



**Investigating the Induction and Measurement of Equine
Positive Affective State through a Personality-centred Lens**

Ilona Lucy Loftus

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

Doctor of Philosophy

School of Natural and Environmental Sciences

June 2024

Copyright statement

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual or institution for the purposes of research and private study only. The thesis may not be reproduced elsewhere without the permission of the author.

Abstract

This thesis explores the induction and measurement of positive affective state in equines via an exploratory approach, beginning with a three-round Delphi consultation of ninety-three experts worldwide. The Delphi worked towards achieving consensus on the development of a reliable assessment for equine personality and the determination of methods that could be used to induce and measure positive affective state in equines.

Further work developed and tested the expert-devised personality traits assessment (a four phase, iterative process) to develop an equine personality questionnaire, the Equine Personality Eight-Factor Model.

A pilot study was undertaken to refine the methodological design, and a subsequent full study was undertaken which utilised four potential induction methods for equine positive affective state, in a within-subjects, repeated measures, experimental design. Selected physiological measures (heart rate and heart rate variability, respiratory rate, eye and ear temperature and salivary cortisol) and behavioural measures (ear position, eye bias and position, Qualitative Behaviour Assessment and frustration responses) were utilised to attempt to measure affective state.

Exploratory data analysis identified relationships between measures that may indicate the experience of positive affective state, whilst accounting for core personality traits; as well as measures that may indicate the presence of negative affective states. Suggested personality-based differences in responses to the four induction measures were also identified within the exploratory analyses. Results were considered within the framework of a four-quadrant arousal-valence model, providing suggested behavioural and physiological indicators of the affective states of reward acquisition, contentment, low motivation, and punishment avoidance.

The conclusions drawn from this work provide a starting point for further explanatory research in this subject area, to enhance understanding of how we may optimise equine welfare through the provision of ethologically relevant positive experiences for equines and evaluation of their affective state alongside consideration of individual personality traits.

Acknowledgements

This journey has been a long and deep dive into the vast sphere of equine affective state and personality. Throughout the process I have been continually amazed at how little we can be certain of regarding the minds and emotions of the equines we spend so much time with. The journey started with an aim to delve into the relatively novel area of study that was the mental domain of equine welfare, back in 2015 when discussions regarding the preposition of this work began. At this time, the study of promoting positive emotional welfare in equines had been given a relatively cursory glance in the literature. However, through my experiences over twenty years working with equines, and my development as a professional in the clinical animal behaviour sphere, I was steadfast in my belief that we should at least attempt to understand what makes our beloved equines happy and how we may be able to identify if they are so.

Through this process I have been supported and encouraged by many people who have influenced my progression, some of whom may not realise the impact they have had. Firstly, my supervisor Dr Matt Leach who encouraged me to believe in myself and my topic and who has been a constant source of support, reflection and motivation. His welfare lectures during my time as an MSc student inspired me to consider the presence of positive welfare as a true measure of good welfare and his confidence giving nature, filled with good humour, and insight has been truly inspirational. My second supervisor, Professor Lucy Asher, the 'stats queen', kindly agreed to join this exploratory project and for that I am truly grateful. Her knowledge, attention to detail and calm, consistent support has been very gratefully received throughout. I have gratefully received advice from my Annual Progression Review panel each year, thanks to pastoral member Dr Catherine Douglas, and academic members Dr Fritha Langford and formerly Dr Jonathan Guy for their helpful feedback and supportive comments.

I also owe a debt of gratitude to my family who have always been encouraging, pragmatic, and have endeavoured to help whenever possible with the juggling of self-funding a PhD whilst working full-time and caring for what were two very young children as a single mother when this process began. My father Dr David Cornwell has been a continual supporter, reminding me to keep going, as well as helping me to fund some of my studies. My mother Professor Tink Palmer is a constant inspiration and role model to whom I am eternally

grateful. My dear sister, Tamsin Cornwell, always a quiet but fierce cheerleader offering love and support and lastly to my stepdad Simon Palmer, unwaveringly supportive of me.

My new husband Mark Williams, who joined me during this journey, has shared the daily highs and lows with me but his unswerving belief that ‘you can do this!’ has spurred me on through these last years of the project and kept me grounded when my inner worry threatened to overtake me.

To my two darling children, Amalie and Jesse Loftus, thank you for your patience when ‘Mummy has to do some more work now’ became a constant refrain in our home. Your kindness and generosity of spirit has been immense and overwhelming at times.

To my dear friend and colleague, sadly no longer with me to celebrate this milestone, Dave Brown, who picked me up from many a moment of self-doubt and was a true critical friend, you are much missed, and I hope you are proud. To Dr Deirdre Rooney, my former line manager and all-round inspirational human – someone who I truly looked up to and who gave me the confidence and support to begin this process, thank you.

I am truly grateful to all the hundreds of people and equines who have contributed to this research, without them none of this would have been possible. They have taken me on a journey of learning, critical thinking and insight to an extent that I never dreamed possible. Their generosity of spirit and time will not be forgotten. Special thanks go to the Delphi participants who contributed many hours of time to this project and to my MSc students who contributed to parts of the data collection Amy Newman (Chapter 5) and Chloe Murchie (Chapter 4); and Laura Pannewitz with whom I published the behavioural indicators of frustration utilised within Chapter 6. Thank you also to Becky Smith, Gemma Hogg, Ali Durban, Mary Howell and Kate Johnstone for their support – true friends.

Lastly to the equines, who have always been my passion, I can only hope that this work contributes in some small way to your happiness and wellbeing. In the words of José Ortega y Gasset “Tell me what you pay attention to, and I will tell you who you are” a phrase that has stayed with me alongside that of Margaret Bonanno who said “It is only possible to live happily ever after on a daily basis” which, I think, is the fundamental principle for the emotional wellbeing of our beloved equines.

Collaborative work statement

The work contained within this thesis is entirely the work of the candidate, Ilona Loftus, apart from the following collaborative aspects:

Chapter 4 – A study into equine personality

Part three, the validation questionnaire was prepared and administered collaboratively with an MSc student, under the direct supervision of Ilona Loftus, as part of the student's MSc thesis (approximately 20% of the total data presented within this chapter). All data gathered were analysed independently by Ilona Loftus and the write up of the chapter is solely the work of Ilona Loftus. All other aspects of data collection within this chapter were undertaken solely by Ilona Loftus.

Chapter 5 – The pilot study

Data from 9 equines (39% of the data collected) were collected by an MSc student, under the direct supervision of Ilona Loftus, as part of the student's MSc thesis. All data gathered were analysed independently by Ilona Loftus and the write up of the chapter is solely the work of Ilona Loftus. All other aspects of data collection within this chapter were undertaken solely by Ilona Loftus.

Outputs from the work within this thesis

Loftus, L. 2020. Positive affective state in equines. Presentation to the APBC Virtual Equine conference, 28th November 2020, <https://www.apbc.org.uk/wp-content/uploads/bksdjhfsiugkzjshvfkjhbasd.pdf>.

Loftus, L. 2023. Assessing equine personality: the development of a new tool. Webinar presentation for the Association of Pet Behaviour Counsellors, 20th July 2023, <https://www.apbc.org.uk/Events/assessing-equine-personality-the-development-of-a-new-tool/>.

Loftus, L. 2023. The development of a new personality assessment for equines, Behaviour 2023 Conference, Bielefeld, Germany, 17th August 2024, https://www.uni-bielefeld.de/fakultaeten/biologie/forschung/veranstaltungen/behaviour2023/programme/Behaviour_schedule_Thursday_17th.pdf.

Loftus, L. 2023. Investigating the effect of positive interventions on behavioural and physiological lateralisation parameters in equines: a pilot study, Behaviour 2023 Conference, Bielefeld, Germany, 17th August 2024, https://www.uni-bielefeld.de/fakultaeten/biologie/forschung/veranstaltungen/behaviour2023/programme/Behaviour_schedule_Thursday_17th.pdf.

Loftus, L. 2023. Methods to induce and measure positive affective state in equines: a Delphi consultation, Behaviour 2023 Conference, Bielefeld, Germany, 19th August 2024, https://www.uni-bielefeld.de/fakultaeten/biologie/forschung/veranstaltungen/behaviour2023/programme/Behaviour_schedule_Saturday_19th.pdf.

Loftus, L. 2023. Inducing and measuring positive affective state in horses: a Delphi approach, Webinar presentation for FAB Clinicians, 6th September 2023, <https://fabclinicians.org/webinars/>.

Outputs in preparation

Loftus, L. 2024. Inducing and measuring positive affective state in horses through a personality-based lens. Webinar presentation for the Fellowship of Animal Behaviour Clinicians, 6th November 2024, <https://fabclinicians.org/webinars/>.

Chapters 3, 4 and 6 are in preparation for journal publication.

Table of Content

Contents

Copyright statement	ii
Abstract.....	iii
Acknowledgements.....	iv
Collaborative work statement	vi
Outputs from the work within this thesis.....	vii
Outputs in preparation	vii
Table of Content	viii
List of figures.....	xiii
List of tables	xxi
Chapter 1 Introduction	1
1.1 Affective state and welfare	1
1.1.1 Defining emotional welfare and affective state.....	1
1.1.2 Positive emotional welfare	1
1.1.3 Historical context of animal emotion	3
1.2 Understanding affective states in animals.....	5
1.2.1 The presence of affective states in animals.....	5
1.2.2 Valence and arousal.....	8
1.2.3 Studying negative and positive affective state	10
1.3 Inducing affective states	15
1.3.1 Creating instances of emotion	15
1.3.2 Making decisions, predictions and prediction error	16
1.3.3 Reinforcers and punishers in affective state.....	22
1.4 The role of learning and memory in affective state.....	23
1.4.1 Learning.....	24
1.4.2 Memory.....	28
1.5 Measuring affective states.....	33
1.5.1 Behavioural measures to evaluate positive affect in equines	33
1.5.2 Physiological measures to evaluate positive affect in equines.....	38
1.6 The effect of personality on affective state	42
1.6.1 Defining personality	47
1.6.2 Development of research on animal personality.....	48
1.7 Conclusion.....	50
1.8 Research aims and thesis outline.....	50

Chapter 2 General methodologies.....	52
2.1 Introduction	52
2.1.1 Overview of the experimental design.....	52
2.1.2 Overview of the population/sample and sampling strategy	53
2.1.3 Study location	54
2.1.4 Ethical approval.....	54
2.2 Delphi consultation methodology.....	55
2.2.1 Participant recruitment and selection	55
2.2.2 Survey development and question design.....	58
2.2.3 Demographic data collection	61
2.2.4 Questions on personality	61
2.2.5 Questions on emotional processing and affective state.....	62
2.2.6 Questions on inducing and measuring positive affective state	63
2.2.7 The Delphi process	64
2.2.8 Descriptive and thematic analysis.....	68
Chapter 3 Inducing and measuring positive affective state in equines: a Delphi consultation.....	69
3.1 Introduction	69
3.2 Methodology.....	73
3.3 Results.....	75
3.3.1 Retention rate	75
3.3.2 Personality	75
3.3.3 Round one.....	75
3.3.4 Round two.....	92
3.3.5 Round three	99
3.3.6 Overall result synopsis	115
3.4 Discussion.....	118
3.4.1 Demographic factors.....	118
3.4.2 Expert consensus on emotions in equines.....	119
3.4.3 Expert consensus on inducing positive affect in equines	125
3.4.4 Expert consensus on behavioural measures of positive affect in equines	131
3.4.5 Expert consensus on physiological measures of positive affect in equines	137
3.5 Limitations of the study	144
3.6 Recommendations	147
3.7 Conclusion.....	147
Chapter 4 Development of the Equine Personality Eight-Factor Model (EPEM).....	149
4.1 Introduction	149

4.1.1 Prior studies on equine personality	149
4.1.2 Aims and objectives	154
4.2 Methodology.....	154
4.2.1 Part one.....	155
4.2.2 Part two.....	158
4.2.3 Part three	159
4.2.4 Part four	160
4.3 Results.....	161
4.3.1 Part one – October 2019.....	161
4.3.2 Part two – January 2020 – July 2022.....	171
4.3.3 Part three – October – March 2022.....	178
4.3.4 Part four – September 2022 – June 2023	190
4.4 Discussion.....	192
4.4.1 The assessment instrument	192
4.4.2 The sample	196
4.4.3 EPEM in the context of other personality models	198
4.5 Limitations.....	199
4.6 Conclusion.....	200
Chapter 5 Pilot study investigating the effect of positive stimuli application on behavioural and physiological measurement parameters in equines.....	201
5.1 Introduction	201
5.1.1 Emotional processing and affective state	201
5.1.2 Positive stimuli for equines.....	202
5.1.3 Measurements of affective state	203
5.1.4 Aims and Objectives.....	207
5.2 Methodology.....	208
5.2.1 Ethical approval.....	208
5.2.2 Pilot study focus and design overview.....	208
5.2.3 Subjects	210
5.2.4 Habituation to the study area and equipment	210
5.2.5 Test procedure	211
5.2.6 Behavioural and physiological measures.....	213
5.2.7 Statistical analysis	215
5.3 Results.....	216
5.3.1 Behavioural measures to evaluate positive affect in equines	216
5.3.2 Physiological measures to evaluate positive affect in equines.....	221

5.3.3 Relationship between eye bias, eye position, eye temperature, ear position, and heart rate	224
5.4 Discussion.....	229
5.4.1 The methodological design	230
5.4.2 Induction methods for positive affect in equines	231
5.4.3 Behavioural measures of positive affect in equines	231
5.4.3 Physiological measures of positive affect in equines.....	235
5.4.4 Limitations.....	238
5.4.5 Recommendations	239
5.5 Conclusion.....	239
Chapter 6 Inducing and measuring positive affective state in equines: an exploratory study.....	241
6.1 Introduction	241
6.1.1 Outcomes from the pilot study.....	241
6.1.2 Positive induction stimuli utilised within this study.....	242
6.1.3 Behavioural measures to evaluate positive affect in equines	247
6.1.4 Physiological measures to evaluate positive affective state in equines	251
6.1.5 Concurrent measures of affective state.....	254
6.1.6 Aims and objectives	255
6.2 Methodology.....	256
6.2.1 Trial preparation	256
6.2.2 Subjects	257
6.2.3 Methods.....	259
6.2.4 Positive induction stimuli utilised within this study.....	267
6.2.5 Behavioural measures of positive affect in equines utilised within this study	269
6.2.6 Physiological measures of positive affect in equines utilised within this study	273
6.2.7 Personality assessment.....	276
6.2.8 Data processing and statistical analysis	276
6.3 Results	282
6.3.1 Reliability of measures	282
6.3.2 Outlier removal and correlation of measures.....	284
6.3.3 Linear Discriminant Analysis (LDA)	292
6.3.4 Principal Component Analysis (PCA).....	293
6.3.5 Linear modelling of PCA components by stimulus and trial	296
6.3.6 Further LDA and PCA analysis	297
6.3.7 Linear modelling of PCA components by stimulus and trial	305
6.3.8 Linear modelling of PCA components by personality group.....	307

6.3.9 Further Linear Discriminant Analyses (LDA) on companion, food and training stimuli.....	310
6.4 Discussion.....	312
6.4.1 Reliability of measures	313
6.4.2 Data reduction	314
6.4.3 Identifying potential markers of affective state – valence and arousal domains.....	315
6.4.4 Consideration of seeking and conflicted emotions.....	319
6.4.5 Limitations of the study	321
6.4.6 Suggestions for further study	328
6.4.7 Conclusion.....	329
Chapter 7 General discussion	331
7.1 Summary	331
7.2 The Delphi consultation	332
7.3 Equine personality	333
7.4 Inducing and measuring positive affective state in equines.....	334
7.5 Limitations.....	340
7.6 Recommendations	340
7.7 Concluding remarks	341
8. References	342
Appendix A – Chapter 3 - Delphi participant information and consent form.....	411
Appendix B – Chapter 3 - Delphi rounds questions and glossary of terms.....	414
Appendix C – Chapter 4 - Hominoid Personality Questionnaire from Weiss <i>et al.</i> , 2011	463
Appendix D – Chapter 4 - Equine Personality Eight-Factor Model and Abridged Questionnaire	476
Appendix E – Chapter 6 - Study participant information sheet.....	485
Appendix F – Chapter 6 - Pre-trial questionnaire	488
Appendix G – Chapter 6 - Trial order randomisation results.....	495
Appendix H – Chapter 6 - Cowlog instructions for raters	496
Appendix I – Chapter 6 - Additional graphs	504

List of figures

Figure 1.1 Modern Biopsychological view (adapted from Pinel, 2009).	4
Figure 1.2a Core affect represented in a two-dimensional model Italics denote discrete emotions with Q1 and 2 positive affective state and Q3 and 4 negative affective states. Arrows denote presumed biobehavioural systems associated with reward acquisition (green) and punishment avoidance (red). Adapted from Mendl <i>et al.</i> (2010a), Russell (2003) and Burgdorf and Panksepp (2006), and 1.2b The 12-point Affect Circumplex adapted from Yik <i>et al.</i> (2011) and Russell (1980).....	9
Figure 1.3 Neurological, behavioural and physiological measures that may be utilised to investigate affective state (Boissy <i>et al.</i> , 2007a).	14
Figure 1.4 Schema of the four-systems network control model proposed by Swanson (2012). One system (motor) controls somatic responses (i.e. behaviour). The motor system, in turn, is controlled by three inputs from the sensory, cognitive and behavioural state systems. Direct sensory inputs to the motor system mediate reflexive behaviours, inputs from the cognitive system mediate voluntary behaviour and influences from the state system mediate state control influences. The motor system influences the external environment via skeletomotor changes and physiological responses within the body (visceral stimuli and responses). These influences on the external and internal environment (dashed black lines) provide feedback signals to the nervous system (dashed light grey line).....	18
Figure 1.5 The process of interoceptive predictions adapted from Barrett (2017) and Barbas (2015).....	19
Figure 1.6 Schematic illustration of the hypothesised role of core affect in model-free decision making. A high punishment mood-like core affective state, low blood glucose and prior memories determine the horse's initial situation (left hand box). Its mood state and internal status influence attention (thick red and black arrows in left hand box) resulting in detection of new sensory information (food availability with potential conflict from other horses). The new situation is matched to memory of similar past situations. Sensory information is unambiguous (likely to provide a reliable basis for prediction of outcomes), it has a stronger influence on memory retrieval (thick green arrow in left hand box) than mood or internal status (thin black and red arrows in left hand box). The retrieved memory is associated with	

previous actions including evolved Pavlovian responses (action options = approach, avoid, ignore) and their outcomes (action values) which within core affect space represent affective options. Through a process of comparison, the most positively valenced option (avoid) is selected. This now becomes an affective prediction (a prediction of the affective consequence of the selected action in this situation). The physical outcome is probabilistic (thick black arrow to Null (safe)) and the associated affective outcome is determined by the direct experience of reward or punishment and their presence or absence in relation to the affective prediction. Physical outcomes feedback to influence internal status whilst affective outcomes act as prediction errors to update situation-dependent action values. They also update mood-like core affect. Within the core affect diagrams the angled arrows denote Reward Acquisition System (RAS) and Punishment Avoidance System (PAS) and the red and green circles indicate negative and positive valence respectively. Grey circles indicate a hypothetical integrated location in core affect space. Line drawings authors own, diagram adapted from Mendl and Paul (2020a).....20

Figure 1.7 Depicting the alignment between Mendl *et al's.* (2010) framework and the quadrants of reinforcement and punishment in behavioural theory. Quadrants 1-4 are identified along with Mendl *et al's.* predicted events/stimuli that will generate affective state (**in bold**), the affective state achieved (red or green text) and the applicable quadrant of reinforcement or punishment (*in italics*). Authors own.....28

Figure 1.8 The relation of adjectives to dimensions according to Plutchik's theory of emotions. Adapted from (Bucik *et al.*, 1995). *Rectangles denote EPI adjectives; ovals denote dimensions or constructs of emotions.....44

Figure 1.9 Comparison of Zuckerman and Cloniger's findings in relation to the neurochemistries of personality.....46

Figure 2.1 The research process (Daniel and Harland, 2017).....53

Figure 2.2 Field study locations (Editable map from Softonic, 2022).....54

Figure 2.3 Literature searches undertaken to develop initial Delphi themes and questions...59

Figure 3.1 Conceptual map of the equine industry and upstream supply categories, adapted from DEFRA (2004).....69

Figure 3.2 Summarised diagram of the Delphi process.....	74
Figure 3.3 Demographic data surveying participants area of expertise.....	75
Figure 3.4 World map showing location of participants and number of participants per country (Authors own image).....	76
Figure 3.5 Bar chart showing the qualification status of participants.....	77
Figure 3.6 The level of importance participants placed on the characteristics of valence, arousal and surgency in relation to equine emotional state.....	78
Figure 3.7 Stacked bar chart depicting level of agreement on methods to induce positive affect in equines.....	80
Figure 3.8 Aggregated ranking of the induction methods for positive affect in equines across each round. Rank scores were reversed to allow for clear visualisation of the most popular methods suggested by participants to induce positive affect in equines.....	115
Figure 3.9 Aggregated ranking of the behavioural methods for positive affect in equines across each round. Rank scores were reversed to allow for clear visualisation of the most popular methods suggested by participants to behaviourally measure positive affect in equines.....	116
Figure 3.10 Aggregated ranking of the physiological methods to measure positive affect in equines across each round. Rank scores were reversed to allow for clear visualisation of the most popular methods suggested by participants to physiologically measure positive affect in equines.....	117
Figure 3.11 Visualising the fourth dimension, a conceptualisation of measures of emotional welfare over time. Authors own diagram adapted from Radiya-Dixit (2017).....	122
Figure 3.12 Representation of the interplay of emotional processing components (authors own).....	123
Figure 4.1 An overview of the study methods, the full Delphi methodology can be found in Chapter 2.....	155
Figure 4.2 Inter-rater agreement per trait (count data). Absolute agreement shown in the lightest shade (top of each bar), differences in scores of plus or minus one point shown in mid	

shade (middle of each bar) and differences of more than one point shown in dark shade (bottom of relevant bars, only 9 instances across the 54 traits).	162
Figure 4.3 Mean scores correlation between observer one and observer two.....	165
Figure 4.4 Principal components scree plot identifying 10 PCs with eigenvalues greater than 1 (1 signified by the red line).....	166
Figure 4.5 Loading plot for principal components 1 and 2. Dashed lines represent three distinct large clusters and continuous lines represent the subdivided three larger and five smaller clusters.....	167
Figure 4.6 Schema depicting the internal consistency levels of each original group and how each large group was divided into small groups with the new group internal consistency ascribed.....	170
Figure 4.7 Clustered bar chart displaying expert rated scores for each trait across each characteristic with a total score line displayed in yellow.....	174
Figure 4.8 Principal components scree plot identifying 2 PCs with eigenvalues greater than 20 (20 signified by the red line).....	181
Figure 4.9 Trait influences on the first and second components. Groupings denoted by '}' are the original groupings identified in Part two.....	182
Figure 5.1 Overview of the pilot study methodology.....	209
Figure 5.2 Cross-tied subject with heart rate monitor fitted under Polypad and surcingle...	211
Figure 5.3 Eye and ear positions recorded.....	214
Figure 5.4 Anatomical recording position (medial posterior palpebral border of the lower eyelid and lacrimal caruncle) for eye temperature.....	215
Figure 5.5 Mean duration (seconds) with SEM, spent in each eye bias position across treatments and phases.....	217
Figure 5.6 Mean duration (seconds) with 95% confidence intervals, spent in eye bias across treatments and phases.....	217
Figure 5.7 Mean duration (seconds) with SEM, spent in each eye position across treatments and phases.....	218

Figure 5.8 Mean time (seconds) with 95% confidence intervals, spent in eye positions across treatments and phases.....	219
Figure 5.9 Ear asymmetrical with 95% confidence intervals, during all phases and treatments.....	220
Figure 5.10 Ear backward with 95% confidence intervals, during all phases and treatments.....	220
Figure 5.11 Ear forward with 95% confidence intervals, during all phases and treatments.....	221
Figure 5.12 Eye temperature (°C) with 95% confidence intervals, across the three treatments.....	221
Figure 5.13 Eye temperature (°C) with 95% confidence intervals, during all phases and treatments.....	222
Figure 5.14 Heart rate (beats per minute) with 95% confidence intervals, during all phases and treatments.....	223
Figure 5.15 Relationship between left eye bias and temperature during Food treatment..	224
Figure 5.16 Relationship between left eye bias and temperature after Companion treatment.....	224
Figure 5.17 Relationships between left eye bias duration and heart rate for all phases and all treatments with Spearman rank correlation results. C = Companion, F = Food, and P = Praise.....	226
Figure 5.18 Relationships between left eye bias duration and eye position duration for all phases and all treatments with Spearman rank correlation results. C = Companion, F = Food, and P = Praise.....	227
Figure 5.19 Correlations between ear position ‘backwards’ duration (seconds) and eye position (eye white on the left and eyes open on the right) duration (seconds).....	228
Figure 6.1 Schema depicting the component model of emotion adapted from Scherer (2009) and Mohammadi <i>et al.</i> (2019).....	254
Figure 6.2 Examples of positive responses to wither scratching a) head and neck extension with mouth position changes and b) grooming of the human scratcher during wither scratching.....	259

Figure 6.3 The three views generated by camera placement at two of the five study locations.....	261
Figure 6.4 Placement of the Polar Equine Heart rate monitor using a Polar equine band....	263
Figure 6.5 Flowchart depicting an outline of the data collection method.....	264
Figure 6.6. Timings of each trial ensuring that there would be at least a six-hour wash-out period between trials.....	265
Figure 6.7 Affiliative conspecific interactions during the conspecific trial.....	266
Figure 6.8. Mid-way through a food trial.....	267
Figure 6.9. The wither area permitted for scratching circled in red, according to individual equine preference, adapted from Dittrich (1898).....	267
Figure 6.10 Example of target training with a) the horse touching the target with their muzzle and b) receiving a low value food reinforcer (Dengie Hi-Fi chaff).....	268
Figure 6.11 Salivary cortisol collection and LFD analysis process.....	274
Figure 6.12 Data reading positions for the left eye (Sp1), right eye (Sp2), left ear (Sp3) and right ear (Sp4). A marker line indicated the slope across the base of each ear to aid data spot recording accuracy. Eye temperature was taken from the medial canthus of the eye as suggested by Kim and Cho (2021).....	277
Figure 6.13 CowLog coding set-up.....	278
Figure 6.14 Flow chart detailing the statistical analysis methodology undertaken.....	280
Figure 6.15 Correlation plot of LFD and ELISA salivary cortisol results with box and whisker to depict the mean and range of both data sets. Outliers are depicted by *	281
Figure 6.16a Corrplot of ear and eye temperature (number method).....	283
Figure 6.16b Corrplot of ear positions (number method).....	284
Figure 6.16c Corrplot of eye position and eye bias (number method).....	284
Figure 6.17 Corrplot of QBA measures (number method).....	285
Figure 6.18 Corrplot of Frustration measures (number method).....	286

Figure 6.19 Corplot of heart rate indices (number method).....	287
Figure 6.20 Interval plot of cortisol data by trial phase and treatment with confidence intervals.....	291
Figure 6.21 Interval plot of respiratory data by trial phase and treatment with confidence intervals.....	291
Figure 6.22 Separation of variables between the 'before' and 'during' phases of the scratch trial. With correlations of greater than 0.7 only identified between variables 'Low frequency power' and 'Low frequency/high frequency ratio'	292
Figure 6.23 Scree plot of the first ten principal components.....	293
Figure 6.24 Loading plot of variables showing clustering of variables across negative and positive dimensions of principal components 1 and 2. Coloured lines represent the amount of each variable that is represented within each given component with green and yellow lines indicating higher levels of representation than brown and black lines.....	294
Figure 6.25a Boxplot of trial and stimulus on principal component 1, and 6.25b Boxplot of trial and stimulus on principal component 2.....	295
Figure 6.26 Separation of variables between the 'before' and 'during' phases of the scratch trial. With correlations of greater than 0.7 identified between variables 'SDRR' and 'RMSSD'	298
Figure 6.27 Correlation between all variables shown on a correlation plot. Positive correlation between SDRR and RMSSD identified by the bright red squares slightly above and below the central within variable correlation line.....	299
Figure 6.28 Scree plot of the first ten components from the PCA.....	300
Figure 6.29 Proposed model for measures of valence and affect in equines based upon the framework of Mendl et al. (2010) following evaluation of measurement variable relationships. Dimension 1 is proposed as the arousal dimension and dimension 2 as the valence dimension.....	301
Figure 6.30 Loadings of each measure on principal components one and two.....	302

