997 target with the muzzle during WR. Since all horses had previously performed several trial 3998 blocks in a similar task (object recognition test, ORT, Chapter 5), it could be that the animals 3999 were resistant to extinction, hence kept responding to the target despite the lack of reward 4000 (Williams et al., 2004). All horses had previously experienced unrewarded trials (although no 4001 more than three consecutive trials) if they contacted the incorrect object in the discrimination 4002 test. Maybe the number of unrewarded test trials was too low to induce significant shifts in 4003 horses' emotional states. In Williams et al. (2004), horses performed on average 137 trials of 4004 an operant task (nose touch traffic cone) without receiving food as reinforcement. On the 4005 other hand, it could be that my test horses formed a habitual response to the target object 4006 following the performance of a high number of training and test trials during ORT. Hence, the 4007 animals might have performed the conditioned behaviour in an automated manner (i.e. 4008 having formed an association between an action and its preceding stimulus, but not with its 4009 consequences (Lansade et al., 2017)). However, if this was the case, I would not have expected 4010 to find any changes in SDNN and LF/HF ratio during the reward phases, but these indices 4011 significantly increased suggesting that sympathovagal balance shifted towards higher SNS 4012 activity. Maybe the horses perceived having to perform the conditioned response as stressful. 4013 To better understand this, it would have been beneficial to collect HRV data while horses were 4014 fed carrots without needing to contact the object. Moreover, the presence of the familiar 4015 experimenter might have induced changes in arousal. Horses establish long-lasting positive 4016 associations with humans during positive reinforcement training (Sankey et al., 2010), hence 4017 the sight of the experimenter alone might prompted anticipation of food causing an increase 4018 in arousal. It has been argued that physiological changes associated with 'stress response' 4019 actually reflect adaptations in metabolic and cardiovascular requirements associated with behavioural activity instead of reflecting stimulus valence (Knoch et al., 2022), which might 4020

4021 explain further explain why I found no difference in the physiological profiles measured during4022 IR and WR.

4023 Nonetheless, I analysed changes in L-R EyeMax as a potential indicator of valence as 4024 previously suggested by Telkanranta et al. (2019) and Ramirez Montes De Oca (2021). In yard 4025 2 horses, WR induced higher left than right eye temperatures. This corresponds with Ramirez 4026 Montes De Oca's (2021) findings that a negative situation (*i.e.* watching a conspecific being 4027 disbudded after being disbudded) causes higher left than right eye temperatures in calves. 4028 Hence my results might support the theory that right hemispheric dominance during negative 4029 emotion processing causes a greater activity of associated sensory organs (*i.e.* left eye). 4030 Interestingly, the eye on the side the handler was standing was warmer than the opposite 4031 side which might further suggest enhanced sensory activity of the eye directly viewed the 4032 handler (i.e. the source of reward). Furthermore, blood flow in the area around the eye, 4033 including the medial canthus, might have increased to enhance muscular activity for rapid eye 4034 movement and heightened sensory processing (Pavlidis et al., 2001 cited in Ioannou et al., 4035 2015), which ultimately elevated the emitted thermal radiation in this area. In humans, startle 4036 induced warming of periorbital areas (*i.e.* area around the eyes), while simultaneously a 4037 reduction in cheek temperatures was recorded (Pavlidis et al., 2001). Yet, average L-R eye 4038 temperatures did not significantly differ from baseline and no significant difference in left to 4039 right eye temperatures were observed in yard 1 horses. Surface skin temperatures are 4040 considered 'interface' temperatures determined by core temperature and surrounding 4041 environmental factors (e.g. ambient temperature, humidity, windspeed (Okada et al., 2013)). 4042 That is, eye temperature is positively correlated with ambient temperature (Okada et al., 4043 2013). At yard 1, the maximum ambient temperature (24°C) was close to horses' thermal 4044 neutral zone (approx. 25 °C (Morgan, 1998)) meaning that ceiling effects might have masked 4045 eye temperature asymmetries. Overall, the suitability of asymmetric temperature profiles as valence indicators, and how external factors (e.g. climate, presence of humans) might affect 4046 4047 these, needs further investigation.

4048 It should be noted that heat radiation of body surfaces areas of endotherm animals can be 4049 caused by various homeostatic processes (*e.g.* thermoregulation, circadian changes in 4050 physiology, metabolic rate, behavioural activities), besides emotional regulation (Ioannou et 4051 al., 2014; Tattersall, 2016). Hence, in order to detect significant changes in thermal profiles

4052 associated with the physiological changes of interest, it is necessary for these to have a 4053 greater magnitude than 'other' physiological processes occurring at the same time, but 4054 controlling for this might be difficult and might require repeated recording of thermographic 4055 measurements over a longer period of time (Knoch et al., 2022).

4056 I found no significant effects of age or sex on any of the explored physiological markers. 4057 Though, a higher welfare score (supposedly reflecting a more positive emotional state) 4058 significantly predicted greater SDNN, RMSSD and HF values. This is in line with previous reports in humans, and other animals, that more positive emotional states are associated 4059 4060 with greater HRV (von Borrell et al., 2007). Similarly, horses receiving fewer trial blocks to 4061 reach the learning criterion (supposedly having a higher learning ability) had significantly 4062 higher SDNN, RMSDD, and HF values compared to horses needing more trial blocks, 4063 suggesting that the former were more relaxed during testing than the latter. Langbein et al. 4064 (2004) proposed that learning and restricted reward delivery might act as stressors potentially 4065 causing long-term effects on HRV. They observed that visual discrimination training can have 4066 reducing effects on RMSSD in goats, probably caused by the mental effort, resulting in higher 4067 stress levels ultimately leading to PNS withdrawal (Langbein et al., 2004).

4068 None of the explored individual factors (*i.e.* age, welfare score, learning ability) significantly 4069 affected the thermal measurements. This implies that HRV changes might provide more 4070 sensitive indices for detecting individual variations in physiological response profiles 4071 measured during emotional situations. Due to time constraints, thermal data were extracted 4072 from three frames selected within the test period when the horses' head was in the required 4073 position. However, emotion-induced surface temperature profiles might be characterised by 4074 rapid dynamic changes (Stewart et al., 2008a,b, 2010). Hence, it might be more insightful to 4075 continuously sample thermal changes. This would also allow the magnitude of temperature 4076 changes to be measured, which is linked to the intensity of perceived stress level 4077 (Wongsaengchan et al., 2022). Therefore, I plan to investigate this further in a future 4078 collaboration with C. Wongseangchan.

4079 Overall, my findings show that IR and WR altered horses' level of arousal, but it is less clear 4080 whether testing induced the anticipated changes in valence, or the shift in emotional states 4081 was not reflected in the parameter analysed here. To manipulate affective states more 4082 effectively, motivation tests could help identify what resources horses value, which might

vary between individuals. Testing animals' willingness to work for a specific resource (*e.g.*social, enrichment, food) enables us to understand what resources animals want and how
much they want them (Kirkden & Fraser, 2006), the latter aspect being particularly relevant
in addressing the limitations of this study.

4087 **6.5 Conclusion**

In this study, horses were exposed to varying reward schedules following a conditioned behaviour to assess the physiological profiles of positive (immediate reward) and less positive (withheld reward) emotional states. Physiological parameters associated with arousal changed significantly during both reward phases compared to baseline, but not between IR and WR. Against my predictions, IR and WR did not result in any significant differences in valence indices, though this may be because of experimental limitations or failure to induce the desired shifts in horses' emotional states.

Table 6.2 Model estimates from LMER analysis of HRV indices. WR withheld reward, IR immeditate reward.

	Mean HR				SDNN			RMSSD			HF		LF/HF ratio		
Predictors	Estimates	CI	р	Estimates	CI	р	Estimates	CI	р	Estimates	CI	р	Estimates	CI	р
(Intercept)	41.49	39.58 – 43.40	<0.001	-19.14	-73.42 –	0.490	-33.89	-	0.390	-4094.98	-	0.116	-0.63	-0.91 —	<0.001
					35.15			111.20 -			9206.10 -			-0.36	
								43.42			1016.13				
Reward	2.08	1.09 - 3.07	<0.001	6.37	0.20 – 12.54	0.043	-1.28	-9.73 –	0.767	-408.57	-	0.257	0.52	0.22 – 0.83	0.001
[WR]								7.18			1114.64 –				
											297.50				
Reward	2.67	1.86 - 3.48	<0.001	8.72	3.69 – 13.76	0.001	4.27	-2.64 –	0.226	238.83	-337.67 –	0.417	0.32	0.08 – 0.57	0.010
[IR]								11.17			815.33				
Learning ability				-2.29	-3.43 –	<0.001	-3.25	-4.88 –	<0.001	-195.04	-302.28 –	<0.001			
					-1.15			-1.63			-87.80				
Welfare				7.51	3.79 –	<0.001	10.15	4.85 –	<0.001	620.30	270.75 –	0.001			
Score					11.22			15.44			969.86				
Random effects															
σ ²	3.35			128.74			242.00			1687086.9			0.31		
τ _{00 id}	21.40			208.30			426.50			1732369.3			0.19		
ICC	0.86			0.62			0.64			0.51			0.38		
N id	26			26			26			26			26		
Observat-ions	130			130			130			130			130		
Marginal R ² /	0.042 /			0.433 /			0.417 /			0.344 /			0.054 /		
Conditional R ²	0.870			0.784			0.789			0.676			0.415		

		EyeMax		EarMax				WhorlMax			L-R EyeMax difference		
Predictors	Estimates	CI	р	Estimates	CI	р	Estimates	CI	p	Estimates	CI	р	
(Intercept)	34.25	34.02 - 34.48	<0.001	30.09	29.49 – 30.69	<0.001	30.28	29.84 - 30.72	<0.001	0.02	-0.09 - 0.12	0.725	
Reward	-0.32	-0.420.22	<0.001	-0.58	-0.92 – -0.24	0.001	-0.49	-0.63 – -0.34	<0.001	0.01	-0.09 – 0.11	0.821	
[WR]													
Reward	-0.34	-0.43 – -0.26	<0.001	0.12	-0.18 – 0.43	0.421	-0.34	-0.480.21	<0.001	-0.06	-0.14 - 0.03	0.202	
[IR]													
Handler_side	0.10	0.03 - 0.17	0.005										
[same]													
ROI_side [right]	-0.03	-0.10 - 0.04	0.468	-0.16	-0.33 - 0.02	0.078							
Handler_side	0.08	-0.02 - 0.18	0.097										
[same] *													
ROI_side [right]													
Random Effects													
σ ²	0.14		1.73				0.33			0.14			
τ _{00 ID}	0.20		1.20				0.79			0.02			
ICC	0.59		0.41				0.71			0.12			
N ID	17		17				17			17			
Observations	894		894				894			894			
Marginal R ² /	0.042 / 0.605		0.030 /				0.014 /			0.006 /			
Conditional R ²			0.427				0.709			0.121			

Table 6.3 Model estimates from LMER analysis of IRT indices recorded at Yard 1. WR withheld reward, IR immeditate reward, ROI region of interest.
Chapter 6 | Changes in HRV and IRT in response to immediate and withheld reward

	EyeMax			EarMax			WhorlMax			L-R EyeMax difference		
Predictors	Estimates	CI	р	Estimates	CI	р	Estimates	CI	р	Estimates	CI	р
(Intercept)	32.72	32.38 - 33.05	<0.001	24.70	23.40 – 25.99	<0.001	26.91	25.76 – 28.05	<0.001	0.07	-0.12 – 0.26	0.446
Reward	-0.16	-0.31 – -0.02	0.024	-0.38	-1.01 – 0.24	0.231	-0.61	-0.84 – -0.38	<0.001	0.03	-0.10 - 0.16	0.671
[WR]												
Reward	-0.32	-0.44 – -0.19	<0.001	-0.37	-0.93 – 0.19	0.192	-0.73	-0.94 – -0.53	<0.001	-0.09	-0.20 – 0.03	0.138
[IR]												
Handler_side	0.10	0.03 - 0.18	0.006									
[same]												
ROI_side [right]												
Handler_side												
[same] *												
ROI_side [right]												
Random Effects												
σ ²	0.23			4.55			0.61			0.19		
τ _{00 ID}	0.30			4.36			4.00			0.08		
ICC	0.57			0.49			0.87			0.29		
N ID	12			12			12			12		
Observations	636			636			636			636		
Marginal R ² /	0.025 / 0.579			0.001/			0.010/			0.009 /		
Conditional R ²				0.490			0.868			0.295		

Table 6.4 Model estimates from LMER analysis of IRT indices recorded at Yard 2. WR withheld reward, IR immeditate reward, ROI region of interest.

4095

95 Chapter 7 | General discussion

4096 Summary

4097 The assessment of cognitive components associated with emotional experiences has been 4098 promoted as a possible method of gaining a better understanding of affective valence in 4099 animals. Measuring valence is crucial when judging whether welfare is good or bad, but it is 4100 more difficult to objectively assess compared to arousal. Changes in cognitive processing 4101 involving attention have been linked to negative affective states in humans, showing that 4102 affective valence can be inferred from attentional biases in response to negative stimuli. 4103 Previous studies suggested that emotional states also have modulating effects on how 4104 animals attended to external stimuli. Therefore, the main motivation in this thesis was to 4105 investigate and validate AB as a novel cognitive proxy of affective states in horses. Moreover, 4106 I aimed to assess physiological changes associated with short-term changes in emotional 4107 states as additional markers of affective states. In my general discussion, I will assess to what 4108 extent the thesis aims were achieved and my hypotheses were confirmed. I will also evaluate 4109 the implications of my findings in terms of providing recommendations for future research.

4110 **7.1** How should AB be assessed in animals?

4111 In chapter 2, I reviewed the existing literature on AB and humans and how AB tests have been 4112 translated into animal studies. I found discrepancies between study findings as animals might 4113 show attention to, or away, from a threat stimulus. The heterogeneity of existing AB 4114 observations might largely be due to differences in experimental designs, including the type 4115 and presentation protocols of test stimuli used, and consequently how AB is measured. Based 4116 on my literature review, I developed a visual AB test in Chapter 4, using competing (dual) 4117 presentation of biologically meaningful stimuli (conspecific faces) supposedly differing in 4118 emotional content without varying in any other perceptual aspects, which could influence 4119 horses' attention. This is the first study testing AB in horses. Contrary to my predictions, my 4120 findings indicated that horses in general showed no significant preferential attention 4121 allocation to the two differently valenced stimuli, as all face images were attended to for a 4122 similar length of time. Combined with my observations from the image recognition test 4123 (Chapter 5) suggesting that not all horses' are able to recognise on-screen images, my findings 4124 highlight the need to further examine what type of stimuli are suitable to test AB in horses. It

4125 is possible that biologically meaningful (innate) stimuli are processed differently to 4126 conditioned stimuli (e.g. objects associated with reward/punishment), which might be why 4127 the findings from Chapter 5 do not correspond with previous reports of image recognition in 4128 horses (e.g. Wathan et al., 2016). Horses' attention to the negative stimuli was not completely 4129 random, as horses with poorer welfare scores did not look at the negative faces as long as 4130 horses with better scores, opposite to my predictions, which might suggest that face images 4131 might be appropriate stimuli to some aspects of AB in horses after all. Nonetheless, the 4132 influence of other information conveyed by faces, such as familiarity or identity, and how 4133 these are best visually presented needs further assessment. Ultimately, my findings 4134 emphasise that testing AB might have future applications in assessing affective states in 4135 animals, but current test paradigms have some experimental limitations and therefore 4136 require further refinement.

4137 One possible step towards further enhancement in utilising cognitive proxies such as AB might 4138 be to integrate physiological measures when testing cognitive processes. My findings in 4139 Chapter 6 indicated that it might not be possible to differentiate between the direction of 4140 emotional valence using indices of HRV and IRT, further supporting the argument that valence 4141 is more difficult to assess via physiological parameters (Mendl & Mason, 1993; Paul et al., 4142 2005). However, the observed changes in HRV and IRT indices suggested that these 4143 parameters could be sensitive measures of arousal. Combining HRV and/or IRT indices with 4144 AB assessment might provide valuable insights into how animals perceive different test 4145 stimuli in terms of their emotional intensity. Moreover, a combined approach might also 4146 reveal why certain types of stimuli are attended to more than others, or explain individual 4147 differences in responses to test stimuli. Maybe some individuals react more strongly (i.e. are 4148 more aroused) to test stimuli compared to others. How different arousal levels affect AB 4149 responses in animals is still unknown, but in humans, emotional arousal conveyed by test 4150 stimuli modulates attention (Lundqvist et al., 2014), and arousal level of the observer also 4151 influences cognitive processes (e.g. attention and decision-making; Mendl & Paul, 2020). 4152 Hence, measuring the physiological parameters I used in Chapter 6 during cognitive testing 4153 could yield a better understanding of not only how animals respond to AB test stimuli, but 4154 also why.

4155 **7.2** Is it sensible to estimate affective state from welfare measurements?

4156 Investigating the relationship between horses' welfare state and cognitive measures (i.e. AB 4157 in Chapter 4; learning ability in Chapters 5 and 6) was a core part of my assessment of how 4158 emotional states influence cognitive processes. However, there is no direct method to assess 4159 animal welfare (Mendl & Mason, 1993). I adopted the view of welfare as a multifaceted 4160 subjective state determined by an animal's perception of its own circumstances (Broom, 4161 1986). Measuring external circumstances, it is theoretically possible to infer animals' 4162 subjective state by evaluating their current living conditions in light of existing scientific 4163 evidence of how these different factors relate to welfare (e.g. preference testing showed that 4164 horses prefer straw over shavings (Mills et al., 2000); and digesting straw might ameliorate 4165 behaviour problems associated with feeding low-fibre diets to horses (McGreevy et al., 4166 1995b)). Based on this, I calculated an individual welfare score representing individuals' 4167 potential to experience positive welfare. As I already discussed in Chapter 4, this score can 4168 only be considered an approximation for welfare state. Still, the question remains whether 4169 this theoretical welfare state is comparable with affective state. One particular difficulty I see 4170 in doing so is the fact that evaluating welfare reliably is hampered by the lack of knowledge 4171 of how different welfare factors (e.g. social or environmental resources) contribute to animal 4172 well-being. Some factors might be more important to welfare than others, but there is 4173 currently no dynamic model of welfare to assess this (Appleby, 2007).

4174 Others (Bracke et al., 1999a,b; Botreau et al., 2007a,b) have proposed evaluating animal 4175 welfare using a hierarchical aggregation process, which assigns different aspects of animals' 4176 environment different values, or rankings, based on animals needs and their relevance to 4177 welfare as predicted by experts' opinions. Still, a welfare measure aggregation approach also 4178 has constraints (Sandøe et al., 2019), including that the assigned value of measures in relation 4179 to animal welfare is likely to be dependent on peoples' attitude to what is important (Botreau 4180 et al., 2007b), hence the central problem, understanding what is truly important to an 4181 animal's welfare remains. Furthermore, individual animals might vary in their preference for 4182 environmental resources, but it can be challenging to evaluate how individual choices for 4183 particular resources are related to other physiological measures of welfare (Nicol et al., 2009). 4184 Here, I evaluated horses' environmental conditions based on scientific evidence that certain 4185 conditions are preferred to others, or some have greater health implications than others. In

case of social contact, it might be of value to further investigate how horses' opportunity for
social interactions might influence their attention when presented with the social cues in the
AB test.

4189 Certain welfare aspects might be more difficult to measure, like the presence of pain 4190 (Sneddon et al., 2014). I planned to assess horses' potential for experiencing pain using HGS 4191 and pressure algometry, but the tests were unreliable so could not be integrated into my final 4192 welfare assessment protocol. Manipulating horses' emotional state, e.g. by exposing animals 4193 to situations that have been shown to induce positive or negative emotions immediately 4194 before or during testing could have been a better approach in investigating how (short-term) 4195 emotions influence cognitive measures like AB. I had intended to conduct experiments of this 4196 kind, but my data collection plans were disrupted by COVID-19. However, as my findings from 4197 Chapter 6 illustrate, affect manipulations can be challenging. Contrary to my expectation, 4198 reward and withheld reward did not seem to induce the anticipated shifts in emotional states 4199 in horses. Although assessing the direction of emotional valence is crucial for measuring 4200 animal emotions, the importance of determining arousal levels should not be 4201 underestimated. Evaluating whether certain situations (negative or positive) cause relatively 4202 mild or intense emotional experiences can be decisive when considering how animals should 4203 be treated. In this case, the HRV and IRT indices I investigated in Chapter 6 could provide 4204 valuable insights into how animals perceive specific situations.

4205 Overall, an integrative approach that combines cognitive, physiological and behavioural 4206 measures is likely to result in the most accurate scientific evaluation of animal welfare and 4207 affective states. To this end, further validation of the cognitive and physiological indices 4208 addressed in this thesis could have substantial value.

4209 **7.3 Future research questions**

The findings of my thesis highlighted the need for further investigation of theoretical and experimental questions before AB might be used as a potential proxy of animal affect. In my view, one particularly important issue to address further is the investigation of whether AB is limited to negative emotional experiences and stimuli. So far, attentional responses to negative stimuli have predominantly been tested in animals, but there is little investigation whether animals exhibit AB in response to positive stimuli. I observed no AB when horses

4216 were presented with positive-neutral stimulus pairs but this might be because horses did not 4217 perceive these stimuli as positive and neutral as discussed in Chapter 4. Likewise, little 4218 research has addressed how positive emotional states modulate attention. I found that 4219 horses with better welfare scores were more likely to attend to the face images for longer 4220 compared to horses with poorer welfare scores. Maybe instead of narrowing attentional 4221 scope (i.e. focussing on particular information or "not seeing the forest for the trees"), 4222 attentional scope might be broadened during positive emotional states, as suggested in 4223 humans (Fredrickson, 2001). Indeed in dogs, attentional scope is altered by emotional state, 4224 but the influence of emotions on attention seems to be stimulus-dependent as shifts in 4225 attentional scope were observed only following affect manipulation with food and not social 4226 stimuli (Hamlaoui et al., 2022). Again, this highlights the need to further investigate the effect 4227 of different types of stimuli on attention. Similarly, this raises the question of whether AB is 4228 stimulus-specific or exhibited consistently across diverse stimuli. If the former were true, this 4229 might have implications for providing animals with suitable welfare conditions, as the 4230 absence/presence of these stimuli (e.g. social) might be particularly influential on animals' 4231 perception of their own environment. This knowledge might promote future applications of 4232 AB, for example, to modulate emotional states during specific situations. That is, if animals 4233 show AB to specific positive stimuli, the presentation of these whilst animals undergo aversive 4234 situations might have alleviating effects.

Since the use of AB testing has been proposed as an alternative to JBT for practical reasons,
future research should address how attention and judgment bias co-vary as scientific
evidence that AB testing can substitute JBT is presently very limited.

4238 Moreover, experimental refinements in the assessment of dynamic changes in cardiac activity 4239 and surface temperatures in relation to emotional states might open up new ways to measure 4240 animal emotions in future, for instance, by pairing these techniques with image processing or 4241 sensor technologies to record physiological changes automatically on a larger scale.

4242 **7.4 Concluding remarks**

Investigating the link between animal emotions and attentional processing, my thesis findings
suggest that, providing the experimental limitations I highlighted in my literature review
(Chapter 2) are resolved, measuring AB could be valuable for the assessment of affective

4246 states in animals, including horses. My work in Chapter 4 described the first investigation of 4247 emotion-dependent preferential attention allocation in horses, but the findings were 4248 inconclusive. In addition to cognitive measures, I explored changes in HRV and IRT parameters in response to emotional experiences associated with immediate and withheld food rewards 4249 4250 (Chapter 6). My results imply that HRV and IRT measures may not be appropriate to 4251 differentiate the direction of emotional valence, yet, their use as arousal indices should not 4252 be dismissed. I hope that the work described in this thesis contributes to future studies in the 4253 area of animal cognition and emotion.

- Ahloy-Dallaire, J., Espinosa, J., & Mason, G. (2018). Play and optimal welfare: Does play indicate the presence of positive affective states? *Behavioural Processes*, *156*, 3–15.
- Allritz, M., Call, J., & Borkenau, P. (2016). How chimpanzees (Pan troglodytes) perform in a modified emotional Stroop task. *Animal Cognition*, *19*(3), 435–449.
- Amici, F. (2019). Memories of emotional expressions in horses. *Learning & Behavior*, 47(3), 191–192.
- Amting, J. M., Greening, S. G., & Mitchell, D. G. V. (2010). Multiple Mechanisms of Consciousness: The Neural Correlates of Emotional Awareness. *Journal of Neuroscience*, 30(30), 10039– 10047.
- Anderson, C. H., Van Essen, D. C., & Olshausen, B. A. (2005). Chapter 3—Directed Visual Attention and the Dynamic Control of Information Flow. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 11–17). Academic Press.
- Anderson, D. J., & Adolphs, R. (2014). A Framework for Studying Emotions across Species. *Cell*, 157(1), 187–200.
- Anon-BETA. (2006). The National Equestrian Survey 2005/06: Structural Report. SMRC Sportswise Ltd
- Anselme, P. (2015). Incentive salience attribution under reward uncertainty: A Pavlovian model. *Behavioural Processes*, 111, 6–18.
- Appleby, M. C., Olsson, I. A. S., & Galindo, F. (Eds.) (2018). Animal Welfare (3rd ed.). CABI.
- Archer, D. C., Pinchbeck, G. L., French, N. P., & Proudman, C. J. (2008). Risk factors for epiploic foramen entrapment colic: An international study. *Equine Veterinary Journal*, 40(3), 224– 230.
- Archer, J. (1973). Tests for emotionality in rats and mice: A review. *Animal Behaviour*, 21(2), 205–235.
- Arduini, A., Redaelli, V., Luzi, F., Dall'Olio, S., Pace, V., & Nanni Costa, L. (2017). Relationship between Deck Level, Body Surface Temperature and Carcass Damages in Italian Heavy Pigs after Short Journeys at Different Unloading Environmental Conditions. *Animals*, 7(12), 10.
- Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A metaanalytic review and synthesis. *Clinical Psychology Review*, *32*(8), 704–723.
- Asher, L., Friel, M., Griffin, K., & Collins, L. M. (2016). Mood and personality interact to determine cognitive biases in pigs. *Biology Letters*, *12*(11), 20160402.
- Auer, U., Kelemen, Z., Engl, V., & Jenner, F. (2021). Activity Time Budgets—A Potential Tool to Monitor Equine Welfare? *Animals*, *11*(3), 850.
- Aust, U., & Huber, L. (2006). Picture-object recognition in pigeons: Evidence of representational insight in a visual categorization task using a complementary information procedure. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(2), 190–195.
- AWIN. (2015). AWIN welfare assessment protocol for horses.
- Bachmann, I., Bernasconi, P., Herrmann, R., Weishaupt, M. A., & Stauffacher, M. (2003). Behavioural and physiological responses to an acute stressor in crib-biting and control horses. *Applied Animal Behaviour Science*, 82(4), 297–311.
- Baciadonna, L., & McElligott, A. (2015). The use of judgement bias to assess welfare in farm livestock. *Animal Welfare*, 24(1), 81–91.

- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007).
 Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study.
 Psychological Bulletin, 133(1), 1–24.
- Barker, T. H., Howarth, G. S., & Whittaker, A. L. (2016). The effects of metabolic cage housing and sex on cognitive bias expression in rats. *Applied Animal Behaviour Science*, *177*, 70–76.
- Barnard, S., Wells, D. L., Milligan, A. D. S., Arnott, G., & Hepper, P. G. (2018). Personality traits affecting judgement bias task performance in dogs (Canis familiaris). *Scientific Reports*, 8(1), 6660.
- Barnett, J. L., & Hemsworth, P. H. (1990). The validity of physiological and behavioural measures of animal welfare. *Applied Animal Behaviour Science*, *25*(1), 177–187.
- Barstow, A., & Dyson, S. (2015). Clinical features and diagnosis of sacroiliac joint region pain in 296 horses: 2004–2014. *Equine Veterinary Education*, *27*(12), 637–647.
- Bartolomé, E., Sánchez, M. J., Molina, A., Schaefer, A. L., Cervantes, I., & Valera, M. (2013). Using eye temperature and heart rate for stress assessment in young horses competing in jumping competitions and its possible influence on sport performance. *Animal*, *7*(12), 2044–2053.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using Ime4. *Journal of Statistical Software*.
- Bateson, M. (2016). Optimistic and pessimistic biases: A primer for behavioural ecologists. *Current Opinion in Behavioral Sciences*, *12*, 115–121.
- Bateson, M., Brilot, B., & Nettle, D. (2011). Anxiety: An Evolutionary Approach. *The Canadian Journal* of Psychiatry, 56(12), 707–715.
- Bateson, M., & Nettle, D. (2015). *Development of a cognitive bias methodology for measuring low mood in chimpanzees*. 25.
- Bateson, P., Mendl, M., & Feaver, J. (1990). Play in the domestic cat is enhanced by rationing of the mother during lactation. *Animal Behaviour*, 40(3), 514–525.
- Bateson, P., & Young, M. (1981). Separation from the mother and the development of play in cats. *Animal Behaviour, 29*(1), 173–180.
- Baumgartner, M., Boisson, T., Erhard, M. H., & Zeitler-Feicht, M. H. (2020). Common Feeding Practices Pose A Risk to the Welfare of Horses When Kept on Non-Edible Bedding. *Animals*, 10(3), 411.
- Beauchamp, G., Li, Z., Yu, C., Bednekoff, P. A., & Blumstein, D. T. (2021). A meta-analysis of the group-size effect on vigilance in mammals. *Behavioral Ecology*, *32*(5), 919–925.
- Bekoff, M. (2000). Animal Emotions: Exploring Passionate Natures. BioScience, 50(10), 861.
- Bellegarde, L. G. A., Haskell, M. J., Duvaux-Ponter, C., Weiss, A., Boissy, A., & Erhard, H. W. (2017a). Face-based perception of emotions in dairy goats. *Applied Animal Behaviour Science*, 193, 51–59.
- Bellegarde, L. G. A., Erhard, H. W., Weiss, A., Boissy, A., & Haskell, M. J. (2017b). Valence of Facial Cues Influences Sheep Learning in a Visual Discrimination Task. *Frontiers in Veterinary Science*, 4, 188.
- Benhajali, H., Richard-Yris, M.-A., Leroux, M., Ezzaouia, M., Charfi, F., & Hausberger, M. (2008). A note on the time budget and social behaviour of densely housed horses. *Applied Animal Behaviour Science*, 112(1–2), 196–200.

- Bethell, E. (2009). *Cognitive bias in rhesus macaques (Macaca mulatta), a novel measure of animal welfare*. University of Roehampton.[Thesis]
- Bethell, E., Holmes, A., MacLarnon, A., & Semple, S. (2016). Emotion Evaluation and Response
 Slowing in a Non-Human Primate: New Directions for Cognitive Bias Measures of Animal
 Emotion? *Behavioral Sciences*, 6(1), 2.
- Bethell, E. J. (2015). A "How-To" Guide for Designing Judgment Bias Studies to Assess Captive Animal Welfare. *Journal of Applied Animal Welfare Science*, *18*(sup1), S18–S42.
- Bethell, E. J., Holmes, A., MacLarnon, A., & Semple, S. (2012). Evidence That Emotion Mediates Social Attention in Rhesus Macaques. *PLoS ONE*, *7*(8), e44387.
- Bethell, E. J., Cassidy, L. C., Brockhausen, R. R., & 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, 1051.
- Bethell, E., Kemp, C., Thatcher, H., Schroeder, J., Arbuckle, K., Farningham, D., Witham, C., Holmes, A., MacLarnon, A., & Semple, S. (2019b). Heritability and maternal effects on social attention during an attention bias task in a non-human primate, Macaca mulatta [Preprint].
 EcoEvoRxiv. [preprint]
- Billman, G. E., Huikuri, H. V., Sacha, J., & Trimmel, K. (2015). An introduction to heart rate variability: Methodological considerations and clinical applications. *Frontiers in Physiology*, 6.
- Boggiani, L., Addessi, E., & Schino, G. (2018). Receiving aggression triggers attention bias in tufted capuchin monkeys. *Animal Behaviour*, *146*, 173–180.
- Boissy, A. (1995). Fear and Fearfulness in Animals. *The Quarterly Review of Biology*, 70(2), 165–191.
- Boissy, A., Manteuffel, G., Jensen, M. B., Moe, R. O., Spruijt, B., Keeling, L. J., Winckler, C., Forkman, B., Dimitrov, I., Langbein, J., Bakken, M., Veissier, I., & Aubert, A. (2007). Assessment of positive emotions in animals to improve their welfare. *Physiology & Behavior*, *92*(3), 375–397.
- Botreau, R., Bracke, M. B. M., Perny, P., Butterworth, A., Capdeville, J., Van Reenen, C. G., & Veissier,
 I. (2007a). Aggregation of measures to produce an overall assessment of animal welfare.
 Part 2: Analysis of constraints. *Animal*, 1(8), 1188–1197.
- Botreau, R., Bracke, M. B. M., Perny, P., Butterworth, A., Capdeville, J., Van Reenen, C. G., & Veissier,
 I. (2007b). Aggregation of measures to produce an overall assessment of animal welfare.
 Part 2: Analysis of constraints. *Animal: An International Journal of Animal Bioscience*, 1(8), 1188–1197.
- Bouwknecht, A., Olivier, B., & Paylor, R. E. (2007). The stress-induced hyperthermia paradigm as a physiological animal model for anxiety: A review of pharmacological and genetic studies in the mouse. *Neuroscience & Biobehavioral Reviews*, *31*(1), 41–59.
- Bovet, D., & Vauclair, J. (2000). Picture recognition in animals and humans. *Behavioural Brain Research*, 109(2), 143–165.
- Bracke, M. B. M., Metz, J. H. M., & Spruijt, B. (1999). Overall animal welfare reviewed. Part 2: Assessment tables and schemes. *Assessment of Animal Welfare Collection*.
- Bracke, M. B. M., Spruijt, B., & Metz, J. H. M. (1999). Overall animal welfare assessment reviewed. Part 1: Is it possible? *Assessment of Animal Welfare Collection*.
- Bradley, B. P., Mogg, K., White, J., Groom, C., & de Bono, J. (1999). Attentional bias for emotional faces in generalized anxiety disorder. *The British Journal of Clinical Psychology*, *38*(3), 267–278.

- Briefer, E. F., & McElligott, A. G. (2013). Rescued goats at a sanctuary display positive mood after former neglect. *Applied Animal Behaviour Science*, *146*(1), 45–55. 7
- Briefer, E. F., Maigrot, A.-L., Mandel, R., Freymond, S. B., Bachmann, I., & Hillmann, E. (2015).
 Segregation of information about emotional arousal and valence in horse whinnies. *Scientific Reports*, 5(1), 9989.
- Briefer Freymond, S., Briefer, E. F., Zollinger, A., Gindrat-von Allmen, Y., Wyss, C., & Bachmann, I.
 (2014). Behaviour of horses in a judgment bias test associated with positive or negative reinforcement. *Applied Animal Behaviour Science*, *158*, 34–45.
- Brilot, B. O., Asher, L., & Bateson, M. (2009a). Water bathing alters the speed–accuracy trade-off of escape flights in European starlings. *Animal Behaviour*, 78(4), 801–807.
- Brilot, B. O., & Bateson, M. (2012). Water bathing alters threat perception in starlings. *Biology Letters*, 8(3), 379–381.
- Brilot, B. O., Normandale, C. L., Parkin, A., & Bateson, M. (2009b). Can we use starlings' aversion to eyespots as the basis for a novel 'cognitive bias' task? *Applied Animal Behaviour Science*, 118(3–4), 182–190.
- Broadbent, D. (1958). Perception and Communication. Pergamon Press.
- Brooks, M., Kristiansen, K., van Benthem, K., Magnusson, A., Berg, C., Nielsen, A., Skaug, H.,
 Maechler, M., & Bolker, B. (2017). GlmmTMB Balances Speed and Flexibility Among
 Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, 9(2), 378–400.
- Broom, D. M. (1986). Indicators of poor welfare. British Veterinary Journal, 142(6), 524–526.
- Broom, D. M. (1995). A Useable Definition of Animal Welfare. In *The Thinking Horse* (A. Clarke, pp. 66–75, Guelph: Equine Research Centre).
- Broom, D. M. (1996). Animal Welfare Defined in Terms of Attempts to Cope with the Environment. *Acta Agriculturae Scandinavica Section A, Animal Science*. 27, 22–28.
- Broom, D. M. (2001). Coping, stress and welfare. In *Coping with Challenge: Welfare in Animals including Humans, Proceedings of Dahlem Conference*, ed. D.M.Broom 1–9. Berlin:Dahlem University Press.
- Broom, D. M., & Johnson, K. G. (1993). Stress and Animal Welfare. Kluwer Academic Publishers.
- Brosch, T., Scherer, K., Grandjean, D., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*.
- Browning, H. (2019). *If I Could Talk to the Animals: Measuring Subjective Animal Welfare*. The Australian National University[Thesis]
- Browning, H. (2022a). The Measurability of Subjective Animal Welfare. *Journal of Consciousness Studies*, *29*(3), 150–179.
- Browning, H. (2022b). Assessing measures of animal welfare. *Biology & Philosophy*, 37(4), 36.
- Buckner, J. D., Maner, J. K., & Schmidt, N. B. (2010). Difficulty Disengaging Attention from Social Threat in Social Anxiety. *Cognitive Therapy and Research*, *34*(1), 99–105.
- Burke, D., & Sulikowski, D. (2013). The Evolution of Holistic Processing of Faces. *Frontiers in Psychology*, *4*.
- Cabanac, M. (1992). Pleasure: The common currency. *Journal of Theoretical Biology*, 155(2), 173–200.