Figure 6.31 Loadings of measures on PCs 3 and 4. Physiological variables show a tendency to load onto both PC 3 and 4 either positively or negatively per measure apart from mean heart rate. Behavioural measures appear to load appositionally on each principal component apart from QBA friendly, which has very low contribution to each component; QBA uneasy, eye sclera and ears gently back also load negatively across both principal components.....	303
Figure 6.32 Mapped affect markers overlaid onto the affective circumplex with quadrants identified (Q1-4) and green arrows denoting increased pleasure and red aroused denoting decreased arousal. Positive affect markers are placed within green boxes and negative affect markers within peach boxes. Original 12-measure affect diagram from Yik <i>et al.</i> (2011) overlaid with the suggested affect markers from this study.....	304
Figure 6.33a Box and whisker plot displaying loadings on PC1 accounting for the effect of stimulus and trial, and 6.33b Box and whisker plot displaying loadings on PC2 accounting for the effect of stimulus and trial.....	305
Figure 6.34 Separation of variables between the ‘before’ and ‘during’ phases of the training trial. With correlations of greater than 0.7 only identified between variables ‘Mean heart rate’, ‘SDHR’, ‘Minimum heart rate’ and ‘Maximum heart rate’.....	310
Figure 6.35 Separation of variables between the ‘before’ and ‘during’ phases of the companion trial. No correlations of greater than 0.7 were identified between variables.....	310
Figure 6.36 Separation of variables between the ‘before’ and ‘during’ phases of the food trial. With correlations of greater than 0.7 identified between variables ‘Mean heart rate’, ‘Minimum heart rate’ and ‘Maximum heart rate’; ‘Maximum heart rate’ and ‘SDHR’; ‘Left eye temperature’ and ‘Left ear temperature’; ‘Left ear’ and ‘Right ear’ temperatures.....	311
Figure 7.1 Overview of potential relationships between valence and arousal: (a) independence, (b) positive linear relation, (c) symmetric V-shaped relation, (d) asymmetric V-shaped relation including positivity offset and negativity bias. Figure adapted from Kuppens <i>et al.</i> (2017).....	337

List of tables

Table 1.1 Proposed links between environmental conditions (reinforcers and punishers), resulting predominant core affective mood state and biases in expectation that drive optimistic or pessimistic decision making. Quadrant 1 (Q1) = high arousal in positive valence (reward acquisition), Quadrant 2 (Q2) = low arousal in positive valence (calm, content), Quadrant 3 (Q3) = low arousal in negative valence (depressed) and Quadrant 4 (Q4) = high arousal in negative valence (punishment avoidance). Adapted from Mendl <i>et al.</i> (2010a).....	23
Table 1.2 The classical conditioning reflex (Baldwin and Baldwin, 2001).....	25
Table 2.1 Methods of recruitment for different types of professional.....	55
Table 2.2 Inclusion and Exclusion criteria for recruitment of Delphi consultation participants.....	57
Table 2.3 Demographic data collection fields.....	61
Table 2.4 Delphi consultation round one presentation.....	64
Table 2.5 Delphi consultation round two presentation.....	66
Table 2.6 Delphi consultation round three presentation.....	67
Table 3.1 Levels of agreement relating to questions surrounding affective state in equines. Values reaching consensus are placed in bold.....	79
Table 3.2 Expert rankings of methods to induce positive affect in equines. Overall rank = the ranking for Round one. Mean (rank) = mean of all rater rankings (1 is low and 5 is high rank) with SD. Mode and median values are also represented. Higher mean, median and mode values equal higher agreement with the measure. Overall rank = sum of all ranking values by raters and associated preference rank.....	81
Table 3.3 Themes, subthemes and counts of suggestions made by participants when invited to add free text responses.....	82
Table 3.4 Results for expert ranking of behavioural methods to measure positive affect in equines. Overall rank = rank in this round with 1 being high and 8 being low. Mean rank =	

lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD. Mode and median values identified.....	85
Table 3.5 Themes, subthemes and counts of suggestions made by participants when invited to add free text responses.....	86
Table 3.6 Results for expert ranking of physiological methods to measure positive affect in equines. Rank = rank in this round. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD. Mode and median values identified.....	88
Table 3.7 Themes, subthemes and counts of suggestions made by participants when invited to add free text responses.....	90
Table 3.8 List of agreed statements, percentage agreement and additional feedback supplied by participants. Only statements achieving agreement consensus are included within the table (S1-11, S13).....	93
Table 3.9 Ranking of participant suggested methods to induce positive affect in equines. Overall rank = rank in this round with 1 being high and 8 being low. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD. Mode and median values identified.....	97
Table 3.10 Ranking of participant suggested behavioural methods to measure positive affect in equines. Overall rank = rank in this round with 1 being high and 8 being low. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD. Mode and median values identified.....	98
Table 3.11 Ranking of participant suggested physiological methods to measure positive affect in equines. Overall rank = rank in this round with 1 being high and 8 being low. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD. Mode and median values identified.....	99
Table 3.12 Agreement on methods that could induce positive affect in equines.....	100
Table 3.13 Thematic analysis of round three methods to induce positive affect in equines...	102

Table 3.14 Agreement on behavioural methods that could be used to measure positive affect in equines.....	105
Table 3.15 Round three thematic analysis of feedback relating to potential behavioural methods to measure positive affect in equines.....	108
Table 3.16 Agreement on physiological methods that could be used to measure positive affect in equines.....	110
Table 3.17 Round three thematic analysis of feedback relating to potential physiological methods to measure positive affect in equines.....	113
Table 4.1 Examples of the same question within the hominoid and equine questionnaires and the Likert scale definitions given to participants.....	156
Table 4.2 Demographic information for the 35 test equines. Definitions: Mare – entire female, Gelding – castrated male, Stallion – entire male 4 years old or over, Colt – entire male less than 4 years old or Filly – entire female less than 4 years old. Turned out – spends 24 hours a day in the field, Stabled – is not turned out in a field, Both – spends time in both the stable and a field in each 24 hour period. Light work – up to an hour a day a few times a week, Medium – daily exercise which includes predominant work in trot or canter with some jumping or schooling, Hard work – daily exercise with predominance of faster work with additional competitions or None – not ridden.....	157
Table 4.3 Inter-rater agreement of individual personality traits. Fair to moderate significant agreement in bold, fair and above non-significant agreement in italics, differences between observer means of less than 0.5 underlined. $nd > 1$ = number of observations per trait with disagreement greater than one point. nAA = number of pairs of ratings with absolute agreement per trait, out of 54 total ratings per trait.....	163
Table 4.4 New personality groupings arising from PCA and internal consistency analysis.....	171
Table 4.5 <i>Post hoc</i> Mann-Whitney analysis to identify differences in expert scores for the clarity, presence and reliability of the traits in equines (significant pairings highlighted in bold) with the median for Reliable significantly lower than the medians for Clarity and Presence.....	173

Table 4.6 Version two personality groupings utilised in the owner survey and presented to experts in Round three of the Delphi.....	178
Table 4.7 One-way random effects Intraclass Correlation Coefficient (ICC (1, k) for inter-rater agreement across 14 horses and 2, 3 or 4 raters, significant agreement highlighted in bold and good or higher agreement in italics.....	180
Table 4.8 Expert derived group names and traits from the Delphi consultation with traits and groups removed following backwards stepwise analysis (strikethrough) and traits added (bold) after forward stepwise comparison analysis of the owner-rater generated data. Expert agreement level from Round three of the Delphi (consensus level reached highlighted in bold). Group level internal consistency is reported via Cronbach's Alpha value (excellent agreement in bold, very good agreement in italics and moderate agreement in standard text).....	183
Table 4.9 Expert feedback on the second-round traits and trait groupings.....	185
Table 4.10 Comparison of the final proposed Equine personality eight-factor model (EPEM) in this study with the Big Five (NEO PI-R) model.....	188
Table 4.11 Group level correlation between full questionnaire scores and short form (abridged scores). Significant associations in bold.....	190
Table 4.12 Agreement between abridged and unabridged questionnaire scores in twenty equines.....	191
Table 4.13 Construct validity tested via Cronbach's alpha between presumed opposing personality type groupings.....	191
Table 6.1 Ranked induction methods from the Delphi consultation with exclusion rationale for excluded methods.....	241
Table 6.2 Ranked behavioural measures from the Delphi consultation with exclusion rationale for excluded methods.....	246
Table 6.3 Ranked physiological measures from the Delphi consultation with exclusion rationale for excluded methods.....	250
Table 6.4 Demographic data for the twenty subjects.....	257

Table 6.5 Preference test results for wither scratching area and food preference. Results for each of the three tests per induction method are presented. LF = Left-hand side fast, LS = left-hand side slow, RF = right-hand side fast, RS = right-hand side slow, TF = top of wither fast, TS = top of wither slow. H = Hay, N = Nuggets, L = Lick.....	260
Table 6.6 The ear and eye positions with descriptors and images as utilised by the video scorers. Ethogram developed from Lansade <i>et al.</i> (2018).....	269
Table 6.7 Eye bias positions with images and descriptors. Ethogram developed from Des Roches <i>et al.</i> (2008) and Austin and Rogers (2007).....	270
Table 6.8 QBA names and descriptors from AWIN (2015).....	271
Table 6.9 Indicators of frustration from Pannewitz and Loftus (2023).....	272
Table 6.10 ICC results for behavioural measures for each equine.....	282
Table 6.11 Measures taken forward for further analysis. Rationale for taking each measure forward is presented along with the results of the Pearson's product-moment correlation, the correlation estimate and whether the item positively or negatively correlated with the QBA measures.....	288
Table 6.12 Rationale for adding further measures into principal component analysis.....	297
Table 6.13 Results of linear models examining the effect of personality traits on responses to stimuli and trial phase. Results in bold indicate that the entire model was significant.....	307
Table 7.1 Proposed measures that may be indicative of positive, neutral and negative valence in high, neutral or low arousal states.....	335
Figure 1.1 Modern Biopsychological view (adapted from Pinel, 2009).	4
.....	xiii

Chapter 1 Introduction

1.1 Affective state and welfare

1.1.1 Defining emotional welfare and affective state

The Brambell Report (Brambell, 1965) gave rise to the formal discipline of animal welfare and the establishment of it as a 'young science' (Carenzi and Verga, 2009). Defining the term welfare has been an ongoing process to establish a definition which can be assessed scientifically, used within legislation and is meaningful to the stakeholders to whom it is relevant. Three main approaches were taken to the defining of 'animal welfare', 1) a focus on the biological function of the animal including health and behaviour, 2) a focus on the psychological approach including stress and eustress, and 3) a focus on natural living and the ability to display species typical behaviours (Carenzi and Verga, 2009; Hemsworth *et al.*, 2015). Indeed, perhaps the ultimate definition is that seen in the initial Brambell report which stated:

“welfare is a wide term that embraces both the physical and mental well-being of the animal. Any attempt to evaluate welfare, therefore, must take into account the scientific evidence available concerning the feelings of animals that can be derived from their structure and functions and also from their behaviour” (Brambell, 1965).

Over time various approaches to animal welfare have been proposed including that of Dockès and Kling-Eveillard (2006) who proposed that welfare should be considered across four main themes: 1) biological and technical approaches, 2) regulatory approaches for law making and guidance, 3) philosophical approaches considering the status of animals in human society, and 4) communication between human and non-human animals. From a scientific perspective the study of an animal's welfare is an assessment of the animal's subjective experience measured via behavioural and physiological indices to place the animal within the welfare continuum of very bad to very good (Hemsworth *et al.*, 2015). This thesis focusses on the affective component of welfare, namely the induction and measurement of positive affective states in equines.

1.1.2 Positive emotional welfare

Mendl and Paul (2020b) define '*Affect*' as an overarching term for states that have valence as a property (i.e. they are positive or negative) and include both emotions and moods but also

the valenced components of sensations (such as the pleasure of scratching an itch), and non-conscious components such as behavioural indicators and neural changes. Following the recognition of animals as sentient beings within the Amsterdam Treaty of 1997 (Carenzi and Verga, 2009), their emotional welfare has become an increasingly important consideration within more holistic assessments of welfare, due to their capacity for affective experiences such as stress and fear or conversely the enjoyment of positive feelings (Mellor *et al.*, 2009; Mellor, 2015b). A rise in welfare standards, coupled with the requirement to consider both physiological and psychological welfare within welfare assessment, has led to a renewed focus on emotions in animals and on positive welfare states including positive affective state (Reimert *et al.*, 2023); with the increase in the focus on this topic reviewed by Lawrence *et al.* (2019). The recent development of Social Licence to Operate (SLO) in the equestrian industry has also invigorated the study of equine emotions with reference to promoting positive affect in domesticated equines to enhance welfare (Douglas *et al.*, 2022; Heleski, 2023).

Evidence within the equine, wider animal, and human literature has identified the impact of both long- and short-term positive and negative affective states on health and wellbeing; for example much research has focussed on the development of pre-pathological and pathological states across species arising from chronic stress and anxiety (Ohl *et al.*, 2008; Moberg, 2000; Koolhaas *et al.*, 1999; Wiepkema and Koolhaas, 1993; Weiss, 1972) and specifically in equines (Lesimple, 2020; Kovac *et al.*, 2022; Broom and Kirkden, 2004; Pawluski *et al.*, 2017). The effects of trigger stacking (repeated exposure to negative emotional stimuli) have been widely covered in the clinical animal behaviour literature (Loftus, 2023; Durston, 2022; Henley, 2019) with reference to its physiological and psychological effects on short- and longer-term behaviour. Latterly a shift in focus to consider the impact of positive emotions in animal welfare has been identified in the literature, including methods to provide an optimal environment to engender positive emotional welfare in animals under management conditions. These include methods to monitor positive emotions and associated welfare (Boissy *et al.*, 2007a; Mendl and Paul, 2020b), considerations for providing positive anticipation, enrichment and control over the environment for animals (Boissy *et al.*, 2007b), the importance of joy states in animals and the evolution of research addressing these aspects (Lawrence *et al.*, 2019), providing socially for the needs of domesticated animals (Spinka,

2012) and providing opportunities for animals to engage in activities that produce positive affective states (Mellor, 2015a).

1.1.3 Historical context of animal emotion

The concept of emotion in non-human animals can be seen retrospectively within the literature of the 19th century with the notable publication of '*The expression of emotion in man and animals*' (Darwin, 1872) and the work of Charles Bell (Bell, 1847) amongst other researchers publishing in this sphere. In his 1872 volume Darwin compared the production of emotions and associated behaviours in a variety of species and developed a theory regarding the evolution of emotional expression composed of three key ideas:

1. Expressions of emotion evolve from behaviours that indicate what an animal is likely to do next,
2. If signals provided by such behaviours benefit the animal that displays them, they will evolve in ways that enhance their communicative function, and their original function may be lost,
3. Opposite messages are often signalled by opposite movements and postures, an idea termed the *principle of antithesis* (Darwin, 1872).

Following on from Darwin's work, the James-Lange theory (1884) emerged (Lange and James, 1922). This proposed that sensory stimuli which could induce an emotion were received and interpreted by the cortex, triggering changes in visceral organ activity via the autonomic nervous system and skeletal muscle (seen as behavioural) changes via the somatic nervous system. These autonomic and somatic responses were thought by the theorists to separately cause emotional experiences in the brain (James, 1884; Lange, 1922). Prior to this theory the default view was that a fear inducing stimulus would cause a feeling of fear which would lead to a physiological reaction (Pinel, 1990). Cannon (1927), who had reviewed the James-Lange theory presented an alternative hypothesis which was then built upon and promoted by Bard (Bard, 1934) culminating in the Cannon-Bard theory (Guthrie Yarwood, 2022). This theory postulated that emotional stimuli have two independent excitatory effects stimulating both the feeling of emotion in the brain (which they suggested may occur prior to physiological changes) and the expression of emotion in the autonomic and somatic nervous systems; that is, they can be thought of as parallel processes with no direct causal relation.

Papez (1937) built on the work of Broca (1878) to theorise the concept of the limbic system and proposed that this system of several interconnected neural structures controlled emotional expression (Pessoa and Hof, 2015). They purported that this occurred through the action of other limbic structures on the hypothalamus and that these emotional states were experienced through the action of the limbic structures on the cortex (Rajmohan and Mohandas, 2007). Over time emotion theory developed and Paul MacLean coined the term the “limbic system” in 1952 and proposed and defined the Triune concept of the brain (MacLean, 1990). These theories were appraised by Dror (2014) with reference to greater current evidence, who postulated that in fact neither the James-Lange, Cannon-Bard nor the traditional views were correct. Subsequently Dror and other researchers put forward the currently accepted ‘Modern biopsychological view’ (Reisenzein and Stephan, 2014; Dror, 2014). The current theory proposes that the interplay between stimulus perception, autonomic and sensory responses and affective experience is a reciprocal and cyclical feedback phenomenon (Figure 1.1); a concept which is embedded into the work within this thesis.

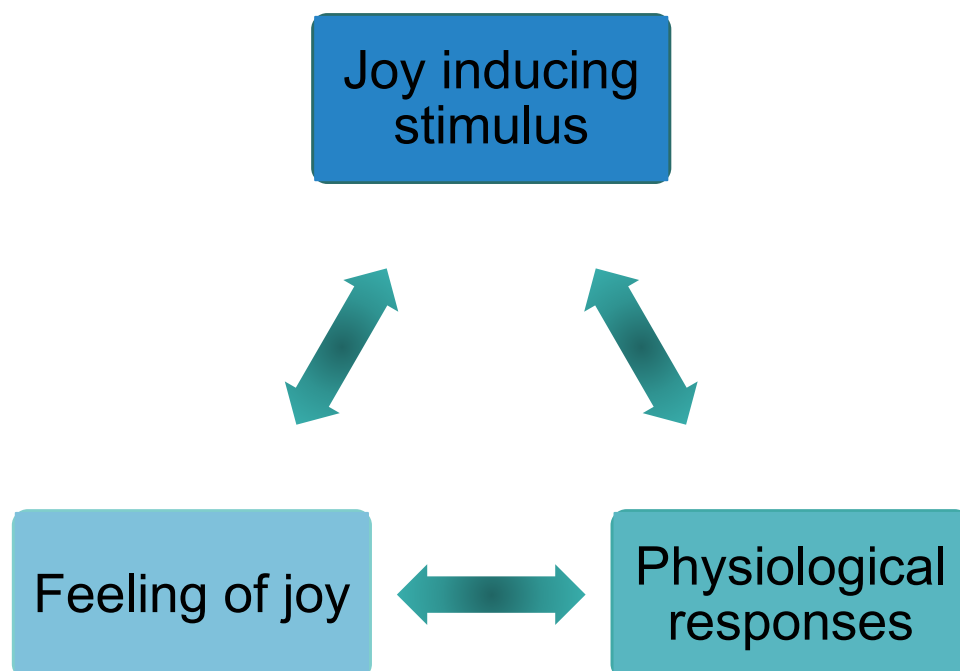


Figure 1.1 Modern Biopsychological view (adapted from Pinel, 2009)

1.2 Understanding affective states in animals

Mendl and Paul (2020b) present a clear definition of affective states in animals:

“Animal affective states are elicited by rewards and punishers or their predictors. A reward is anything for which an animal will work, and a punisher is anything that it will work to escape or avoid. Rewards or the absence of punishers, and associated predictions thereof, induce positive affect. Punishers or the absence of rewards, and associated predictions thereof, induce negative affect. Short-term emotion-like states follow immediately from individual rewarding or punishing events, whilst cumulative experience of events influences longer-term mood-like states.”

Therefore, it follows that it is likely we can utilise rewards and punishers as methods of inducing affective states and that we can monitor behaviour to understand how an animal feels about a stimulus (whether they actively attempt to engage or avoid it) as a proxy measure of their affective state at the time; both of which are concepts that are explored throughout this thesis.

1.2.1 The presence of affective states in animals

Emotions (short-term valenced states) (Mendl and Paul, 2020b) in non-verbal animals are challenging to study; however, developing research has progressed our understanding of how animals perceive their environment and how we can assess their emotional experiences (Mendl *et al.*, 2010a; Boissy and Lee, 2014). Human research has shown that emotion is bidirectionally linked to cognition (Quadt *et al.*, 2022; Storbeck and Clore, 2007) and further cross-species research has identified the neural homology between human and many non-human animal species including equines, farm animals, poultry and farmed fish (Shettleworth, 2009; Boissy and Lee, 2014). By studying cognition (the mechanisms by which animals acquire, process, store and act upon information from the environment) it is possible to explore affective experiences; with cognitive processes providing insight into the scientific study of emotional events (acute states) in animals as well as longer-lasting affect (chronic states) termed moods (Boissy *et al.*, 2007a; Mendl and Paul, 2020b). Reported evidence for the presence of affective states in non-human animals includes studies of neurologic similarities during drug abuse in animals and humans (Drevets *et al.*, 2001; Volkow *et al.*, 2002); neural imaging data showing cross-species (human and non-human animal) anatomical homologies for the major loci of control for affective experience (Liotti, 2004; Damasio *et al.*, 2000) and

brain stimulation studies indicating alignment between human and non—human animals when investigating regions of the brain generating affect (Heath, 1996; Panksepp, 1998).

Early welfare research focussed on identifying and resolving disturbances in biological function arising from changes within the animal's internal and external environment and identifying relationships between negative emotional experiences and the functional states and environmental factors that generated them (Mellor et al., 2009). These identified relationships allowed for later validation of indices which could be used to demonstrate the likely presence or absence of the associated emotional experiences. Original questions around the experience of emotion in animals, asked by Darwin, James and others, have been subsequently discussed by researchers, including Burgdorf and Panksepp (2006), with increasing interest over the last 25 years. Panksepp (2011a) published '*The basic emotional circuits of mammalian brains: Do animals have affective lives?*' a review of research and evidence published to that date. In this review Panksepp described the technology now available to consider affective states through evaluation of the basic brain mechanisms underlying more highly cognitive processes, in addition to the evaluation of behavioural and physiological evidence. Panksepp (2011b) and others (Mendl et al., 2010a; Nesse and Ellsworth, 2009; Rolls, 2005; Watson et al., 1999) have proposed that these '*primal manifestations of mind*' are raw, affective experiences which are natural functions of mammalian brains and are the entry points for our understanding of the use of reinforcement and punishment (discussed later in this chapter). They suggest that the 'higher' aspects of the human mind are strongly linked to the 'lower' basic neuropsychological processes of animal minds with much of the proliferation in neuroscience research centring on the topic of the mind as a manifestation of brain activity. The limitations of this experimental research are clear however, in that we have no first-hand access to the minds of animals who cannot convey their feelings verbally to us; therefore we must use their behavioural outputs, paired with physiological data as proxy measures for their feelings. There are however strategies that can be implemented to overcome these limitations such as measuring natural emotional behaviours demonstrated by animals and modelling these behaviours alongside imaged brain activity in a hypothetico-deductive research methodology (Messina et al., 2021; Dolcos et al., 2011) and using markers of biological functioning such as cardiovascular and neuroendocrine measures (Hemsworth et al., 2015).

The production of emotion is hugely complex and involves wide interplay between numerous circuits within the brain for production of each instance of emotion. The degeneracy principle (a many to one principle) where many different combinations of neurological activities can produce the same outcome, means that variation doesn't necessarily lead to incorrect outcomes. Conversely the core brain system (an example of a one-to-many principle) enables a single brain area or network to contribute to many different mental states and activities (Barrett, 2017a). The classical view, of one-to-one brain systems (the Triune brain), presented by scientists such as Jaak Panksepp, describes dedicated physiological functions of discrete brain areas, a theory which has been superseded by the more recent advances in emotion science. Likewise previous theories of equipotentiality, the theory that every neuron does the same thing, have since been disproven (Barrett, 2017). Neurons are multipurpose, there is currently no known fingerprint for emotion nor a defined discrete location for emotion production within the brain (Jungilligens *et al.*, 2022). Indeed emotions, per se, can be considered category headings, which within them hold much variation in their display and feeling. If we take a population view, we consider a species as a collection of diverse individuals that can be summarised statistically, however, this summary is an abstraction that does not actually exist in nature, it does not describe an individual.

When considering neural interactions within major emotional systems of the brain there are several key characteristics found across mammalian species and further, these include that:

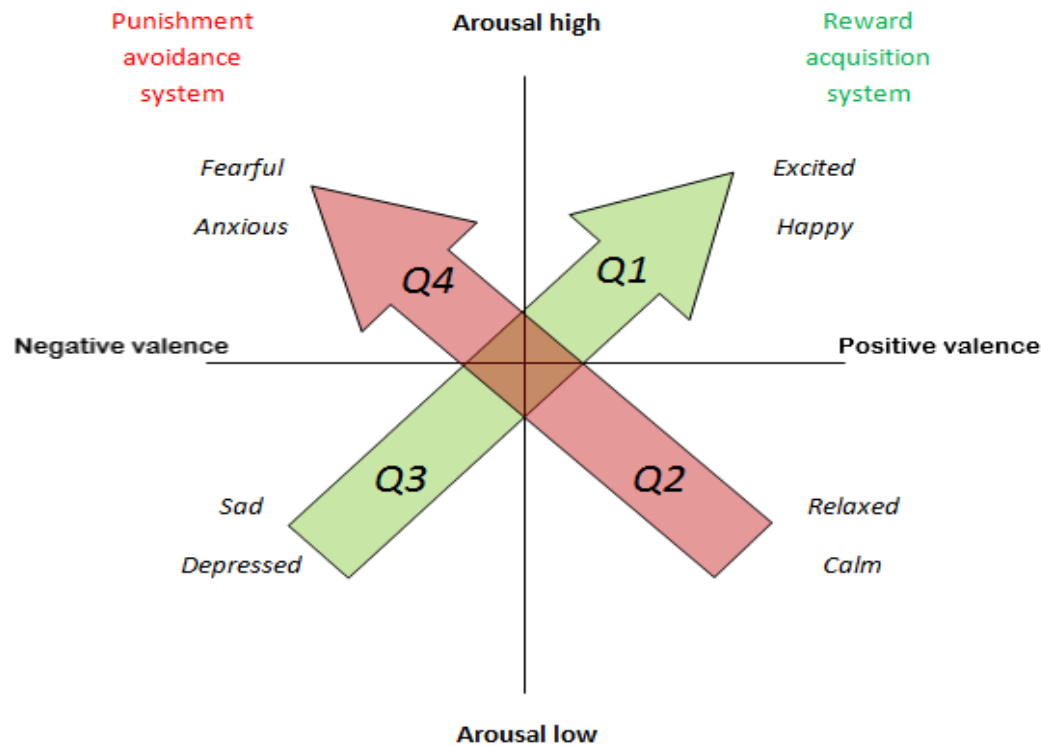
1. Sensory stimuli can unconditionally access emotional systems,
2. Emotional systems can generate instinctive motor outputs,
3. Emotional systems also modulate sensory inputs promoting incentive salience,
4. Emotional systems have positive feedback components which can sustain emotional arousal after precipitating events have passed,
5. Emotional systems can be modulated by cognitive inputs,
6. Emotional systems can modify and channel cognitive activities, thereby modulating incentive salience (Genon *et al.*, 2018; Berridge and O'Doherty, 2014; Burgdorf *et al.*, 2000; Barrett, 2017).

Understanding of the roles of specific brain regions needs to be bootstrapped as we learn more about their functional contributions, repertoire and their interactions with other regions (Barrett, 2017; Pessoa, 2022).

1.2.2 Valence and arousal

The modern definition of core affect was developed by Wilhelm Wundt, referring to the change in the internal environment experienced as feelings of valence and arousal (Blumenthal, 2002), the key underlying dimensions of affective states (Mendl and Paul, 2020b). Affect can be considered a constant current through life, ebbing and flowing, rather than turning on and off (Paul *et al.*, 2020), however we cannot be certain that animals share identical emotional states to humans, experience them subjectively or have the capacity for conscious experience of them (Mendl and Paul, 2020b). Therefore, the use of human-centric emotion words such as happiness, fear and anxiety can prove problematic when lay-usage may convey a presumption around conscious experience of emotional states (LeDoux, 2017) and therefore Mendl and Paul (2020b) suggest adding the suffix ‘-like’ to emotion words (for example joy-like) to avoid the presumption of subjective experience of these emotions in animals.

Feelings of pleasure and displeasure are summaries of the body budgeting state, i.e. the sum of the brain tracking interoceptive feedback, with unpleasant affect equalling an unbalanced body budget and vice versa. The term ‘affective niche’ refers to a short-term feeling that is experienced and contributes to longer term affect, and the affective spectrum encompasses intensities of negative to positive affective state with neutral affective state purportedly in the middle of these opposing valences (Gasper *et al.*, 2021). The affective circumplex originated by Russell (1980) and supported by Mendl *et al.* (2010a) does not position neutral affect because Mendl asserts that emotional experiences are valenced and are therefore perceived as positive or negative, rewarding or punishing, pleasant or unpleasant and as such neutral states are not emotional states (Figure 1.2a). In addition, emotional experiences also range in their activation or arousal level allowing for many variations within an affective state in each quartile (Figure 1.2b).



a



b

Figure 1.2a Core affect represented in a two-dimensional model. Italics denote discrete emotions with Q1 and 2 positive affective state and Q3 and 4 negative affective states. Arrows denote presumed biobehavioural systems associated with reward acquisition (green) and punishment avoidance (red) (adapted from Mendl *et al.* (2010a), Russell (2003) and Burgdorf and Panksepp (2006)). Figure 1.2b The 12-point Affect Circumplex (adapted from Yik *et al.* (2011) and Russell (1980)).

Also of note is that there can be a range of emotions within each quartile depending on the stimulus involved, as such a highly aroused, negatively valenced state may give rise to a flight reaction or a fight reaction. Barrett (2017) disagrees with Mendl *et al.* (2010) and suggests that neutral affect is commonly experienced and likens it to sitting quietly. The consideration of neutral affect is therefore an aspect put forward for discussion within the Delphi consultation of experts reported in Chapter 3 and within later discussion in Chapters 6 and 7.

Cabanac (2002) proposes that emotion is a mental experience with four dimensions T, X, Y and Z axes expressed as:

$$\Psi = f(x[t], y[t], z[t])$$

Where the X-axis is the qualitative dimension (the emotion state e.g. fear), a non-parametric infinite number of discrete items; Y-axis the intensive dimension (urgency), a quantitative parametric variable starting from zero; Z-axis the hedonic categorical dimension (pleasant/unpleasantness and indifference) and T-axis, the time dimension where emotion is limited in the time that follows the stimulus that aroused it. With these suppositions in place Cabanac goes on to assert that an emotion can only be identified as such if it has high intensity and high hedonicity but may have a short or long duration. This model incorporates into one concept the multidimensional aspect of emotion and includes analysis of the intensity of emotion as a fundamental aspect.

Within this thesis the core affect model proposed by (Mendl *et al.*, 2010a) is used within evaluations of potential affective states induced through experimental studies and the concept of neutral affect is explored within a Delphi consultation of experts.

1.2.3 Studying negative and positive affective state

Negative and positive affective states are described as broad mood factors (overarching emotional states) that are the dominant dimensions in an individual's mood (Mendl and Paul, 2020a). They can be measured as states or events and as such can be both long- and short-lasting (Mendl and Paul, 2020b). Whilst it is often purported that positive and negative affect exists solely on one axis, they can be distinctive orthogonal dimensions independent of each other (Watson *et al.*, 1988) encompassing a wide range of different combinations of valence and arousal states that an individual may experience. However, the terms positive and negative affect have distinct definitions within the human literature as defined below (Zevon

and Tellegen, 1982) and these definitions allow us to pair behavioural outputs indicative of positive (e.g. play) or negative states (e.g. pain) with the definitions given below to enable proxy identification of affective state.

Negative affect is defined as

“...a general factor of subjective distress...subsuming a broad range of negative mood states including fear, anxiety, hostility, scorn and disgust. Mood states related to depression such as sadness and loneliness also have substantial loadings on this factor. At the trait level, Negative Affect is a broad and pervasive predisposition to experience negative emotions”

In contrast positive affect is described as

“...a dimension reflecting (an individual's) level of pleasurable engagement with the environment. High Positive Affect...reflects enthusiasm, energy level, mental alertness, interest, joy and determination. Trait positive affect is a corresponding predisposition to positive emotional experience; it reflects a generalised sense of wellbeing and competence”
(Watson et al., 1988)

Fraser and Duncan (1998) suggest that, in animals, affective states are motivational and may be negative motivational affective states (-MASs) such as fear or thirst, or positive motivational affective states (+MASs) such as enjoyment of eating or playing. They define negative MASs as evolving from need states, to preserve fitness and developing in response to changes within the body and external environment, whereas positive MASs evolve in opportunity situations, where an action becomes advantageous, such as play, because the fitness cost of performing it has decreased. They note that some behaviours may be motivated by both positive and negative affect, for example food consumption may be motivated by the enjoyment of food but also by the need to eat due to hunger.

Cognitive-affective models suggest that trait emotionality (of which positive and negative affect are two aspects) contribute to the development, maintenance and recurrence of short- and long-term affective states with affect as the mediating variable among stimuli, the cognitive process and response behaviour (Martínez Caro and Martínez García, 2007). In addition, research into trait affectivity as a temperamental basis of response styles indicates that in addition to positive or negative affect, individuals can also be predisposed to

‘dampening’ or ‘enhancing’ traits whereby they are more likely to alter responses to events or states to heighten or subdue them. For example, research in humans has shown that individuals with a lower tendency to experience positive emotions are also less likely to employ enhancing responses when faced with positive emotions or events (Nelis *et al.*, 2011).

Historically models of affect in animals were based upon controlled tests, often over long periods of time, which yielded interesting results however these results, and particularly treatments to alter affect arising from the research, did not show efficacy when transferred to human patients (McArthur and Borsini, 2006). This may have been related to differences in cognitive ability or the methodological design of the experiments. Over time methodologies were adapted to allow for more rapid and repeatable tests, however these were still conducted under controlled laboratory conditions rather than being ethologically informed models and as such improved animal models to study affect and affective disorders became a focal area of study (Nestler *et al.*, 2002)

Modelling of negative affective states in animals has been researched for decades, with animal emotions considered as reliable models for human emotions such as depression and anxiety. However, research to develop models in guinea pigs and gerbils did not yield satisfactory results and as such an equine model for depression was proposed (Fureix *et al.*, 2012). The premise of this model was that an ethological approach could be taken to focus on spontaneous behaviours in home environments where equines often live in conditions comparable to human environments that are known to induce depression, such as social and spatial restriction, work stress due to type of work and interpersonal relationships at work. Studies in other species have demonstrated characteristics of ‘apathy’, ‘withdrawal’ and ‘inhibition’ in pigs, sheep and dogs living in social or spatial confinement or subjected to repeated aversive stimuli including chronic stress (Broom, 1987; Fordham *et al.*, 1991; Beerda *et al.*, 1999). Previous studies on horses have demonstrated ‘apathetic’, ‘unresponsive’ or ‘switched-off’ horses in restricted domestic environments or under harsh conditions in lower economic countries (Hall *et al.*, 2008; Burn *et al.*, 2010).

Perea *et al.* (2012) studied the effect of genetic and environmental influences on negative affectivity and corroborated results found by previous researchers that genetic influences combined with environment interactions do predict susceptibility to depression and anxiety in humans (Caspi *et al.*, 2003; Benedetti *et al.*, 2011) and that stress and trauma during early

developmental periods is strongly associated with long lasting negative affectivity (Duman and Monteggia, 2006; Schumacher *et al.*, 2005). Within their study Schumacher *et al.* (2005) investigated the effect of various gene polymorphisms and stress during various life stages and reported a correlation between different allele carriers and negative affectivity at different life stages via multiple logistic regression analysis. Considering the similarity in neural anatomy and emotion generation between humans and non-human animals it is purported that studies of this type have considerable cross-species transference (Panksepp *et al.*, 2017), a rationale which is supported by the work of Fureix *et al.* (2012). Indeed, Cockburn *et al.* (2018) found evidence of negative affective state in Cavalier King Charles Spaniels with Syringomyelia, similarly to Browning *et al.* (2017) who detected negative affective state in rats following concussion. Fureix and Meagher (2015) purported that those animals in low-arousal negative affective state display low activity and increased energy conservation due to perception of loss or lack of reward.

In recent years a research focus around positive affective state in animals has emerged. Wemelsfelder *et al.* (2001) and Burghardt (2005) pioneered the development of approaches to study positive affect in animals. The Boissy *et al.* (2007a) review highlights that good welfare can primarily be assessed by presence of positive experiences, rather than just the absence of negative experiences. The now widely accepted premise that animals have affective lives has invigorated research to investigate the link between the ability to experience and express positive emotions and more persistent positive affective states, such as happiness. To date research into positive emotions in animals has utilised three approaches:

1. Dimensional approaches (pleasant/unpleasant valence and weak/strong excitation),
2. Mechanistic approaches focussing on eliciting mechanisms (neural circuit and basic emotion models),
3. Componential approaches addressing behavioural components (posturing and activities), autonomic components (visceral and endocrine responses) and subjective components (emotional experience or feeling) (Désiré *et al.*, 2002).

Alexandrov and Sams (2005) unified concept of consciousness and cognition and the link between these facets, reviewed by Mendl and Paul (2004), highlight the importance of cognitive science for research on animal emotions and thus animal welfare. Research has

shown that causal links between emotions and cognition are bi-directional in that cognitive manipulations can influence felt emotion and emotional manipulations can influence cognitive functioning (Paul *et al.*, 2005), however it is still unclear the extent to which animals cognitively experience emotions (Mendl and Paul, 2020b).

Methodologies to measure affective state can be developed using a multimodal componential approach via three pathways to assess positive emotions from neurological, behavioural and physiological perspectives whilst considering the dimensional approaches. The factors outlined in Figure 1.3 can be considered within methods to explore affective states.

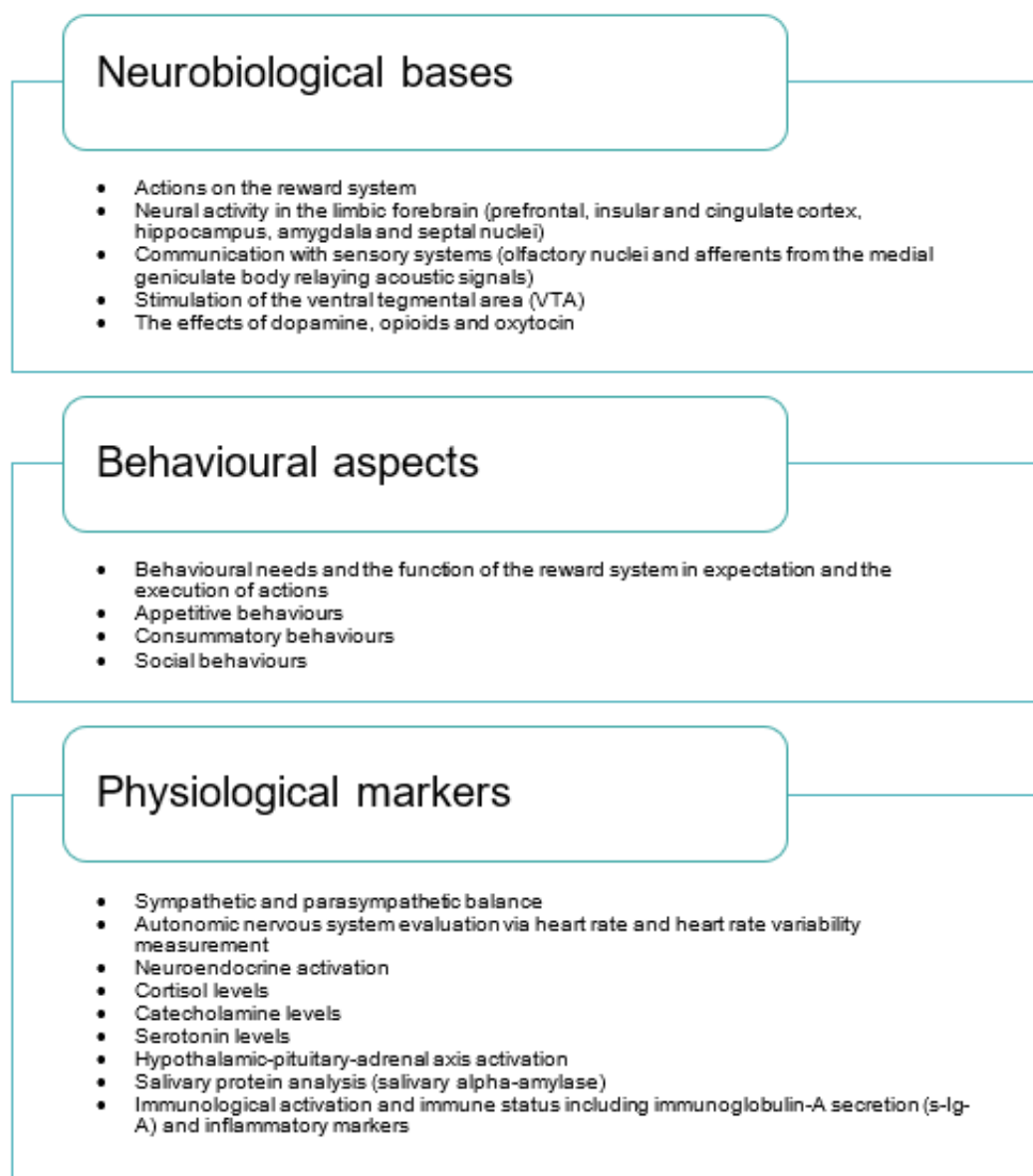


Figure 1.3 Neurological, behavioural and physiological measures that may be utilised to investigate affective state (Boissy *et al.*, 2007a).

1.3 Inducing affective states

To be able to measure affective state we must be able to reliably induce it; therefore, it is important that we identify species specific stimuli that are likely to induce an instance of emotion or a longer lasting state. The ethology of the species combined with past experience, current environment and internal state will contribute to whether a stimulus is perceived as positive or negative in that instance and will influence the emotional event or state produced. Within this thesis expert opinion on methods to induce positive affect are discussed (Chapter 3) and perceived optimal methods taken forward to experimental studies (Chapters 5 and 6).

1.3.1 Creating instances of emotion

An instance of emotion occurs when the brain makes meaning from internal and external stimuli, constructing an instance of that concept at that time; essentially the mind is making sense of the body within the context of the world around it (Adolphs *et al.*, 2019). Emotions are not passive reactions to the world and the being experiencing the emotion is an active constructor of their own emotions; using previous experience and sensory input the brain constructs meaning and prescribes action (Mellor, 2012; Barrett, 2018). Subjective experiences including affective states are often initiated by exteroceptors such as those found within the eyes, ears, nose and skin which identify stimuli and elicit a situational perception related to potential threats or positive interactions (Mellor, 2012). Statistical learning affects responses to stimuli based on these prior experiences and their association with the potential for reward or punishment. The experience of emotion is continually evolving as new experiences add further detail to the previous construction, a phenomenon termed ‘the remembered present’ (Edelman, 1989). Construction occurs via three main pathways:

1. Social - the culture and concepts surrounding the individual,
2. Psychological - the core brain and body systems, and
3. Neuro-construction - pathways created through experience and learning.

The combined effect of the genetics of the individual and its experiences contribute to plasticity and enable the development of construction (Barrett, 2017).

Pleasure and displeasure feel qualitatively different however what is pleasant to one being is not necessarily pleasant to another. These feelings of pleasure and displeasure are produced through interoception; the brain makes representations of all the sensations it is receiving

from the body. The intrinsic brain networks interact with complexity, like a sports team with substitutes, different areas of the system (or players in the team) are used for different tasks but always within the network system; these networks produce every sensation experienced by the individual however component parts of the network cannot necessarily be understood individually because these interactions produce new properties that are not present in the constituent parts alone (Barrett, 2017). Fear, for example, cannot be reductively deduced to one bodily pattern, rather it is an interactive and context-based phenomenon. Likewise, many combinations of neurons can produce the feeling of fear; this is advantageous as it increases robustness (Whitacre and Bender, 2010), but can also contribute to heredity of traits, which may be beneficial or detrimental to the animal (Turkheimer *et al.*, 2014).

1.3.2 Making decisions, predictions and prediction error

Decision making requires an individual to have a neural representation of the world including its own state and that of the environment around it. Within decision making a probability analysis occurs considering the potential actions that could be taken and their relative value or benefit to the individual, guided by reinforcement learning (Redish, 2013) discussed in section 1.4. Decision making occurs through two processes, model-free (habitual) and model-based (goal-directed) control. Model-based control requires higher levels of information processing and working memory capacity because the animal must simulate a model of the current environment and utilise a decision tree process to identify the optimal response to make. In humans this requires conscious reflection; in animals, whilst the neural processing may be similar, it is unknown whether this is accompanied by a conscious experience of future planning. However, there is some evidence to suggest that this may be the case (Raby *et al.*, 2007; Clayton and Dickinson, 1998; Clayton and Dickinson, 1999) as has been evidenced in jays but not across mammalian species to date (Clayton *et al.*, 2003). Model-free (habitual) control does not require simulation of the environment, instead it utilises decision-making based upon the value of actions in the current scenario and experiential learning from the past. Due to the lack of internal model of the world, information regarding the effect of decisions made is not computed meaning that new learning only occurs when the value of the new outcome is experienced making this type of control less flexible within environmental changes (Mendl and Paul, 2020a).

When processing stimuli and computing emotional and physiological responses the brain makes millions of micro predictions every second, based upon sensory stimuli and prior experiences; memory works in a similar way. Within the brain many more neurons are recruited to carry these predictions than are carrying sensory information (Bär *et al.*, 2016). Sometimes prediction errors occur, however, in most instances the brain quickly adjusts its prediction by comparing the prediction to the sensory input and correcting the error. At any one time there are multiple prediction loops running in parallel within the brain, communicating with other networks (Joiner *et al.*, 2017). From a brain locus perspective, the rest of the body is part of the external environment which it must interpret. The brain 'control network' helps to resolve uncertainties in the noise of competing predictions based on all the current sensory inputs (Lynn and Bassett, 2019; Gu *et al.*, 2015). The term control network can be considered as an 'optimiser', which amplifies or subdues the firing rate of neurons to move sensory input into and out of the attentional spotlight of the individual (Barrett, 2017). Via this method the brain can whittle down the most likely 'reality'; a process which aids body budget, enables stable perception and facilitates the launching of actions (Bullmore and Sporns, 2012). The control network can help to differentiate between emotion and non-emotion (for example anxiety and nausea), different emotion concepts (differentiating between excitement and fear), different goals for emotion (flight or fight) and different instances (to vocalise or not, for example) (Pessoa, 2017). Barrett (2017) likens the relationship between the control network and interoceptive network to large airport hubs with information passing efficiently between the different brain networks (airport hubs). The hubs synchronise the brain's activity (much like the airport hubs synchronise flight activity) and may be a prerequisite for consciousness as proposed by Swanson (2012) (Figure 1.4).

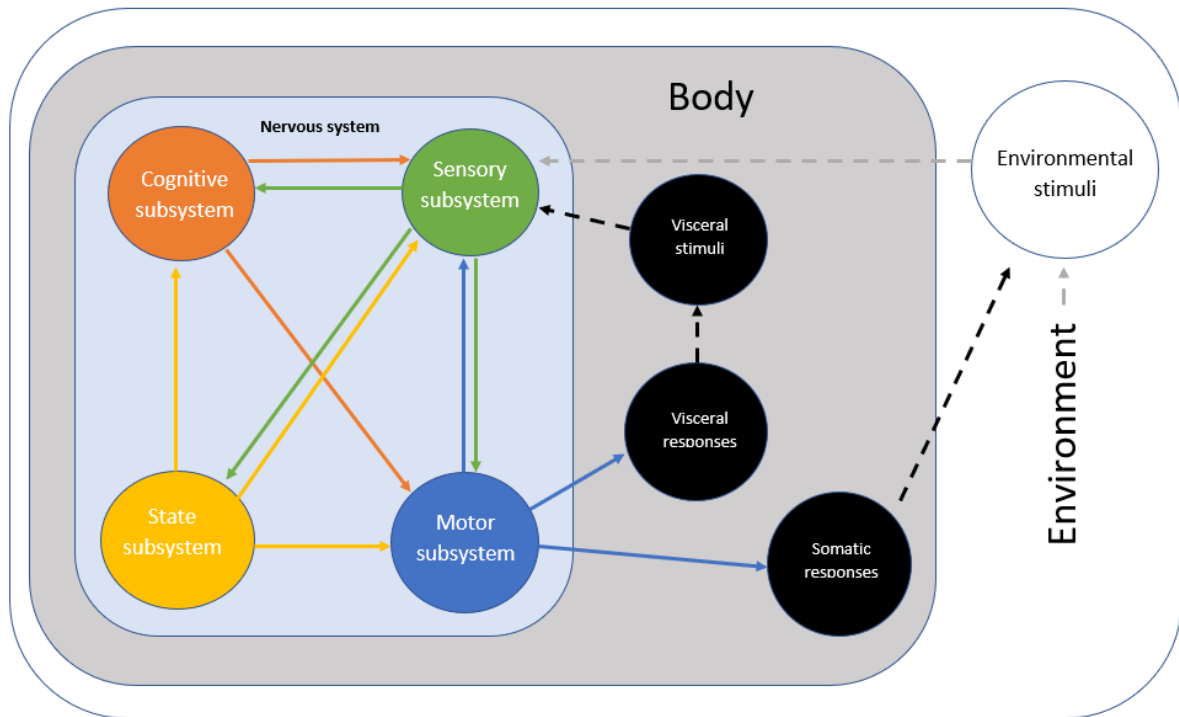


Figure 1.4 Schema of the four-systems network control model proposed by Swanson (2012). One system (motor) controls somatic responses (i.e. behaviour). The motor system, in turn, is controlled by three inputs from the sensory, cognitive and behavioural state systems. Direct sensory inputs to the motor system mediate reflexive behaviours, inputs from the cognitive system mediate voluntary behaviour and influences from the state system mediate state control influences. The motor system influences the external environment via skeletomotor changes and physiological responses within the body (visceral stimuli and responses). These influences on the external and internal environment (dashed black lines) provide feedback signals to the nervous system (dashed light grey line).

Interoceptive predictions occur via the process described in Figure 1.5. The body budgeting region sits within the interoceptive network and does not react in emotion, instead it predicts, intrinsically, to regulate body budget. The brain predicts that the body will need energy and the body predicting regions instruct the adrenal gland in the kidneys to release cortisol. Cortisol, in itself, is not a stress hormone, instead it causes the bloodstream to be flooded with glucose to provide immediate energy to cells, this happens to include at times of stress (Sikorska *et al.*, 2023). The body reacts to the prediction, the individual experiences physiological changes and something about the situation becomes meaningful to the individual. The body can make the prediction as a result of a very small stimulus e.g. seeing a

person, causing a rush of hormones and the sensory feeling (Scopa *et al.*, 2019). No movement has occurred the body is just preparing for flight or fight. The brain predicts the body's response based on prior experience to produce an interoceptive sensation.

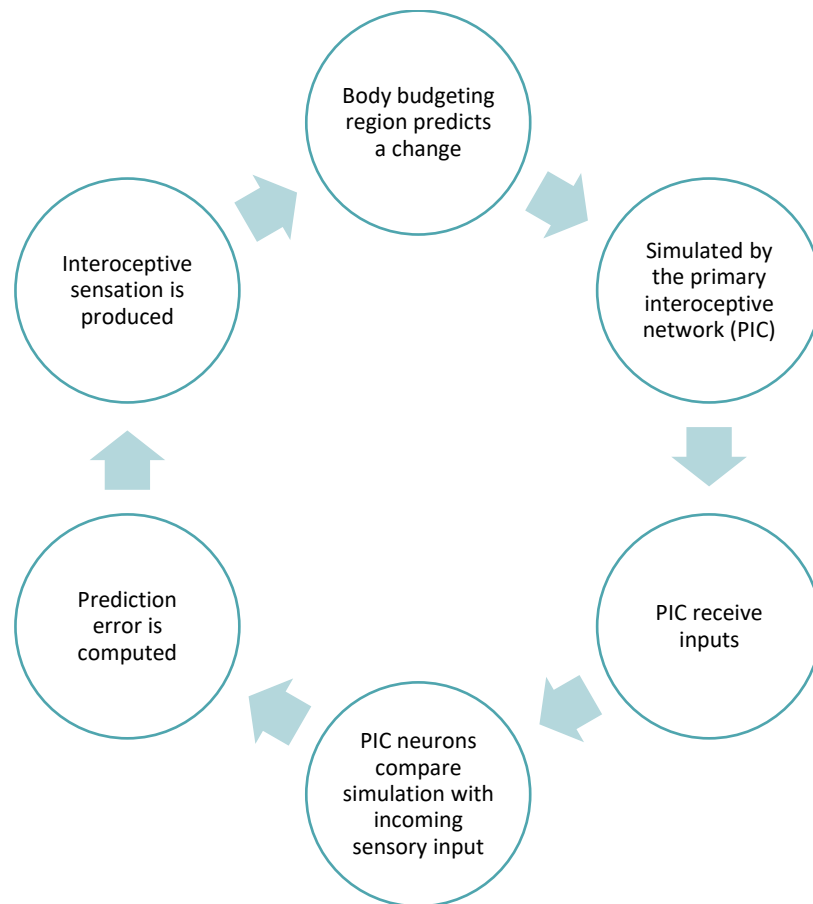


Figure 1.5 The process of interoceptive predictions adapted from Barrett (2017) and Barbas (2015).

Interoceptive predictions interplay with core affect to influence decision making. For example, an equine provided with a bucket of food in a field with other equines must decide whether to approach the food or not. This process involves the processing of information about internal state (level of hunger), mood-like core affect, sensory information from the environment and the individual's evaluation of cost-benefit based upon memory of prior experiences as demonstrated in Figure 1.6.

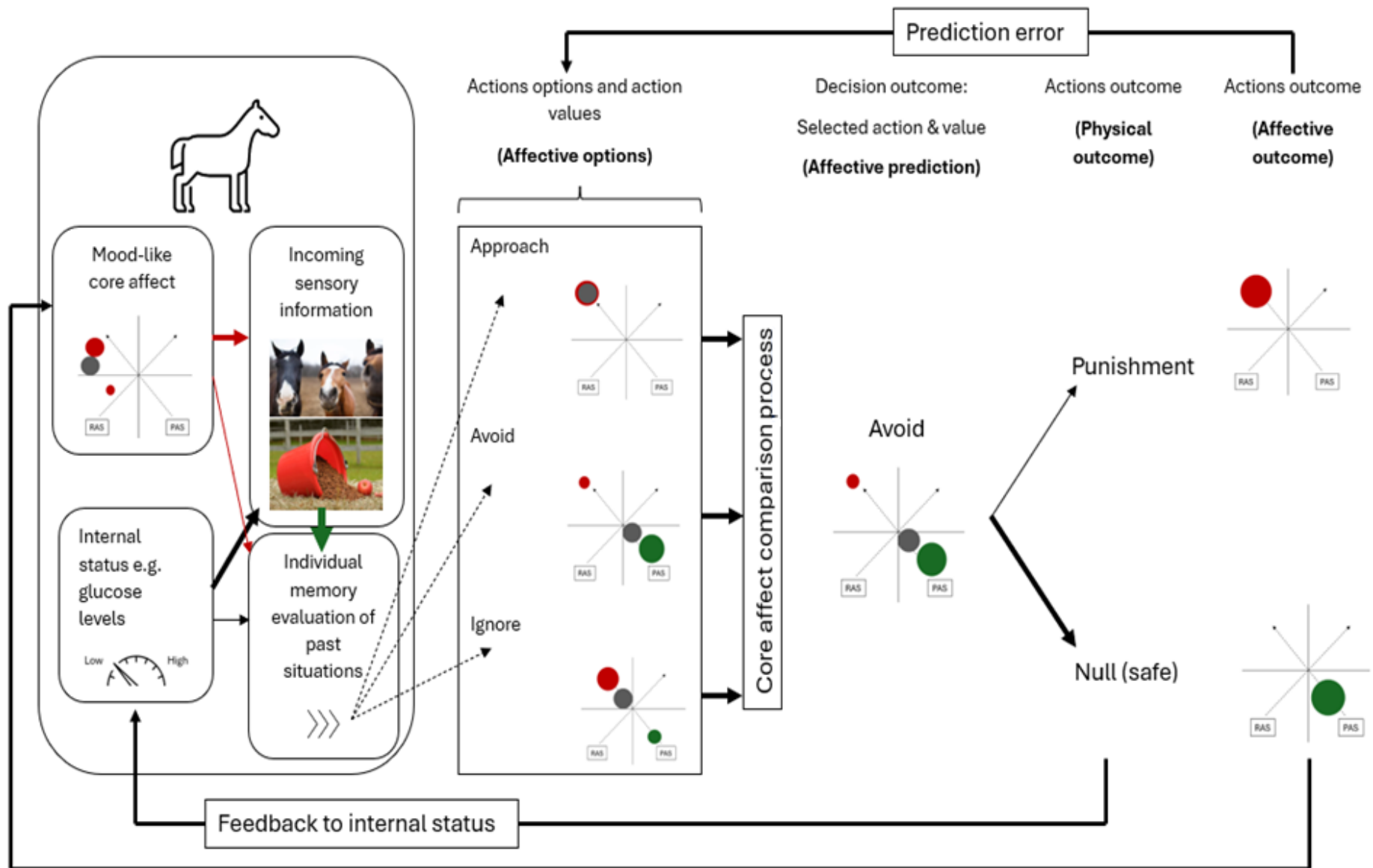


Figure 1.6 Schematic illustration of the hypothesised role of core affect in model-free decision making. A high punishment mood-like core affective state, low blood glucose and prior memories determine the horse's initial situation (left hand box). Its mood state and internal status influence attention (thick red and black arrows in left hand box) resulting in detection of new sensory information (food availability with potential conflict from other horses). The new situation is matched to memory of similar past situations. Sensory information is unambiguous (likely to provide a reliable basis for prediction of outcomes), it has a stronger influence on memory retrieval (thick green arrow in left hand box) than mood or internal status (thin black and red arrows in left hand box). The retrieved memory is associated with previous actions including evolved Pavlovian responses (action options = approach, avoid, ignore) and their outcomes (action values) which within core affect space represent affective options. Through a process of comparison, the most positively valenced option (avoid) is selected. This now becomes an affective prediction (a prediction of the affective consequence of the selected action in this situation). The physical outcome is probabilistic (thick black arrow to Null (safe)) and the associated affective outcome is determined by the direct experience of reward or punishment and their presence or absence in relation to the affective prediction. Physical outcomes feedback to influence internal status whilst affective outcomes act as prediction errors to update situation-dependent action values. They also update mood-like core affect. Within the core affect diagrams the angled arrows denote Reward Acquisition System (RAS) and Punishment Avoidance System (PAS) and the red and green circles indicate negative and positive valence respectively. Grey circles indicate a hypothetical integrated location in core affect space. Line drawings authors own, diagram adapted from Mendl and Paul (2020a).

Within this thesis concepts around reward acquisition with reference to positive affective states will be tested empirically to derive data that may add further information to the current knowledge base around these processes in equines.

1.3.3 Reinforcers and punishers in affective state

Just as discrete emotions are characterised by their underlying valence structure, which locates them in a core affective space, the same is true for internal sensations (pain, hunger, thirst) and motivations (to eat, sleep, reproduce for example) which are rewarding or punishing and as such also occupy a space in the core affect continuum. Location in core affective space acts as a common currency to enable comparative analysis and cost: benefit trade-offs to be undertaken by the animal over time, forming longer term mood states which can be evaluated by researchers via the behavioural outputs of the animal (Cabanac, 2002). From a neural perspective, interplay of the activity of the discrete emotional systems (subcortical structures such as Periaqueductal gray) and those that are activated during wider reinforcing or punishing events (the widely dispersed neuromodulatory systems such as the dopaminergic, serotonergic, noradrenergic and opioidergic systems) as well as higher cortical areas such as the anterior cingulate cortex and orbitofrontal cortex allow for distillation across time of overall positive or negative experience (Spruijt *et al.*, 2001; Panksepp, 2007). It is thought that differences in the sensitivities of these systems may underlie gross personality traits such as optimism and pessimism (discussed in section 1.6) and that these differences may be determined via several routes including genetic and experiential pathways (Corr, 2008) as indicated in Table 1.1.