- Cabe, P. A. (1976). Transfer of discrimination from solid objects to pictures by pigeons: A test of theoretical models of pictorial perception. *Perception & Psychophysics*, *19*(6), 545–550.
- Campbell, D. L. M., Dickson, E. J., & Lee, C. (2019a). Application of open field, tonic immobility, and attention bias tests to hens with different ranging patterns. *Peerj*, *7*, e8122.
- Carbone, L. (2020). Do "Prey Species" Hide Their Pain? Implications for Ethical Care and Use of Laboratory Animals. *Journal of Applied Animal Ethics Research*, 2(2), 216–236.
- Carere, C., & Locurto, C. (2011). Interaction between animal personality and animal cognition. *Current Zoology*, *57*(4), 491–498.
- Carrasco, M. (2011). Visual attention: The past 25 years. Vision Research, 51(13), 1484–1525.
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, 7(3), 308–313.
- Cassidy, L. C., Bethell, E. J., Brockhausen, R. R., Boretius, S., Treue, S., & Pfefferle, D. (2021). The Dot-Probe Attention Bias Task as a Method to Assess Psychological Well-Being after Anesthesia: A Study with Adult Female Long-Tailed Macaques (*Macaca fascicularis*). European Surgical Research.
- Chan, A. H. S., & Courtney, A. J. (1996). Foveal acuity, peripheral acuity and search performance: A review. *International Journal of Industrial Ergonomics*, *18*(2), 113–119.
- Chen, J., van den Bos, E., & Westenberg, P. M. (2020). A systematic review of visual avoidance of faces in socially anxious individuals: Influence of severity, type of social situation, and development. *Journal of Anxiety Disorders*, *70*, 102193.
- Christensen, J. W., Ahrendt, L. P., Lintrup, R., Gaillard, C., Palme, R., & Malmkvist, J. (2012). Does learning performance in horses relate to fearfulness, baseline stress hormone, and social rank? *Applied Animal Behaviour Science*, *140*(1–2), 44–52.
- Christensen, J. W., Zharkikh, T., Ladewig, J., & Yasinetskaya, N. (2002). Social behaviour in stallion groups (Equus przewalskii and Equus caballus) kept under natural and domestic conditions. *Applied Animal Behaviour Science*, 76(1), 11–20.
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A Taxonomy of External and Internal Attention. *Annual Review of Psychology*, 62(1), 73–101.
- Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, *30*(2), 203–216.
- Collier, R. J., Dahl, G. E., & VanBaale, M. J. (2006). Major Advances Associated with Environmental Effects on Dairy Cattle. *Journal of Dairy Science*, *89*(4), 1244–1253.
- Cooper, J. J., & Mason, G. J. (1998). The identification of abnormal behaviour and behavioural problems in stabled horses and their relationship to horse welfare: A comparative review. *Equine Veterinary Journal*, *30*(S27), 5–9.
- Cooper, J. J., McDonald, L., & Mills, D. S. (2000). The effect of increasing visual horizons on stereotypic weaving: Implications for the social housing of stabled horses. *Applied Animal Behaviour Science*, *69*(1), 67–83.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), Article 3.
- Costa, H., Fragoso, S., & Heitor, F. (2019). The relevance of affiliative relationships in horses: Review and future directions. *Pet Behaviour Science*, *8*, 11–26.

- Craig, K., Brown, K., & Baum, A. (1995). Environmental factors in the etiology of anxiety. In *Bloom FE,Kupfer DJ, eds. Psychopharmacology: The fourth Generation of Progress* (pp. 1325–1339). Raven Press.
- Cronin, K. A., Bethell, E. J., Jacobson, S. L., Egelkamp, C., & Hopper, L. M. (2018). Evaluating mood changes in response to anthropogenic noise with a response-slowing task in three species of zoo-housed primates. *Animal Behavior and Cognition*, *5*(2), 209–221.
- Crowell-Davis, S. L., Houpt, K. A., & Burnham, J. S. (1985). Snapping by foals of Equus caballus. *Zeitschrift Für Tierpsychologie*, 69(1), 42–54.
- Crump, A., Arnott, G., & Bethell, E. (2018). Affect-Driven Attention Biases as Animal Welfare Indicators: Review and Methods. *Animals*, 8(8), 136.
- Cussen, V. A., & Mench, J. A. (2014). Personality predicts cognitive bias in captive psittacines, Amazona amazonica. *Animal Behaviour*, *89*, 123–130.
- Dai, F., Cogi, N. H., Heinzl, E. U. L., Dalla Costa, E., Canali, E., & Minero, M. (2015). Validation of a fear test in sport horses using infrared thermography. *Journal of Veterinary Behavior*, 10(2), 128– 136.
- Dalla Costa, E., Minero, M., Lebelt, D., Stucke, D., Canali, E., & Leach, M. C. (2014a). Development of the Horse Grimace Scale (HGS) as a Pain Assessment Tool in Horses Undergoing Routine Castration. *PLoS ONE*, *9*(3), e92281.
- Dalla Costa, E., Murray, L., Dai, F., Canali, E., & Minero, M. (2014b). Equine on-farm welfare assessment: A review of animal-based indicators. *Animal Welfare*, *23*(3), 323–341.
- Darwin, C. (1965). The Expression of the Emotions in Animals and Man. Chicago University Press.
- Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition*, 20(1), 125–151.
- Davidson, R. J. (1995). Brain Asymmetry. MIT Press.
- Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology. I. *Journal of Personality and Social Psychology*, *58*(2), 330–341.
- Dawkins, M. (1980). Animal Suffering. The science of animal welfare (Vol. 1). Springer, Dordrecht.
- Dawkins, M. (2015). Animal Welfare and the Paradox of Animal Consciousness. In Advances in the Study of Behavior (Vol. 47, pp. 5–38). Elsevier.
- Dawkins, M. S. (1990). From an animal's point of view: Motivation, fitness and animal welfare. Behavioral and Brain Sciences, 13, 1–61.
- Dawkins, M. S. (2008). The Science of Animal Suffering. Ethology, 114(10), 937–945.
- Dawkins, M. S. (2017). Animal welfare with and without consciousness. *Journal of Zoology*, 301(1), 1–10.
- De Boyer Des Roches, A., Richard-Yris, M.-A., Henry, S., Ezzaouïa, M., & Hausberger, M. (2008). Laterality and emotions: Visual laterality in the domestic horse (Equus caballus) differs with objects' emotional value. *Physiology & Behavior*, *94*(3), 487–490.
- De Heus, P., Van Oossanen, G., Van Dierendonck, M. C., & Back, W. (2010). A Pressure Algometer Is a Useful Tool to Objectively Monitor the Effect of Diagnostic Palpation by a Physiotherapist in Warmblood Horses. *Journal of Equine Veterinary Science*, *30*(6), 310–321.
- de Ruiter, C., & Brosschot, J. F. (1994). The emotional Stroop interference effect in anxiety: Attentional bias or cognitive avoidance? *Behaviour Research and Therapy*, *32*(3), 315–319.

- de Vere, A. J., & Kuczaj II, S. A. (2016). Where are we in the study of animal emotions? *WIREs Cognitive Science*, 7(5), 354–362.
- de Waal, F. B. M. (2011). What is an animal emotion? What is an animal emotion? Annals of the New York Academy of Sciences, 1224(1), 191–206.
- D'Eath, R. B. (1998). Can video images imitate real stimuli in animal behaviour experiments? Biological Reviews of the Cambridge Philosophical Society, 73(3), 267–292.
- Denny, E. B., & Hunt, R. R. (1992). Affective valence and memory in depression: Dissociation of recall and fragment completion. *Journal of Abnormal Psychology*, *101*, 575–580.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review* of Neuroscience, 18, 193–222.
- Desire, L., Boissy, A., & Veissier, I. (2002). Emotions in farm animals: A new approach to animal welfare in applied ethology. *Behavioural Processes*, 16.
- Di Giminiani, P., Brierley, V. L. M. H., Scollo, A., Gottardo, F., Malcolm, E. M., Edwards, S. A., & Leach, M. C. (2016). The Assessment of Facial Expressions in Piglets Undergoing Tail Docking and Castration: Toward the Development of the Piglet Grimace Scale. *Frontiers in Veterinary Science*, *3*.
- Dougherty, D. M., & Lewis, P. (1991). Stimulus Generalization, Discrimination Learning, and Peak Shift in Horses. *Journal of the Experimental Analysis of Behavior*, *56*(1), 97–104.
- Dougherty, L. R., & Guillette, L. M. (2018). Linking personality and cognition: A meta-analysis. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1756), 20170282.
- Douma, J. C., & Weedon, J. T. (2019). Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods in Ecology and Evolution*, *10*(9), 1412–1430.
- Doyle, R. E., Vidal, S., Hinch, G. N., Fisher, A. D., Boissy, A., & Lee, C. (2010). The effect of repeated testing on judgement biases in sheep. *Behavioural Processes*, *83*(3), 349–352.
- Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*(1), 53–78.
- Dukas, R., & Clark, C. W. (1995). Sustained vigilance and animal performance. *Animal Behaviour*, 49(5), 1259–1267.
- Duke-Elder, S. (1958). System of ophthalmology: Vol. I. The eye in evolution. Henry Kimpton.
- Duncan, I. J. H. (2005). Science-based assessment of animal welfare: Farm animals, Rev. sci. *Tech. Off. Int. Epiz., Vol*, 483–492.
- Dwyer, C. M. (2004). How has the risk of predation shaped the behavioural responses of sheep to fear and distress? *Animal Welfare*, *13*, 269–281.
- Dyson, S., Berger, J., Ellis, A. D., & Mullard, J. (2018). Development of an ethogram for a pain scoring system in ridden horses and its application to determine the presence of musculoskeletal pain. *Journal of Veterinary Behavior*, *23*, 47–57.
- Edgar, J. L., Nicol, C. J., Pugh, C. A., & Paul, E. S. (2013). Surface temperature changes in response to handling in domestic chickens. *Physiology & Behavior*, *119*, 195–200.
- Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, 48, 269–297.
- Ekman, P. (2003). Emotions Revealed: Understanding Faces and Feelings. Cambridge University Press.

- Ekman, P. (2016). What Scientists Who Study Emotion Agree About. *Perspectives on Psychological Science*, *11*(1), 31–34.
- Evangelista, M. C., Monteiro, B. P., & Steagall, P. V. (2022). Measurement properties of grimace scales for pain assessment in nonhuman mammals: A systematic review. *Pain*, 163(6), e697– e714.
- Eysenck, M. W., & Keane, M. T. (2002). Attention and performance limitations (p. 398). MIT Press.
- Fagot, J. (2000). Picture Perception in Animals. Psychology Press.
- Fagot, J., & Parron, C. (2010). Picture Perception in Birds: Perspective from Primatologists. *Comparative Cognition & Behavior Reviews*, *5*, 132–135.
- Fagot, J., Thompson, R. K. R., & Parron, C. (2010). How to read a picture: Lessons from nonhuman primates. *Proceedings of the National Academy of Sciences*, *107*(2), 519–520.
- Fazio, E., & Ferlazzo, A. (2003). Evaluation of stress during transport. *Veterinary Research Communications*, 27 Suppl 1, 519–524.
- Findlay, J. M., & Gilchrist, I. D. (2003). Active Vision: The Psychology of Looking and Seeing. Oxford University Press.
- Flannery, B. (1997). Relational discrimination learning in horses. *Applied Animal Behaviour Science*, 54(4), 267–280.
- Fleishman, L. J., & Endler, J. A. (2000). Some comments on visual perception and the use of video playback in animal behavior studies. *Acta Ethologica*, *3*(1), 15–27.
- Fletcher, K., Cameron, L., & Freeman, M. (2021). Contemplating the Five Domains model of animal welfare assessment: UK horse owner perceptions of equine well-being. *Animal Welfare*, 30(3), 259–268.
- Flethøj, M., Kanters, J. K., Pedersen, P. J., Haugaard, M. M., Carstensen, H., Olsen, L. H., & Buhl, R. (2016). Appropriate threshold levels of cardiac beat-to-beat variation in semi-automatic analysis of equine ECG recordings. *BMC Veterinary Research*, 12(1), 266.
- Forkman, B., Boissy, A., Meunier-Salaün, M.-C., Canali, E., & Jones, R. B. (2007). A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiology & Behavior*, 92(3), 340– 374.
- Fortin, M., Valenchon, M., Lévy, F., Calandreau, L., Arnould, C., & Lansade, L. (2018). Emotional state and personality influence cognitive flexibility in horses (Equus caballus). *Journal of Comparative Psychology*, 132(2), 130–140.
- Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General*, 130(4), 681– 700.
- Fox, J., & Weisberg, S. (2019). An R Companion to Applied Regression (3rd ed.). Sage.
- Franchi, G. A., Herskin, M. S., & Jensen, M. B. (2020). Do dietary and milking frequency changes during a gradual dry-off affect feed-related attention bias and visual lateralisation in dairy cows? *Applied Animal Behaviour Science*, *223*, 104923.
- Frantz, R. L. (1958). Pattern vision in young infants. *The Psychological Record*, 8, 43–47.
- Fraser, A. F. (2009). The Behaviour and Welfare of the Horse (2nd ed.). Cabi.
- Fraser, D. (1995). Science, Values and Animal Welfare: Exploring the 'Inextricable Connection.' Animal Welfare, 4(2), 103–117.

- Fraser, D. (2003). Assessing Animal Welfare at the Farm and Group Level: The Interplay of Science and Values. *Animal Welfare*, *12*(4), 433–443.
- Fraser, D. (2007). Understanding animal welfare. Acta Veterinaria Scandinavica, 50(Suppl 1), S1. https://doi.org/10.1186/1751-0147-50-S1-S1
- Fraser, D., Weary, D. M., Pajor, E. A., & Milligan, B. N. (1997). A Scientific Conception of Animal Welfare that Reflects Ethical Concerns. *Animal Welfare*, *6*, 187–205.
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-andbuild theory of positive emotions. *American Psychologist*, *56*, 218–226.
- Friedman, B. H., & Thayer, J. F. (1998). Autonomic balance revisited: Panic anxiety and heart rate variability. *Journal of Psychosomatic Research*, 44(1), 133–151.
- Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, 43, 349–358.
- Frohnwieser, A., Murray, J. C., Pike, T. W., & Wilkinson, A. (2016). Using robots to understand animal cognition. *Journal of the Experimental Analysis of Behavior*, *105*(1), 14–22.
- Fureix, C., Beaulieu, C., Argaud, S., Rochais, C., Quinton, M., Henry, S., Hausberger, M., & Mason, G. (2015). Investigating anhedonia in a non-conventional species: Do some riding horses Equus caballus display symptoms of depression? *Applied Animal Behaviour Science*, *162*, 26–36. 7
- Fureix, C., Bourjade, M., Henry, S., Sankey, C., & Hausberger, M. (2012a). Exploring aggression regulation in managed groups of horses Equus caballus. *Applied Animal Behaviour Science*, 138(3–4), 216–228.
- Fureix, C., Hausberger, M., Seneque, E., Morisset, S., Baylac, M., Cornette, R., Biquand, V., & Deleporte, P. (2011). Geometric morphometrics as a tool for improving the comparative study of behavioural postures. *Naturwissenschaften*, *98*(7), 583–592.
- Fureix, C., Jego, P., Henry, S., Lansade, L., & Hausberger, M. (2012b). Towards an Ethological Animal Model of Depression? A Study on Horses. *PLoS ONE*, 7(6), e39280.
- Fureix, C., Menguy, H., & Hausberger, M. (2010). Partners with Bad Temper: Reject or Cure? A Study of Chronic Pain and Aggression in Horses. *PLoS ONE*, *5*(8), e12434.
- Furtado, T., Perkins, E., McGowan, C., & Pinchbeck, G. (2021a). Equine Management in UK Livery Yards during the COVID-19 Pandemic—"As Long As the Horses Are Happy, We Can Work Out the Rest Later." Animals, 11(5), 1416.
- Furtado, T., Perkins, E., Pinchbeck, G., McGowan, C., Watkins, F., & Christley, R. (2021b). Hidden in Plain Sight: Uncovering the Obesogenic Environment Surrounding the UK's Leisure Horses. *Anthrozoös*, 34(4), 491–506.
- Gabor, V., & Gerken, M. (2012). Cognitive testing in horses using a computer based apparatus. *Applied Animal Behaviour Science*, 139(3–4), 242–250.
- Gaidica, M., & Dantzer, B. (2020). Quantifying the Autonomic Response to Stressors—One Way to Expand the Definition of "Stress" in Animals. *Integrative and Comparative Biology*, 60(1), 113–125.
- Geor, R. J. (2008). Metabolic Predispositions to Laminitis in Horses and Ponies: Obesity, Insulin Resistance and Metabolic Syndromes. *Journal of Equine Veterinary Science*, *28*(12), 753–759.
- Godoy, L. D., Rossignoli, M. T., Delfino-Pereira, P., Garcia-Cairasco, N., & de Lima Umeoka, E. H.
 (2018). A Comprehensive Overview on Stress Neurobiology: Basic Concepts and Clinical Implications. *Frontiers in Behavioral Neuroscience*, 12.