Table 1.1 Proposed links between environmental conditions (reinforcers and punishers), resulting predominant core affective mood state and biases in expectation that drive optimistic or pessimistic decision making. Quadrant 1 (Q1) = high arousal in positive valence (reward acquisition), Quadrant 2 (Q2) = low arousal in positive valence (calm, content), Quadrant 3 (Q3) = low arousal in negative valence (depressed) and Quadrant 4 (Q4) = high arousal in negative valence (punishment avoidance). Adapted from Mendl *et al.* (2010a).

	Low reward-opportunity environment (leads to Q3 state)	Intermediate reward-opportunity environment	High reward-opportunity environment (leads to Q1 state)
Low-threat environment (leads to Q2 state)	Q2/Q3 Decreased expectation of negative and positive	Q2 Decreased expectation of negative and <i>optimism</i> about positive	Q1/Q2 Increased expectation of positive and decreased expectation of negative <i>Full optimism</i>
Intermediate-threat environment	Q3 Decreased expectation and <i>pessimism</i> about positive	Neutral state – no (or baseline) bias	Q1 Increased expectation of positive and <i>optimism</i> about positive
High-threat environment (leads to Q4 state)	Q3/Q4 Decreased expectation of positive and increased expectation of negative <i>Full pessimism</i>	Increased expectation of negative and <i>pessimism</i> about negative	Q1/Q4 Increased expectation of both positive and negative

1.4 The role of learning and memory in affective state

As highlighted in section 1.3, learning and memory have a significant influence on decision making in animals and in the production of affective states. Reward and punishment feed into predictions and body budgeting and are fundamental in the formation of memories. Throughout this thesis consideration of whether stimuli presented are positive or negative to

the individual, evaluated with reference to the animals internal and external environment and prior experiences, is a recurring theme within the Delphi consultation (Chapter 3) and later experimental studies (Chapters 5 and 6). These studies utilise the principles of learning theory and memory formation to attempt to reliably induce positive affective state.

1.4.1 Learning

Primal emotions (seeking, rage, fear, lust, care, grief and play) are very important aspects of animal's mental lives and enable anticipation of the future (Panksepp, 2011a; Janczarek and Kędzierski, 2011; Brucks *et al.*, 2022). These primal emotions may be augmented using the learning theory quadrants of positive punishment, negative punishment, positive reinforcement and negative reinforcement (Mendl *et al.*, 2010b). External application of reinforcers or punishers mediate learning by regulating affective neural substrates in the brain (Panksepp, 2011b). Reinforcers and punishers may be applied intentionally during training or more passively as a consequence of experiences or behaviours displayed.

Learning in animals, including equines, occurs via three learning pathways:

- Non-associative learning
 - Habituation
 - Sensitisation
- Associative learning
 - Classical conditioning
 - Operant conditioning
- Complex learning or insight (McLean and Christensen, 2017; Murphy and Arkins, 2007)

Non-associative learning is often a passive process occurring over time and is based around exposure to stimuli. In habituation the repetition of a stimulus causes it to lose its novelty. The speed at which habituation occurs is variable with more complex, unfamiliar and variable stimuli necessitating a longer habituation process. In addition, the individual characteristics of the equine (personality, prior experiences, health status) can affect the length of time to habituation and the propensity for recovery following a period of non-exposure to the stimulus (Baldwin and Baldwin, 2001). Sensitisation is effectively the opposite of habituation whereby following repeated presentations of the stimulus there is an increase in response; in

most cases sensitisation occurs when the animal is presented with an aversive stimulus and in these cases, sensitisation can override habituation (McGreevy, 2012).

In associative learning Classical Conditioning (also known as Pavlovian or Respondent conditioning) is the process by which a 'neutral stimulus' becomes a 'conditioned stimulus' because it precedes an 'unconditioned stimulus' (Baldwin and Baldwin, 2001) as shown in Table 1.2 below.

Table 1.2 The classical conditioning reflex (Baldwin and Baldwin, 2001).

	Stimulus	Response
Before conditioning - an unconditioned reflex	Food (unconditioned stimulus)	Salivation (unconditioned response)
During conditioning	Bell (neutral stimulus) + Food (unconditioned stimulus)	Salivation (unconditioned response)
After conditioning	Bell (conditioned stimulus)	Salivation (conditioned response)

Operant conditioning (also known as Instrumental Conditioning) is generally expressed in a simple $A:B \rightarrow C$ continuum. Where A = Antecedent cues, B = Behaviour and C =Consequences. Antecedent cues do not cause the behaviour, they are cues that set the occasion for the behaviour and signal the availability to give a behavioural response to achieve the consequence. The behaviour causes the consequence, and the consequence influences the frequency of the behaviour and the ability of the antecedent cues to set the occasion for that behaviour in the future (Baldwin and Baldwin, 2001). The operant (unit of behaviour) can be strengthened or weakened depending on whether the consequences are reinforcers or punishers.

Within operant conditioning there are three types of stimuli that precede and follow behaviour; these are S^R , S^D and S^A . S^R is a reinforcing stimulus that comes after the behaviour and may be intrinsic or extrinsic in form. When a behaviour is followed by an S^R both the behaviour and any antecedent stimuli are reinforced and the antecedent stimuli become

discriminative stimuli (S^D). As such the simple ABC equation is re-written as $S^D: B \rightarrow S^R$. When the behaviour (B) is followed by S^R in one context but not another the antecedent context cue becomes an S^D and helps set the occasion for future responses. When a behaviour is not followed by reinforcement the stimuli that best predict non-reinforcement become S^Δ , that is, discriminative stimuli that inhibit responding. They signal that the behaviour is not likely to be reinforced. In simple terms S^D are the green lights and S^Δ are the red lights (Baldwin and Baldwin, 2001).

Reinforcement and punishment, whether conscious, unconscious, extrinsic or intrinsic, influence the psychological affect of the equine and the associations it makes with individual environments, people, practices and other animals. These associations become discriminative stimuli helping the equine to predict situations which may be pleasurable or unpleasurable and enabling action to be taken by the individual to actively engage or retreat from situations as the discriminative stimuli become apparent (Mendl and Paul, 2020a). These behavioural actions are performed daily by equines in all environments to ensure survival and facilitate a comfortable existence. In feral and semi-domesticated equines these behaviours are often performed from a survival perspective to see off threats from other herd members or animals of other species and to ensure access to resources such as food, water and mates. In domesticated equines the actions are much more frequently observed in response to different stimuli such as confinement, companionship availability and interactions with humans. Many of the behaviours seen in domesticated equines which are deemed to be problem or negative behaviours are normal feral equine behaviours performed in a different (domesticated) context where they are often perceived as unacceptable by their owners or carers (Goodwin, 2007; Hockenhull and Creighton, 2014). The production of avoidance and defensive behaviours are normal, biologically essential survival behaviours in equines and in many cases they are performed in response to a situation or procedure undertaken within the horse-human dyad for which there is no naturally occurring analogue available for reference within the horse-horse dyad, making the process a behaviourally and biologically foreign one for the equine (McGreevy *et al.*, 2009; Hockenhull and Creighton, 2013).

Insight learning is the process whereby equines learn in a more complex manner than just responding to single cues or making simple associations. In these more complex processes, the equine utilises insight (the capacity to observe and understand relationships or patterns

in the environment and to apply this understanding to new situations) to arrive at a solution without the use of trial-and-error learning. Whilst the capacity of equines to undertake insight learning is still debated, they have capacity to utilise general as well as specific features of their environment as cues in discrimination learning (Cooper, 1998). Likewise, social learning in equines is still widely debated. Rørvang *et al.* (2018) conclude that learning which is often deemed to be undertaken through social learning processes (goal emulation and imitation) may in fact occur through social transmission *id est* through social facilitation, local and stimulus enhancement processes; a conclusion subsequently supported by McVey *et al.* (2018). This topic is important as assumptions regarding the capacity of equines to engage in social learning could have wide ranging welfare implications especially in relation to the perceived mimicking of repetitive behaviours and consequent isolation of such equines from conspecifics to avoid ‘passing on’ of the behaviour to others.

There is surprisingly little research into basic learning abilities such as habituation and sensitisation in equines however more literature can be identified on themes around operant conditioning including reinforcement schedules, avoidance responses and discrimination learning (Nicol, 2002). Research has shown that horses have the cognitive capability to form categories (Hanggi, 1999) but their capacity is thought, thus far, to be less than that of primates and dolphins (Leslie, 2001) and their ability to form concepts is still a topic of enormous theoretical interest (Lea, 2010; Nicol, 1996; Flannery, 1997).

Within the sphere of measuring animal emotion, it is important to consider the variance with which individual animals may perceive a situation or stimulus and the resultant decrease in consistency of the behavioural and physiological responses detected (Forkman *et al.*, 2007; Panksepp, 2006; Davis *et al.*, 2014). To combat this Mendl *et al.* (2010a) suggests formulating *a priori* hypotheses utilising their framework for the types of stimuli that will elicit affective states in each of the four quadrants, a factor studied at length within this thesis. The hypothesis is very closely aligned with the four quadrants of reinforcement and punishment as demonstrated in the Figure 1.7.

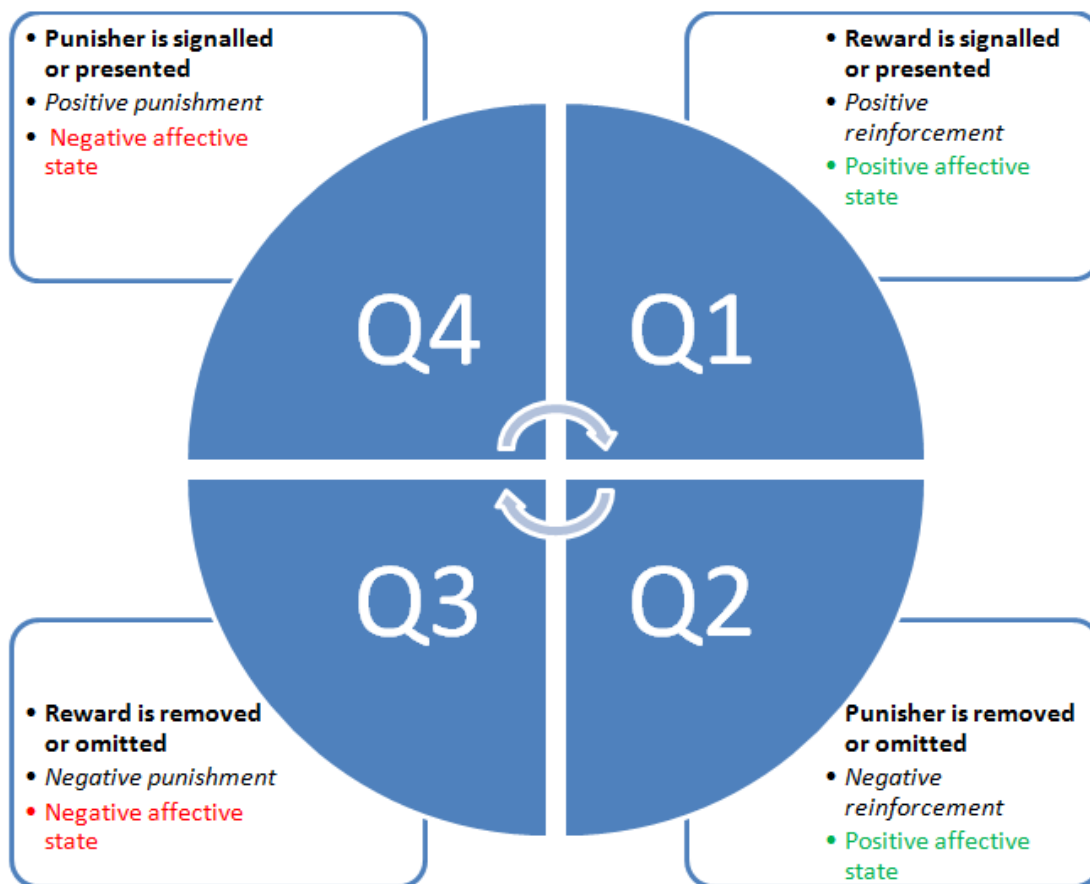


Figure 1.7 Depicting the alignment between Mendl *et al*'s. (2010) framework and the quadrants of reinforcement and punishment in behavioural theory. Quadrants 1-4 are identified along with Mendl *et al*'s. predicted events/stimuli that will generate affective state (**in bold**), the affective state achieved (red or green text) and the applicable quadrant of reinforcement or punishment (*in italics*). Authors own.

Throughout this thesis quadrants one and two are explored with reference to inducing positive affective state by the provision of environmental stimuli that are predicted to elicit positive emotions in equines. During the measurement of the consequent behavioural and physiological outputs consideration is given to the possible elicitation of negative emotional states and as such the statistical modelling used is designed to tease apart data that potentially indicate the presence of negative versus positive emotional state as well as considering whether arousal is high or low in specific scenarios.

1.4.2 Memory

It is well established that memory plays an important part in decision making and has an impact on the core affect of the animal (Mendl and Paul, 2020a). Equines have long been

identified as mammals with excellent memories (Nicol, 2002; Hanggi and Ingersoll, 2009; Sankey *et al.*, 2010). Early work into equine memory identified that equines could learn their way through mazes within 1 week of training (Marinier and Alexander, 1994) and remember operant responses to behaviours such as opening food chests, even performing the response more quickly after a month's interval from the learning of the behaviour (Wolff and Hausberger, 1996).

Over the past decades a larger body of work has been produced in the field of equine memory, this has involved research into both short- and long-term memory and into task specific memory as well as emotional memory. Hanggi (2010) conducted a study whereby horses' memory responses to feed delivered into one of two feed receptacles was tested. The horses were released to identify the correct feed containing vessel under a variety of contexts (immediate release and delayed release of 5, 10, 15, 20 and 30 seconds). Results showed that horses remembered the correct vessel in all contexts and observational data identified that the horses were using reliance on memory rather than orienting cues to identify the correct response. In addition, as previous trials in the study were rewarded or not rewarded in various locations, horses could not rely on outcomes of earlier trials or locations but had to rely on memory to achieve the task. This study was supported by the findings of Murphy (2009) whose horses correctly responded in a similarly structured test but was contradicted by previous studies including that of McLean (2004) where the horses failed the 10-second delay test, although the methodology was significantly different with no other delay tests or instant release tests performed. The methodological design of these tests appears important in success rates, in that tests where horses were successful were those where the horses were presented with the tests in their normal environments with access to external stimuli. During the tests undertaken by McLean the horses were placed in areas devoid of external stimuli and the author reported that some horses appeared anxious during the testing process indicating that perhaps stress has an impact on memory function (McLean, 2004) or potentially that the horse's attentional spotlight is focussed elsewhere within this environment (Barrett, 2017). There is little consensus in the literature regarding the effect of stress on learning and memory; however, more recent review studies indicate that rapid catecholamine and glucocorticoid actions appear to interact with the basolateral amygdala to facilitate the storage and retention of stressful experiences into long-term memory however

the timing of the stressor is important. If the individual is exposed to stress before learning this can alter encoding of memory which, depending on the circumstance, may enhance or impair memory and this is mediated by the valence of the situation. Consolidation of memory is often enhanced by mild stressors however retrieval of memory tends to be impaired by stress (Schwabe *et al.*, 2012) which may explain the outcome in the McLean trial. In addition, during the McLean trial horses were restrained by handlers during the delivery of the stimulus whereas in the Murphy and Hanggi trials there was no restraint undertaken and minimal human observers were used indicating that this may have contributed to stress in the McLean trial and led to biased responses as a result of unintentional human gestural cues which have been shown to alter equine responses (Proops and McComb, 2010).

Valenchon *et al.* (2013) supported the results of Hanggi and Murphy by demonstrating that horses could correctly identify the location of food in a disappearing food test during various delays up to and exceeding 20 seconds however they also went on to investigate the effect of the addition of an acute stressor, during the test, on memory function. Their data showed a decrease in memory response to less than 12 seconds indicating a stress induced memory impairment. Also identified was the effect of temperament on memory response (as previously identified in corvids (Range *et al.*, 2008; Scheid and Noë, 2010); primates (Hare, 2007); and horses (Lansade and Simon, 2010; Valenchon *et al.*, 2013), with more fearful horses performing poorly on memory tests when stressors were involved than horses categorised as having a less fearful temperament, as previously identified in rodents (Brinks *et al.*, 2007) and humans (Sorg and Whitney, 1992). Valenchon *et al.* (2013) discussed their data in the context of the Yerkes-Dodson law of stress and performance which states that, at low levels, stress can increase performance but as level of stress increases cognitive performance decreases. The impact of this on fearful and non-fearful temperament equines means that all equines are likely to suffer cognitive impairment under stress; however fearful horses are likely to show greater impairments than non-fearful ones; therefore personality is a facet that should be considered within experimental design of studies intending to measure equine performance in learning and memory tasks.

Van Den Berg *et al.* (2016) reported that horses primarily rely on sampling behaviour over spatial cue related memory in a forage preference test scenario with orosensory cues superseding the spatial cues through a process of biological associative importance, a facet

that is considered within the experimental design in Chapter 6. The authors identified however that the flexibility of memory circuits in equines appears to be dependent on the variability of the environment and stimuli presented providing implications for further equine studies investigating working memory in training contexts. Similarly, Whishaw *et al.* (2009) identified that horses use place-object memory to provide guidance during movement with the memory of object placement lasting for over 15 minutes in a trial situation and for hours if the item contains further sensory information (Whishaw and Burke, 2021). In addition, the researchers found that place-object memory overshadowed working memory in the test with horses stepping higher over places where objects had previously been placed even when no object remained in place.

Another consideration in training is the memory processes within the use of a secondary reinforcer in equines. Lansade and Calandreau (2018) found that a secondary (conditioned) reinforcer such as an auditory signal did not help to maintain a response to a cue and as such did not alter the process of extinction in a learned task. Many equine trainers use conditioned secondary reinforcers such as verbal cues (good, yes) and clickers to reinforce behaviour; this study indicates that whilst the auditory response to a behaviour may be an effective marker of the performance of the correct behaviour it may, in itself, not be effective as a reinforcing stimulus thus presenting a challenge to some training techniques and their efficacy.

Long-term memory is an adaptive process which avoids the need for repeated information processing of the same stimuli and engages higher order cognitive skills which are advantageous as they enable novel problem solving without the need for prior experience (Valenchon *et al.*, 2013; Reichmuth Kastak and Schusterman, 2002). Studies into long-term memory in animals are not extensive in any species due to limitations of maintaining testing conditions and available subjects over a long period of time and this is certainly true of equine research into long-term memory (Gabor and Gerken, 2018). In recent years several controlled experiments have demonstrated that equines can remember learned instrumental and spatial tasks after prolonged periods of weeks (Marinier and Alexander, 1994; Wolff and Hausberger, 1996), months (Sankey *et al.*, 2010) and years (Valenchon *et al.*, 2013). Long-term memory of complex higher-order tasks has not been well studied in horses however Giebel (1958) and Dixon (1970) reported that horses could remember correct answers for multiple, two-choice, discrimination tasks for several months following exposure. Hanggi and

Ingersoll (2009) reported long-term memory of categorisation and concept learning lasting for over ten years in some horses and for a minimum of seven years for all in the cohort.

Gabor and Gerken (2018) conducted a trial into higher cognition capabilities in Shetland ponies with ability to remember tasks tested one year after initial learning of the tasks. The tests involved a set of motor skills which had to be performed in sequence and then a quantity discrimination task based on a matching to sample design. The ponies initially learned and successfully completed both parts of the test however when tested one year later the ponies immediately completed the motor skill memory test but could not reliably complete the matching test. The authors purported that this may be due to memory formation being very complex involving multiple memory systems that are encoded, stored and recalled in different ways. The ponies appeared to be successful at the non-declarative or procedural memory tasks involving motor skills which are dependent on the cerebellum and basal ganglia but less successful at the spatial and declarative learning tests requiring hippocampal involvement. Aspects of this study contrast with that of Hanggi and Ingersoll (2009) who confirmed long-term memory for complex tasks in equines, however what was clear from the findings was that horses tend to naturally orientate towards procedural problem-solving rather than complex learning to solve memory tasks and gain a reward where available as a strategy.

What is apparent across numerous studies is the effect of emotion on learning and memory and the effect of prior learning and memory on consequent affective state. Both positive and negative associations with events and tasks are shown to heighten memory (Sankey *et al.*, 2010; Waring, 1983; McGreevy, 2012) due to mediation by the amygdala, nucleus accumbens, prefrontal cortex and dopaminergic afferents to all three regions (Cardinal *et al.*, 2002). Heleski and Bello (2010) identified that horses exposed to a leading learning task with high levels of positive reinforcement remembered the task and completed it more calmly than those exposed to learning via standard negative reinforcement protocols for leading. This study agrees with Sankey *et al.* (2010) who found that positive associations during training led to long-term memory formation and further positive interactions with the initial stimuli; but crucially also to novel stimuli within the same overall group of stimuli. Whilst the evidence for heightened learning and memory in situations of high emotional arousal is building there is however only scant evidence that horses learn and remember poorly under passive

conditions and as such a direct and clear comparison has yet to be made (Murphy and Arkins, 2007).

Within the experimental design of the field trials reported within this thesis consideration is given to prior learning, past experiences and memory formation and how these factors may have affected the responses seen in the horses included within the experimental protocols. Physical and affective action outcomes are assessed in light of the information gathered for each individual equine within the sample with the aim of providing insight into the affective predictions generated by the group and consequently how the experimental interventions alter the animals positioning in core affective space.

1.5 Measuring affective states

As discussed in section 1.2, behavioural and physiological measures have long been used as proxies for identification of affective state in animals. These outputs can be considered in terms of opportunities for pleasure (positive affect) or avoidance of threats (negative affect) (Mendl *et al.*, 2010b). The plethora of literature investigating behavioural and physiological markers of negative affective states in equines has provided scientists with a wide array of proxy measures for negative core affect; however there is a much more limited selection of validated proxies for positive core affect in equines and as such this thesis seeks to explore further potential behavioural and physiological markers of positive affect through a consultation of experts (Chapter 3) and via experimental trials (Chapters 5 and 6).

1.5.1 Behavioural measures to evaluate positive affect in equines

Assessment and monitoring of animal welfare through the study of positive behavioural expressions such as play behaviour, affiliative behaviour, self-grooming, vocalisations and information gathering activities is suggested to be a valid process to provide for optimal animal welfare (Boissy *et al.*, 2007a). These behaviours are indicatively positive behaviours mediated by the seeking, care, and play systems (Panksepp, 2011a) and therefore unlikely to be identified in animals in negative affect (Harding *et al.*, 2004; Mendl *et al.*, 2009; Mendl *et al.*, 2010b; Mendl and Paul, 2020b; Paul *et al.*, 2005).

Hall *et al.* (2018) published a review on assessing equine emotional state which presented current evidence on the physiological, behavioural and cognitive components of equine emotion and endeavoured to evaluate how questions around how equines feel could be

answered. Their review includes a precis of much of the information presented within this review and suggests some additional methodologies for assessing affect in equines including physiological measures such as infrared thermography (IRT) of specific areas of the face to identify temperature changes associated with emotional response, along with electrodermal responses, core temperature evaluation, urinary oxytocin measurement, electroencephalography, electromyography of specific facial muscles and use of the EquiFACS system to monitor facial movements of the equine. The suggested novel methodologies were compiled through interrogation of the literature surrounding the assessment of emotional state in humans and the cumulative evidence for the accuracy of these techniques. In addition they suggested a number of behavioural measures including analysis of behavioural resilience (the degree to which a behaviour can resist being compressed in the time budget of an equine that has been forced to spend its time on another behaviour), analysis of changes in behavioural patterns such as sleep patterns and recumbency, analysis of exploratory behaviours, post-inhibitory rebound behaviour, detailed analysis of individual features of facial expression such as wrinkles around the eye and analysis of approach and avoidance behaviour to indicate valence following appraisal of stimuli.

The AWIN project (Dalla Costa *et al.*, 2016a) developed a welfare assessment protocol for horses within which the following factors are assessed:

1. Social interaction
2. Stereotypy performance (or absence of)
3. Fear (or absence of)
4. Human-animal relationships assessed via avoidance distance, voluntary animal approach and forced human approach tests
5. Qualitative behaviour assessment

Researchers involved in the AWIN project identified that one of the major objectives in equine welfare currently is the development of a scientifically sound process for the assessment of animal welfare that is valid (the methods accurately measure what it is intended that they measure), reliable (can be repeated with similar results generated) and relatively easy to undertake. Within the AWIN project several behavioural variables were assessed alongside physiological and environmental variables and highlighted the importance of equine behaviour as a proxy for emotional state in a holistic measurement of welfare. Whilst the

AWIN protocol does not solely look at affect, several measurements are designed to give a broad overview of potential affective state through identification of the presence of behaviours indicative of anticipated reward acquisition (activation of the RAS system) and absence of behaviours indicative of punishment avoidance (PAS activity reduction) (Dalla Costa *et al.*, 2016a).

Villas-Boas *et al.* (2016) used a mixed-model to assess behavioural, endocrine and cardiac autonomic responses to a model of startle in horses. They assessed the response of horses to sudden visual startle stimulus (opening of an umbrella) using ultra-short-term heart rate variability (HRV) analysis, heart rate (HR) measurement, reactivity assessment (flight latency, duration and distance), attention bias analysis and blood cortisol measures. Their study aimed to assess average startle responses within the test group as well as individual differences between horses as a form of addressing suitability for use in situations where the horses may encounter startle stimuli (such as police horses). Whilst their results did not show significant correlations between cortisol levels, behaviour and HR/HRV assessed via multi-comparison statistical methods, a correlation between behavioural startle response and HR was detected as was a pattern of cardiac activity reported by other authors investigating cardiac responses to startle; this being a tachycardic-bradycardic-tachycardic cycle mediated initially by the parasympathetic system and by both the parasympathetic and sympathetic systems working reciprocally to restore balance. The authors suggested that the use of HRV is a valid measure of short-term affectivity (within their study a measure of ultra-short term negative affect during the startle response) in horses. Mixed-model methodologies have also been used to assess temperament in horses. Visser *et al.* (2003) studied the correlation between rider assessment of horse temperament, results of behavioural tests (novel object and handling), horse heart rate (HR) and heart rate variability (HRV) and concluded that high heart rates were correlated with more anxious behaviours and HRV correlated positively with exploratory behaviours but negatively with freeze behaviours. In this study the 16 riders assessed 10 temperament traits with significant agreement ($p < 0.001$) and principal component analysis revealed two underlying components to the level of agreement between raters; principally traits implying the response of the horse to the environment and those implying response of the horse to the handler. Rider ratings correlated significantly with HRV and handling tests but not with novel object tests however the authors suggest that this may be due to reduced

behavioural responses from the mature horses in the sample (potentially mediated by prior learning in response to novel objects or the experience of being unable to avoid novel experiences) (Mendl and Paul, 2020a). These learned responses may have belied their actual autonomic response assessed via HRV, potentially a demonstration of learned helplessness in the mature horse sample (Hall *et al.*, 2008). The prior literature around the use of mixed methodologies in assessment of affective state has influenced the design of the experimental studies reported within this thesis where pairing of behavioural and physiological measures is used to identify relationships between variables which may indicate the presence of positive affective states in equines.

Other forms of behavioural assessment have been utilised by a plethora of researchers including appraisal theories: in essence this process is a series of 'stimulus checks' assessing the key characteristics of the stimulus - valence, predictability and familiarity - which consequently generates an emotion which organises a short-term response to the stimulus to facilitate survival (Mendl *et al.*, 2010a; Boissy and Lee, 2014). Qualitative Behaviour Assessment (QBA) is a form of assessment that is widely used in the study of animal temperament and personality (Wemelsfelder *et al.*, 2001) and generally uses either free choice profiling (FCP) or fixed lists (FL) both of which have benefits and limitations (Fleming *et al.*, 2016). QBA was included in the 2004-2009 European Commission's Welfare Quality Audit and identified as the only measure which captures positive welfare (Temple *et al.*, 2011) and was validated by Minero *et al.* (2018) as a reliable method of identifying affective state in horses achieving good inter-observer reliability and sensitivity to the quality of human-horse interaction. The use of QBA is discussed later within this thesis (Chapter 3) and explored experimentally in Chapter 6.

Boissy and Lee (2014) also developed methodologies to assess longer-lasting affective states (moods) such as agitation and depression which can negatively affect welfare based again on prior human psychological methodologies. They utilised cognitive bias testing to determine how affective states altered the ways in which animals perceived, evaluated and interpreted information from their environment following on from research by the pioneering authors (Paul *et al.*, 2005) in other species including rats (Harding *et al.*, 2004), starlings (Matheson *et al.*, 2008), dogs (Burman *et al.*, 2011), sheep (Doyle *et al.*, 2010; Destrez *et al.*, 2014) and pigs (Douglas *et al.*, 2012). Such judgement bias research has shown that negative past

experiences such as the removal of environmental enrichment, repeated exposure to aversive or unpredictable events and the administration of a serotonin inhibitor to induce depression, elicit a negative bias whereas the administration of the benzodiazepine Diazepam (an anxiolytic) or the opioid agonist morphine or housing in an enriched environment produces a positive bias (Boissy and Lee, 2014). This method of judging long-term affective states shows promise however does require lengthy training of the animals (or administration of psychopharmacy) and as such alternative cognitive bias methods such as attention bias or threat perception may be more practicable. Attention bias testing was developed following numerous studies on judgement bias. Unlike judgement bias studies, attention bias investigation presents two or more stimuli simultaneously and records how the animal utilises and directs attentional resources to the different stimuli. Often the stimuli presented are of differing valences to ascertain information regarding affective state and motivations. This form of testing gives the researcher indications around the animal's current goals, previous experience and expectations (Mendl *et al.*, 2009).