- Goodwin, D. (2002). Horse behaviour: Evolution, domestication and feralisation. In *N. Waran (ed), The Welfare of Horses* (pp. 1–18). Springer.
- Gottlieb, D. A. (2005). Acquisition With Partial and Continuous Reinforcement in Rat Magazine Approach. Journal of Experimental Psychology: Animal Behavior Processes, 31(3), 319–333.
- Gregory, N. J., Bolderston, H., & Antolin, J. V. (2019). Attention to faces and gaze-following in social anxiety: Preliminary evidence from a naturalistic eye-tracking investigation. *Cognition and Emotion*, 33(5), 931–942.
- Griffin, A. S., Blumstein, D. T., & Evans, C. S. (2000). Training Captive-Bred or Translocated Animals to Avoid Predators. *Conservation Biology*, 14(5), 1317–1326.
- Grzimek, B. (1943). Begrussung zweier Pferde: Das Erkennen von Phantomen und Bildern. 5, 465–480.
- Gupta, R. (2019). Positive emotions have a unique capacity to capture attention. In *Progress in Brain Research* (Vol. 247, pp. 23–46). Elsevier.
- Gygax, L. (2014). The A to Z of statistics for testing cognitive judgement bias. *Animal Behaviour*, 95, 59–69.
- Hamlaoui, A., Keeling, L., Burman, O., & Verbeek, E. (2022). Investigating attentional scope as a novel indicator of emotional state in animals. *Scientific Reports*, *12*(1), Article 1.
- Hanggi, E. (2001). Can horses recognize pictures? *Proceedings of the Third International Conference of Cognitive Science*, 52–56.
- Hanggi, E. B. (1999). Categorization learning in horses (Equus caballus). *Journal of Comparative Psychology*, *113*(3), 243–252.
- Hanggi, E. B. (2003). Discrimination learning based on relative size concepts in horses (Equus caballus). *Applied Animal Behaviour Science*, *83*(3), 201–213.
- Hanggi, E. B., & Ingersoll, J. F. (2009a). Stimulus discrimination by horses under scotopic conditions. *Behavioural Processes*, 82(1), 45–50.
- Hanggi, E. B., & Ingersoll, J. F. (2009b). Long-term memory for categories and concepts in horses (Equus caballus). *Animal Cognition*, *12*(3), 451–462.
- Hanggi, E. B., Ingersoll, J. F., & Waggoner, T. L. (2007). Color vision in horses (Equus caballus): Deficiencies identified using a pseudoisochromatic plate test. *Journal of Comparative Psychology*, 121(1), 65–72.
- Hansen, R. M. (1976). Foods of Free-Roaming Horses in Southern New Mexico. *Journal of Range Management*, *29*(4), 347–347.
- Harding, E. J., Paul, E. S., & Mendl, M. (2004). Cognitive bias and affective state. *Nature*, 427(6972), 312–312.
- Harris, P. A. (1999). Review of equine feeding and stable management practices in the UK concentrating on the last decade of the 20th Century. *Equine Veterinary Journal*, *31*(28), 46–54.
- Harrison, R. (1964). Animal machines; the new factory farming industry. Stuart.
- Hartig, F., & Lohse, L. (2022). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models (0.4.6).
- Hausberger, M., Fureix, C., Bourjade, M., Wessel-Robert, S., & Richard-Yris, M.-A. (2012). On the significance of adult play: What does social play tell us about adult horse welfare? *Naturwissenschaften*, *99*(4), 291–302.

- Hausberger, M., Stomp, M., Sankey, C., Brajon, S., Lunel, C., & Henry, S. (2019). Mutual interactions between cognition and welfare: The horse as an animal model. *Neuroscience & Biobehavioral Reviews*, *107*, 540–559.
- Haussler, K. K. (2020). Pressure Algometry for the Detection of Mechanical Nociceptive Thresholds in Horses. *Animals*, *10*(12), 2195.
- Haussler, K. K., & Erb, H. N. (2006). Pressure algometry for the detection of induced back pain in horses: A preliminary study. *Equine Veterinary Journal*, *37*(1), 76–81.
- Hayes, M. H. H. (1978). Veterinary Notes for Horse Owners. In *McGreevy et al., Management factors* associated with stereotypic and redirected behaviour in the Thoroughbred horse, Equine vet. *J.* (1995) 27 (2) 86-91. Stanley Paul and Co.
- He, Z. (2020). The control mechanisms of heart rate dynamics in a new heart rate nonlinear time series model. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-61562-6
- Hebel, R. (1976). Distribution of retinal ganglion cells in five mammalian species (pig, sheep, ox, horse, dog). *Anatomy and Embryology*, *150*(1), 45–51.
- Held, S. D. E., & Špinka, M. (2011). Animal play and animal welfare. *Animal Behaviour*, *81*(5), 891–899.
- Henry, S., Fureix, C., Rowberry, R., Bateson, M., & Hausberger, M. (2017). Do horses with poor welfare show 'pessimistic' cognitive biases? *The Science of Nature*, *104*(1–2), 8.
- Herborn, K. A., Graves, J. L., Jerem, P., Evans, N. P., Nager, R., McCafferty, D. J., & McKeegan, D. E. F. (2015). Skin temperature reveals the intensity of acute stress. *Physiology & Behavior*, 152, 225–230.
- Herborn, K. A., Jerem, P., Nager, R. G., McKeegan, D. E. F., & McCafferty, D. J. (2018). Surface temperature elevated by chronic and intermittent stress. *Physiology & Behavior*, 191, 47–55.
- Herron, M. A., Martin, J. E., & Joyce, J. R. (1978). Quantitative study of the decussating optic axons in the pony, cow, sheep, and pig. *American Journal of Veterinary Research*, *39*(7), 1137–1139.
- Hintze, S., & Schanz, L. (2021). Using the Judgment Bias Task to Identify Behavioral Indicators of Affective State: Do Eye Wrinkles in Horses Reflect Mood? *Frontiers in Veterinary Science*, 8.
- Hintze, S., Smith, S., Patt, A., Bachmann, I., & Würbel, H. (2016). Are Eyes a Mirror of the Soul? What Eye Wrinkles Reveal about a Horse's Emotional State. *PLOS ONE*, *11*(10), e0164017.
- Hockenhull, J., Birke, L., & Creighton, E. (2010). The Horse's Tale: Narratives of Caring for/about Horses. *Society & Animals*, *18*(4), 331–347.
- Hockenhull, J., & Creighton, E. (2014a). Management practices associated with owner-reported stable-related and handling behaviour problems in UK leisure horses. *Applied Animal Behaviour Science*, *155*, 49–55.
- Hockenhull, J., & Creighton, E. (2014b). Pre-feeding behaviour in UK leisure horses and associated feeding routine risk factors. *Animal Welfare*, 23(3), 297–308.
- Hockenhull, J., & Creighton, E. (2015). The day-to-day management of UK leisure horses and the prevalence of owner-reported stable-related and handling behaviour problems. *Animal Welfare*, *24*(1), 29–36.
- Hockenhull, J., & Furtado, T. (2021). Escaping the gilded cage: Could COVID-19 lead to improved equine welfare? A review of the literature. *Applied Animal Behaviour Science*, 237, 105303.

- Hockenhull, J., & Whay, H. R. (2014). A review of approaches to assessing equine welfare: A review of approaches to assessing equine welfare. *Equine Veterinary Education*, 26(3), 159–166.
- Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J.-H., & Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*, *81*(7), 2288–2303.
- Hopper, L. M., Allritz, M., Egelkamp, C. L., Huskisson, S. M., Jacobson, S. L., Leinwand, J. G., & Ross, S. R. (2021). A Comparative Perspective on Three Primate Species' Responses to a Pictorial Emotional Stroop Task. *Animals*, *11*(3), 588.
- Horseman, S. V., Buller, H., Mullan, S., & Whay, H. R. (2016). Current Welfare Problems Facing Horses in Great Britain as Identified by Equine Stakeholders. *PLOS ONE*, *11*(8), e0160269.
- Hotchkiss, J. W., Reid, S. W. J., & Christley, R. M. (2007). A survey of horse owners in Great Britain regarding horses in their care. Part 1: Horse demographic characteristics and management. *Equine Veterinary Journal*, *39*(4), 294–300.
- Hothersall, B., & Casey, R. (2012). Undesired behaviour in horses: A review of their development, prevention, management and association with welfare: Undesired behaviour in horses. *Equine Veterinary Education*, *24*(9), 479–485.
- Hothersall, B., Gale, E. V., Harris, P., & Nicol, C. J. (2010). Cue use by foals (Equus caballus) in a discrimination learning task. *Animal Cognition*, *13*(1), 63–74.
- Houpt, K. A., & Houpt, T. R. (1988). Social and Illumination Preferences of Mares. *Journal of Animal Science*, 66(9), 2159.
- Howarth, E. R. I., Kemp, C., Thatcher, H. R., Szott, I. D., Farningham, D., Witham, C. L., Holmes, A., Semple, S., & Bethell, E. J. (2021). Developing and validating attention bias tools for assessing trait and state affect in animals: A worked example with Macaca mulatta. *Applied Animal Behaviour Science*, 234, 105198.
- Hughes, A. (1977). The Topography of Vision in Mammals of Contrasting Life Style: Comparative Optics and Retinal Organisation. In F. Crescitelli (Ed.), *The Visual System in Vertebrates: Vol.* 7(5), 613–756. Springer Berlin Heidelberg.
- Huskisson, S. M., Ross, S. R., & Hopper, L. M. (2021). Do zoo visitors induce attentional bias effects in primates completing cognitive tasks? *Animal Cognition*, *24*(3), 645–653.
- Ille, N., Erber, R., Aurich, C., & Aurich, J. (2014). Comparison of heart rate and heart rate variability obtained by heart rate monitors and simultaneously recorded electrocardiogram signals in nonexercising horses. *Journal of Veterinary Behavior*, 9(6), 341–346.
- Ioannou, S., Chotard, H., & Davila-Ross, M. (2015). No strings attached: Physiological monitoring of rhesus monkeys (Macaca mulatta) with thermal imaging. *Frontiers in Behavioral Neuroscience*, 9.
- Ioannou, S., Gallese, V., & Merla, A. (2014). Thermal infrared imaging in psychophysiology: Potentialities and limits. *Psychophysiology*, *51*(10), 951–963.
- Jacobs, G. H., Deegan, J. F., & Neitz, J. (1998). Photopigment basis for dichromatic color vision in cows, goats, and sheep. *Visual Neuroscience*, *15*(03).
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014). An Introduction to Statistical Learning Springer Texts in Statistics An Introduction to Statistical Learning (Vol. 103). Springer Publishing Company, Incorporated.
- James, W. (1890). The principles of psychology (Vol. 2). Henry Holt and Company.

- Jonckheer-Sheehy, V. S. M., Delesalle, C. J., van den Belt, A. J. M., & van den Boom, R. (2012). Bad behavior or a physical problem? Rearing in a Dutch Warmblood mare. *Journal of Veterinary Behavior*, 7(6), 380–385.
- Jonckheer-Sheehy, V. S. M., & Houpt, K. A. (2015). Management methods to improve the welfare of horses used in research. *Lab Animal*, *44*(9), 350–358.
- Joy, A., Taheri, S., Dunshea, F. R., Leury, B. J., DiGiacomo, K., Osei-Amponsah, R., Brodie, G., & Chauhan, S. S. (2022). Non-invasive measure of heat stress in sheep using machine learning techniques and infrared thermography. *Small Ruminant Research*, *207*, 106592.
- Kappel, S., Mendl, M. T., Barrett, D. C., Murrell, J. C., & Whay, H. R. (2017). Lateralized behaviour as indicator of affective state in dairy cows. *PLOS ONE*, *12*(9), e0184933.
- Katayama, M., Kubo, T., Mogi, K., Ikeda, K., Nagasawa, M., & Kikusui, T. (2016). Heart rate variability predicts the emotional state in dogs. *Behavioural Processes*, *128*, 108–112.
- Kellough, J. L., Beevers, C. G., Ellis, A. J., & Wells, T. T. (2008). Time course of selective attention in clinically depressed young adults: An eye tracking study. *Behaviour Research and Therapy*, 46(11), 1238–1243.
- Kendrick, K. M., Atkins, K., Hinton, M. R., Heavens, P., & Keverne, B. (1996). Are faces special for sheep? Evidence from facial and object discrimination learning tests showing effects of inversion and social familiarity. *Behavioural Processes*, 38(1), 19–35.
- King, H. M., Kurdziel, L. B., Meyer, J. S., & Lacreuse, A. (2012). Effects of testosterone on attention and memory for emotional stimuli in male rhesus monkeys. *Psychoneuroendocrinology*, 37(3), 396–409.
- Knoch, S., Whiteside, M. A., Madden, J. R., Rose, P. E., & Fawcett, T. W. (2022). Hot-headed peckers: Thermographic changes during aggression among juvenile pheasants (*Phasianus colchicus*). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1845), 20200442.
- Knudsen, E. I. (2020). Evolution of neural processing for visual perception in vertebrates. *Journal of Comparative Neurology*, *528*(17), 2888–2901.
- Koda, H., Sato, A., & Kato, A. (2013). Is attentional prioritisation of infant faces unique in humans?:
 Comparative demonstrations by modified dot-probe task in monkeys. *Behavioural Processes*, 98, 31–36.
- Koster, E. H. W., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy*, 42(10), 1183–1192.
- Krauzlis, R. J., Wang, L., Yu, G., & Katz, L. N. (2021). What is attention? WIREs Cognitive Science.
- Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, 84(3), 394–421.
- Kremer, L., Bus, J., Webb, L., Bokkers, E., Engel, B., van der Werf, J., Schnabel, S., & van Reenen, C. (2021). Housing and personality effects on judgement and attention biases in dairy cows. SCIENTIFIC REPORTS, 11(1).
- Kremer, L., Klein Holkenborg, S. E. J., Reimert, I., Bolhuis, J. E., & Webb, L. E. (2020). The nuts and bolts of animal emotion. *Neuroscience & Biobehavioral Reviews*, *113*, 273–286.
- Kret, M. E., Jaasma, L., Bionda, T., & Wijnen, J. G. (2016). Bonobos (Pan paniscus) show an attentional bias toward conspecifics' emotions. *Proceedings of the National Academy of Sciences of the United States of America*, 113(14), 3761–3766.

- Krueger, K., Esch, L., Farmer, K., & Marr, I. (2021). Basic Needs in Horses?—A Literature Review. Animals, 11(6), 1798.
- Krueger, K., Farmer, K., & Heinze, J. (2014). The effects of age, rank and neophobia on social learning in horses. *Animal Cognition*, 17(3), 645–655.
- Kuwahara, M., Hashimoto, S., Ishii, K., Yagi, Y., Hada, T., Hiraga, A., Kai, M., Kubo, K., Oki, H., Tsubone, H., & Sugano, S. (1996). Assessment of autonomic nervous function by power spectral analysis of heart rate variability in the horse. *Journal of the Autonomic Nervous System*, 60(1–2), 43–48.
- Kyrios, M., & Iob, M. A. (1998). Automatic and strategic processing in obsessive-compulsive disorder: Attentional bias, cognitive avoidance or more complex phenomena? *Journal of Anxiety Disorders*, 12, 271–292.
- Lacreuse, A., Schatz, K., Strazzullo, S., King, H. M., & Ready, R. (2013). Attentional biases and memory for emotional stimuli in men and male rhesus monkeys. *Animal Cognition*, *16*(6), 861–871.
- Lagisz, M., Zidar, J., Nakagawa, S., Neville, V., Sorato, E., Paul, E. S., Bateson, M., Mendl, M., & Løvlie, H. (2020). Optimism, pessimism and judgement bias in animals: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 118, 3–17.
- Lampe, J. F., & Andre, J. (2012). Cross-modal recognition of human individuals in domestic horses (Equus caballus). *Animal Cognition*, *15*(4), 623–630.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, Attention, and the Startle Reflex. *Psychological Review*, *97*(3), 377–395.
- Lang, P. J., Davis, M., & Ohman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, *61*(3), 137–159.
- Langbein, J., Nürnberg, G., & Manteuffel, G. (2004). Visual discrimination learning in dwarf goats and associated changes in heart rate and heart rate variability. *Physiology & Behavior*, *82*(4), 601–609.
- Lange, A., Bauer, L., Futschik, A., Waiblinger, S., & Luerzel, S. (2020). Talking to Cows: Reactions to Different Auditory Stimuli During Gentle Human-Animal Interactions. *Frontiers in Psychology*, 11, 579346.
- Lange, A., Waiblinger, S., van Hasselt, R., Mundry, R., Futschik, A., & Lürzel, S. (2021). Effects of restraint on heifers during gentle human-animal interactions. *Applied Animal Behaviour Science*, 243, 105445.
- Langford, D. J., Bailey, A. L., Chanda, M. L., Clarke, S. E., Drummond, T. E., Echols, S., Glick, S., Ingrao, J., Klassen-Ross, T., LaCroix-Fralish, M. L., Matsumiya, L., Sorge, R. E., Sotocinal, S. G., Tabaka, J. M., Wong, D., van den Maagdenberg, A. M. J. M., Ferrari, M. D., Craig, K. D., & Mogil, J. S. (2010). Coding of facial expressions of pain in the laboratory mouse. *Nature Methods*, 7(6), Article 6.
- Lansade, L., Colson, V., Parias, C., Reigner, F., Bertin, A., & Calandreau, L. (2020a). Human Face Recognition in Horses: Data in Favor of a Holistic Process. *Frontiers in Psychology*, *11*, 575808.
- Lansade, L., Colson, V., Parias, C., Trösch, M., Reigner, F., & Calandreau, L. (2020b). Female horses spontaneously identify a photograph of their keeper, last seen six months previously. *Scientific Reports*, 10(1), 6302.

- Lansade, L., Marchand, A. R., Coutureau, E., Ballé, C., Polli, F., & Calandreau, L. (2017). Personality and predisposition to form habit behaviours during instrumental conditioning in horses (Equus caballus). *PLOS ONE*, *12*(2), e0171010.
- Larose, C., Richard-Yris, M.-A., Hausberger, M., & Rogers, L. J. (2006). Laterality of horses associated with emotionality in novel situations. *Laterality*, 11(4), 355–367.
- LeDoux, J. (1996). The emotional brain. Simon & Schuster.
- Lee, C., Cafe, L. M., Robinson, S. L., Doyle, R. E., Lea, J. M., Small, A. H., & Colditz, I. G. (2018). Anxiety influences attention bias but not flight speed and crush score in beef cattle. *Applied Animal Behaviour Science*, 205, 210–215.
- Lee, C., Verbeek, E., Doyle, R., & Bateson, M. (2016). Attention bias to threat indicates anxiety differences in sheep. *Biology Letters*, *12*(6), 20150977.
- Lee, J., Floyd, T., Erb, H., & Houpt, K. (2011). Preference and demand for exercise in stabled horses. *Applied Animal Behaviour Science*, 130(3–4), 91–100.
- Leinwand, J. G., Fidino, M., Ross, S. R., & Hopper, L. M. (2022). Familiarity mediates apes' attentional biases toward human faces. *Proceedings Biological Sciences*, *289*(1973), 20212599.
- Leliveld, L. M. C., Langbein, J., & Puppe, B. (2013). The emergence of emotional lateralization: Evidence in non-human vertebrates and implications for farm animals. *Applied Animal Behaviour Science*, *145*, 1–14.
- Leopold, D. A., & Rhodes, G. (2010). A comparative view of face perception. *Journal of Comparative Psychology*, *124*(3), 233–251.
- Lesimple, C. (2020). Indicators of Horse Welfare: State-of-the-Art. Animals, 10(2), 294.
- Lesimple, C., Fureix, C., Biquand, V., & Hausberger, M. (2013). Comparison of clinical examinations of back disorders and humans' evaluation of back pain in riding school horses. *BMC Veterinary Research*, 9(1), 209.
- Lesimple, C., Fureix, C., De Margerie, E., Sénèque, E., Menguy, H., & Hausberger, M. (2012). Towards a Postural Indicator of Back Pain in Horses (Equus caballus). *PLoS ONE*, 7(9), e44604.
- Lesimple, C., Fureix, C., Menguy, H., & Hausberger, M. (2010). Human Direct Actions May Alter Animal Welfare, a Study on Horses (Equus caballus). *PLoS ONE*, *5*(4), e10257.
- Lesimple, C., Gautier, E., Benhajali, H., Rochais, C., Lunel, C., Bensaïd, S., Khalloufi, A., Henry, S., & Hausberger, M. (2019). Stall architecture influences horses' behaviour and the prevalence and type of stereotypies. *Applied Animal Behaviour Science*, *219*, 104833.
- Lesimple, C., Poissonnet, A., & Hausberger, M. (2016). How to keep your horse safe? An epidemiological study about management practices. *Applied Animal Behaviour Science*, 181,
- Levenson, R. W. (2014). The Autonomic Nervous System and Emotion. *Emotion Review*, 6(2), 100–112.
- Lewis, L. S., Kano, F., Stevens, J. M. G., DuBois, J. G., Call, J., & Krupenye, C. (2021). Bonobos and chimpanzees preferentially attend to familiar members of the dominant sex. *Animal Behaviour*, *177*, 193–206.
- Lima, S. L., & Dill, L. M. (1990). Behavioral decisions made under the risk of predation: A review and prospectus. *Canadian Journal of Zoology*, *68*(4), 619–640.
- Liu, Y., Fu, Q., & Fu, X. (2009). The interaction between cognition and emotion. *Chinese Science Bulletin*, 54(22), 4102.

- Löckener, S., Reese, S., Erhard, M., & Wöhr, A.-C. (2016). Pasturing in herds after housing in horseboxes induces a positive cognitive bias in horses. *Journal of Veterinary Behavior*, 11, 50–55.
- Lockhofen, D. E. L., Hübner, N., Hemdan, F., Sammer, G., Henare, D., Schubö, A., & Mulert, C. (2021). Differing Time Courses of Reward-Related Attentional Processing: An EEG Source-Space Analysis. *Brain Topography*, *34*(3), 283–296.
- Long, K., McGowan, C. M., & Hyytiäinen, H. K. (2020). Effect of Caudal Traction on Mechanical Nociceptive Thresholds of Epaxial and Pelvic Musculature on a Group of Horses With Signs of Back Pain. *Journal of Equine Veterinary Science*, 93, 103197.
- Lonsdorf, E. V., Engelbert, L. M., & Howard, L. H. (2019). A competitive drive? Same-sex attentional preferences in capuchins. *American Journal of Primatology*, *81*(6), e22998.
- Love, E. J., Murrell, J., & Whay, H. R. (2011). Thermal and mechanical nociceptive threshold testing in horses: A review. *Veterinary Anaesthesia and Analgesia*, *38*(1), 3–14.
- Low, P. (2012). Cambridge Declaration on Consciousness.
- Ludwig, N., Gargano, M., Luzi, F., Carenzi, C., & Verga, M. (2007). Technical note: Applicability of infrared thermography as a non invasive measurements of stress in rabbit. *World Rabbit Science*, *15*(4), Article 4.
- Luna, D., Gonzalez, C., Byrd, C. J., Palomo, R., Huenul, E., & Figueroa, J. (2021). The Effect of Demonstrator Social Rank on the Attentiveness and Motivation of Pigs to Positively Interact with Their Human Caretakers. *Animals*, 11(7), 2140.
- Lundqvist, D., Juth, P., & Öhman, A. (2014). Using facial emotional stimuli in visual search experiments: The arousal factor explains contradictory results. *Cognition and Emotion*, *28*(6), 1012–1029.
- Luo, L., Reimert, I., de Haas, E. N., Kemp, B., & Bolhuis, J. E. (2019). Effects of early and later life environmental enrichment and personality on attention bias in pigs (Sus scrofa domesticus). *Animal Cognition*, 22(6), 959–972.
- Luthersson, N., Nielsen, K. H., Harris, P., & Parkin, T. D. H. (2009). The prevalence and anatomical distribution of equine gastric ulceration syndrome (EGUS) in 201 horses in Denmark. *Equine Veterinary Journal*, *41*(7), 619–624.
- MacFarland, D. (1987). The Oxford Companion to Animal Behaviour. Oxford University Press.
- Mack, A., & Rock, I. (1998). Inattentional Blindness. Bradford Book.
- MacLeod, A. K., & Byrne, A. (1996). Anxiety, depression, and the anticipation of future positive and negative experiences. *Journal of Abnormal Psychology*, *105*, 286–289.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.
- MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*(1), 15–20.
- Mandalaywala, T. M., Parker, K. J., & Maestripieri, D. (2014). Early Experience Affects the Strength of Vigilance for Threat in Rhesus Monkey Infants. *Psychological Science*, *25*(10), 1893–1902.
- Mandel, R., Whay, H. R., Nicol, C. J., & Klement, E. (2013). The effect of food location, heat load, and intrusive medical procedures on brushing activity in dairy cows. *Journal of Dairy Science*, *96*(10), 6506–6513.

- Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The Effects of Fear and Anger Facial Expressions on Approach- and Avoidance-Related Behaviors. *Emotion*, *5*, 119–124.
- Marzouki, Y., Gullstrand, J., Goujon, A., & Fagot, J. (2014). Baboons' Response Speed Is Biased by Their Moods. *PLoS ONE*, *9*(7), e102562.
- Mason, G. (2006). Stereotypic behaviour in captive animals: Fundamentals and implications for welfare and beyond. In G. Mason & J. Rushen (Eds.), *Stereotypic animal behaviour: Fundamentals and applications to welfare* (2nd ed., pp. 325–356). CABI.
- Mason, G., Cooper, J., & Clarebrough, C. (2001). Frustrations of fur-farmed mink. 410(6824), 35-36.
- Mason, G., & Latham, N. (2004). Can't stop, won't stop: Is stereotypy a reliable animal welfare indicator? *Animal Welfare*, 13.
- Mason, G., & Mendl, M. (1993). Why is there no simple way of Measuring Animal Welfare? *Animal Welfare*, 19.
- Mason, G., & Rushen, J. (Eds.). (2006). Stereotypic animal behaviour: Fundamentals and applications to welfare (2nd ed). CABI Pub.
- Mathews, A., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy*, *23*, 563–569.
- Mathews, A., & MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. Annual Review of Psychology, 45, 25–50.
- Mathews, A., & MacLeod, C. (2002). Induced processing biases have causal effects on anxiety. *Cognition and Emotion*, *16*, 331–354.
- McBride, S. D., & Hemmings, A. (2005). Altered mesoaccumbens and nigro-striatal dopamine physiology is associated with stereotypy development in a non-rodent species. *Behavioural Brain Research*, *159*(1), 113–118.
- McBride, S. D., Hemmings, A., & Robinson, K. (2004). A preliminary study on the effect of massage to reduce stress in the horse. *Journal of Equine Veterinary Science*, *24*(2), 76–81.
- McBride, S. D., & Morton, A. J. (2018). Visual attention and cognitive performance in sheep. *Applied Animal Behaviour Science*, *206*, 52–58.
- McBride, S., & Hemmings, A. (2009). A Neurologic Perspective of Equine Stereotypy. *Journal of Equine Veterinary Science*, 29(1), 10–16.
- McCafferty, D. J., Gallon, S., & Nord, A. (2015). Challenges of measuring body temperatures of freeranging birds and mammals. *Animal Biotelemetry*, *3*(1), 33.
- McGreevy, P. (2012). Equine Behavior (2nd ed.). Elsevier.
- McGreevy, P., French, N., & Nicol, C. (1995a). The prevalence of abnormal behaviours in Dressage, Eventing and Endurance horses in relation to stabling. *The Veterinary Record*, *137*, 36–37.
- Mcgreevy, P. D., Cripps, P. J., French, N. P., Green, L. E., & Nicol, C. J. (1995b). Management factors associated with stereotypic and redirected behaviour in the Thoroughbred horse. *Equine Veterinary Journal*, 27(2), 86–91.
- McGreevy, P. D., Richardson, J. D., Nicol, C. J., & Lane, J. G. (1995c). Radiographic and endoscopic study of horses performing an oral based stereotypy. *Equine Veterinary Journal*, 27(2), 92–95.
- McLennan, K. M., Rebelo, C. J. B., Corke, M. J., Holmes, M. A., Leach, M. C., & Constantino-Casas, F. (2016). Development of a facial expression scale using footrot and mastitis as models of pain in sheep. *Applied Animal Behaviour Science*, *176*, 19–26.

- Mellor, D. (2017). Operational Details of the Five Domains Model and Its Key Applications to the Assessment and Management of Animal Welfare. *Animals*, 7(12), 60.
- Mendl, M. (1999). Performing under pressure: Stress and cognitive function. *Applied Animal Behaviour Science*, 65(3), 221–244.
- Mendl, M., Burman, O. H. P., Parker, R. M. A., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, 118(3–4), 161–181.
- Mendl, M., Burman, O. H. P., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, 277(1696), 2895–2904.
- Mendl, M., Neville, V., & Paul, E. S. (2022). Bridging the Gap: Human Emotions and Animal Emotions. *Affective Science*.
- Mendl, M., & Paul, E. (2004). Consciousness, emotion and animal welfare: Insights from cognitive science. *Animal Welfare*, *13*(1), 17–25.
- Mendl, M., & Paul, E. S. (2020). Animal affect and decision-making. *Neuroscience & Biobehavioral Reviews*, *112*, 144–163.
- Menke, E. S., Blom, G., van Loon, J. P. A. M., & Back, W. (2016). Pressure Algometry in Icelandic Horses: Interexaminer and Intraexaminer Reliability. *Journal of Equine Veterinary Science*, *36*, 26–31.
- Miller, R. E., & Murphy, J. V. (1964). Influence of the spatial relationships between the cue, reward, and response in discrimination learning. *Journal of Experimental Psychology*, 67(2), 120–123.
- Mills, D. S., Eckley, S., & Cooper, J. J. (2000). Thoroughbred bedding preferences, associated behaviour differences and their implications for equine welfare. *Animal Science*, *70*(1), 95–106.
- Minero, M., & Canali, E. (2009). Welfare issues of horses: An overview and practical recommendations. *Italian Journal of Animal Science*, 8(sup1), 219–230.
- Mitchell, K. J. (2019). Equine Electrocardiography. *Veterinary Clinics of North America: Equine Practice*, *35*(1), 65–83.
- Mitchell, K. J., & Schwarzwald, C. C. (2021). Heart rate variability analysis in horses for the diagnosis of arrhythmias. *The Veterinary Journal*, *268*, 105590.
- Moe, R. O., Stubsjøen, S. M., Bohlin, J., Flø, A., & Bakken, M. (2012). Peripheral temperature drop in response to anticipation and consumption of a signaled palatable reward in laying hens (Gallus domesticus). *Physiology & Behavior*, 106(4), 527–533.
- Mogg, K., & Bradley, B. P. (2016). Anxiety and attention to threat: Cognitive mechanisms and treatment with attention bias modification. *Behaviour Research and Therapy*, *87*, 76–108.
- Mogg, K., Millar, N., & Bradley, B. P. (2000). Biases in Eye Movements to Threatening Facial Expressions in Generalized Anxiety Disorder and Depressive Disorder. *Journal of Abnormal Psychology*, *109*(4), 695–704.
- Monk, J. E., Doyle, R. E., Colditz, I. G., Belson, S., Cronin, G. M., & Lee, C. (2018a). Towards a more practical attention bias test to assess affective state in sheep. *Plos One*, *13*(1), e0190404.
- Monk, J. E., Belson, S., Colditz, I. G., & Lee, C. (2018b). Attention Bias Test Differentiates Anxiety and Depression in Sheep. *Frontiers in Behavioral Neuroscience*, *12*, 246.

- Monk, J. E., Belson, S., & Lee, C. (2019). Pharmacologically-induced stress has minimal impact on judgement and attention biases in sheep. *Scientific Reports*, *9*, 11446.
- Monk, J. E., Lee, C., Dickson, E., & Campbell, D. L. M. (2020). Attention Bias Test Measures Negative But Not Positive Affect in Sheep: A Replication Study. *Animals*, *10*(8), 1314.
- Morgan, K. (1998). Thermoneutral zone and critical temperatures of horses. *Journal of Thermal Biology*, 23(1), 59–61.
- Morin, E. L., Howell, B. R., Meyer, J. S., & Sanchez, M. M. (2019). Effects of early maternal care on adolescent attention bias to threat in nonhuman primates. *Developmental Cognitive Neuroscience*, *38*, 100643.
- Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393(6684), Article 6684.
- Müller, C. A., Riemer, S., Rosam, C. M., Schößwender, J., Range, F., & Huber, L. (2012). Brief owner absence does not induce negative judgement bias in pet dogs. *Animal Cognition*, 15(5), 1031–1035.
- Murphy, J., Waldmann, T., & Arkins, S. (2004). Sex differences in equine learning skills and visuospatial ability. *Applied Animal Behaviour Science*, 87(1–2), 119–130.
- Murray, M. J., Schusser, G. F., Pipers, F. S., & Gross, S. J. (1996). Factors associated with gastric lesions in thoroughbred racehorses. *Equine Veterinary Journal*, *28*(5), 368–374.
- Nakamura, K. (2015). Neural circuit for psychological stress-induced hyperthermia. *Temperature*, 2(3), 352–361.
- Nakayama, K., Goto, S., Kuraoka, K., & Nakamura, K. (2005). Decrease in nasal temperature of rhesus monkeys (Macaca mulatta) in negative emotional state. *Physiology & Behavior*, *84*(5), 783– 790.
- Nanchen, D. (2018). Resting heart rate: What is normal? *Heart*, 104(13), 1048–1049.
- Negro, S., Bartolomé, E., Molina, A., Solé, M., Gómez, M. D., & Valera, M. (2018). Stress level effects on sport performance during trotting races in Spanish Trotter Horses. *Research in Veterinary Science*, 118, 86–90. 7
- Nettle, D., & Bateson, M. (2012). The Evolutionary Origins of Mood and Its Disorders. *Current Biology*, 22(17), R712–R721.
- Neville, V., Nakagawa, S., Zidar, J., Paul, E. S., Lagisz, M., Bateson, M., Løvlie, H., & Mendl, M. (2020). Pharmacological manipulations of judgement bias: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *108*, 269–286.
- Nicol, C. J., Caplen, G., Edgar, J., & Browne, W. J. (2009). Associations between welfare indicators and environmental choice in laying hens. *Animal Behaviour*, *78*(2), 413–424.
- Nicol, C. J., Davidson, H. P. D., Harris, P. A., Waters, A. J., & Wilson, A. D. (2002). Study of crib-biting and gastric inflammation and ulceration in young horses. *Veterinary Record*, 151(22), 658–
- Nicol, C. J., & Guilford, T. (1991). Exploratory activity as a measure of motivation in deprived hens. Animal Behaviour, 41(2), 333–341.
- Ninomiya, S., Kusunose, R., Sato, S., Terada, M., & Sugawara, K. (2004). Effects of feeding methods on eating frustration in stabled horses. *Animal Science Journal*, *75*(5), 465–469.
- Ninomiya, S., Sato, S., & Sugawara, K. (2007). Weaving in stabled horses and its relationship to other behavioural traits. *Applied Animal Behaviour Science*, *106*(1–3), 134–143. 4

- Ogle, D., Doll, J., Wheeler, P., & Dinno, A. (2022). FSA:Fisheries Stock Analysis [R package version 0.9.3].
- Oh, J., Šlipogor, V., & Fitch, W. T. (2019). Artificial visual stimuli for animal experiments: An experimental evaluation in a prey capture context with common marmosets (Callithrix jacchus). *Journal of Comparative Psychology*, *133*(1), 72–80.
- Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522.
- Okada, K., Takemura, K., & Sato, S. (2013). Investigation of Various Essential Factors for Optimum Infrared Thermography. *Journal of Veterinary Medical Science*, 75(10), 1349–1353.
- Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behavioral and Brain Sciences*, *5*(3), 407–422.
- Panksepp, J. (1994). Evolution constructed the potential for subjective experience within the neurodynamics of the mammalian brain. *In: Ekman, P., Davidson, R.J. (Eds.), The Nature of Emotion: Fundamental Questions. OUP, Oxford, pp. 396–399.*
- Panksepp, J. (2011). Toward a cross-species neuroscientific understanding of the affective mind: Do animals have emotional feelings? *American Journal of Primatology*, 73(6), 545–561.
- Parr, L. A. (2001). Cognitive and physiological markers of emotional awareness in chimpanzees (Pan troglodytes). *Animal Cognition*, 4(3), 223–229.
- Parr, L. A., Modi, M., Siebert, E., & Young, L. J. (2013). Intranasal oxytocin selectively attenuates rhesus monkeys' attention to negative facial expressions. *Psychoneuroendocrinology*, 38(9), 1748–1756. 1
- Parron, C., Call, J., & Fagot, J. (2008). Behavioural responses to photographs by pictorially naïve baboons (Papio anubis), gorillas (Gorilla gorilla) and chimpanzees (Pan troglodytes). *Behavioural Processes*, *78*(3), 351–357.
- Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology*, *52*, 629–651.
- Paul, E. S., Harding, E. J., & Mendl, M. (2005). Measuring emotional processes in animals: The utility of a cognitive approach. *Neuroscience & Biobehavioral Reviews*, *29*(3), 469–491.
- Paul, E. S., & Mendl, M. T. (2018). Animal emotion: Descriptive and prescriptive definitions and their implications for a comparative perspective. *Applied Animal Behaviour Science*, 205, 202–209.
- Paul, E. S., Sher, S., Tamietto, M., Winkielman, P., & Mendl, M. T. (2020). Towards a comparative science of emotion: Affect and consciousness in humans and animals. *Neuroscience & Biobehavioral Reviews*, 108, 749–770.
- Pavlidis, I., Levine, J., & Baukol, P. (2001). Thermal imaging for anxiety detection,. *Proceedings in International Conference on Image Processing*.
- Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety*, *27*(12), 1135–1142.
- Percie du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Karp, N. A., Lazic, S. E., Lidster, K., MacCallum, C. J., Macleod, M., ... Würbel, H. (2020). The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLOS Biology*, *18*(7), e3000410.