Novel object tests are another behavioural measure which have been used to evaluate equine temperament and as part of welfare assessment and can be indicative of the equine's general behaviour and emotional state (Bulens *et al.*, 2015). The use of novel object testing is widespread in behavioural science with researchers assessing responses to novel visual, olfactory and auditory stimuli (Christensen *et al.*, 2021a) shown to elicit significant behavioural and physiological effects indicative of stimulation of the sympathetic nervous system. Leiner and Fendt (2011) found that behavioural and physiological fear responses in horses were correlated and reliable when assessed via novel object testing and heart rate analysis. Considering the response to novel objects from a positive affective perspective, it would be anticipated that horses in positive affect would show a more positive bias towards novel objects (including humans) through demonstration of investigatory or 'seeking' behaviours (Birke *et al.*, 2011; S ndergaard and Halekoh, 2003), a facet explored later in this thesis.

Facial actions have been validated as accurate indicators of pain (grimace scales) in many animal species (Leach *et al.*, 2012) including stabled horses (Gleerup *et al.*, 2015; Dalla Costa *et al.*, 2014; Dalla Costa *et al.*, 2016b) and ridden horses (Dyson *et al.*, 2017; Mullard *et al.*, 2017) when combined with other methods such as composite pain scores. Dalla Costa *et al.*

(2017) investigated the effect of different emotional states on the Horse Grimace Scale (HGS) with their pilot study results indicating that, apart from during fear, the HGS score did not alter in any of the experimental conditions and as such it is purported that the features assessed in the HGS may be indicators of negative valence such as pain and fear and not positive valence. The authors observed that a much larger scale investigation was required to ascertain whether these suggestions are accurate. More recently facial actions have been utilised in the study of positive affective state (Lansade *et al.*, 2022; Lansade *et al.*, 2018; Trösch *et al.*, 2020) and have been suggested to be reliable indicators of positive affect in specific situations (grooming, observation of positive interactions with humans); these methods are discussed in greater detail in Chapter 3 and utilised experimentally in Chapters 5 and 6.

1.5.2 Physiological measures to evaluate positive affect in equines

Physiological methods measure the internal state of the animal predominantly via assessment of the cardiovascular and neuroendocrine systems (Schmidt *et al.*, 2010a; Schmidt *et al.*, 2010b; Sikorska *et al.*, 2023; Villas-Boas *et al.*, 2016; von Lewinski *et al.*, 2013; Munsters *et al.*, 2013; Duran *et al.*, 2017; Rochais *et al.*, 2018; Malmkvist *et al.*, 2012; Mormède *et al.*, 2007; McGreevy and Nicol, 1998).

Traditionally cortisol levels have been utilised as an indicator of stress in horses using samples collected via blood, urine, faeces, hair or saliva. Increases in cortisol levels are reported to reflect activation of a physiological stress response which may indicate the presence of psychological and/or physiological stress according to the methodology utilised. There has been a plethora of research into serum cortisol levels as an indirect measure of stress in horses during transport (Clark *et al.*, 1993; Cavallone *et al.*, 2002; Fazio *et al.*, 2008; Fazio *et al.*, 2013), competition (Covalesky *et al.*, 1992), training (Alexander and Irvine, 1998; Cayado *et al.*, 2006) and stress associated with repetitive behaviours (McGreevy and Nicol, 1998). Peeters *et al.* (2011) compared serum and salivary cortisol concentrations in horses using an adrenocorticotrophic hormone (ACTH) challenge reporting a highly significant correlation between serum and salivary cortisol concentrations confirmed by regression analysis. These results prompted them to advocate for the use of salivary measures over serum measures to provide for a less invasive biomarker of stress in horses. Christensen *et al.* (2012) assessed faecal cortisol metabolites (FCM) in a study investigating fearfulness in horses but did not find

any correlation between behavioural scores of fearfulness and FCM measurements however they did report that elevated FCM levels predicted impaired learning performance under stressful conditions and were also associated with rank (lower rank correlating with higher FCM levels). Malmkvist *et al.* (2012) also identified an association between increased FCM and stress responses in horses with gastric ulcers; their study determined that FCM was sensitive and specific enough to detect changes in circulating cortisol measurements in horses.

Duran *et al.* (2017) undertook a study into the use of hair cortisol concentration (HCC) as a chronic stress biomarker and found that hair from the mane (in preference to hair from the tail) could generate reliable data on basal and elevated hypothalamic-pituitary-adrenal (HPA) activity in horses. The study assessed HCC over a period of months following handling and physiological stress and compared results both over time and between sexes. This methodology could provide a non-invasive, practical way of assessing stress in equines to reveal temporal changes via assessment of different segments of the hair length. However, it should be noted that the investigators did not compare the HCC data with cortisol measures using any of the other three previously investigated methods which would have provided further contextual information to assess reliability and sensitivity within the study. Visser *et al.* (2008) identified that horses housed in stressful conditions such as in social isolation or with frequent changes to housing conditions displayed lower cortisol and adrenocorticotrophic hormone (ACTH) response when exposed to corticotropin releasing factor (CRF) challenge which did not always correlate with their behavioural responses however they asserted that as reported in previous literature in other species (bulls and calves) these results suggested a depression in adrenocortical reactivity to exogenous ACTH caused by a desensitisation of the HPA axis in response to previous stress induced rises in ACTH and cortisol.

Just as increases in cortisol are likely reflective of acute stress-like states, decreases in cortisol may indicate induction of more positive affective states (Buchanan *et al.*, 1999) however cortisol evaluation is not straight forward. Studies have demonstrated that lower levels of faecal cortisol metabolites and plasma cortisol are found in equines with chronically compromised welfare (Pawluski *et al.*, 2017) and increases in cortisol levels can be associated with positive arousal (Ralph and Tilbrook, 2016). It is also important to consider that glucocorticoids, including cortisol, have widespread actions and effects throughout the body

pertaining to normal function as well as stress responses; glucocorticoid secretion is highly variable and pulsatile (within a 90 minute period) and can be significantly impacted by events such as eating; it is therefore imperative that cortisol measures are considered alongside other behavioural and physiological measures in order to elucidate their meaning in a context-specific situation (Ralph and Tilbrook, 2016). As such, methodological designs using cortisol measures should be carefully considered to ensure that the methods used (such as sample type hair, saliva, serum for example) are suitable to identify the measures of interest.

Alterations in glucocorticoid levels have been shown to have a demonstrable effect on the immune system (Budzyńska, 2014). The study of psychoneuroimmunology is a relatively new discipline in animals however researchers have shown dynamic relationships between behavioural factors, the central nervous system, endocrine and immune systems which can cause activation of immune mechanisms (in acute stress), reduction in immune activity back to baseline (in moderately extended stress states) and immunosuppression in chronic stress (Dhabhar, 2009; Bartolomé and Cockram, 2016). Davis *et al.* (2008) outlined a methodology for assessing stress response via analysis of leukocyte profiles resultant of characteristic changes to the cells via heightened glucocorticoid levels. This included assessment of lymphopenia, neutrophilia, eosinopenia, and leukocytosis. In addition, the adrenal hormones were shown to suppress the activity of B cells and cytotoxic T cells, and decrease interleukin-1 and interleukin-2 synthesis with all aspects reliably measured in mammalian species; as such this could be a reliable indicator of stress or eustress in equines however further investigation of these methods alongside other validated methods is required.

Electroencephalography (EEG) is a well-established tool in human medicine and over time has traversed to the field of veterinary medicine including use in equines (Lacombe *et al.*, 2001; Grint *et al.*, 2014). Williams *et al.* (2008) investigated EEG recording of horses in various stages of vigilance using non-invasive electrode placement to gather normative data in a cohort of horses. The authors compared their findings with earlier invasive methodologies using dermal screws to assess the validity of their results. Using EEG, electromyographic data (EMG) of facial movements and electrocardiographic (ECG) data to monitor heart rate, they collated datasets for horses in various sleep and waking patterns which correlated with previous research findings indicating that this method may be a reliable measure in horses provided that a standardisation of electrode placement can be identified. The use of EEG in assessing

attentional state and chronic welfare states (Stomp *et al.*, 2021) has also more recently been validated (Rochais *et al.*, 2018), and a conceptual ambulatory EEG (AEEG) produced and trialled for use in equines (Cousillas *et al.*, 2017). These measures are of interest in the development of methods to measure positive affect in equines however they require specialist equipment and skillsets to gather and reliably analyse the data and as such are more limited in their use than other more practical on-farm methods.

As well as generating data on the physical exertion levels of the equine, heart rate (HR) data has been used to assess emotional components of behaviour. In many cases these measures are combined with heart rate variability (HRV) assessment however some researchers have investigated the potential to capture the dynamic HR responses of the equine and mathematically decompose the HR signal into a physical and an emotional component (Jansen *et al.*, 2009). Using a second-order output-error model the researchers were able to input individual data on activity and the equine's characteristics and arrive at a predictive model for the HR in relation to the activity level undertaken. They then modelled the difference between the predicted HR and actual HR to assign the difference to the emotional component of HR considering individual variation, time variance and the dynamic nature of biological systems and reported that this could be a useful model albeit the sample size was very small. HRV assessment is more routinely used in trials assessing affective state in animals as a measure to evaluate autonomic balance between the sympathetic and parasympathetic nervous systems, a potentially valuable tool for assessing autonomic modulation of cardiac activity at low-moderate exercise (Physick-Sheard *et al.*, 2000) assuming careful interpretation of the data (Von Borell *et al.*, 2007; Stucke *et al.*, 2015; Malik *et al.*, 1996). Adaptability to changing environments is, in part, regulated by the interplay between the heart rate reducing parasympathetic (PNS), and heart rate increasing sympathetic nervous systems (SNS) producing a total effect on HR that may be resultant of either synchronous or independent influence (Loftus *et al.*, 2016). HR and/or HRV (using time-domain analysis and frequency-domain analysis) have been used widely within the assessment of psychological stress for horses undergoing various tasks, including initial training (Fowler *et al.*, 2012; Schmidt *et al.*, 2010a; Janczarek *et al.*, 2017; Munsters *et al.*, 2013), desensitisation to alarming stimuli (Munsters *et al.*, 2012), public competition (von Lewinski *et al.*, 2013; Kinnunen *et al.*, 2006), ridden obstacle tests (Ille *et al.*, 2013), novel object/startle tests (Villas-

Boas *et al.*, 2016; Visser *et al.*, 2002), transportation (Schmidt *et al.*, 2010b; Schmidt *et al.*, 2010c) and sudden individual stabling (Erber *et al.*, 2013). However, HR and HRV are further externally influenced by factors including sex, age, breed, circadian rhythm, time since last meal, physical exertion, pain, and stress and as such these should be considered in the interpretation of data gathered (Von Borell *et al.*, 2007; Stucke *et al.*, 2015; Bowen, 2010).

1.6 The effect of personality on affective state

In order to understand the dimensions of personality it is important to consider the developmental, causal biological bases, genetic and environmental origins as well as function and evolutionary history (Ashton, 2022). In recent years personality research in animals has been conducted using two distinct methodologies:

1. Rating-based research - whereby personality traits (described often using single adjective words) are rated by trained raters, owners or carers of animals. Studies have shown high inter-observer reliability between independent raters of personality traits in animals, comparable to reliability seen in human studies (Gosling *et al.*, 2003a; Gosling *et al.*, 2003b; King and Figueredo, 1997).
2. Experimental test-based research - direct observational research of an individual's behaviour often in experimental tests designed to elicit individual differences in a particular behaviour. To accurately identify a personality trait, behaviours need to be observed in several experimental test situations to enable correlations between multiple behaviour-in-situation scores to be computed. The limitations of this methodology are that behaviours identified may be very context specific therefore showing only a weak influence of the effect of a personality trait (Carter *et al.*, 2013).

Ashton (2022) suggests that Hogan's framework (Hogan, 2002) for the study of behaviour is a good basis for personality study. The framework suggests that behaviours of a given class, for example aggression, are regulated by a behaviour system of three parts: a motor mechanism (producing the movement pattern of the behaviour); a perceptual mechanism (recognising internal and external states to excite or inhibit behaviours) and a central mechanism (which integrates input from the perceptual mechanism and organises motor outputs). This framework maintains a distinction between causal and functional questions around behaviour and allows both questions to be addressed within the assessment of a specific behaviour and consideration of the ontogeny and phylogeny of that behaviour.

Much of the research conducted in the field of personality across human and non-human animal species has centred around the 'Five Factor Model' (FFM) (McCrae and Costa, 1987). This model of personality includes five basic dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness to Experience. It is widely supported by rich experimental data as a model which is comprehensive and applicable across contexts and cultures (McCrae and John, 1992). Over years of personality research in the human psychology field the FFM has been relied upon as the primary reference for the development of many other personality assessments including the Hogan Personality Inventory (HPI) (Hogan, 2002), the Neo Personality Inventory (NEO-PI) (Costa and McCrae, 1989) and has been compared with other measurement metrics including the Eysenck Personality Inventory (Eysenck and Eysenck, 1984).

Other studies have focussed on the Plutchik (1980) theory of human emotions from which the Emotions Profile Index (EPI) questionnaire was developed (Plutchik and Kellerman, 1974). This index uses bipolar attributes such as boldness/shyness whereby a two-alternative forced choice method is employed to give a multidimensional emotional profile of the subject. In human research the index was overwhelmingly utilised as a self-reporting tool however in animal research it is used as a rating or observing instrument (Martau *et al.*, 1985). Plutchik devised a series of eight adjectives to define basic adaptive functions and then linked these to 12 dimensions or constructs of emotions to develop a personality profile (Figure 1.8).

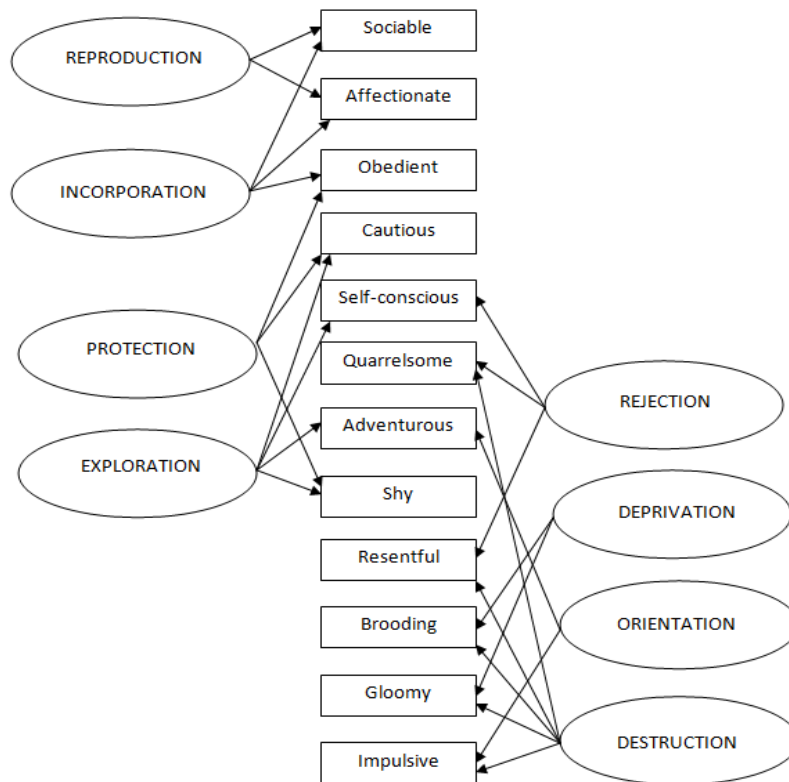


Figure 1.8 The relation of adjectives to dimensions according to Plutchik's theory of emotions. Adapted from (Bucik et al., 1995). *Rectangles denote EPI adjectives; ovals denote dimensions or constructs of emotions.

Gosling (2008)'s review of the current state of personality research in non-human animals reveals that, on the whole, research in this field meets the standards expected in human research trials in this area and satisfies the three main criteria, which are:

- a) Assessments made by individual observers must agree with one another
- b) These assessments must predict behaviours and real-world outcomes
- c) Observer ratings must be shown to reflect genuine attributes of the individuals rated, not merely the observer's implicit theory regarding co-variance of personality traits.

Gosling also reported that there is significant evidence to demonstrate temporal stability in personality traits which is supported by Visser *et al.* (2001) who found that behavioural variables indicative of personality traits were positively correlated over periods of 9, 10, 21 and 22 months of ages in young horses, with flightiness proving to be the most consistent trait over a period of years. Interestingly this review also highlights that whilst the literature available currently indicates that around 70% of animal personality studies are conducted

using behavioural coding's for overt behaviours displayed; in fact direct comparison of methods shows that trait rating for animal personality assessment is superior to behaviour coding as this method is more reliable, not as subjective as first presumed, and more practical (see Vazire *et al.* (2007) for detailed discussion).

Bucik *et al.* (1995) carried out two studies to examine the stability and dimensionality of the EPI. Study one investigated the dimensionality of the structure of the EPI. They found that some dimensions such as 'Destruction' are not on the same hierarchical level as other factors and as such should not carry the same weighting under personality profiling. For example, Destruction seemed to be an overarching factor leading to rejection, deprivation and orientation which were then linked to relevant EPI adjectives. As such they reported that the EPI dimensions may not be wholly reliable as a methodology. In their second study the authors assessed the stability of the structure of emotions specifically the correctness of denotations made, and comprehensiveness of adjectives used in their relation to the various EPI dimensions. Results indicated that the language used contributed to a high degree of bias in self-reporting subjects and that subjects found some dimensions more difficult to comprehend than others (for example destruction and reproduction). Subjects were asked to attribute dimensions to EPI adjectives on a blank chart such as the one in Figure 1.8 (without arrows) with results showing large differences in the attribution of dimensions to adjectives across subjects with five of the eight dimensions of the EPI model misinterpreted by subjects raising questions about the validity and reliability of the method.

Zuckerman (1998) highlights the role of brain chemistry in his biological basis of personality theory supported by Cloninger (1986) as shown in the Figure 1.9. These theories link with Panksepp (2005)'s work on the role of the neuroendocrine system in emotion and subsequent behavioural responses and due to the aforementioned homology seen between human and non-human neurochemistry it is likely that the link found between brain chemistry and personality in human studies is cross-transferable to non-human work.

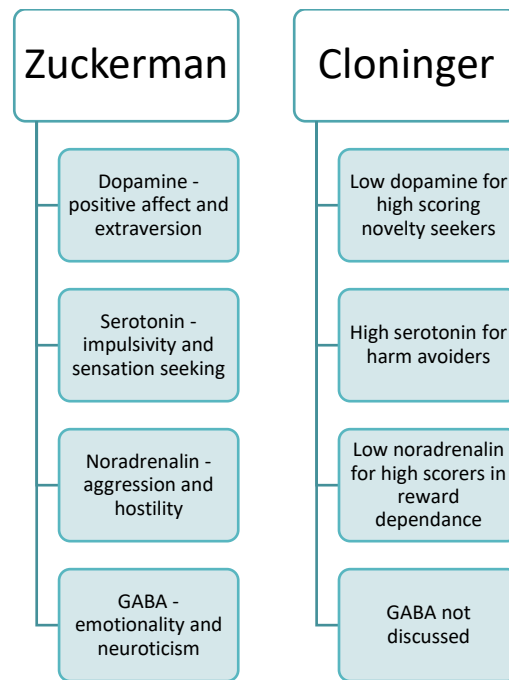


Figure 1.9 Comparison of Zuckerman and Cloninger's findings in relation to the neurochemistries of personality.

There are however a great number of covariates in the production or suppression of neurochemical markers making achieving high correlations between psychometric measures and biochemical markers an area of difficulty as discussed in detail by Cacioppo and Tassinari (1990). It appears likely that to formulate a rounded view of equine personality, measurements such as physiological parameters, learning capacity and caretaker personality profiling could be used in conjunction with one another to facilitate an accurate representation of the individual (Duberstein and Gilkeson, 2010).

Another facet of personality, empathy, has been little studied in equines, a factor which Edgar *et al.* (2012) noted within their review of empathy in animals. Most research has been undertaken in primate species with rats being the predominant research species for non-primate-based empathy research to date. Cozzi *et al.* (2010) examined post-conflict affiliation in a small stable group of horses and found that on almost 60% of post conflict occasions a third-party horse acted either to appease the aggressor or console the victim. Further research in this area of personality and affect is required, however, with the inclusion of controls and measures of valence or distress (of both observer and observee) to identify if these behaviours are truly empathic and /or prosocial.

It is thought that an individual's personality may impact on the intensity of affect experienced. Within the human psychology literature there is a multitude of evidence supporting the effect of personality on mood, emotion regulation and psychiatric disorders (Peterson *et al.*, 2007; Zuckerman, 1998; Akiskal, 2013; Gunnar Götestam *et al.*, 2008). Hughes *et al.* (2020) reviewed the role of personality traits in shaping the intra and interpersonal regulation styles of individuals with reference to the Big Five personality traits (McCrae and Costa, 1987) and the Gross (2015) extended process model of emotional regulation. They identified that personality traits played a significant role in the three stages of emotion regulation: identifying which emotions to regulate, selecting the regulation approach, and implementing regulatory tactics. Within the non-human animal field, Kremer *et al.* (2020) reviewed animal affect, including personality and considered the implications of personality traits on both short-term emotions and longer-term mood identifying that it is currently difficult to disentangle emotion and mood practically and that individual personality likely has effects on both aspects of affective state.

1.6.1 Defining personality

Within ethology the terms 'personality', 'temperament' and 'behavioural syndromes' are often used interchangeably when looking at individuals within a population and how they vary in their behavioural responses to stimuli. MacKay and Haskell (2015) produced a review of terminology used in the field of 'consistent individual behavioural variation' science and presented a working framework of definitions for the three most used terms cited above; these being:

1. Personality - the individual's behavioural variation in reference to the personality dimensions found in the population,
2. Temperament - the animal's behavioural response in a single context measured on some biological scale, for example flight speed,
3. Behavioural syndromes - used to describe variation within populations, not individuals. They link two or more variables across a population and describe the distribution of personalities within a population.

Gosling (2008) suggests that, to satisfy all researchers, personality should be described as 'those characteristics of individuals that describe and account for temporally stable patterns of affect, cognition and behaviour'. Whilst Gosling acknowledges that personality historically

has been seen as an anthropomorphic term it does encompass a range of different labels within the phenomena and in recent times has been adopted as the preferred term by a range of established journals including '*Nature*' and '*Proceedings of the Royal Society*'.

1.6.2 Development of research on animal personality

Over the past twenty years there has been a surge in interest in the study of animal personality with the number of articles published in that time (n=200+) exceeding that published in the entire preceding century (n=118) (Sih *et al.*, 2015; Gosling, 2008). Much of the increased interest appears to have stemmed from the foundational research preceding it which reported several key findings:

1. Personality exists and can be measured in animals,
2. Personality traits have been identified across a wide range of taxa,
3. Some traits show more cross-species generality than others (Gosling, 2008).

There has been an accumulation of evidence for significant correlations and repeatability of effects between various behavioural traits such as boldness, sociability, fearfulness and explorative behaviours which vary across contexts partially as a result of ecology (Bell, 2007; Bell *et al.*, 2009; Garamszegi *et al.*, 2012), and partly due to early experience and the development of personality structure over time (Sih, 2011; Buwalda *et al.*, 2013). Nicol (2002) reported large individual differences between subjects in equine learning and memory experiments, a factor also reported by numerous other authors (Fiske and Potter, 1979; McCall *et al.*, 1981; Marinier and Alexander, 1994). These authors suggested the differences, especially in matched-pair experimental design, may be due to differences in personality as well as gender, life stage and prior handling variables.

Sih *et al.* (2015) present a framework aiming to unify the logic of the numerous recently described personality models cited by other authors. The framework attempts to enable rigorous testing of models including modelling of feedback loops and integration of proximate and ultimate mechanisms over ecological, developmental, and evolutionary timescales to produce predictive models of the variation in animal personalities. The premise of this work is that whilst behaviour can be infinitely flexible the repeatability and differentiation in behaviour between individuals can be explained by changes in variables combined with adaptive state-dependant behaviour. *Exempli gratia*, individuals differ in behaviour because

they differ in state, therefore they adjust their behaviour to adapt to these differences. The individual's state will include a cost-benefit analysis of any behavioural actions and consider the focal individuals characteristics (such as energy reserves, reproductive status, condition, age, size) as well as characteristics from the social environment (density, social partnerships) and the ecological environment (predation, competition, parasites). Taking these variables into account it is then possible to model these changeable behaviours and states to explain the co-development of state and behaviour that drives a persistent divergence of longer-term differences in both states and behaviours between individuals resulting in consistent among-individual variation in behaviour *ergo* 'personality'.

Bergner (2020), in his longitudinal review of the field of human personality assessment highlights that it is important to avoid considering personality traits as 'the whole of the being' and instead, they suggest, we should utilise personality traits as a part of a more informative 'parametric analysis of persons'. Reflecting on the work of Ossorio (2006), Bergner suggests that personality scientists should consider not just personality but the wider 'person characteristics' which Ossorio defines as follows:

$$<PC> = <Ds,P,Dr>$$

Where PC = person characteristics, Ds = Dispositions (including personality traits but also attitudes and interests), P = Powers (abilities, knowledge and values) and Dr = Derivatives (states such as ill health, capacities to learn and embodiment – the physical characteristics of the individual).

Interestingly within the field of Clinical Animal Behaviour these considerations are at the forefront when assessing the predisposing, initiating, and maintaining factors for behavioural conditions in animals to undertake differential diagnostic processes as they are suggested to be interlinked as contributing factors to behavioural production (Mills, 2022).

Within this thesis equine personality is explored through both a Delphi consultation of experts (Chapter 3) and an iterative development of an equine personality questionnaire (Chapter 4), following which the findings are utilised within the experimental design of the field-based trial which focusses on inducing and measuring positive affective state in equines whilst considering the effect of personality on affective state identified through behavioural and physiological variables (Chapter 6).

1.7 Conclusion

The study of equine emotion, as with other human and non-human animal emotion study, is an ongoing process with new research and applications frequently developing this field of study. Whilst there is almost universal agreement that equines experience affective states there is great discussion regarding how these states occur and even less consensus on how they may be induced and measured. The human literature lends a broad and deep knowledge base for the study of affect in animals, and this is also true when considering the study of animal personality. Strong evidence from the human literature and developing evidence from the animal literature suggests that personality has a significant impact upon affective state and as such should be considered within methodologies which intend to study animal emotions. Due to the complex nature of the brain and its interactions with the body and external environment, consideration must also be given to the effect of learning and memory on expression of affective state. The impact of prior learning and memory on equine responses to experimental and non-experimental stimuli should not be underestimated within methodological design, nor should the consideration of an individual's personality, health, and wider environment.

Initial research into assessing positive affect in equines has highlighted some potentially useful behavioural and physiological measures which could be utilised within a framework to assess affective state, however, there is limited validation of many measures to date and very little focus on the impact of personality on experience of affective state in equines. There is a clear need for further consensus within the equine scientific community regarding the validity and reliability of measures to both induce and measure positive affective state in equines and a requirement for further empirical studies to test a range of more novel methods, alongside those which have been previously validated, combined with modelling of personality traits, to expand our understanding of which methods may be useful in an applied context.

1.8 Research aims and thesis outline

This thesis encompasses several interlinked aims which explore facets of both equine personality and affective state; these are:

1. To gather the views of experts on equine personality traits and their interplay with affective state,
2. To develop a valid personality assessment for equine use,

3. To gather expert opinion on the experience of affective states in equines and methods to both induce and measure positive affective state,
4. To test, experimentally, the expert derived methods through the attempted induction and measurement of positive affective state in a cohort of equines,
5. To analyse the data gathered in an exploratory manner, incorporating the effect of assessed personality traits within data modelling, to identify potential behavioural and physiological markers of positive affective state and distinguish these from markers of negative affective state.

This thesis intends to build on prior work in the field of equine emotion and personality, outlined in this chapter, by examining expert opinion on equine personality and affective state. It presents and discusses experimental data relating to personality assessment in equines and evaluates methods to induce and measure affective state through the lens of individual personality.

Chapter two presents the methodologies utilised within the study which are interlinked. There is cross-over between chapters for example, where elements of a subsequent chapter include results from the Delphi consultation within their methodology.

Chapter three presents the Delphi consultation of experts into equine personality and methods to induce and measure positive affective state in equines.

Chapter four presents the newly devised Equine Personality Eight-factor Model and includes data from the Delphi consultation and three survey rounds utilised for questionnaire development.

Chapter five presents a pilot study which was undertaken to trial the methods and materials for the final study and Chapter six presents the full final study which aimed to induce and measure positive affective state in equines and distinguish from negative affective state.

The final chapter, chapter seven, presents a general discussion of the work contained within the thesis.

Chapter 2 General methodologies

2.1 Introduction

This thesis explores the potential validity and reliability of methods to measure and induce positive affective state in equines through both qualitative and quantitative lenses. This chapter outlines and places into context the overarching methodological approaches taken within the research process to achieve the research objectives.