- Pergamin-Hight, L., Naim, R., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Bar-Haim, Y. (2015). Content specificity of attention bias to threat in anxiety disorders: A meta-analysis. *Clinical Psychology Review*, 35, 10–18.
- Peters, S. M., Bleijenberg, E. H., van Dierendonck, M. C., van der Harst, J. E., & Spruijt, B. M. (2012). Characterization of anticipatory behaviour in domesticated horses (Equus caballus). *Applied Animal Behaviour Science*, 138(1–2), 60–69.
- Pfungst, O., & Rahn, C. L. (1911). Clever Hans: The horse of Mr. Von Osten. A contribution to experimental animal and human psychology. Holt, Rineheart, and Winston.
- Phillips, C. J. C., Oevermans, H., Syrett, K. L., Jespersen, A. Y., & Pearce, G. P. (2015). Lateralization of behavior in dairy cows in response to conspecifics and novel persons. *Journal of Dairy Science*, 98(4), 2389–2400.
- Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2016). Attentional bias for positive emotional stimuli: A meta-analytic investigation. *Psychological Bulletin*, 142(1), 79–106.
- Popescu, S., & Diugan, E.-A. (2013). The Relationship Between Behavioral and Other Welfare Indicators of Working Horses. *Journal of Equine Veterinary Science*, 33(1), 1–12.
- Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology*, 74(2), 116–143.
- Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(03).
- Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25.
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.
- Powers, A., Fani, N., Murphy, L., Briscione, M., Bradley, B., Tone, E. B., Norrholm, S. D., & Jovanovic, T. (2019). Attention bias toward threatening faces in women with PTSD: Eye tracking correlates by symptom cluster. *European Journal of Psychotraumatology*, *10*(1), 1568133.
- Proctor, H. S., & Carder, G. (2015). Nasal temperatures in dairy cows are influenced by positive emotional state. *Physiology & Behavior*, *138*, 340–344.
- Proops, L., Grounds, K., Smith, A. V., & McComb, K. (2018). Animals Remember Previous Facial Expressions that Specific Humans Have Exhibited. *Current Biology*, *28*(9), 1428-1432.e4.
- Proops, L., & McComb, K. (2010). Attributing attention: The use of human-given cues by domestic horses (Equus caballus). *Animal Cognition*, *13*(2), 197–205.
- Proops, L., McComb, K., & Reby, D. (2009). Cross-modal individual recognition in domestic horses (Equus caballus). Proceedings of the National Academy of Sciences, 106(3), 947–951.
- Proops, L., Rayner, J., Taylor, A. M., & McComb, K. (2013). The Responses of Young Domestic Horses to Human-Given Cues. *Plos One*, *8*(6), e67000.
- Putman, R. J., Pratt, R. M., Ekins, J. R., & Edwards, P. J. (1987). Food and Feeding Behaviour of Cattle and Ponies in the New Forest, Hampshire. *Journal of Applied Ecology*, *24*(2), 369–380.
- Quaranta, A., Siniscalchi, M., & Vallortigara, G. (2007). Asymmetric tail-wagging responses by dogs to different emotive stimuli. *Current Biology*, *17*(6), R199–R201.
- R Core team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

- Ramirez Montes De Oca, M. A. (2021). Temperature asymmetries in facial areas as indicators of affective state in dairy calves and horses. University of Bristol[Thesis]
- Raoult, C. M. C., & Gygax, L. (2018). Valence and Intensity of Video Stimuli of Dogs and Conspecifics in Sheep: Approach-Avoidance, Operant Response, and Attention. *Animals*, 8(7), Article 7.
- Raoult, C. M. C., & Gygax, L. (2019). Mood induction alters attention toward negative-positive stimulus pairs in sheep. *Scientific Reports*, *9*(1), 7759.
- Reefmann, N., Bütikofer Kaszàs, F., Wechsler, B., & Gygax, L. (2009). Ear and tail postures as indicators of emotional valence in sheep. *Applied Animal Behaviour Science*, 118(3–4), 199– 207.
- Reefmann, N., Muehlemann, T., Wechsler, B., & Gygax, L. (2012). Housing induced mood modulates reactions to emotional stimuli in sheep. *Applied Animal Behaviour Science*, 136(2–4), 146–155.
- Ricci-Bonot, C., Romero, T., Nicol, C., & Mills, D. (2021). Social buffering in horses is influenced by context but not by the familiarity and habituation of a companion. *Scientific Reports*, *11*(1), 8862.
- Riemer, S., Assis, L., Pike, T. W., & Mills, D. S. (2016). Dynamic changes in ear temperature in relation to separation distress in dogs. *Physiology & Behavior*, *167*, 86–91.
- Rietmann, T. R., Stuart, A. E. A., Bernasconi, P., Stauffacher, M., Auer, J. A., & Weishaupt, M. A. (2004). Assessment of mental stress in warmblood horses: Heart rate variability in comparison to heart rate and selected behavioural parameters. *Applied Animal Behaviour Science*, 88(1–2), 121–136.
- Rinck, M., Reinecke, A., Ellwart, T., Heuer, K., & Becker, E. S. (2005). Speeded detection and increased distraction in fear of spiders: Evidence from eye movements. *Journal of Abnormal Psychology*, 114(2), 235–248.
- Roberts, K., Hemmings, A. J., McBride, S. D., & Parker, M. O. (2017). Causal factors of oral versus locomotor stereotypy in the horse. *Journal of Veterinary Behavior*, *20*, 37–43.
- Robin, C. A., Ireland, J. L., Wylie, C. E., Collins, S. N., Verheyen, K. L. P., & Newton, J. R. (2015).
 Prevalence of and risk factors for equine obesity in Great Britain based on owner-reported body condition scores. *Equine Veterinary Journal*, 47(2), 196–201.
- Rochais, C., Henry, S., Fureix, C., & Hausberger, M. (2016a). Investigating attentional processes in depressive-like domestic horses (Equus caballus). *Behavioural Processes*, *124*, 93–96.
- Rochais, C., Fureix, C., Lesimple, C., & Hausberger, M. (2016b). Lower attention to daily environment: A novel cue for detecting chronic horses' back pain? *Scientific Reports*, 6(1), 20117.
- Rochais, C., Henry, S., & Hausberger, M. (2017a). Spontaneous attention-capture by auditory distractors as predictor of distractibility: A study of domestic horses (Equus caballus). *Scientific Reports*, 7(1), 15283.
- Rochais, C., Sébilleau, M., Houdebine, M., Bec, P., Hausberger, M., & Henry, S. (2017b). A novel test for evaluating horses' spontaneous visual attention is predictive of attention in operant learning tasks. *The Science of Nature*, *104*(7–8), 61.
- Rochais, C., Stomp, M., Sébilleau, M., Houdebine, M., Henry, S., & Hausberger, M. (2022). Horses' attentional characteristics differ according to the type of work. *PLOS ONE*, *17*(7), e0269974.
- Roelofs, S., Boleij, H., Nordquist, R. E., & van der Staay, F. J. (2016). Making Decisions under Ambiguity: Judgment Bias Tasks for Assessing Emotional State in Animals. *Frontiers in Behavioral Neuroscience*, 10.

- Rolls, E. T. (1999). The Brain and Emotion. Oxford University Press.
- Rolls, E. T. (2005). Emotion explained. Oxford University Press.
- Rørvang, M. V., Nielsen, B. L., & McLean, A. N. (2020). Sensory Abilities of Horses and Their Importance for Equitation Science. *Frontiers in Veterinary Science*, *7*, 633.
- Rosenberger, K., Simmler, M., Langbein, J., Keil, N., & Nawroth, C. (2021). Performance of goats in a detour and a problem-solving test following long-term cognitive test exposure. *Royal Society Open Science*, 8(10), 210656.
- Rowe, C., & Healy, S. D. (2014). Measuring variation in cognition. *Behavioral Ecology*, 25(6), 1287–1292.
- Rushen, J. (1991). Problems associated with the interpretation of physiological data in the assessment of animal welfare. *Applied Animal Behaviour Science*, *28*(4), 381–386.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172.
- Salles, M. S. V., da Silva, S. C., Salles, F. A., Roma, L. C., El Faro, L., Bustos Mac Lean, P. A., Lins de Oliveira, C. E., & Martello, L. S. (2016). Mapping the body surface temperature of cattle by infrared thermography. *Journal of Thermal Biology*, 62, 63–69.
- Salzer, Y., Honig, H. H., Shaked, R., Abeles, E., Kleinjan-Elazary, A., Berger, K., Jacoby, S., Fishbain, B., & Kendler, S. (2021). Towards on-site automatic detection of noxious events in dairy cows. *Applied Animal Behaviour Science*, 236, 105260.
- Sanchez, A., Vazquez, C., Marker, C., LeMoult, J., & Joormann, J. (2013). Attentional disengagement predicts stress recovery in depression: An eye-tracking study. *Journal of Abnormal Psychology*, *122*(2), 303–313.
- Sankey, C., Richard-Yris, M.-A., Leroy, H., Henry, S., & Hausberger, M. (2010). Positive interactions lead to lasting positive memories in horses, Equus caballus. *Animal Behaviour*, 79(4), 869– 875.
- Saslow, C. A. (2002). Understanding the perceptual world of horses. *Applied Animal Behaviour Science*, *78*(2–4), 209–224.
- Sato, A., Koda, H., Lemasson, A., Nagumo, S., & Masataka, N. (2012). Visual Recognition of Age Class and Preference for Infantile Features: Implications for Species-Specific vs Universal Cognitive Traits in Primates. *PLOS ONE*, 7(5), e38387.
- Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In *Scherer, K.R., Ekman, P. (Eds.), Approaches to Emotion. Erlbaum, Hillsdale* (pp. 293–317).
- Scherer, K. R. (1999). Appraisal theory. In *Handbook of cognition and emotion* (pp. 637–663). John Wiley & Sons Ltd.
- Schmidt, A., Möstl, E., Wehnert, C., Aurich, J., Müller, J., & Aurich, C. (2010). Cortisol release and heart rate variability in horses during road transport. *Hormones and Behavior*, 57(2), 209– 215.
- Schmidt, M. J., Knemeyer, C., & Heinsen, H. (2019). Neuroanatomy of the equine brain as revealed by high-field (3Tesla) magnetic-resonance-imaging. *PLOS ONE*, *14*(4), e0213814.
- Scopa, C., Greco, A., Contalbrigo, L., Fratini, E., Lanatà, A., Scilingo, E. P., & Baragli, P. (2020). Inside the Interaction: Contact With Familiar Humans Modulates Heart Rate Variability in Horses. *Frontiers in Veterinary Science*, 7, 582759.

Sears, C. R., Thomas, C. L., LeHuquet, J. M., & Johnson, J. C. S. (2010). Attentional biases in dysphoria: An eye-tracking study of the allocation and disengagement of attention. *Cognition and Emotion*, 24(8), 1349–1368. https://doi.org/10.1080/02699930903399319

Selye, H. (1976). Stress in Health and Disease. Butterworths.

Shettleworth, Sara J. (1998). Cognition, Evolution and Behaviour.Oxford University Press.

- Shibasaki, M., & Kawai, N. (2009). Rapid detection of snakes by Japanese monkeys (Macaca fuscata): An evolutionarily predisposed visual system. *Journal of Comparative Psychology*, *123*(2), 131–135.
- Shields, R. W. J. (1993). Functional Anatomy of the Autonomic Nervous System. *Journal of Clinical Neurophysiology*, *10*(1), 2–13.
- Sigurjónsdóttir, H., & Haraldsson, H. (2019). Significance of Group Composition for the Welfare of Pastured Horses. *Animals*, *9*(1), 14.
- Silberman, E. K., & Weingartner, H. (1986). Hemispheric lateralization of functions related to emotion. Brain and Cognition, 5(3), 322–353.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, 28(9), 1059–1074.
- Singer, J. W., Bobsin, N., Bamka, W. J., & Kluchinshi, D. (1999). Horse pasture management. *Journal* of Equine Veterinary Science, 19.
- Siniscalchi, M., d'Ingeo, S., & Quaranta, A. (2021). Lateralized emotional functioning in domestic animals. *Applied Animal Behaviour Science*, 237, 105282.
- Siniscalchi, M., Sasso, R., Pepe, A. M., Vallortigara, G., & Quaranta, A. (2010). Dogs turn left to emotional stimuli. *Behavioural Brain Research*, 208(2), 516–521.
- Smith, A. V., Proops, L., Grounds, K., Wathan, J., & McComb, K. (2016). Functionally relevant responses to human facial expressions of emotion in the domestic horse (*Equus caballus*). *Biology Letters*, 12(2), 20150907.
- Sneddon, L. U., Elwood, R. W., Adamo, S. A., & Leach, M. C. (2014). Defining and assessing animal pain. *Animal Behaviour*, *97*, 201–212.
- Somppi, S., Törnqvist, H., Hänninen, L., Krause, C., & Vainio, O. (2012). Dogs do look at images: Eye tracking in canine cognition research. *Animal Cognition*, *15*(2), 163–174.
- Somppi, S., Tornqvist, H., Kujala, M. V., Hanninen, L., Krause, C. M., & Vainio, O. (2016). Dogs Evaluate Threatening Facial Expressions by Their Biological Validity—Evidence from Gazing Patterns. *Plos One*, *11*(1), e0143047.
- Søndergaard, E., Jensen, M. B., & Nicol, C. J. (2011). Motivation for social contact in horses measured by operant conditioning. *Applied Animal Behaviour Science*, *132*(3–4), 131–137.
- Spelke, E. S. (1985). Preferential-looking methods as tools for the study of cognition in infancy. In *Measurement of audition and vision in the first year of postnatal life: A methodological overview* (pp. 323–363). Ablex Publishing.
- Spruijt, B., Bos, R. van den, & Pijlman, F. (2001). A concept of welfare based on reward evaluating mechanisms in the brain: Anticipatory behaviour as an indicator for the state of reward systems. *Adaptive Behavior Collection*.
- Steimer, T. (2002). The biology of fear- and anxiety-related behaviors. *Dialogues in Clinical Neuroscience*, 4(3), 231–249.

- Steinman, S. B., & Steinman, B. A. (1998). Vision and attention. I: Current models of visual attention. Optometry and Vision Science: Official Publication of the American Academy of Optometry, 75(2), 146–155.
- Stewart, M., Stafford, K. J., Dowling, S. K., Schaefer, A. L., & Webster, J. R. (2008a). Eye temperature and heart rate variability of calves disbudded with or without local anaesthetic. *Physiology & Behavior*, 93(4–5), 789–797.
- Stewart, M., Schaefer, A., Haley, D., Colyn, J., Cook, N., Stafford, K., & Webster, J. (2008b). Infrared thermography as a non-invasive method for detecting fear-related responses of cattle to handling procedures. *Animal Welfare*, 7.
- Stewart, M., Verkerk, G. A., Stafford, K. J., Schaefer, A. L., & Webster, J. R. (2010a). Noninvasive assessment of autonomic activity for evaluation of pain in calves, using surgical castration as a model. *Journal of Dairy Science*, 93(8), 3602–3609.
- Stewart, M., Webster, J. R., Stafford, K. J., Schaefer, A. L., & Verkerk, G. A. (2010b). Technical note: Effects of an epinephrine infusion on eye temperature and heart rate variability in bull calves. *Journal of Dairy Science*, *93*(11), 5252–5257.
- Stewart, M., Webster, J., Schaefer, A., Cook, N., & Scott, S. (2005). Infrared thermography as a noninvasive tool to study animal welfare. *Animal Welfare*, 7.
- Stoehr, A. M. (1999). Are significance thresholds appropriate for the study of animal behaviour? Animal Behaviour, 57(5), F22–F25.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.
- Tamioso, P. R., Maiolino Molento, C. F., Boivin, X., Chandèze, H., Andanson, S., Delval, É., Hazard, D., da Silva, G. P., Taconeli, C. A., & Boissy, A. (2018). Inducing positive emotions: Behavioural and cardiac responses to human and brushing in ewes selected for high vs low social reactivity. *Applied Animal Behaviour Science*, 208, 56–65.
- Taskforce of the European Society of Cardiology the North American Society Society of Pacing Electrophysiology. (1996). Heart Rate Variability Standard of Measurement, Physiological Interpretation, and Clinical Use. *Eur Heart Journal*, *17*, 354–381.
- Tatemoto, P., Broom, D. M., & Zanella, A. J. (2022). Changes in Stereotypies: Effects over Time and over Generations. *Animals*, *12*(19), Article 19.
- Tattersall, G. J. (2016). Infrared thermography: A non-invasive window into thermal physiology. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 202,* 78–98.
- Taylor, P. M., Crosignani, N., Lopes, C., Rosa, A. C., Luna, S. P., & Filho, J. N. P. (2016). Mechanical nociceptive thresholds using four probe configurations in horses. *Veterinary Anaesthesia and Analgesia*, 43(1), 99–108.
- Telkanranta, H., Paul, E., Ramirez Montes De Oca, M. A., Caplan, G., Held, S., Whay, H. R., & Mendl, M. (2019, April 7). Lateral Differences in Eye and Nostril Temperature: A Novel Approach to Developing Thermographic Methods to Measure Emotions. *Advancing Animal Welfare Science: How Do We Get There? Who Is It Good For?* UFAW Internation Animal Welfare Science Symposium, Bruges, Belgium.
- Thayer, J. F., & Brosschot, J. F. (2005). Psychosomatics and psychopathology: Looking up and down from the brain. *Psychoneuroendocrinology*, *30*(10), 1050–1058.

- Thayer, J. F., Hahn, A. W., Pearson, M. A., Sollers, J. J., Johnson, P. J., & Loch, W. E. (1997). Heart rate variability during exercise in the horse. *Biomedical Sciences Instrumentation*, *34*, 246–251.
- Timney, B., & Macuda, T. (2001). Vision and hearing in horses. *Journal of the American Veterinary Medical Association*, 218(10), 1567–1574.
- Tomonaga, M., Kumazaki, K., Camus, F., Nicod, S., Pereira, C., & Matsuzawa, T. (2015). A horse's eye view: Size and shape discrimination compared with other mammals. *Biology Letters*, 11(11), 20150701.
- Travain, T., Colombo, E. S., Grandi, L. C., Heinzl, E., Pelosi, A., Prato Previde, E., & Valsecchi, P. (2016).
 How good is this food? A study on dogs' emotional responses to a potentially pleasant event using infrared thermography. *Physiology & Behavior*, 159, 80–87.
- Travain, T., & Valsecchi, P. (2021). Infrared Thermography in the Study of Animals' Emotional Responses: A Critical Review. *Animals*, *11*(9), 2510.
- Trevarthen, A. C., Kappel, S., Roberts, C., Finnegan, E. M., Paul, E. S., Planas-Sitjà, I., Mendl, M. T., & Fureix, C. (2019). Measuring affect-related cognitive bias: Do mice in opposite affective states react differently to negative and positive stimuli? *PLOS ONE*, *14*(12), e0226438.
- Trösch, Cuzol, Parias, Calandreau, Nowak, & Lansade. (2019). Horses Categorize Human Emotions Cross-Modally Based on Facial Expression and Non-Verbal Vocalizations. *Animals*, 9(11), 862.
- Trösch, M., Pellon, S., Cuzol, F., Parias, C., Nowak, R., Calandreau, L., & Lansade, L. (2020). Horses feel emotions when they watch positive and negative horse–human interactions in a video and transpose what they saw to real life. *Animal Cognition*, *23*(4), 643–653.
- Tuyttens, F. A. M., de Graaf, S., Heerkens, J. L. T., Jacobs, L., Nalon, E., Ott, S., Stadig, L., Van Laer, E., & Ampe, B. (2014). Observer bias in animal behaviour research: Can we believe what we score, if we score what we believe? *Animal Behaviour*, 90, 273–280. https://doi.org/10.1016/j.anbehav.2014.02.007
- Tyler, S. J. (1972). The Behaviour and Social Organization of the New Forest Ponies. *Animal Behaviour Monographs*, *5*, 87–196.
- Tyrrell, L. P., Butler, S. R., Yorzinski, J. L., & Fernández-Juricic, E. (2014). A novel system for bi-ocular eye-tracking in vertebrates with laterally placed eyes. *Methods in Ecology and Evolution*, *5*(10), 1070–1077.
- Valenchon, M., Lévy, F., Prunier, A., Moussu, C., Calandreau, L., & Lansade, L. (2013). Stress Modulates Instrumental Learning Performances in Horses (Equus caballus) in Interaction with Temperament. *PLOS ONE*, 8(4), e62324.
- Valera, M., Bartolomé, E., Sánchez, M. J., Molina, A., Cook, N., & Schaefer, A. (2012). Changes in Eye Temperature and Stress Assessment in Horses During Show Jumping Competitions. *Journal* of Equine Veterinary Science, 32(12), 827–830.
- Vallortigara, G., Cozzutti, C., Tommasi, L., & Rogers, L. J. (2001). How birds use their eyes: Opposite left-right specialization for the lateral and frontal visual hemifield in the domestic chick. *Current Biology*, *11*, 29–33.
- Van Dierendonck, M. C. (2006). The importantce of social relationships in horses. Utrecht University. [Thesis]
- van Loon, J. P. A. M., & Van Dierendonck, M. C. (2015). Monitoring acute equine visceral pain with the Equine Utrecht University Scale for Composite Pain Assessment (EQUUS-COMPASS) and the Equine Utrecht University Scale for Facial Assessment of Pain (EQUUS-FAP): A scaleconstruction study. *Veterinary Journal (London, England: 1997), 206*(3), 356–364.

- van Rooijen, R., Ploeger, A., & Kret, M. E. (2017). The dot-probe task to measure emotional attention: A suitable measure in comparative studies? *Psychonomic Bulletin & Review*, 24(6), 1686–1717.
- Vandenheede, M., & Bouissou, M. F. (1994). Fear reactions of ewes to photographic images. *Behavioural Processes*, 32(1), 17–28.
- Varcoe-Cocks, K., Sagar, K. N., Jeffcott, L. B., & Mcgowan, C. M. (2006). Pressure algometry to quantify muscle pain in racehorses with suspected sacroiliac dysfunction. *Equine Veterinary Journal*, 38(6), 558–562.
- Veissier, I., & Boissy, A. (2007). Stress and welfare: Two complementary concepts that are intrinsically related to the animal's point of view. *Physiology & Behavior*, *92*(3), 429–433.
- Verbeek, E., Colditz, I., Blache, D., & Lee, C. (2019). Chronic stress influences attentional and judgement bias and the activity of the HPA axis in sheep. *Plos One*, *14*(1), e0211363.
- Verbeek, E., Ferguson, D., & Lee, C. (2014). Are hungry sheep more pessimistic? The effects of food restriction on cognitive bias and the involvement of ghrelin in its regulation. *Physiology & Behavior*, *123*, 67–75.
- Versace, E., Morgante, M., Pulina, G., & Vallortigara, G. (2007). Behavioural lateralization in sheep (Ovis aries). *Behavioural Brain Research*, 184(1), 72–80.
- Vianna, D. M. L., & Carrive, P. (2005). Changes in cutaneous and body temperature during and after conditioned fear to context in the rat. *European Journal of Neuroscience*, *21*(9), 2505–2512.
- Visser, E. K., Ellis, A. D., & Van Reenen, C. G. (2008). The effect of two different housing conditions on the welfare of young horses stabled for the first time. *Applied Animal Behaviour Science*, 114(3–4), 521–533.
- Visser, E., Vanreenen, C., Vanderwerf, J., Schilder, M., Knaap, J., Barneveld, A., & Blokhuis, H. (2002). Heart rate and heart rate variability during a novel object test and a handling test in young horses. *Physiology & Behavior*, *76*(2), 289–296.
- Vögeli, S., Wolf, M., Wechsler, B., & Gygax, L. (2015). Housing conditions influence cortical and behavioural reactions of sheep in response to videos showing social interactions of different valence. *Behavioural Brain Research*, 284, 69–76.
- von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant-Forde, J., Marchant-Forde, R., Minero, M., Mohr, E., Prunier, A., Valance, D., & Veissier, I. (2007). Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals—A review. *Physiology & Behavior*, *92*(3), 293–316.
- Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist*, 47(3), 373–388.
- Wang, H.-M., & Huang, S.-C. (2012). SDNN/RMSSD as a Surrogate for LF/HF: A Revised Investigation. *Modelling and Simulation in Engineering*, 2012, 1–8.
- Waran, N. (Ed.). (2007). The Welfare of Horses. Springer.
- Waring, G. H. (2003). Horse Behaviour (2nd ed.). Noyes Publications.
- Wascher, C. A. F. (2021). Heart rate as a measure of emotional arousal in evolutionary biology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1831), 20200479.
- Wathan, J., Proops, L., Grounds, K., & McComb, K. (2016). Horses discriminate between facial expressions of conspecifics. *Scientific Reports*, *6*(1), 38322.

- Watkins, P. C., Mathews, A., Williamson, D. A., & Fuller, R. D. (1992). Mood-congruent memory in depression: Emotional priming or elaboration? *Journal of Abnormal Psychology*, 101(3), 581– 586.
- Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology*, 77(1), 97–108.
- Weary, D. M., Droege, P., & Braithwaite, V. A. (2017). Chapter Two Behavioral Evidence of Felt Emotions: Approaches, Inferences, and Refinements. In M. Naguib, J. Podos, L. W. Simmons, L. Barrett, S. D. Healy, & M. Zuk (Eds.), *Advances in the Study of Behavior* (Vol. 49, pp. 27– 48). Academic Press.
- Webb, L. E., Veenhoven, R., Harfeld, J. L., & Jensen, M. B. (2019). What is animal happiness? *Annals of the New York Academy of Sciences*, 1438(1), 62–76.
- Weisman, R., & Spetch, M. (2010). Determining When Birds Perceive Correspondence Between Pictures and Objects: A Critique. *Comparative Cognition & Behavior Reviews*, *5*, 117–131.
- Weiss, J. M. (1972). Psychological factors in stress and disease. Scientific American, 226, 104–113.
- Westermann, S., Buchner, H. H. F., Schramel, J. P., Tichy, A., & Stanek, C. (2013). Effects of infrared camera angle and distance on measurement and reproducibility of thermographically determined temperatures of the distolateral aspects of the forelimbs in horses. *Journal of the American Veterinary Medical Association*, 242(3), 388–395.
- Whittaker, A. L., & Barker, T. H. (2020). A consideration of the role of biology and test design as confounding factors in judgement bias tests. *Applied Animal Behaviour Science*, 232.
- Wickens, C. L., & Heleski, C. R. (2010). Crib-biting behavior in horses: A review. Applied Animal Behaviour Science, 128(1–4), 1–9.
- Williams, D. C., Aleman, M., Holliday, T. A., Fletcher, D. J., Tharp, B., Kass, P. H., Steffey, E. P., & LeCouteur, R. A. (2008). Qualitative and Quantitative Characteristics of the Electroencephalogram in Normal Horses during Spontaneous Drowsiness and Sleep. *Journal* of Veterinary Internal Medicine, 22(3), 630–638.
- Williams, J. L., Friend, T. H., Nevill, C. H., & Archer, G. (2004). The efficacy of a secondary reinforcer (clicker) during acquisition and extinction of an operant task in horses. *Applied Animal Behaviour Science*, 88(3–4), 331–341.
- Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*(1), 3–24.
- Winters, S., Dubuc, C., & Higham, J. P. (2015). Perspectives: The Looking Time Experimental Paradigm in Studies of Animal Visual Perception and Cognition. *Ethology*, *121*(7), 625–640.
- Wiśniewska, M., Janczarek, I., Wilk, I., & Wnuk-Pawlak, E. (2019). Use of Music Therapy in Aiding the Relaxation of Geriatric Horses. *Journal of Equine Veterinary Science*, *78*, 89–93.
- Wolfe, J. M. (1999). Inattentional amnesia. In *Fleeting memories: Cognition of brief visual stimuli* (pp. 71–94). The MIT Press.
- Wolter, R., Stefanski, V., & Krueger, K. (2018). Parameters for the Analysis of Social Bonds in Horses. *Animals*, 8(11), 191.
- Wongsaengchan, C., McCafferty, D. J., Evans, N. P., McKeegan, D. E. F., & Nager, R. G. (2022). Body Surface Temperature of Rats Reveals Both Magnitude and Sex Differences in the Acute Stress Response. SSRN Electronic Journal.

- Wylie, C. E., Ireland, J. L., Collins, S. N., Verheyen, K. L. P., & Newton, J. R. (2013). Demographics and management practices of horses and ponies in Great Britain: A cross-sectional study. *Research in Veterinary Science*, *95*(2), 410–417.
- Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology. Human Perception and Performance*, 25(3), 661–676.
- Yeates, J. W., & Main, D. C. J. (2008). Assessment of positive welfare: A review. *The Veterinary Journal*, 175(3), 293–300.
- Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition & Emotion*, 24(1), 3–47.
- Yorzinski, J. L., Patricelli, G. L., Babcock, J. S., Pearson, J. M., & Platt, M. L. (2013). Through their eyes: Selective attention in peahens during courtship. *Journal of Experimental Biology*, *216*(22), 4310–4310.
- Yuan, J., Tian, Y., Huang, X., Fan, H., & Wei, X. (2019). Emotional bias varies with stimulus type, arousal and task setting: Meta-analytic evidences. *Neuroscience & Biobehavioral Reviews*, 107, 461–472.
- Zebunke, M., Langbein, J., Manteuffel, G., & Puppe, B. (2011). Autonomic reactions indicating positive affect during acoustic reward learning in domestic pigs. *Animal Behaviour*, *81*(2), 481–489.
- Ziemssen, T., & Siepmann, T. (2019). The Investigation of the Cardiovascular and Sudomotor Autonomic Nervous System—A Review. *Frontiers in Neurology*, *10*, 53.
- Zupan, M., Buskas, J., Altimiras, J., & Keeling, L. J. (2016). Assessing positive emotional states in dogs using heart rate and heart rate variability. *Physiology & Behavior*, *155*, 102–111.
Appendix A (Chapter 2)

Table A1 Overview of publications investigating AB in animals. NegAM Negative affect manipulation, PosAM Positive affect manipulation, NA no affect manipulation applied

Publication	Species	Stimuli	Affect manipulation	Findings				
	(Final sample		(individual modulators)					
	Dot-probe task							
Cassidy et	Long-tailed	negative (threat) vs.	At baseline, AB to threat for 100 ms, but not 1000 ms					
al., (2021)	macaques	neutral conspecific faces		presentation duration;				
	(N=8)			AB away from threat one day following negative affect manipulation (anesthia).				
Koda et al.,	Japanese	infant vs. adult faces	NA	No evidence of AB toward newborn faces				
(2013)	macaques (N=2)							
Kret et al.,	Bonobos	Negative, positive and	NA	AB toward positive but not negative images;				
(2016)	(N=4)	neutral social scenes	(post hoc analysis of	One monkey exhibited significantly more frustration-related				
			behaviours expressed	behaviours (nose wipe) on congruent trials (i.e. probe location				
			during testing)	same as emotional stimulus location)				
Leinwand et	Chimpanzees,	Unfamiliar and familiar	NA	AB toward unfamiliar neutral human faces compared				
al., (2022)	Gorillas	human faces in neutral or		with familiar neutral human faces with and without a surgical				
	(N=12)	surprised expression; or		face mask;				
		masked		no significant AB with (un)familiar surprised expression human				
				faces				
King et al.,	Rhesus	Negative, neural or	Increased arousal via	AB to negative paired neutral faces, but not non-social stimuli at				
(2012)	macaques	positive conspecific faces	testosterone administration	baseline;				
	(N=6)	& non-social stimuli (e.g.		exogenous testosterone did not enhance AB to negative social				
		food, capture glove)		threat				
Lareuse et	Rhesus	Negative, positive or	NA	AB towards threat faces but not non-social stimuli;				
al., (2013)	macaques	neutral conspecific faces		No evidence of AB to positive stimuli				
	(N=6)	& non-face images						

Publication	Species (Final sample	Stimuli	Affect manipulation (individual modulators)	Findings
	size)			
Morin et al.,	Rhesus	social + non-social	NegAM (maltreated infants	Maltreated animals showed AB to social threat
(2019)	macaques		vs. infacts receiving	than animals with competent maternal care;
	(N=25)		competent maternal care)	exposure to higher levels of prenatal cortisol linked to AB away
				from threat in maltreated and control animals
Parr et al.,	Rhesus	neutral conspecific faces	NA	OT reduced AB to negative facial expression, but not neutral or
(2013)	monkeys	with direct or averted	(Enhance social motivation	non-social images
	(N=6)	gaze, facial expressions &	via	
		clip art images	Intra-nasal oxytocin (OT)	
			adminstratio)	
			Emotional-Stroop task	
Allritz et al.,	Chimpanzees	Negative, neutral or	NegAM	Animals performed worse with negative stimuli (vet image)
(2016)	(N=6)	positive social (human)	(anaesthesia)	compared to control trials (object images); no significant
		images & non-social		difference between vet and neutral (stranger image) or positive
		images		(care taker image) trials
Hopper et	Chimpanzees,	Negative, neutral or	NA	More errors were made with positive (food) or negative (snakes)
al., (2021)	Gorillas,	positive non-social		images as stimuli compared to neutral (objects) images
	Japanese	images		
	macaques			
	(N=19)			
			Response slowing task	
Bethell et	Rhesus	Target stimulus (grey	NegAM	Following NegAM, monkeys were slower to touch target stimuli
al., (2016a)	macaques	square) embedded	(vet check)	when these embedded direct gaze (mild threat) face images
	(N=10)	distractor stimuli of		relatve to baseline (no distractor embedded in target stimulus)
		conspecific face with		
		direct or averted gaze		
Bethell et	Rhesus	Target stimulus (grey	NA	More fearful monkeys were less likely to touch target stimulus
al., (2016b)	macaques	square) embedded	(AB in relation to	embedding direct-gaze faces (threat stimulus) than
	(17)	distractor stimuli of	fearfulness temperament)	

Publication	Species	Stimuli	Affect manipulation	Findings
	(Final sample		(individual modulators)	
	size)			
		negative conspecific face		control trials (no distractor stimuli), i.e. showed greater
		with direct or averted		behavioral inhibition to threat
		gaze; image of mask		
		previous seen in fear test		
		& object images		
Cronin et al.,	Chimpanzees,	Target stimulus (grey	NegAM	Macaquesshowed significant increase in latency to touch images
(2018)	Gorillas,	square) embedded	Unpredictable bouts of loud	of conspecific faces relative to control (grey square) during the
	Japanese	distractor stimuli of	noise overhead (flying jets)	days of the loud event;
	macaques	negative conspecific face		no significant effect of noice event on gorialls and chimpanzees
	(N=8)	with direct (threatening)		task performance
		or averted (non-		
		threatening) gaze		
Howarth et	Rhesus	Threat and neutral	NA	AB towards threat
al., (2021)	macaques	conspecific face images		
	(N=7)			
			Visual search task	
Shibasaki &	Japanese	Threat-relevant (e.g.	NA	AB to threat evident in significantly faster detection of threat-
Kawai,	macaques	snake images) and threat-		relevant stimuli than the detection of threat-irrelevant stimuli
(2009)	(N=3)	irrelevant (e.g. flower		among threat stimuli
		images) stimuli		
Marzouki et	Baboons	Conditioned symbol ('T')	NA	Attention biased by underlying mood as RT to detect target >
al., (2014)	(N=6)	as rewarded stimulus	(Affect inferred from	following negatively valenced behaviors than RT following neutral
		shown alongside	expressed positive/negative	and positively valenced behavior
		distractor symbol ('L')	social behaviours)	
		Preferentia	al looking time task (dual stimu	Ilus presentation)
Bethell et	Rhesus	Negative and neutral	NegAM	Initially AB to threat;
al., (2012)	macaques	conspecific face images	(vet check)	NegAM induced AB away from threat
	(N=7)		PosAM	
			(Enriched housing)	