The methodology and approach for the Delphi consultation into positive affective state and personality is presented in detail and relates to the results reported in Chapters 3 and 4.

2.1.1 Overview of the experimental design

The research philosophy for this thesis is based upon a positivism approach whereby the aim is to make objective observations via a hypothetico-deductive model (Park *et al.*, 2020; Alharahsheh and Pius, 2020). A hypothetico-deductive method is used to construct a scientific theory by considering the results obtained through observation and experimentation which can be later verified or questioned through empirical evidence derived from other experiments (Tariq, 2015). This thesis aims to derive theories around the induction and measurement of affective state in equines via exploratory analyses which can later be further verified in empirical studies. Due to the relatively limited empirical research to date on this topic many of the methods employed within this study are inductive and exploratory (Swedberg, 2020) beginning with a survey of expert opinion in order to draw together consensus themes to inform further methodological design as has been adopted widely in the medical field (Hohmann *et al.*, 2018; Duncan *et al.*, 2020; Cary *et al.*, 2023). The study adopts a mixed methods approach with both qualitative and quantitative data gathered throughout the research process (Teddlie and Tashakkori, 2011) and with experimental and grounded theory survey-based approaches (Laxman, 2015) utilised within an overall action research theme (McNiff, 2013) as described in Figure 2.1. Grounded theory is a process whereby conceptual hypotheses are systematically generated to produce a theory about a substantive area, it is an iterative process used to formulate a theory beginning with purposive sampling of participants, coding and thematic analysis of the data obtained and testing of interpretations as the study progresses (Chun Tie *et al.*, 2019), a method utilised throughout the Delphi consultation within this thesis.

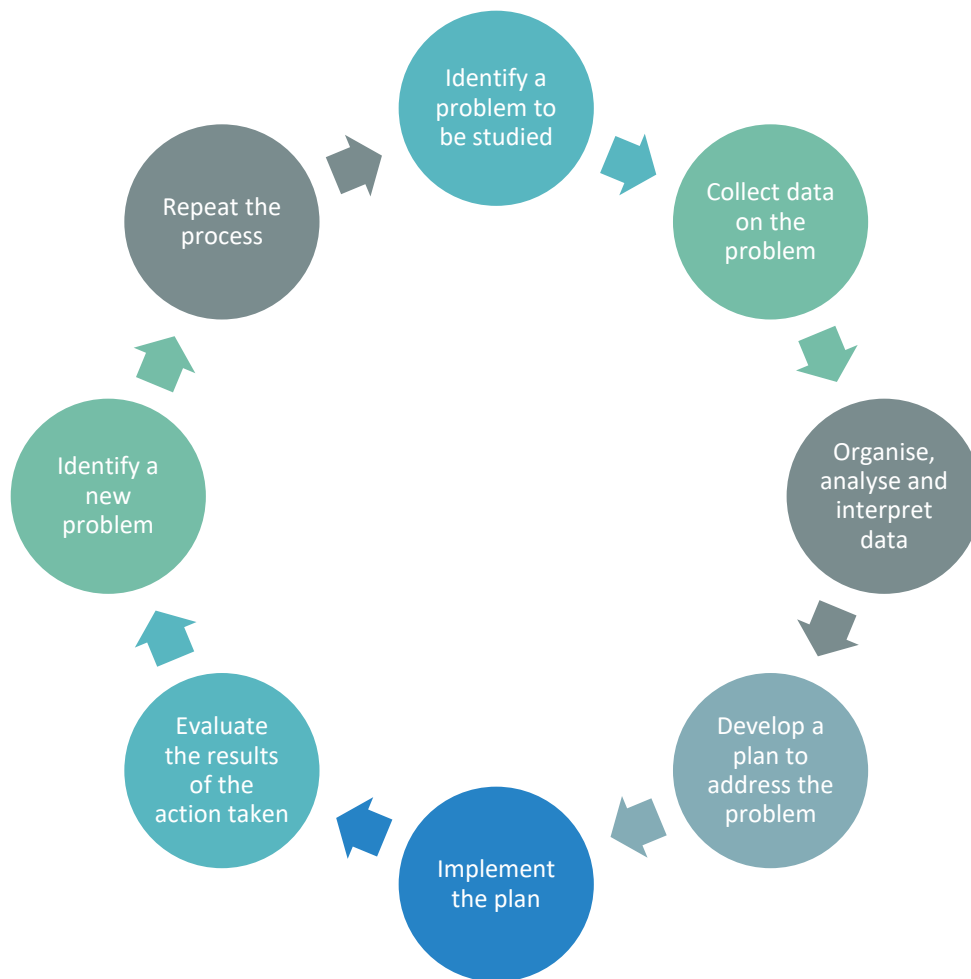


Figure 2.1 The research process (Daniel and Harland, 2017).

The research approach adopted both a cross-sectional and semi-longitudinal time horizon approach with Chapters 5 and 6 adopting a cross-sectional approach (sampling of individuals at a single point in time) and Chapters 3 and 4 utilising a longitudinal iterative design where individuals were sampled repeatedly over time (Rindfleisch *et al.*, 2008).

2.1.2 Overview of the population/sample and sampling strategy

A mixed approach was taken to the sampling strategy which was influenced by the type of data being collected. The Delphi consultation utilised a non-random sampling technique whereby participants were selected and invited to participate based upon a set of pre-ordained criteria relating to their expertise and experience (Mweshi and Sakyi, 2020; Shariff, 2015). Owner survey aspects were undertaken using a probability sampling approach to aid a representative sample (Singh and Masuku, 2014). Within this approach specific inclusion criteria were set *a priori* relating to ownership parameters (such as length of ownership time) to aid in gathering of reliable data.

Field data sampling followed a semi-random sampling methodology whereby inclusion within the trials was open to all, however, following this a set of *a priori* established criteria relating to the individual animal, its environment and its ownership history were employed to facilitate a robust sample for investigation (Marsbøll *et al.*, 2019; Dalla Costa *et al.*, 2014c).

2.1.3 Study location

The Delphi consultation and aspects of the personality investigations were undertaken online via Qualtrics™ (Provo, UT) (Qualtrics, 2020). Field trials were undertaken at seven equestrian facilities across the United Kingdom as shown in Figure 2.2.



Figure 2.2 Field study locations (Editable map from Softonic, 2022).

2.1.4 Ethical approval

The protocol and methods utilised throughout this thesis were approved by the Animal Welfare Ethics Review Board (AWERB) of Newcastle University (Project ID: 1025) and the Ethical Review Committee of University Centre Askham Bryan (Approval number: PGP532019)

prior to study commencement. Permission was also granted by Queen's Veterinary School Hospital, University of Cambridge for data collection at their site.

2.2 Delphi consultation methodology

The Delphi consultation methodology was utilised to collect data pertaining to Chapter 3 and parts of Chapter 4. The methodology is described in full here and summarised in Chapters 3 and 4.

2.2.1 Participant recruitment and selection

A recruitment plan was initiated whereby several routes (Table 2.1) were explored to identify suitable experts to be invited to contribute to the consultation.

Table 2.1 Methods of recruitment for different types of professional

Type of professional	Method of recruitment
Behaviourists who hold any of the following professional accreditations: <i>Certificated Clinical Animal Behaviourist (CCAB) with equine specialism</i> <i>Clinical Animal Behaviourist (CAB) with equine specialism</i> <i>Certified Equine Behaviour Consultant (CEBC)</i> <i>Applied Animal Behaviourist (AAB) with equine specialism</i> <i>Animal Behaviour technician (ABT) with equine specialism</i>	Members of any of the following professional organisations: <i>Animal Behaviour and Training Council (ABTC)</i> <i>CCAB Certification LTD's list of CCAB's (CCAB)</i> <i>Fellowship of Animal Behaviour Clinicians list of equine practitioners (FABC)</i> <i>Association of Pet Behaviour Counsellors list of equine practitioners (APBC)</i> <i>International Association of Animal Behaviour Consultants list of equine practitioners (IAABC)</i>
Equine veterinary surgeons with a behavioural interest specialism	Members of any of the following professional organisations: <i>RCVS Advanced Practitioners in behavioural medicine</i> <i>European College of Veterinary Medicine specialist in Behavioural Medicine (ECVMBM)</i>

	<i>Australia and New Zealand College of Veterinary Scientists (ANZCVS) with behavioural specialism</i> <i>American College of Veterinary Surgeons (ACVS) with a behavioural specialism</i>
Research academics	Review of authors contributing to published peer reviewed journal articles in the field of animal affective state, equine ethology, equine learning theory, animal personality and equine welfare related to the emotional domain within the last ten years
Teaching academics	Those that currently taught on an equine behaviour or ethology module at UK Framework for Higher Education Qualifications (FHEQ) level 4 or above or the equivalent at a recognised institution providing higher education
Equine trainers	ABTC Animal Trainer (AT) and Animal Training Instructor (ATI) registers British Horse Society (BHS) qualified instructors using non-punishment-based training techniques. Invitation of equine trainers using ethical positive reinforcement methods contacted via professional networks in the UK and internationally
Professionals working in equestrian sport within the remit of enhancing and furthering equine welfare and ethical training	Working in an equestrian sport regulated by one of the following bodies: <i>British Horseracing Authority (BHA)</i> <i>British Equestrian Federation (BEF)</i> <i>International Equestrian Federation (FEI)</i>
Those working in the charity and rehabilitation sector with a focus on	Working within any of the following organisations: <i>The Donkey Sanctuary</i>

the applied care and rehabilitation of equines	<i>Redwings</i> <i>Blue Cross</i> <i>The Horse Trust</i> <i>Royal Society for the Prevention of Cruelty to Animals</i> <i>Bransby Horses</i> <i>Horses and Ponies Protection Association</i>

A snowballing process (addition of participants through a series of referrals that are made within a circle of people who know each other) (Browne, 2005) was also undertaken where those invited within the initial round of invitation were asked to suggest and refer further experts known to them in their field who met the inclusion criteria for participation. Inclusion and exclusion criteria for the study participants can be found in Table 2.2.

Table 2.2 Inclusion and Exclusion criteria for recruitment of Delphi consultation participants

Inclusion criteria	Exclusion criteria
Be an equine veterinary surgeon. And/or Be an equine behaviourist. And/or Be an equine trainer. And/or Be an academic published in or qualified to teach in the specific field of equine behaviour, science or welfare. And/or Be an academic published in or qualified to teach in the specific field of personality or affective state in animals. And/or	Less than 5 years' experience in the field. And/or No formal equine or field related scientific or professional qualification. And/or Those who advocate for aversive training methods (positive or negative punishment). And/or Those who have been retired from or not working in the field for more than 5 years. And/or Those who do not fully complete surveys.

Be an accredited member of an equine or professional body such as BEVA, BHA, ABTC, RCVS, HEA, BHS, ABRS, UKCC, BVNA, APBC, ASAB, ISES, ABTC, ARCVS, BE, BD, BSJA, AAEP the Jockey Club. And Have at least 5 years of experience in their field of expertise.	And/or Those below the age of 18.

One hundred and eighty-four internationally distributed experts were identified initially via the recruitment plan (Table 2.1) and invited to contribute their expertise to the consultation process. Email invitations were sent to all identified experts in December 2019 explaining the rationale for and remit of the consultation and attaching a full participant information sheet and consent form (Appendix A). Within four weeks of the initial dissemination of participation invitations all contacts identified through snowballing were assessed and those meeting the inclusion criteria were also invited to participate. At two and four weeks post invite further reminders were circulated to non-responders providing a deadline (24th January 2020) for inclusion in the consultation.

Positive responders were asked to read the participant information sheet and return their signed consent form, this was countersigned by the researcher and a copy returned to the participant for their records. All participant data were stored electronically on a password protected device backed up to an encrypted cloud storage platform.

In total one hundred and two participants (55% recruitment rate) were recruited to participate in the study by the end of January 2020.

2.2.2 Survey development and question design

A survey was developed using Qualtrics software™ (Provo, UT) (Qualtrics, 2022) to explore expert opinion on several facets surrounding positive affective state in equines and factors that may affect this according to the literature, including:

- Capacity for equines to experience affective states,
- The effect of individual personality on affective state and

- Potential methods to induce and measure affective state in equines.

The survey was developed following a systematic review of the literature on equine affective state, personality and measuring behaviour and physiology related to emotional state in equines as shown in Figure 2.3. The survey was piloted through sharing of a link to the draft version of the Qualtrics survey amongst the professional peer group (n=5), consisting of the PhD supervision team and three colleagues holding PhD's themselves, to receive feedback highlighting errors and omissions and to ensure correct branching was made available within the Qualtrics software.

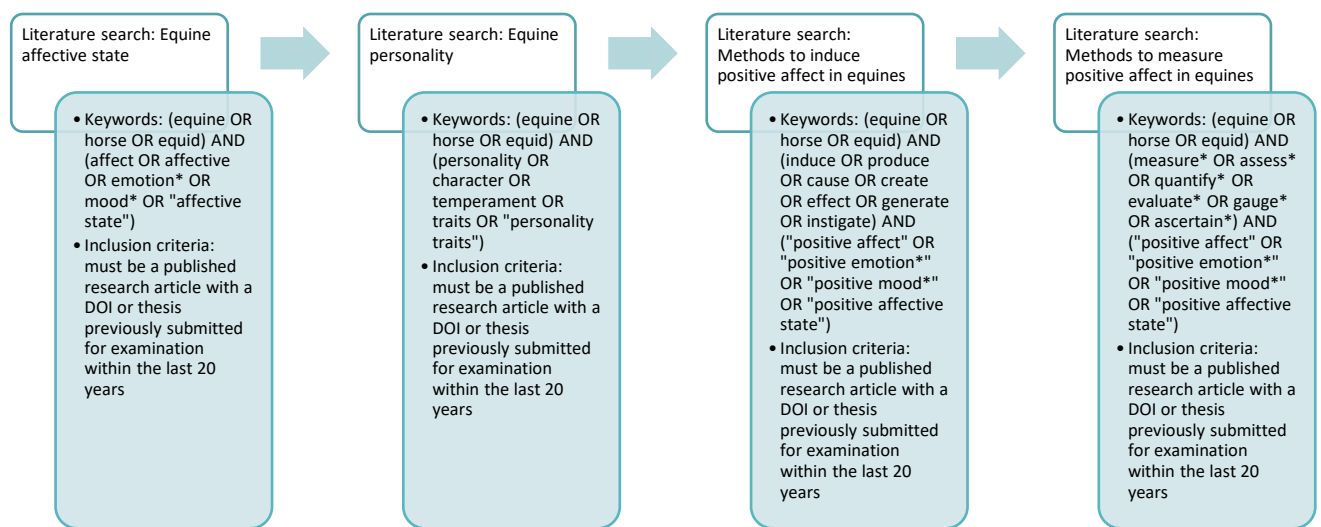


Figure 2.3 Literature searches undertaken to develop initial Delphi themes and questions.

The Delphi consultation was a three-round process where data collected in each round were summarised and fed back to participants within the next round and questions iteratively developed utilising participant feedback. Questions and statements achieving consensus in rounds one and two were referred back to participants to confirm their acceptance of the consensus in rounds two and three. Questions and statements not achieving consensus were refined and presented back to participants for further rating and commentary.

Survey one consisted of an initial demographic data collection section followed by sections on equine personality, emotional processing and inducing and measuring positive affect in equines. Survey two consisted of an initial section presenting the results from survey one and asking participants to confirm their acceptance of the consensus achieved and to provide further commentary where they did not accept the results reported. This was followed by

questions regarding personality traits and trait groupings and then questions asking participants to rank methods to induce and measure positive affect state in equines. Survey three consisted of an initial section presenting the results of the personality trait and trait grouping consensus and asking participants to confirm their acceptance of the consensus achieved and to provide further commentary where they did not accept the results reported. It then presented participants with finalised lists of induction and measurement methods for positive affective state in equines and asked them to select methods that they supported and provide additional commentary on the measures suggested. All surveys can be viewed in Appendix B.

Question design included a range of multiple-choice questions such as in Round one Question 21 'Do you consider a multi-measure approach (which assesses behavioural and physiological indices) is optimal for measuring affective state in equines? Yes, No , Unsure'; this style was utilised in order to attempt to increase accuracy of responses which may be reduced if a binary forced-choice approach was taken (Young, 2015). However there is some evidence within the literature to show that participants may think more deeply about the question when forced to give one of two binary responses (Smyth *et al.*, 2006).

Ranking questions, such as Round one Question 22 'Please drag and drop to rank the following behavioural measures for positive affective state in equines in order of preference', were utilised as it has been demonstrated to be the optimal method for collecting ranked data in web surveys (Blasius, 2012).

Percentage agreement scales, such as Round one Question 15 'Please rate your percentage agreement with the following statement where 0% indicates no agreement and 100% indicates complete agreement' were utilised in a visual analogue format to give respondents complete flexibility in rating their level of agreement and have shown to be both robust and practical for use on mobile devices (Toepoel and Funke, 2018).

Free text response questions, such as Round one Question 20 'Please identify in the free text box below any methods that you feel can induce positive affect in equines that have not been described in question 20 above', were used to gather both depth and breadth of participant

knowledge and experience in these areas generating rich qualitative data for thematic analysis (Rich *et al.*, 2013; Young, 2015).

2.2.3 Demographic data collection

Questions were designed to collect demographic data from participants that were considered potentially useful in the future analysis of results obtained. Demographic data collected can be seen in Table 2.3.

Table 2.3 Demographic data collection fields.

Demographic data category	Categories provided
Field of work (multiple selections allowed)	Academia Behaviour Charity Competition Other Rehabilitation Training Veterinary
Number of years of experience in the sector	Sliding scale from 5 years (minimum inclusion criteria) up to 100 years
Age	Sliding scale from 20 to 100 years
Gender	Male, Female, Other
Country of work	Free text box
Level of qualification (multiple selections allowed)	No formal higher education qualification Undergraduate degree at UK level 6 or equivalent Postgraduate degree at UK level 7 or equivalent Postgraduate degree at UK level 8 or equivalent Professional qualification (e.g., British Horse Society qualification) Clinical Animal Behaviour Accreditation (e.g., CAB, CCAB, AAB) Veterinary licence to practise

2.2.4 Questions on personality

In round one, questions on the following topics were presented: personality definitions, traits used to describe equine personality, groupings of personality traits and the participants view on the reliability (robustness in terms of intra and inter-rater reliability) (Meng, 2020),

presence in equines (that the trait can be identified in equines) (Stamps and Groothuis, 2010; Weiss, 2018) and clarity (that the name and definition of each trait is clear) (Koski, 2011) of each trait in order to strengthen empirical research presented in subsequent chapters of this thesis. Questions posed can be found in Appendix B.

Initially it was deemed important to confirm whether participants felt that equines possess individual personality traits as this was prerequisite to the study. Then the aim was to determine a consensus on the definition of the term 'personality' due to the variability of definitions used within the literature (Koski, 2011). To ensure that all participants were working from the same definition of personality participants were asked to rate their level of agreement of the terms personality, temperament, and behavioural syndromes. Following this, participants were asked to rate each trait individually in terms of whether they felt that they were clearly defined, were present in equines and were reliably identifiable. Participants were also asked to identify whether they felt that an individual equine's personality could impact upon the level of affect experienced.

In round two, participants were presented with items which had reached consensus in round one and asked to confirm final agreement. They were then asked to rate their level of agreement with 13 groups of personality traits derived from a separate experiment (Chapter 4) and to suggest revisions to the trait group names and the group placement of individual traits as has been shown in prior personality research in humans and other non-human animal species (Miller *et al.*, 2014; Borman, 1977; Letzring, 2008). Participants were asked to select whether they agreed, or not, with the revised group names and groupings in part three of the consultation and provide any feedback they felt important in relation to these final groupings as per prior methodologies in the human literature (Samuel *et al.*, 2012).

2.2.5 Questions on emotional processing and affective state

Participants were asked to state whether they felt that equines experience affective states and therefore whether emotional welfare is an important consideration within holistic equine welfare; this was a fundamental prerequisite for the remaining questions in this section and important given the evidence in the literature suggestion that people's recognition of affective state in equines is poor (Bell *et al.*, 2019a).

Further questions included gathering expert opinion on whether valence (feelings of goodness and badness, positive and negative), arousal (a state of physiological alertness and readiness for action) and power/surgency (intensity) were all important metrics within assessment of affective state as previously suggested in the literature (Mendl *et al.*, 2010); whether experts perceived that it is difficult to distinguish between cognitive and affective components of affect as previously discussed by Netzer *et al.* (2018) and Panksepp (2003); the experts view on the presence of negative, neutral and positive affective state in equines with particular consideration of the disagreement within in the literature relating to the presence of neutral affective state (Gasper *et al.*, 2019).

Within the second round all aspects reaching consensus in round one were presented to participants for them to confirm final agreement. Definitions of non-consensus items were refined according to participant feedback from round one and questions re-presented to participants for comment.

2.2.6 Questions on inducing and measuring positive affective state

Questions were posed asking participants to separately rank (from most preferred to least preferred) methods of firstly inducing, then measuring, positive affective state in equines including both physiological and behavioural measurements which were listed as separate questions. Participants were also invited to give free text feedback after each question and to suggest other induction and measurement methods to be taken forward to the next round(s).

The methods presented in the initial lists were derived, as described in Figure 2.3, from review of the literature into cited methods to induce and measure positive affect in a range of non-human animal species (Whittaker and Marsh, 2019; Boissy *et al.*, 2007b; Paul *et al.*, 2005a; Mendl and Paul, 2020).

In round two, participants were presented with the ranked preference results from round one which included ranking of preferred methods to induce positive affect, preferred behavioural measures of positive affect and preferred physiological measures of positive affect. They were asked to state whether they agreed with the ranking or to indicate their preferred ranking for each selection. Participants were also presented with all new measures which had been identified by participants in round one and were asked to rank (most preferred to least preferred) these newly suggested methods.

In round three, participants were asked to consider both the original methods and participant suggested methods as a combined list for each section (positive affect induction, behavioural measures and physiological measures) and to rate methods which should be included in the final experimental design, which should not be included and which they were unsure or equivocal about.

2.2.7 The Delphi process

Round one

Round one of the Delphi consultation was released on 31st January 2020 and remained open for four weeks initially, requests from some participants to be given extended time to complete the survey were granted and the final closure date for round one was 31st March 2020. Participants were emailed individually with personal links to access the survey and provided with a password to access and complete their contribution. At this stage participants were provided with the closing date of round 1 and sent a further reminder email one week before the close of the survey.

Round one comprised 21 questions divided into three sections (Table 2.4).

Table 2.4 Delphi consultation round one presentation.

Survey section	Theme	Number of questions	Objective
1	Demographic data	6	1. Ascertain areas of work/specialism, years of experience, age, gender, geographical location, qualification status.
2	Equine personality	5	1. Understand participant views on equine personality, personality traits and the impact of personality on affective state. 2. To gather free-text feedback from participants
3	Emotional processing and inducing and	10	1. Understand participant views on emotional processing in equines and the presence of varying affective states

	measuring PA in equines		2. To rank methods to induce PA in equines 3. To rank methods to measure PA in equines 4. To gather free-text feedback from participants

On survey closure the results were exported from Qualtrics into a visualisation PDF file and a CSV file to facilitate further qualitative and quantitative analysis (section 2.2.8). In addition to the quantitative data collected in discrete answers to forced choice questions, free text qualitative data were reviewed using CATMA (Chiavacci, 2020; Koch *et al.*, 2022); themes and subthemes were identified, extracted and coded and then frequency counted to identify the relative support for each theme and subtheme in order to facilitate the design of Round two survey questions.

Round two

Due to the significant effects of the Covid-19 pandemic and national and regional lockdowns worldwide it was decided (following consultation with participants and advisors) to pause the release of round two of the consultation to a time when most participants would have capacity to engage with the survey. Round two was subsequently released in December 2021 and closed one month later in early January 2022 with eight extension requests granted, taking the final closure date to the end of February 2022.

In round two, participants were presented with 23 questions divided into theme-based sections as shown in Table 2.5.

Table 2.5 Delphi consultation round two presentation.

Survey section	Theme	Number of questions	Objective
1	Presentation of results from previous round	15	1) Ascertain agreement with consensus statements (n=12) from round one. 2) State their level of agreement with the outcome of the three ranking exercises: a) Ranking preference of methods to induce PA in equines b) Ranking preference of physiological methods to measure PA in equines, c) Ranking preference of behavioural methods to measure PA in equines) undertaken in round one.
2	Presentation of rephrased questions using feedback from R1	4	1) Achieve consensus agreement on rephrased questions from round one. 2) State their level of agreement with and suggest improvements to a set of personality groupings and trait names derived from responses in round one and the results of a previous field study testing personality traits in equines. 3) Free text questions were also included to gather further expert opinion in the form of qualitative data.
3	Presentation of participant suggested methods from R1 to induce	4	1) To rank further methods to induce PA in equines based on suggestions presented by participants in round one.

	and measure PA in equines		2) To rank further methods to measure PA in equines based on suggestions presented by participants in round one.

On survey closure the data were treated in the same way as in round one.

Round three

Round three was released in June 2022 and closed at the end of July 2022. Eight questions were posed in round three (Table).

Table 2.6 Delphi consultation round three presentation.

Survey section	Theme	Number of questions	Objective
1	Equine personality	2	1) State level of agreement with the revised personality group names based on suggestions presented by participants in R2 2) Provide free-text feedback for suggested group revisions
2	Methods to induce PA in equines	2	1) Participants asked to state which of the listed (from R1 and R2) methods of inducing PA in equines they felt were suitable for potential inclusion in a further field-based study 2) Provide free-text feedback for suggested group revisions
3	Methods to measure PA in equines	4	1) Participants asked to state which of the listed (from R1 and R2) physiological methods of measuring PA in equines they

			felt were suitable for potential inclusion in a further field-based study 2) Participants asked to state which of the listed (from R1 and R2) behavioural methods of measuring PA in equines they felt were suitable for potential inclusion in a further field-based study 3) Provide free-text feedback for objectives 1 and 2.

Participants were asked whether methods should or should not be included within later experimental design in order to enable identification of potentially useful methods for further field-based studies whilst reducing bias in responding (Watrin *et al.*, 2019). This method has been shown to be simple, stable and valid for measuring opinion and beliefs (Dolnicar and Leisch, 2012). On survey closure the data were treated in the same way as in rounds one and two. Copies of all questionnaires can be found in Appendix B alongside the glossary of terms provided for participants in each round to ensure that all contributors were working from the same definitions.

2.2.8 Descriptive and thematic analysis

All data were viewed and descriptively analysed in Microsoft Excel (Microsoft Excel V16). Demographic data were analysed qualitatively. Free text responses were analysed and coded using CATMA™ (Guis *et al.*, 2024), according to themes identified within the answers to individual questions. Total occurrences of comments within each theme were recorded and sample quotes from participants were selected for reporting where the content of the quote was replicated several times across participants making it a persistent theme (Lochmiller, 2021). Demographic data were graphed in Microsoft Excel (Microsoft Excel V16) and both count and percentage response rates were calculated for each question. Percentage response rates to questions were evaluated to identify whether the 70% consensus threshold had been reached. Ranked data were aggregated into ‘agree’, ‘unsure’, and ‘disagree’ categories and category percentages calculated to identify whether the 70% consensus threshold had been reached.

Chapter 3 Inducing and measuring positive affective state in equines: a Delphi consultation

2.1 Introduction

The equine sector in the UK and internationally has a global value of around \$300billion, employing millions of people worldwide (Equine Business Association, 2022; BETA, 2019) and incorporating an estimated 116 million equids globally (The Brooke, 2022). Consequently, deepening our knowledge regarding equine emotional state could be of significant benefit to equine welfare across sectors through enabling a greater understanding of positive affective states to enhance emotional welfare and improve overall quality of life across the diverse equine sectors (Figure 3.1).

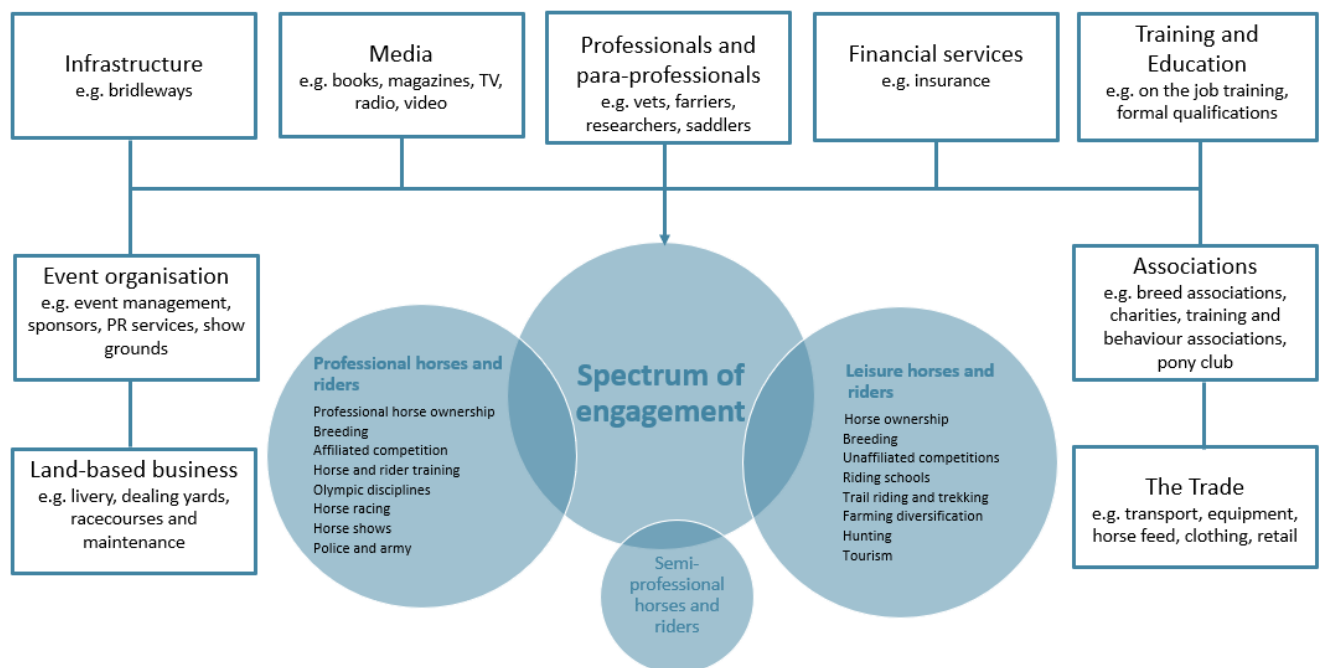


Figure 3.1 Conceptual map of the equine industry and upstream supply categories, adapted from DEFRA (2004).

The significant reach of the equine industry highlights the importance of the standards set within it, including equine welfare standards. There is a plethora of evidence within the literature relating to behavioural and physiological measures of negative affective state (a general factor of subjective distress) and poor welfare (adverse effects that reduce fitness) in horses (von Borstel *et al.*, 2017; Pearson *et al.*, 2021). Over the last twenty years the definition

of good welfare has advanced from the ‘absence of negative welfare’ to the aim of identifying the presence of positive welfare. Indeed, in 2019 the journal *Animals* published a special issue entitled “Positive Aspects of Animal Welfare” with a clear mission to develop research focusing on the development and validation of positive welfare indicators, where it stated:

“...Most of the indicators developed so far focus on negative aspects of animal welfare (e.g. lameness, lesions, diseases, presence of abnormal behaviour, high levels of stress hormones, and many more), and only very few indicators are presently available to highlight positive aspects of animal welfare. However, the lack of negative welfare conditions represents just a minimum standard that should be guaranteed to animals, and it does not necessarily mean that animals are in good welfare conditions and have a good quality of life. To guarantee high welfare standards, animals should experience positive conditions, that allow them to live a life that is really worth living.

For these reasons, the new frontiers of animal welfare should aim to reach this high quality of life, and research should now focus on the development and validation of positive welfare indicators.”

Key aspects for consideration within the enhancement of positive welfare are:

- The strong evidence supporting the experience of affective states in equines,
- Potential methods to induce and measure affective states in equines,
- The effect of individual personality on affective states.

To this end the development of the ‘Five freedoms or needs’ model (Webster, 2005) to the ‘Five Domains’ model (Mellor, 2016) advanced prior work in this area to consider the affective experiences domain which encompasses both positive and negative affective states. This development built on prior work establishing the capacity of non-human animals to experience affective states and highlighted that in addition to physical health and wellbeing, psychological health is fundamental to holistic ‘wellness’ (Hemsworth *et al.*, 2015), as discussed in detail in Chapter 1. Whilst there is significant literature in this area which suggests agreement on the importance of psychological health in non-human animals (Mendl *et al*, 2010; Mellor, 2012, 2016; Hemsworth *et al*, 2015; Mendl and Paul, 2020; Mellor *et al*, 2020), in the equine literature there is little evidence surrounding the level of consensus agreement

across experts on aspects of equine affective state, particularly positive affective state (McGreevy *et al.*, 2018).

Hall *et al.* (2018) highlighted the importance of assessing positive equine emotional state in equine welfare, and evaluated methods previously used to attempt to measure affective states in equines; whilst simultaneously suggesting potential future methods requiring further validation and consensus agreement, stating that both behavioural and physiological correlates were required to enable accurate identification and validation of affective state.

To our knowledge very little empirical research has been undertaken into methods to reliably measure positive affect in equines despite the literature highlighting the importance of this emotional state for good welfare since the early 21st century (Mellor *et al.*, 2020; Mellor, 2012; Boissy and Erhard, 2014; Phillips, 2007). Lawrence *et al.* (2019) highlighted the paucity of positive animal welfare (PAW) literature identifying just ten papers across species in the core literature, with many authored by the same research groups. They identified that the PAW literature encompassed four common features:

1. Positive emotions – the most widely acknowledged aspect within the literature,
2. Positive affective engagement – functional linking of positive emotions to goal-directed behaviour,
3. Quality of life – situates PAW within the context of finding a good balance of positive over negative,
4. Happiness – considering the animal's full-life perspective.

In addition, personality (the combination of characteristics or qualities that form an individual's distinctive character) and how this may affect valence (feelings of positive and negative emotion), arousal (level of psychological and physiological stimulation) and intensity (the strength of the emotional state) of equine emotional state has not been considered, in relation to positive affect, within the equine literature to date; although it has been discussed by Boissy and Erhard (2014) in farm animals. It is likely that personality impacts upon emotional state (Boissy and Erhard, 2014) and that these two factors are salient within the cognitive processing (mental operations the brain performs to process information) of the individual. This may lead to cognitive bias (such as optimism or pessimism) (Henry *et al.*, 2017; Hintze and Schanz, 2021; Löckener *et al.*, 2016; Mendl *et al.*, 2009; Paul *et al.*, 2005b) that

could affect decision making and emotional welfare as demonstrated in other non-human animal species (Robinson *et al.*, 2021; Robinson *et al.*, 2016; Gosling, 2008).

Before attempting to measure positive affective state, it is important that it can be induced with a sufficient level of certainty that the measurement methods being tested are assessing the intended emotional state to ensure a high level of specificity (Hall *et al.*, 2018). It is important to be able to differentiate between different emotional states of varying valence and arousal to determine the presence of positive valence within a range of arousal levels. To facilitate this it is vital that individual preferences for stimuli believed to be capable of inducing positive affect are considered within experimental designs through a systematic preference testing approach such as that used by Holzer *et al.* (2022).

When considering methods to induce and measure positive affective state in equines it is important to ensure that the techniques employed are deemed to be reliable (the degree to which the result of a measurement, calculation, or specification can be depended on to be accurate), repeatable (low variation in measurements taken by a single instrument or person under the same conditions), sensitive (how well a test can detect a specific state in animals who are actually experiencing that state) and specific (the frequency with which a test shows a true negative result among individuals who are not in that state). This ensures that positive and negative predictive values reported are accurate, i.e. close to true value (Trevethan, 2017).

To satisfy these aims a Delphi consultation (Taghipoorreynah, 2023) of experts in the field of equine behaviour, welfare and affective state was undertaken to derive consensus agreement on definitions used within the consideration of positive affective state in equines and the methods perceived as suitable to use to induce and measure these states in later experimental investigations (Chapter 6). A Delphi study was chosen over other methods such as a simple survey, focus group or interviews to investigate the opinions of experts internationally in an anonymous manner through a multi-staged iterative process. The Delphi technique is a widely accepted method for generating consensus of agreement between experts on a particular topic (Keeney *et al.*, 2006; Boulkedid *et al.*, 2011) and has been used in the field of equine behaviour and welfare previously (Pannewitz and Loftus, 2023; Pearson *et al.*, 2021; Rioja-Lang *et al.*, 2020). The objectives of a Delphi study are to provide a group communication process (which is usually anonymised) to examine and discuss a subject area,

to set goals, predict events or investigate policy. The aim is a 'what could/should be' rather than a 'what is' type of approach (Hsu and Sandford, 2007b). Within the current study this approach was undertaken due to the novelty of the area of research and the lack of literature available at the time to devise a reliable empirical study into measuring positive affective state in equines. Prior studies utilising the Delphi methodology have set a range of consensus levels ranging from 51-80%, with no universal consensus level agreed within the literature (Yeoman *et al.*, 2017; Hotchkiss *et al.*, 2006; Green *et al.*, 1999; McKenna, 1994) with recent literature citing a robust consensus level of 70% agreement (Pannewitz and Loftus, 2023).

3.2 Methodology

A three-round Delphi consultation of experts was conducted between January 2020 and June 2022 with the extended timeframe attributable to the Covid-19 pandemic. The full methodology can be reviewed in detail in Chapter 2 however here the methods are summarised in Figure 3.2.

The consultation was designed to focus on three main areas:

1. Expert opinion on the validity of 54 selected personality traits and their development into an equine personality questionnaire (results reported in Chapter 4),
2. Expert opinion on the experience of affective state in equines including positive, negative and neutral affective states,
3. Expert opinion on potential methods that could be used to induce and measure positive affective state in equines.

Each Delphi round was an iterative process where the results of the previous round were presented to participants to ratify the agreed consensus and items not achieving consensus were adjusted following expert feedback throughout the entirety of the consultation. All Delphi participant's suggestions were added to the presented results to ensure that all viewpoints were represented and rated by all participants. After each round the results were qualitatively and quantitatively analysed and presented back to the participants as ranked items or percentage agreement items. At every stage all participants had the opportunity to add their opinions and suggestions through the provision of free-text boxes with the length

of text unlimited to allow for full expression of opinion and expertise. All qualitative data were thematically analysed at each stage and used to refine subsequent rounds of the consultation.

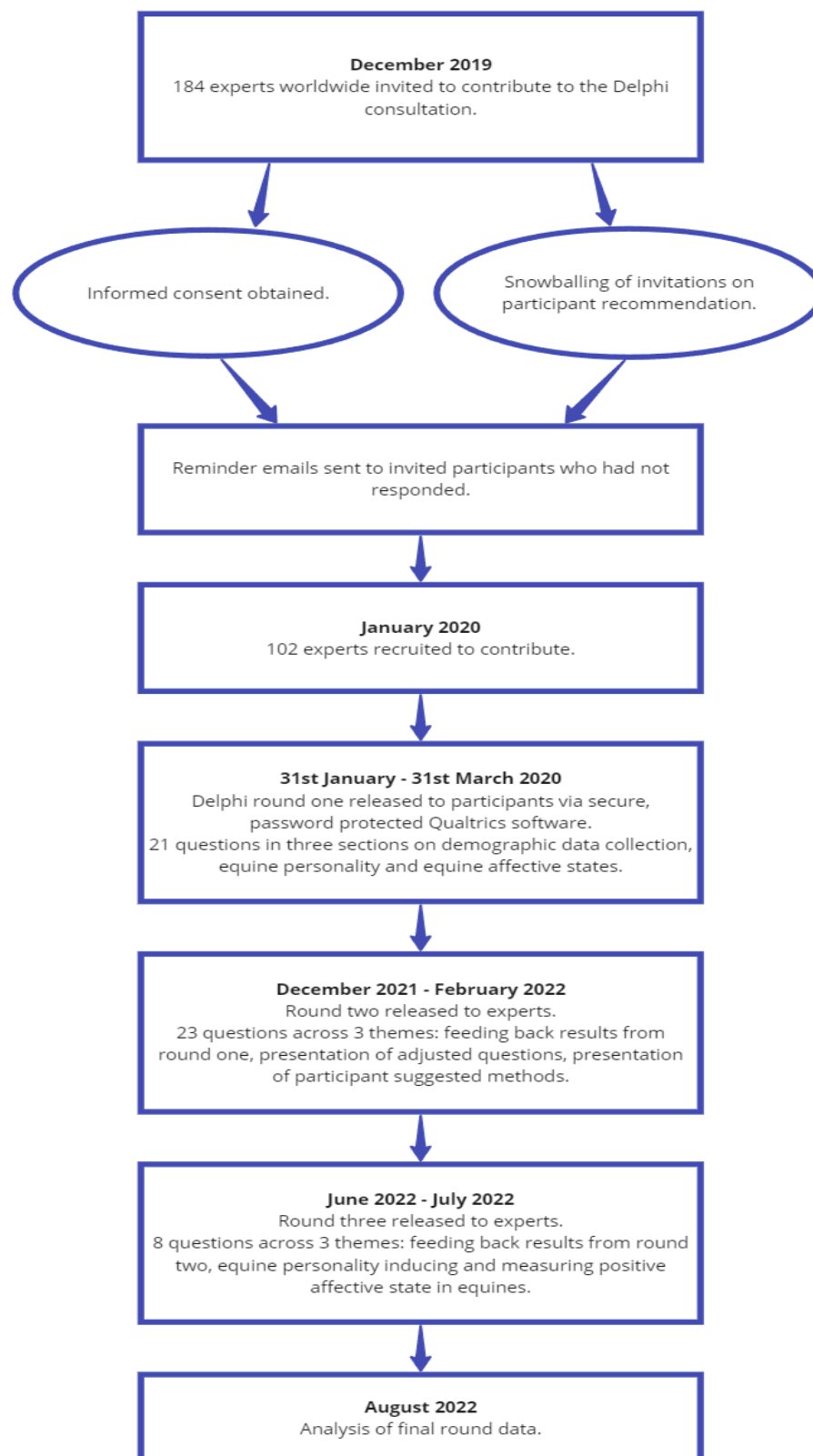


Figure 3.2 Summarised diagram of the Delphi process.

3.3 Results

3.3.1 Retention rate

Within this consultation retention rates were high from round to round. Ninety-three respondents completed round one from 103 who confirmed participation (90% response rate), 89% (n = 83) continued from round one to round two and 88% (n = 73) from round two to round three (89% overall).

3.3.2 Personality

Results pertaining to personality questions are presented and discussed in Chapter 4. This chapter focusses on results surrounding affective states in equines and inducing and measuring positive affect in equines.

3.3.3 Round one

Following the closure of round one, 93 responses were received (90% response rate). Some participants did not answer every question throughout the survey leading to variety in the numbers of responses to each question ranging from n = 71 to n = 93.

Demographic data

Participants were drawn from a wide range of fields within the equine sector (Figure 3.3). Some participants selected more than one area of expertise with some selecting up to five areas in total.

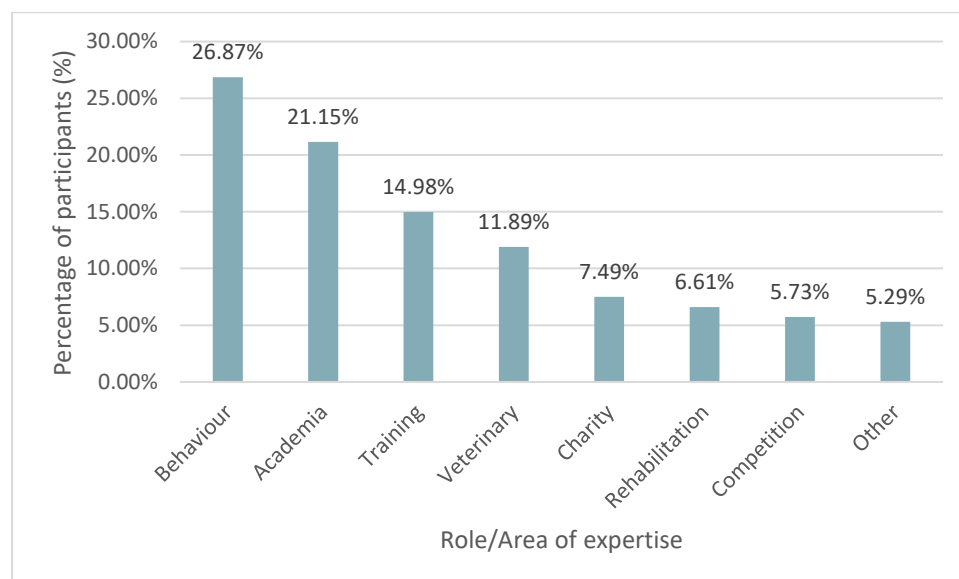


Figure 3.3 Demographic data surveying participants area of expertise.

Those selecting the 'other' option, in relation to their area of expertise, reported roles such as breeding (n=1), farriery (n=1), consultant (n=5), horse-boarding (n=1), scientist (n=1), coach (n=2) and animal assisted interventions (n=1).

Mean years of experience in their area of expertise was 21.75 (SD10.92, range 5-58 years). Mean age of participants was 47.31 years (SD11.45, range 20-81 years). Male respondents (n=10) made up 19% of the group with females (n=83) making up the remaining 81% of participants; no further gender options were selected.

Participants were located across thirteen countries as shown in Figure 3.4 with the modal number of participants in the United Kingdom (n=59).

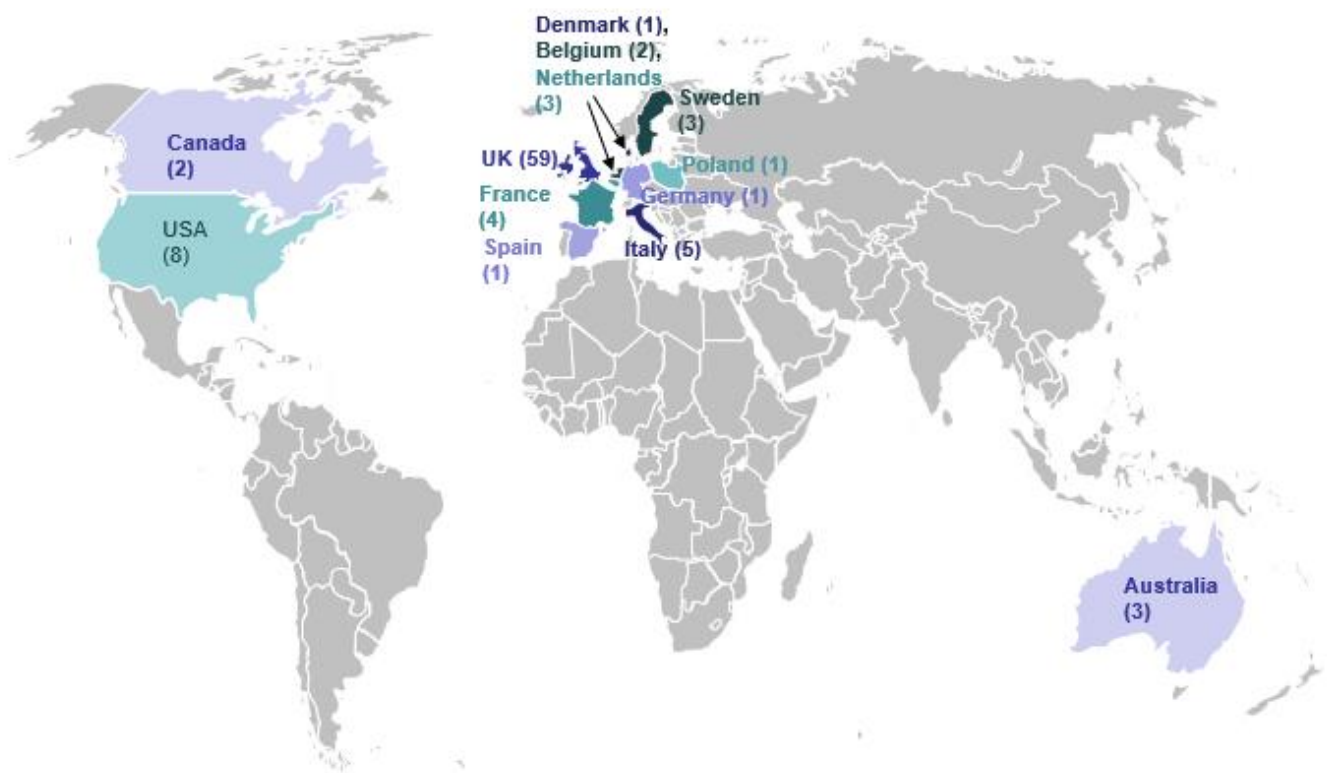


Figure 3.4 World map showing location of participants and number of participants per country (Authors own image with editable map from Softrionics).

Most participants held higher level qualifications and additional accreditations (Figure 3.5).

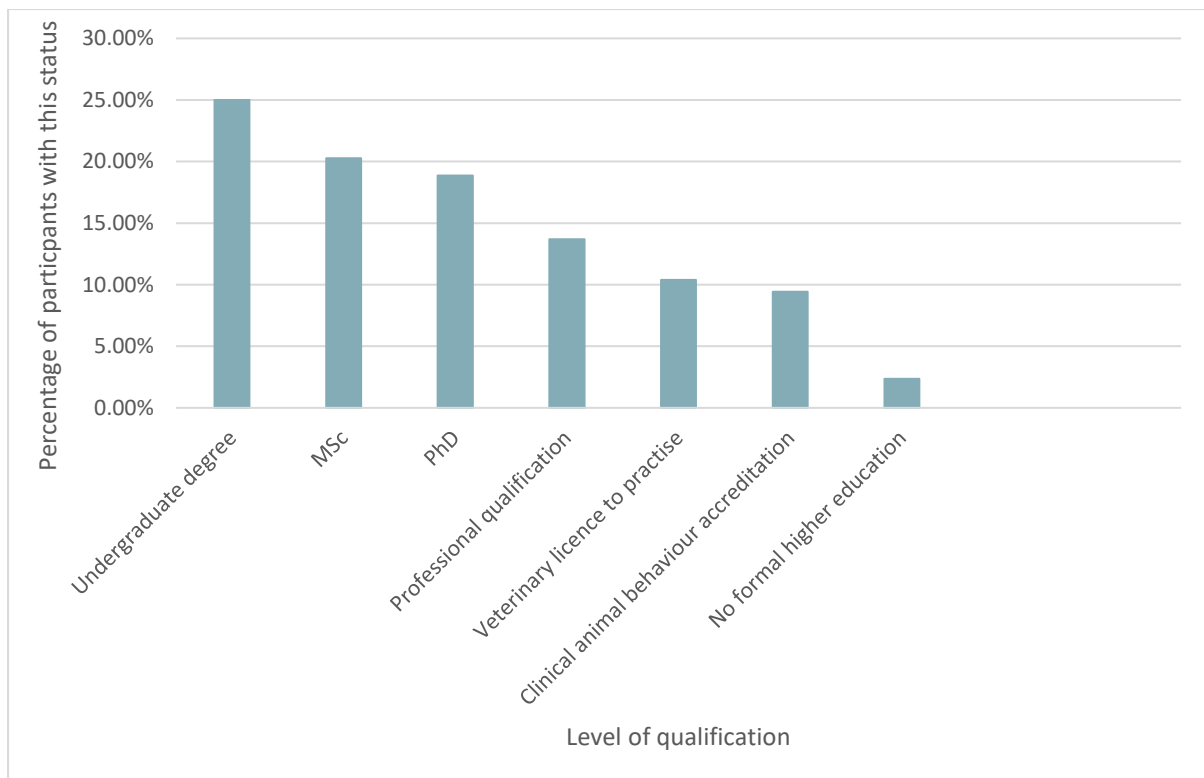


Figure 3.5 Bar chart showing the qualification status of participants.

Emotions in equines

Importance of emotional welfare

Ninety seven percent of participants ($n = 73$) agreed that emotional welfare is an important facet of equine welfare by rating its importance above 70% agreement on a sliding percentage agreement scale. Most respondents (73%, $n=55$) stated a percentage agreement of 90% or higher. Only two (2.6%) respondents reported a percentage agreement lower than 70% (20 and 34% respectively) and the remaining 18 (24%) rated their percentage agreement between 70-89%.

Importance of valence, arousal and surgency in equine emotional welfare

Respondents rated the aspects 'valence' (89.9% agreement) and 'arousal' (81.335% agreement) in affective state as very important in equine welfare with consensus reached on these two aspects. The importance of surgency (or power) of the emotional state did not reach consensus agreement (56.76%) at this stage (Figure 3.6).

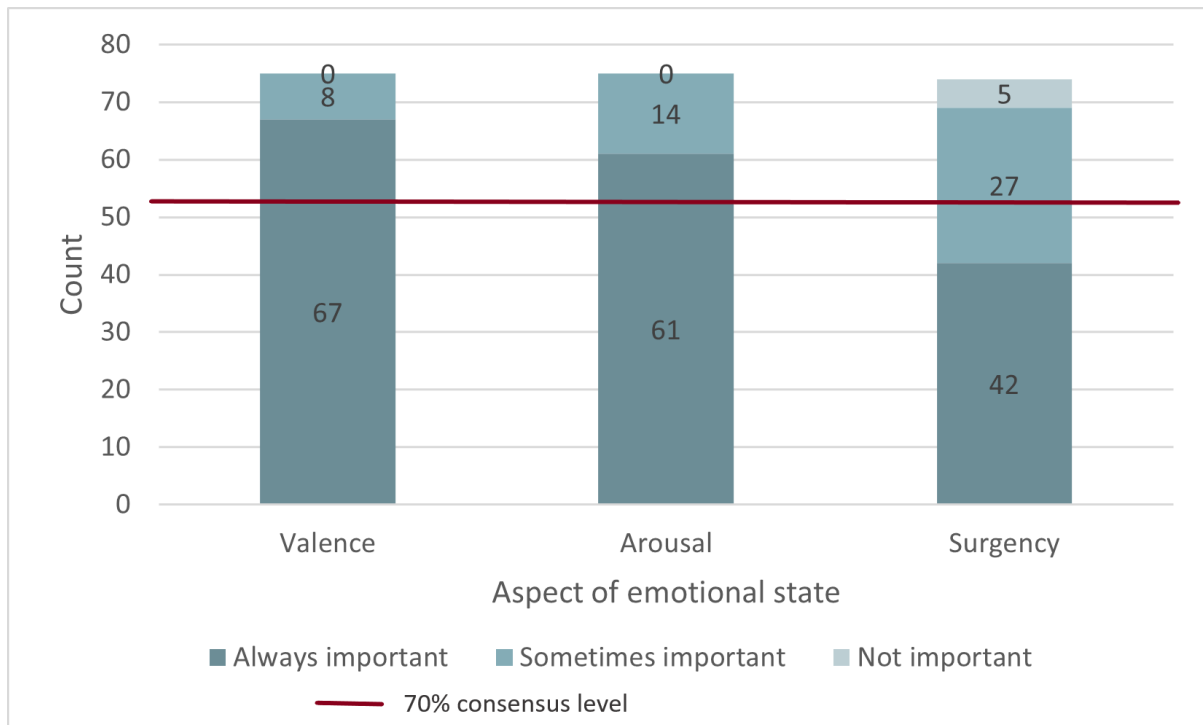


Figure 3.6 The level of importance participants placed on the characteristics of valence, arousal and surgency in relation to equine emotional state.

Contributors to affective state, distinguishing between cognitive and affective components and experience of different valences of affect

Participants demonstrated high levels of agreement on systemic contributors to affective state and agreed that it is difficult to distinguish between the cognitive and affective components of behavioural outputs in equines (Table 3.1). The consensus level of 70% agreement was also exceeded in relation to the ability of equines to experience positive, negative, and neutral affective states.

Table 3.1 Levels of agreement relating to questions surrounding affective state in equines. Values reaching consensus are placed in bold.

Statement	Agree % (n)	Disagree % (n)	Unsure % (n)
Emotional processes include motor-expressive, sensory-perceptual, autonomic-hormonal, cognitive-attentional and affective-feeling aspects	89.33 (67)	8.00 (6)	2.67 (2)
It is difficult to distinguish between the cognitive and affective aspects unambiguously when undertaking field-based research	70.67 (53)	13.33 (10)	16.00 (12)
Equines can experience negative affect (negative emotional state)	93.06 (67)	5.56 (4)	1.39 (1)
Equines can experience neutral affect (neutral emotional state)	80.55 (58)	11.12 (8)	8.33 (6)
Equines can experience positive affect (positive emotional state)	91.66 (66)	5.56 (4)	2.78 (2)

Participants agreed (72%, n=54) that an equine's individual personality would influence the intensity of emotional state experienced; however, 28% (n=21) of respondents were unsure.

Inducing positive affective state in equines

When asked to state their level of agreement (strongly disagree to strongly agree) with proposed methods to induce positive affect in equines participants reached consensus on eight of twelve methods (Figure 3.7). Results were aggregated into 'agree' (included strongly agree and somewhat agree), 'disagree' (included strongly disagree and somewhat disagree) and 'unsure' categories to establish if the overall level of agreement reached the 70% consensus threshold for each measure.