Publication	Species (Final sample	Stimuli	Affect manipulation (individual modulators)	Findings
	size)		(
Bethell et	Rhesus	Negative and neutral	NA	Attention for social information and AB to social threat are
al., (2016)	macaques	conspecific face pairs	(Maternal effects and	heritable and shaped by maternal environmental effects
[preprint]	(N=67)		heritability)	
Raoult &	Sheep	Videos of threat (dog)	NA	No evidence of AB based or looking time behaviours was found
Gygax.	(N=28)	and positive		
(2018)	() /	(conspecifics) stimuli		
Lewis et al.,	Bonobos,	Familiar and unfamiliar	NA	AB to familiar rather than unfamiliar face images when stimulus
(2021)	Chimpanzees	female and male		animal belonged to the sex usually being dominant in that
	(N=29)	conspecific face images		species; no AB to familiar or unfamiliar subordinate conspecific
				faces
Mandalaywa	Rhesus	Negative and neutral	NA	No AB in 3-month old infants;
la et al.,	macaques	conspecific face images	(Maternal rank and	9-months old animals of high ranking and highly protective
(2014)	(N=91)		protective of infant test	mothers showed greater vigilance to threat
			animals)	
		Visual at	tention test using single stimul	us presentation
Brilot et al.,	Starlings	Visual stimuli mimicking	Anxiety induced by	No evidence that white noise and alarm calls generated anxiety,
(2009)	(N=32)	eye spots; conspecific	conspecifics calls	and that eyespots were aversive
		calls and white noise to		
		induce anxiety		
Franchi et	Cattle	Novel object presented	Feed restriction (dry-off	No evidence of AB to feed-related cues in feed restricted cows
al., (2020)	(N=24)	near food source	period) inducing increased	
			food motivation	
Kremer et	Cattle	Threat stimulus (dog	NegAM & PosAM through	Against predictions, cows in the negative conditions attended to
al., (2021)	(N=16)	model) placed within	housing condition	the
		runway to positive	(in interaction with	threat later than cows in the positive or baseline conditions;
		stimulus (food bucket)	personality (i.e. fearfullness,	
			socialness)	

Publication	Species	Stimuli	Affect manipulation	Findings	
	(Final sample		(individual modulators)		
	size)				
				Non-fearful non-social heifers spent more time eating and less	
				time attending to the threat than fearful social, fearful non-social	
				and fearful social animals	
Lee et al.,	Sheep	Threat (live dog) in	NegAM	Increased attention towards the threat	
(2017)	(N=60)	presence of food	Pharmacologically induced	in m-CCP sheep compared to control animals and those receiving	
			states of increased anxiety	drug to reduced anxiety;	
			(m-CCP) ⁴ or reduced anxiety	increased vigilance in anxious compared to less-anxious sheep	
			(diazepam)		
Lee et al.,	Cattle	Threat (live dog) in	NegAM	No significant differences in attention duration to threat between	
(2018)	(N=36)	presence of food	Pharmacologically induced	treatment groups; m-CCP steers looked longer towards the	
			states of increased anxiety	direction of the threat and showed higher vigilance compared	
			(m-CCP) or reduced anxiety	with control and diazepam animals	
			(diazepam)		
Lou et al.,	Pigs	Threat stimulus (sudden	NegAM & PosAM	Current but not early life experience influence AB to threat whi	
(2019)	(N=128)	noise, ovement and	Via barren/enriched	dog was present;	
		flashing light) in	housing condition during	No significant effect of early housing or coping style effects on	
		presence of food	early and current life	time spent attending to the threat location (following dog	
			experience	removal), vigilance, and eating	
			(in interaction with	following removal of threat	
			personality(i.e. coping style)		
McBride &	Sheep	Positively symbols	NA	both high and low levels of fearfulness negatively effect visual	
Morton,	(N=47)	signalling reward	(breed differences in	attention (i.e. increase likelihood to ake erros in visual	
(2018)			fearfulness)	discrimination task)	
Monk et al.,	Sheep	Threat (live dog) in	NegAM	No significant differences in attention to threat between	
(2018a)	(N=60)	presence of food		treatment groups;	

⁴ 1-methy-chlorophenylpiperazine

Publication	Species	Stimuli	Affect manipulation	Findings
	(Final sample		(individual modulators)	
	size)			
			Pharmacologically induced	enhanced vigilance and increased latency to eat in m-CCP sheep
			states of increased anxiety	
			(m-CCP) or reduced anxiety	
			(diazepam); control sheep	
			(saline threatment)	
Monk et al.,	Sheep	Threat (live dog) & image	NegAM	Depressed sheep attended significantly more to the threat and
(2018b)	(N=39)	of conspecific	Pharmacologically induced	less toward the image than control animals;
			states of increased anxiety	Anxious sheep showed AB away from threat as they paid more
			(m-CCP) or	attention toward the image and less
			Depression ⁵ ;	toward the threat than control animals;
			control sheep (saline	Anxious and depressed sheep
			threatment)	were more vigilant than controls
Monk et al.,	Sheep	Threat (live dog) & image	NegAM	No significant differences between treatment groups for duration
(2019)	(N=80)	of conspecific	Pharmacologically induced	of vigilance or looking behaviors;
			states of increased anxiety	Anxious sheep tended to be more vigilant than control animals
			(m-CCP) or Depression ² &	
			PosAM (morphine)	
			control sheep (saline	
			threatment)	
Monk et al.,	Sheep	Threat (live dog) & image	NegAM	No significant effect of treatment on looking behaviors;
(2020)	(N=80)	of conspecific	Pharmacological treatments	Anxious were more vigilant than control animals;
			to induce anxious (m-CCP)	no signifincant difference in vigilance in calm or happy sheep
			&	compared to control sheep
			PosAM	
			Calm (diazepam) , happy	
			(morphine); control (saline	
			threatment)	

⁵ induced with para-chlorophenylalanine

Publication	Species	Stimuli	Affect manipulation	Findings
	(Final sample		(individual modulators)	
	size)			
Verbeek et	Sheep	Positive stimulus (i.e.	NegAM via	LF sheep showed AB to food-related cues (i.e. attented to food
al., (2014)	(N=41)	food bucket)	low feeding level (LF);	bucket more and spend longer time interacting with the bucket)
			PosAM	compared to HF animals
			high feeding level (HF)	
Verbeek et	Sheep	Threat (live dog) in	NegAM via long-term	the chronic stress group showed reduced vigilance behaviour
al., (2019)	(N=60)	presence of food	environmental challenge	towards threat and a faster approach to
			(chronic stress)	the food compared to control animals
Rochais et	Horses	Novel stimulus (laser	NA	Riding school horses exhibited longer and less fragment
al., (2022)	(N=60)	beam)	(interaction with work type	sequences of attention to laster beam than horses used for sport
			and individual	competitions living in the same conditions
			characteristics)	
		Γ	Distractibility tests	
Rochais et	Horses	social & non-social	NA	
al., (2017a)	(N=12)	auditory cues as		
		distactors		
Trevarthen	Mice	Negative (flashing light)	NegAM	Tail-handled mice tended to approach both stimuli faster than
et al., (2019)	Mus musculus	and positive (food)	via tail handling and	tunel handlined mice but no significant difference in response to
	(N=62)	stimuli as distractors	comparatively less negative	negative and positive stimuli
			tunnel handling	
			Free viewing tests	
Luna et al.,	Pigs	Social	NA	Pigs showed AB toward dominant rather than subordinate
(2021)	(N=40)		(Social rank)	conspecifics while these were gently handled
Somppi et	Dogs	Negative, neutral and	NA	AB to threatening conspecifics' faces evoked;
al., (2016)†	(N=31)	positive conspecific face		AB away from threatening
		and human face images		human faces

Publication	Species	Stimuli	Affect manipulation	Findings	
	(Final sample		(individual modulators)		
	size)				
Somppi et	Dogs	Conspecific and human	NA	Preference of conspecific faces images over other images; longer	
al., (2012)†	(N=6)	face images, object		attention to familiar than novel stimuli regardless of	
		images		the category	
Wathan et	Horses	Images on conspecific	NA	Horses looked and also approached the positive attention stimuli	
al. , (2016)†	(N=48)	showing positive,		significantly more often than the agonistic stimuli; no significant	
		negative or neutral facial		differences in attention responses to positive vs. neutral images	
		expressions			
			Approach-avoidance based t	ests	
Cussen et	captive		NA	More neurotic parrots showed greater AB for environmental	
al., (2014)	Psittacines	foraging task based on	(personality traits of	Stimuli;	
	Amazona	visual food-related cues	'neuroticism' and	Negative relationship between attention bias	
	amazonica		'extraversion')	and performance on the foraging task	
	(N=12)				
Raoult &	Sheep	Videos of negative (dog)	NA	No evidence of AB based on looking time behaviours was found	
Gygax,	(N=24)	and positive (conspecific)			
(2018)		stimuli			
Boggiani et	Tufted	Images of familiar	NegAM	Monkeys approached the reward near the threatening stimulu	
al. , (2018)	capuchin	threatening stimuli	Receiving social aggression	faster and more frequently in the post-aggression condition	
	monkeys	(capture glove, net and		compared to non-aggression condition	
	(N=15)	handler); neutral stimulus			
		(familiar human)			
			Other test designs		
Huskisson et	Japanese	Match-to-sample task	NegAM	macaques were significantly more accurate in the presence of	
al., (2021)	macaques	with	Effect of zoo visitor number	small crowds	
	(N=14)		on cognitive performance	than medium or large crowds	
Bellegarde	Sheep	Visual discrimination test	NA	Sheep associate a negative image with a food reward	
et al.,	(N=16)	using conspecific faces		learned faster than animals associating a neutral image with a	
(2017a)†				reward	

Appendix A

Publication	Species	Stimuli	Affect manipulation	Findings	
	(Final sample		(individual modulators)		
	size)				
Bellegarde	Goats	Visual discrimination test	NA	Goats spent more time with their ears forward when the negative	
et al.,	Capra hircus	using conspecific faces		was shown compared	
(2017b)†	(N=32)			to the positive suggesting greater interest in negative faces	
Rochais et	Horses	Distraction test with	NA	Non-withdrawn horses	
al., (2016)†	(N=103)	Novel auditory stimuli	(inferred from housing	Attended to sounds longer than withdrawn horses on the first	
			condition and horse	day of sound presentation; habituation effects in non-withdrawn	
			behaviour)	in response to sounds, no change in behaviour in non-withdrawn	
				horses	
Rochais et	Horses	Visual attention test with	NA	Attentional characteristics shown in response to stimulus	
al., (2017a)†	(N=12)	a novel stimulus (laser	(interaction with learning	consistent across time and different task requiring attention	
		beam)	ability and attention in		
			cognitive task)		

⁺ these studies measured attention to emotional stimuli, but did not address AB

Appendix B (Chapter 3)

Table B1 Scan sampling ethogram.

Behaviour	Description
Normal horses behave	iours
Feeding hay	The horse digests hay.
Feeding concentrate	The horse digests concentrate.
Feeding grass	The horse is grazing.
Feeding other	The horse digests other feed (e.g. food toy, brambles, tree bark)
Drinking	The horse is drinking water.
Attentive	All four feet are on the ground and weight-bearing is evenly distributed, eyes are open, blinking, ears are forward, the head/neck is carried in relaxed posture (not lower than chest and not raised above withers).
Resting	The horse is usually supported by only 3 legs while one hind leg cocked. The neck is sloped and rounder than when attentive and alert; ears are rotate globally (together) forward, laterally or backwards, the eyelids and lips are droopy and the eyes partially of fully closed. In some case, the crest of the neck may drop 20° or more below horizontal, with
	the dorsal surface of the head sometimes reaching vertical. [Fureix et al., 2012; Waring 2003]
Alert	The horses' head is carried high, eyes wide open, ears forward, nostrils dilated, tail possibly raised (head looking out of stall or with head/body in stall)
Withdrawn	The horses is standing with eyes open, stretched neck (open jaw-neck angle) and similar height between neck and back [Fureix et al., 2012] for at least 15 s.
Lying lateral	Recumbent lying position, either lateral thoracic area parallel to and in contact with the ground, head immobile and in contact with the ground, or legs extended; facial and skeletal musculature relaxes [Hartman et al., 2019;Waring 2003]
Lying sternal	Recumbent lying position, with sternum in contact with the ground, legs folded beneath the body,
Rolling	The horse is lying on ground and rolling over its back from side to side.
Vocalising	The horse emits whinnies, nickers, squeals or snorts.
Exploring	The horse is sniffing surroundings (e.g. bedding, walls, stall furniture) with its nose.
Walking	The horse is walking stepping forward in stall - not repetitively (e.g. horse is walking towards door in sight of person) or walking in field for more 3 steps.
Maintenance	The horse is defecating, urinating, scratching or self-grooming.
behaviours	
Social positive	The horse is in affiliative interaction with another horse (e.g. grooming, playing).
Behaviours used as sig	gns of performance of abnormal (repetitive) behaviour
Licking	Repetitive licking of the same object in its environment (except feeding-associated
	items e.g. trough or bucket) for more than 3 times.
Biting	Repetitive biting of the same object in its
	Environment (except the trough) for more than 3 times in a row ().
Head tossing	Repeated movement of the head (e.g. up and down) for more than 3 times repetitively.

Appendix B

Description
The horse is expresses threats aimlessly (e.g. kicking, biting) alone in its box,
observed at least three times scan time (threats not aimed at observer).
The mouth is open with a lateral movement of its neck observed at least 3 times
repeatedly.
Horses rubs, grinds or chatters its teeth more than 3 times repeatedly. The animals
claps its lips or moves its tongue for more than 3 times repeatedly independently
from feeding.
The horse bits pieces of wood out of stall wall/door.
The horse consumes wood shavings.
The horse bits into edge of door/through with the incisor teeth and bolus of air is
swallowed down to stomach while making a grunt noise, observed at least once at
scan.
The horse nods its head in preparation of swallowing air in which head is jerked up
and the horse opens its mouth, air forcefully swallowed down, grunting noise
occurs, observed at least once at scan.
The horse is swaying its head, neck and/or forelimbs from side to side, observed at
three times in a row (waving from one leg to the other and back) at scan.
Repetitive movement for walking in stall; when appears to be present, the horse is
observed for 20s to assess repetitiveness of behaviour (if stops under 20 s $ ightarrow$ scored
as walking).
Persistent pawing in stall (independent from food delivery), for more than 10s.
The horse is pulling its ears back and threatening the observer, teeth possibly
visible, head possibly moving towards the observer.
If any behaviour is observed not specific by the above, list as other and describe
behaviour.

Appendix C (Chapter 4)

Table C1 Images of the four stimuli horses used in the AB test. Following Wathan et al., 2016, I photographed the test stimulus horses in three different situations (i.e. positive – food anticipation, negative – antagonistic interaction with conspecific with conspecific walking past the horses' stall, or comparative more neutral – at rest in stall).

Situation	Negative stimulus	Neutral stimulus	Positive stimulus
Horse A			
Horse B			
Horse C			
Horse D			

Appendix D (Chapter 5)

Table D1 Overview of individual performance during ODT and ORT. Number of trial blocks horses needed to reach learning criterion (LC) during training (i.e. learning ability, LA), and number of pre-test and pre-screen tests horses received before advancing to the object recognition test with on-screen images. Animals received 10 real-object trials per trial block.

ID Horse	Yard	Sex	Target	N trial blocks LC	N pre-test trial blocks	N Pre-screen test trial blocks
1	1	m	kong	11	1	1
2	1	m	kong	7	1	2
3	1	f	kong	4	1	4
4	1	m	kong	3	1	3
5	1	f	ring	12	1	1
6	1	m	kong	7	1	3
7	1	f	kong	7	1	1
8	1	m	ring	11	1	3
9	1	m	kong	11	1	1
10	1	f	ring	7	1	6
11	1	m	ring	7	1	1
12	1	m	ring	11	1	1
13	1	f	ring	15	1	1
14	1	m	ring	19	1	1
15	1	f	ring	15	1	1
16	1	m	ring	11	1	1
17	2	f	ring	16	1	1
18	2	m	kong	8	1	1
19	2	m	kong	19	1	1
20	2	m	ring	20	1	1
21	2	f	kong	8	1	1
22	2	m	ring	12	1	2
23	2	m	ring	19	1	1
24	2	f	kong	12	1	1
25	2	m	ring	19	1	1
26	2	m	ring	12	1	2
27	2	f	ring	12	1	2

Appendix D

Table D2: Overview of individual performance results in the screen test. Numbers in bold indicate correct responses above the level of chance (binomial test, p<0.05)

ID	N trials with image contacts	N trials correct responses	Percentage of correct responses (out of trials with screen contacts)	Percentage of correct responses out of all trials
1	5	3	0.6	0.3
2	6	3	0.5	0.3
3	10	5	0.5	0.5
4	9	7	0.777778	0.7
5	5	2	0.4	0.2
6	8	7	0.875	0.7
7	7	4	0.571429	0.4
8	10	7	0.7	0.7
9	3	2	0.666667	0.2
10	6	4	0.666667	0.4
11	10	7	0.7	0.7
12	6	3	0.5	0.3
13	6	4	0.666667	0.4
14	10	9	0.9	0.9
15	8	8	1	0.8
16	6	4	0.666667	0.4
17	6	6	1	0.6
18	9	5	0.555556	0.5
20	5	2	0.4	0.2
21	7	5	0.714286	0.5
22	7	1	0.142857	0.1
23	6	6	1	0.6
26	8	5	0.625	0.5
27	4	1	0.25	0.1
19	1	0	0	0
24	3	0	0	0
25	1	0	0	0
26	0	0	0	0
27	0	0	0	0

Appendix D