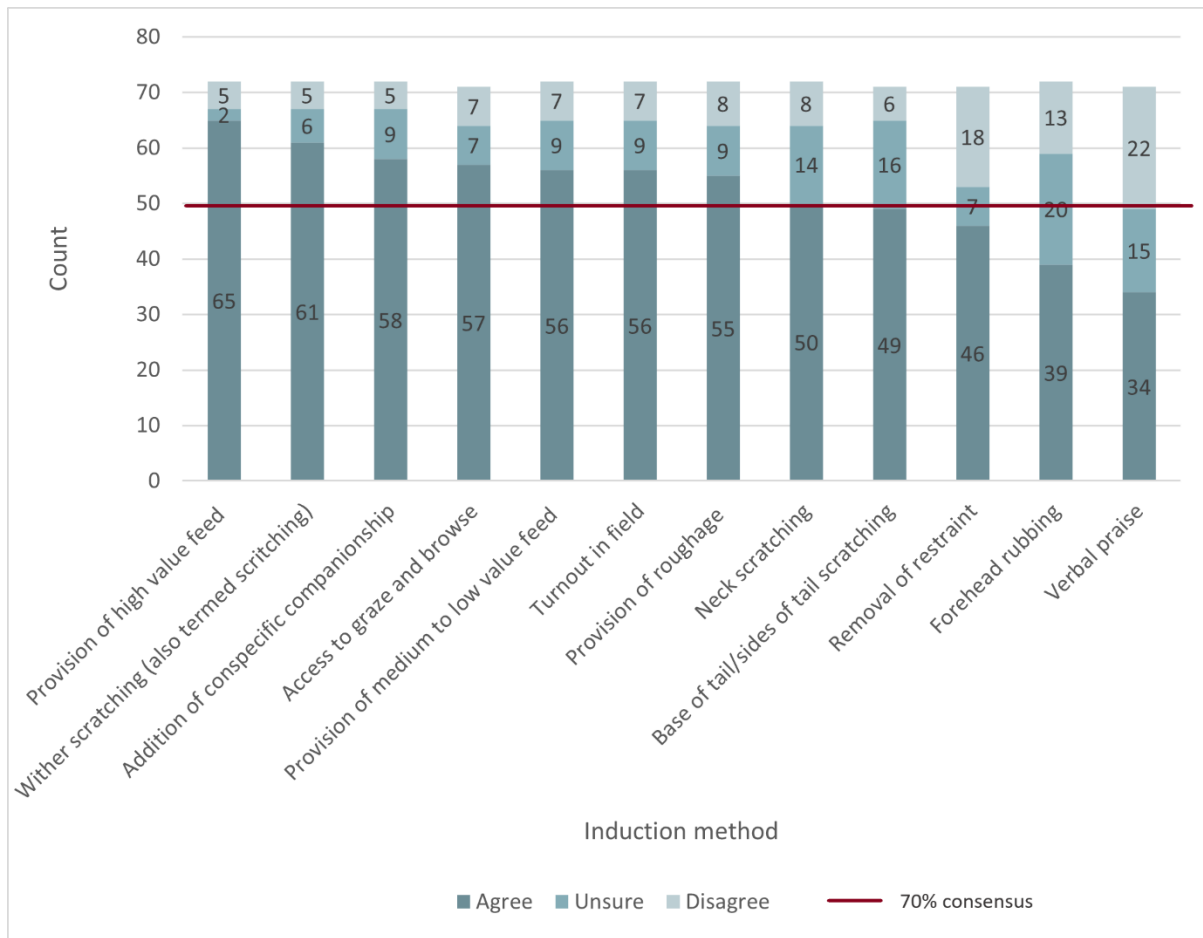


Figure 3.7 Stacked bar chart depicting level of agreement on methods to induce positive affect in equines.

Overall ranking results reported four induction methods with the highest level of ‘strongly agree’ answers, these were ‘provision of a high value concentrate feed’ (48.61% strongly agreed), ‘access to graze and browse’ (40.85% strongly agreed), ‘addition of conspecific companionship’ (34.72% strongly agreed) and ‘withers scratching’ (25% strongly agreed) (Table 3.2).

Table 3.2 Expert rankings of methods to induce positive affect in equines. Overall rank = the ranking for Round one. Mean (rank) = mean of all rater rankings (1 is low and 5 is high rank) with SD. Mode and median values are also represented. Higher mean, median and mode values equal higher agreement with the measure. Overall rank = sum of all ranking values by raters and associated preference rank.

Method to induce positive affect	Overall Rank	Mean (rank)	SD	Mode	Median
Provision of high value (as perceived by the individual animal) feed e.g. Carrots, apples, concentrate feed such as coarse mix	1	4.31	0.88	5	4
Access to graze and browse	2	4.07	1.05	5	4
Addition of conspecific companionship	3	4.06	0.94	4	4
Wither scratching (also termed “scritchng”)	4	4.00	0.87	4	4
Turnout in field	5	3.97	1.03	4	4
Provision of roughage feed e.g. Hay, chaff, haylage	6	3.88	0.93	4	4
Provision of medium to low value (as perceived by the individual animal) concentrate feed e.g. Pony nuts, high fibre cubes	7	3.85	0.86	4	4
Base of tail/sides of tail scratching	8	3.69	0.9	4	4
Neck scratching	9	3.67	0.88	4	4
Removal of restraint	10	3.52	1.23	4	4
Forehead rubbing	11	3.43	0.95	4	4
Verbal praise	12	3.15	1.11	4	3

Qualitative analysis of free-text responses in relation to methods to induce positive affect in equines identified twelve main themes (Table. 3.3).

Table 3.3 Themes, subthemes and counts of suggestions made by participants when invited to add free text responses.

Themes & <i>subthemes</i>	Responses (count)	Selected quotes from respondents
Individual factors <i>Preferences</i> <i>Motivation</i> <i>Context</i>	22	“Motivation for many of these resources affects response to access. So if the animal is really motivated to be turned out, then the turnout is a very positive experience. If the animal is not motivated to be turned out, or scratched then the experience will not be positive, or may even be negative”
Physical <i>Access to roll</i> <i>Patting</i> <i>Grooming (human or horse)</i> <i>Temperature</i> <i>Physical therapies</i>	22	“Access to rolling in a substrate such as sand. Grooming. Equine massage” “letting the horse to roll over, get dirty in the mud or sand, play...”
Conspecific access <i>Preferred partners</i> <i>Stable groups</i> <i>Mare-foal interactions</i>	13	“Degree to which an individual seeks out and interacts with conspecifics” “Meeting with favourite conspecific ("best friend")”
Net difference before and after <i>Removal of restraint</i> <i>Return to safe place</i> <i>Provision of a basic need</i> <i>Negative reinforcement</i>	13	“Reunion with horse/human friend” “Return to a safe place after exploring a novel environment”

		"Any change for the 'better' from the point of view of the animal"
Human interactions <i>Positive horse-human relationship</i> <i>Prior negative experiences</i>	10	"Some of the above methods could induce positive affect depending on the relationship between horse and people"
Play <i>Object</i> <i>Oral</i> <i>Solo</i> <i>With companion</i> <i>With human</i>	9	"Opportunity for oral object play, e.g. horses playing with lead ropes, or "toy box" as part of behavioural enrichment"
Training <i>Clicker</i> <i>Positive reinforcement training</i>	7	"Positive reinforcement training"
Giving choice or agency	7	"Offering choice and subsequently allowing the animal to act according to preference shown"
Provision of non-frightening novel objects	6	"Opportunities to explore novel objects /situations (for horses who have relatively sterile / mundane environments)"
Exercise <i>Turnout</i> <i>At liberty work</i> <i>In hand work</i> <i>Exploration</i> <i>Ridden work</i>	6	"Other methods include hacking out under saddle, cross-training, such as free-jumping and other engaging activities"
Mental Enrichment <i>Problem solving</i> <i>Games</i> <i>Puzzles</i>	5	"Cognitive challenges that can be solved (Eureka effect)"

		“Addition of enrichment in the field environment e.g. a yoga ball”
Overall lifestyle <i>Health</i> <i>Nutrition</i> <i>Handling</i> <i>Environment</i> <i>Social</i>	5	“Combine several positives to create a longer-term positive 'mood' (in the field with preferred conspecifics, receiving high value reward (or solve a maze to get a food reward) and opportunities for voluntary grooming etc)” “...the fulfilment of a need is very likely to induce a positive affect”

In addition to these observations many respondents signalled the need for preference testing of some induction methods within each individual equine before using them in the final experimental design.

One participant reported,

“The problem in my opinion is that it is difficult to find one method that works on all horses. You have to identify what each horse perceive[s] as positive which makes it very difficult methodological[ly]”

Whilst another commented,

“Difficult to answer as impact of the approaches above will vary depending on the individual horse and I would suggest any approach needs to be tailored to that individual which is not easy”.

another supported this view stating,

“Unfortunately, I think it is quite individual, as each animal has a perception of the value of the object or interaction. High value feed is likely to be valued by the greatest majority of individuals. Scratching may be highly valued by some individuals, but only tolerated or even disliked by others”.

The provision of a high value feedstuff was reiterated by respondents with comments such as:

“Agree that food is a direct and instant inducement. Other options are too but food gets an immediate and consistent instant response”.

Measuring positive affective state in equines.

Consensus was reached with most participants 90.67% (n = 68) agreeing that a multi-measure approach is optimal for measuring positive affect in equines. A further 4% (n = 3) were unsure and 5.33% (n = 4) disagreed.

Behavioural methods to measure positive affective state in equines.

‘Qualitative Behavioural Assessment’ was the highest ranked behavioural method for measuring positive affect in equines with 29.17% (n = 21) of participants ranking this method at number one. Conversely ‘absence of repetitive behaviours (stereotypies)’ was the lowest ranked method by the highest number (45.83%) of participants (n = 33) (Table 3.4).

Table 3.4 Results for expert ranking of behavioural methods to measure positive affect in equines. Overall rank = rank in this round with 1 being high and 8 being low. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD, mode and median values identified.

Behavioural Method to measure positive affect	Overall rank	Mean rank	SD	Mode	Median
Free choice profiling/Qualitative Behaviour Assessment (QBA)	1	2.93	1.95	1	2
Cognitive bias testing	2	3.79	1.94	4	4
Facial action scoring	3	3.86	2.20	1	4
Human-Animal relationship tests	4	4.17	1.97	4	4
Time budget analysis	5	4.68	2.52	8	5
Attention bias testing	6	4.92	1.81	5	5
Novel object tests	7	5.65	1.77	7	6
Absence of repetitive behaviour (formerly stereotypy)	8	6.00	2.39	8	7

Seventeen main themes were identified within the free text responses on behavioural methods to measure positive affect in equines as shown in Table 3.5. In addition to these observations five respondents reiterated their preference for qualitative behaviour assessment as a behavioural measure of positive affect.

Table 3.5 Themes, subthemes and counts of suggestions made by participants when invited to add free text responses.

Themes & <i>subthemes</i>	Responses (count)	Selected quotes from respondents
Horse-horse <i>Herd-related happiness</i> <i>Horse-Horse interaction</i> Animal-animal interaction Social behaviour	9	“I would look at herd-related happiness first then human-related happiness” “I would look at the horse's choices and behaviours when free in his herd life first, such as companion choice, anxiety about other horse proximity...”
Body language <i>Whole body</i> <i>Body posture</i> <i>Body tension</i>	5	“Would look at overall body language” “Facial expression”
Horse-human <i>Human-related happiness</i> <i>Response to human cues</i> <i>Positivity towards humans</i> <i>Horse-human interaction</i>	5	“Compliance with human trained requests” “One that has not been used much (in terms of positivity toward humans) is the response horses make when you walk through a barn/stable”

Play <i>Object</i>	4	"Faster approach and more touch/play indicating more positive affect"
Free choice <i>Free choice</i> <i>Choice of resources</i> <i>Freedom to choose</i>	3	"Choice tests"
Anticipation <i>Anticipate</i> <i>Anticipatory behaviour</i>	2	"Anticipatory behaviour could perhaps also be useful (but again would need to be validated for positive affect in horses)"
Preference tests	2	"Preference testing"
Variety <i>Resources</i>	1	"Variety of forage"
Cognitive flexibility	1	"Cognitive flexibility[.]Personality change (after extreme experiences)"
Equine Behaviour Assessment and Research Questionnaire (E-BARQ)	1	"Equine Behaviour Assessment and Research Questionnaire (E-BARQ)"
Judgement bias	1	"A reaction to an unexpected visual, tactile or sound stimuli"
Motivation tests	1	"Motivational tests"
Rebound behaviour <i>Recovery time</i>	1	"Rebound behaviour - intensity, latency to return to normal parameters"
Frustration tests	1	"Frustration tests (arousal, rebound, motivation, aggression...)"

State-trait tests	1	“[It] depends if you want to look at immediate state or overall [traits]”
Vocalisation	1	“Recording and analysing vocalisation”
Behaviour change <i>When the environment changes</i>	1	“The horse may react totally differently [...] in his field vs the arena or stable (human areas) or with a human present or not”

Physiological methods to measure positive affective state in equines

Heart Rate Variability (HRV) was the highest ranked physiological method for measuring positive affect in equines with 42.19% of participants ranking this method at number one (mean score 2.48). Heart rate (HR) and salivary cortisol were ranked methods two and three respectively with body temperature the lowest ranked method by 53.13% of participants (mean score 11.67) (Table 3.6).

Table 3.6 Results for expert ranking of physiological methods to measure positive affect in equines. Rank = rank in this round. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD, mode and median values identified.

Physiological Method to measure positive affect	Overall rank	Mean rank	SD	Mode	Median
Heart rate variability (HRV)	1	2.48	2.02	1	2
Heart rate (HR)	2	4.23	3.06	2	3
Salivary cortisol	3	5.61	2.60	3, 4, 5, 7	5
Blood cortisol	4	5.98	2.94	4	6
Serotonin levels	5	6.27	3.85	2	5
Infrared thermography (IRT)	6	6.75	4.71	14	5
Electroencephalogram (EEG)	7	7.16	4.20	1	8
Hair cortisol	8	8.48	3.31	6	8

Urinary cortisol	9	8.64	2.43	8	8
Corticotrophin releasing factor challenge	10	9.00	2.77	10	10
Immune system function tests	11	9.36	3.14	10, 11	10
Catecholamine levels	12	9.68	3.05	11	10
Body temperature	13	11.67	3.43	14	14

There was a high variance across raters in the ranking of some measures including infrared thermography and electroencephalography. Some measures had multiple modal values which likely impacted upon their overall ranking (salivary cortisol and immune system function tests).

Twenty main themes were identified within the free text responses provided by participants around physiological measures to measure positive affect in equines (Table 3.7). In addition, seven respondents commented that they perceived issues with lack of validation or poor validation of physiological methods to measure positive affect with additional issues concerning specificity of methods highlighted. Participants commented that they were concerned that it is difficult to distinguish between positive and negative affect (i.e. valence) when using methods that measure markers of arousal but do not quantify valence.

Participants stated:

“To my knowledge none of these measures have been validated to measure positive affective states”

And

“I feel our physiological measures are more indicative of negative affect than positive; also, to me, the most important positive affect is something that lasts for a while (many [...] might work for an acute, negative situation or acute stressor), I have far less confidence that we can measure a "chronic" state of positive. At some point we must get brave enough to accept qualitative behavior[u]ral assessment may be one of our best tools for assessing positive affect”

Others commented that

“The physiological measures are not good methods for evaluat[ing] positive affective state”

And

“... but often the problem seems to be distinguishing between different types of arousal (positive or negative)”

Six participants reported that they did not feel they had a secure enough level of knowledge around physiological measures to give a valid opinion. These participants declined to answer this question.

Table 3.7 Themes, subthemes and counts of suggestions made by participants when invited to add free text responses.

Themes & subthemes	Responses (count)	Selected quotes from respondents
Time period <i>Acute versus chronic</i> <i>Long versus short term</i>	11	“...nearly all physiological markers of arousal, and acting over very different time periods. Measure chosen would depend on question asked e.g if wanting to measure arousal over seconds or months.”
Movement <i>Patterns</i> <i>Step count</i> <i>Step frequency</i> <i>Biomechanics</i> <i>Accelerometers</i>	6	“Movement patterns, not just time motion analysis, what type of movement is performed, for example, walking in small circles, reluctance to stand still indicates very alert/excited/anxious, very slow walking may indicate pain”
Oxytocin	4	“Blood oxytocin”

Faecal cortisol	2	"Faecal cortisol should be in here too if you're going to have the other cortisol measurements"
Hypothalamus-pituitary-adrenal axis (HPA) axis challenge	1	"...many of these physiological indicators would pick up a negative affective state, but would need to be validated for positive affective states (e.g., HPA axis challenges..."
Respiratory rate	3	"Respiratory rate"
Cortisol <i>Awakening response</i> <i>Diurnal patterns</i>	2	"The diurnal pattern of cortisol and cortisol awakening response may be linked to positive affect"
Opioids and their metabolites	2	"Opioids"
Magnetic resonance imaging (MRI) <i>During cognitive tasks</i>	2	"If functional MRI becomes more mobile & less expensive, I do feel that can play a role" "MRI scans during cognitive tasks (though expensive), have had great results in dogs trained to go inside"
Pain score	2	"Pain score"
Skin conductivity/sweat rate	2	"Skin conductivity/sweat rate has received little investigation and could be useful"
Tryptophan	1	"...the serotonin precursor tryptophan"
Sleep patterns	1	"Measuring sleep patterns"
Eye <i>Measurement of eye twitch</i> <i>Blink rate analysis</i>	1	"I would also like to see some validation of blink rate"

Chew rate	1	“Measuring frustration when eating for example (bite rates and chew rates)”
Lameness <i>Score Assessment</i>	1	“Lameness grade or score”
Eye temperature	1	“Eye ball temperature”
Muscle tension assessment	1	“Biomechanical software to quantify head and neck postures (i.e., low flat depressed posture - or what you would determine for an appropriate posture for positive affective state)”
Computed tomography (CT) analysis	1	“Tomography”
Gut microflora analysis	1	“Gut microflora”

3.3.4 Round two

Eighty-three responses were recorded in round two, an 89% response rate. Again, not all participants answered each question with response rates to questions ranging from n = 63 (76%) to n = 83 (100%).

Confirming consensus agreement

In round two participants confirmed their agreement with statements that had achieved consensus in round one (R2 part one). Consensus was reached on twelve of sixteen statements (Table 3.8).

Table 3.8 List of agreed statements, percentage agreement and additional feedback supplied by participants. Only statements achieving agreement consensus are included within the table (S1-11, S13).

Statement	Consensus level (%)	Additional feedback
S1 - It was agreed that 'equines possess individual personality traits'.	100	None
S2 - It was agreed that 'personality is the individual's general behaviour compared with those found in the general population'.	92.31	"Rephrase to 'the individual's particular behaviour'".
		"It is important to ensure that affect, cognition and motivation are considered."
		"Could be described as just 'individual behavioural tendencies'."
		"Unique characteristics and tendencies to respond in certain ways."
		"Ensure emotions are explicit interacting factors in individual behaviour."
S3 - It was agreed that 'Following the recognition of animals as sentient beings their emotional welfare has become an important issue due to their capacity to experience affective experiences and thus experience stress and fear or conversely enjoy positive feelings.'	97.44	"Context is also important."
		"Mostly agree but this was also appreciated by many before formal recognition (legal) of sentient status."
	97.44	"Mostly yes, but many advocated concern for emotional welfare of animals even before their (formal) recognition as sentient beings. This seems to imply a historical link that I am not sure is accurate."
		"I was torn between ticking yes and no. If we are describing affect as either positive or

S4 - It was agreed that 'valence is always important when assessing affective state'.		negative how is affect then different to valence?"
		"I wouldn't say always?"
S5 - It was agreed that 'level of arousal is always important when assessing affective state'.	100	None
S6 - It was agreed that 'Emotional processes include motor-expressive, sensory-perceptual, autonomic-hormonal, cognitive-attentional and affective-feeling aspects'	100	None
S7 - It was agreed that 'It is difficult to distinguish between the cognitive and affective aspects unambiguously when undertaking field-based research'	100	None
S8 - It was agreed that 'Equines can experience negative affect (negative emotional state)'.	100	None
S9 - It was agreed that 'Equines can experience neutral affect (neutral emotional state)'.	93.59	"What is neutral affect?"
		"This is arguably a value we place, as neutral might be good-OK for some and not so good-OK for others. Perhaps this is an area that is subjective beyond what we might be able to measure and so it is a value judgement we make for the horse."
		"This is very hard to objectively assess. From an evolutionary point of view, events/experiences should be rewarding or punishing to shape

		behaviour towards survival and reproduction, so where does neutral fit in?"
		"Panksepp clearly states there is no "neutral" affective state."
		"I think they could be experiencing lower arousal but still a positive or negative valence rather than a neither positive nor negative state but it would be hard to tell. Panksepp model wouldn't have a neutral."
S10 - It was agreed that 'Equines can experience positive affect (positive emotional state)'.	100	None
S11 - It was agreed that 'an individual equine's personality influences the intensities of different emotional states within the individual'.	92.31	"Experiences can be relative to expectations, so optimists can experience frequent disappointment."
		"It is likely that personality would influence intensity of various emotional states, but little work in this area. Optimistic individuals will not be constantly "happy" as their expectations are continually thwarted. The intensity of emotional states may be largely relative to expectations so optimists may have more positive prior expectations but more disappointment."
		"Not sure if this is the case, have no proof one way or the other."

		“While this might be true, I'm not aware of any studies that have shown it to be this black-and-white. The mechanisms of personality and what it affects are not really known. This seems to be a causative statement - and couldn't it really be an interaction, that goes in the opposite direction, too?”
		“I find the concept of a horse being pessimistic unlikely.”
		“Yes but also influenced by past experiences so is it exclusively shaped by personality or is personality a transient trait?”
S13 - It was agreed that 'a multi-measure approach which assesses both behavioural and physiological indices is optimal for measuring positive affective state'.	98.72	“Yes, but the measures need to be shown to be valid indicators.”

The ranking results for three statements did not achieve consensus in this round: Statement S12, ranking of methods to induce positive affect in equines, (agreement level 43.59%), statement S14 ranking of methods for behavioural measures of positive affect (agreement level 61.54%) and statement S15 physiological measures of positive affect (agreement level 50%).

Statement 16 (S16) asked participants to state their level of agreement with two definitions; part one *“Temperament is the animal’s behavioural response in a single context measured on some form of scale, for example flight distance”* did not achieve consensus agreement in round two (agreement level 48.6%). S16 part two *“Behavioural syndromes describe variation within populations, not individuals. They describe the distribution of personalities within a population, for example horses which crib bite”* did achieve consensus agreement (agreement level 81.9%).

In this round, participants did agree that surgency or power of a feeling was an important factor in assessment of affect (S20) with 63 (95.45%) of respondents agreeing.

R2 Rating of new participant suggested methods to induce positive affect in equines

When participants were asked to rank the suitability of a new list of methods (suggested by the participant group in round one) to induce positive affect in equines there was no consensus level agreement however ranking results are displayed below (Table 3.9).

Table 3.9 Ranking of participant suggested methods to induce positive affect in equines. Overall rank = rank in this round with 1 being high and 8 being low. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD, mode and median values identified.

Method to induce positive affect	Overall rank	Mean	SD	Mode	Median
Access to a substrate to roll in (e.g. sand)	1	3.11	1.57	3	3
Food enrichment toys	2	3.14	2.17	1	3
Human-horse interaction through training (e.g. clicker training)	3	3.7	2.22	1, 2	4
Grooming	4	3.9	2.28	3	3
Equine massage	5	4.62	2.01	3, 4	4
Problem solving activities	6	5.35	1.94	6	6
Play toys (e.g. horseball)	7	6.19	2.05	8	7
Introduction of a non-frightening novel object	8	7.03	2.12	9	8
Aromatherapy	9	7.95	1.33	9	8

R2 Rating of new participant suggested behavioural methods to measure positive affect in equines.

When asked to rank a new list of behavioural methods (provided by participants in round one) to measure positive affect in equines again no consensus was reached but participant ranking is shown in Table 3.10.

Table 3.10 Ranking of participant suggested behavioural methods to measure positive affect in equines. Overall rank = rank in this round with 1 being high and 8 being low. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD, mode and median values identified.

Behavioural Method to measure positive affect	Overall rank	Mean	SD	Mode	Median
Scoring of social behaviours/ interactions	1	2.71	1.98	3	3
Body language scoring	2	3.5 0	2.30	1	3
Scoring of play behaviours	3	4.33	3.02	3	3
Absence of frustration	4	5.61	3.79	1	4
Spontaneous blink rate analysis	5	6.77	2.65	6, 9	7
Eye gaze analysis	6	6.8 0	2.76	6	7
Low facial grimace score	7	7.45	3.21	7, 11	7
Time spent in lateral recumbency	8	7.62	2.31	9	8
Assessment of vocalisations	9	7.98	2.67	8	8
Time spent in REM sleep	10	8.08	2.70	10	8
Biomechanical analysis	11	8.50	3.47	12	9
E-BARQ	12	8.64	3.01	12	9

R2 Rating of new participant selected physiological methods to measure positive affect in equines

When asked to rank a new set of physiological methods (suggested by participants in round one) to measure positive affect in equines consensus was not reached however the rankings are shown in Table 3.11.

Table 3.11 Ranking of participant suggested physiological methods to measure positive affect in equines. Overall rank = rank in this round with 1 being high and 8 being low. Mean rank = lower value indicates a higher ranking – Rank 1 equals most preferred and rank 8 equals least preferred method, with SD, mode and median values identified.

Physiological Method to measure positive affect	Overall rank	Mean	SD	Mode	Median
Blood oxytocin	1	3.41	1.96	1, 4	3
Faecal cortisol	2	3.55	2.74	1	3
Respiratory rate	3	4.05	2.50	1	3
Opioids and metabolites	4	4.73	2.23	3	4
Eye temperature	5	5.88	2.67	8	6
Skin conductivity/ Sweat rate	6	5.95	2.71	8	6
Tryptophan	7	6.33	2.01	6	5
MRI scans	8	6.62	3.33	10	7
Gut microflora assessment	9	7.73	3.00	10	9
Computed tomography	10	8.56	2.27	9	9
Tympanic membrane temperature	11	9.20	2.30	11	10

3.3.5 Round three

Seventy-three responses were collected in round three (88% response rate). Again, not all participants answered each question with response rates to questions ranging from n = 64 (88%) to n = 73 (100%).

R3 Agreement with methods that could induce positive affect in equines.

Participants agreed that all methods presented from rounds one and two (except for Aromatherapy) could be used to induce positive affect in equines. Ranking of induction methods are detailed in Table 3.12.

Table 3.12 Agreement on methods that could induce positive affect in equines.

Method to induce positive affect	Rank	Agree % (n)	Disagree % (n)	Don't know % (n)
Addition of a known, affiliative conspecific (e.g. Field mate, stable mate)	1	100 (64)	0.00 (0)	0.00 (0)
Access to a substrate to roll in at liberty (e.g. Sand)	2	98.44 (63)	1.56 (1)	0.00 (0)
High value food provision fed by human (e.g. Sweet mix, novel food items)	3	96.88 (62)	3.13 (2)	0.00 (0)
Medium value food provision fed by human (e.g. Succulents, horse nuts)	4	95.31 (61)	3.13 (2)	1.56 (1)
Provision of loose roughage not hand fed by human (hay or similar)	4	95.31 (61)	3.13 (2)	1.56 (1)
Wither scratching (human scratching of the wither area with fingers)	6	93.75 (60)	1.56 (1)	4.69 (3)
Provision of turnout (loose in field)	7	93.75 (60)	3.13 (2)	3.13 (2)
Human-horse interaction through positive training methods correctly	8	90.63 (58)	1.56 (1)	7.81 (5)

undertaken (e.g. Clicker training)				
Equine massage (undertaken by a trained professional)	9	89.06 (57)	1.56 (1)	9.38 (6)
Access to graze and browse (in-hand)	10	89.06 (57)	3.13 (2)	7.81 (5)
Scratching of the base or sides of tail (delivered by human)	11	87.50 (56)	3.13 (2)	9.38 (6)
Neck scratching (scratching of the neck area, not withers, by human)	12	87.50 (56)	4.69 (3)	7.81 (5)
Food enrichment toys (e.g. Hanging fruit and vegetables, puzzle feeder)	13	85.94 (55)	6.25 (4)	7.81 (5)
Grooming (standard grooming with appropriate brushes)	14	84.38 (54)	4.69 (3)	10.94 (7)
Low value food provision fed by human (e.g. Long stalk forage, high fibre cubes, chaff)	15	82.81 (53)	6.25 (4)	10.94 (7)
Problem solving activities (e.g. Free work, scent work)	16	79.69 (51)	3.13 (2)	17.19 (11)
Introduction of a non-frightening novel object (placed in the	17	71.88 (46)	7.81 (5)	20.31 (13)

environment for the horse to investigate at liberty)				
Verbal praise (e.g. Good, yes)	18	71.88 (46)	14.06 (9)	14.06 (9)
Play toys (e.g. Horseball)	19	70.31 (45)	7.81 (5)	21.88 (14)
Aromatherapy (such as lavender, provided in the environment)	20	40.63 (26)	15.63 (10)	43.75 (28)

Qualitative feedback was provided by 23 participants and data were analysed thematically as shown in Table 3.13 below.

Table 3.13 Thematic analysis of round three methods to induce positive affect in equines.

Themes & subthemes	Responses (count)	Selected quotes from respondents
Verbal praise <i>Pairing with primary reinforcer</i>	7	"I'm not convinced verbal praise adds much to affect, other than through its possible association with more salient primary reinforcers"
Aromatherapy <i>Evidence</i> <i>Type</i> <i>Delivery</i>	6	"Do not have enough information about evidence for aromatherapy in equines to comment." "Providing odours in the environment can be a form of enrichment... [no scientific evidence is available for this]" "Aromatherapy may add to positive affect if an appropriate oil is provided by trained practitioner, rather than just a random oil provided"

		"Aromatherapy... these may be overpowering rather than pleasant".
Toys and novel objects <i>Sex</i> <i>Age</i> <i>Sexual frustration</i>	3	"Regarding play toys: some horses like it but mostly horses with lack of contact with conspecifics, or young horses (5 years and less, males) and I do not know females interested in toys, so it is gendered; Regarding novel object, I doubt it triggers pleasure..." "Play toys horse ball - you see popular posts on [social media] of a horse playing with a horse ball but I'm not sure if it's not some sort of sexual frustration".
Grooming <i>Methods</i>	2	"Traditional grooming practice could be pleasurable if the groomer is responsive to the horse in terms of adapting practice to behavioural signs of pleasure or displeasure" "Grooming may be pleasurable but probably a lot less so than a proper scratch"
Scratches <i>Human-animal interactions</i>	2	"Scratching per se is not necessarily pleasant for a horse and depends on the link and experience with humans but is linked to the season (some horses will work to obtain a scratch at the withers when they lose their hair but will not specially appreciate it during winter for example)"
Food <i>Low value necessity</i> <i>Food enrichment toys</i>	2	"Some horses can only eat this [low value food] (e.g old horses, being fed softened fibre nuts due to lack of food) so some of this is necessity (sadly)".

		"Food enrichment toys - can get horses very stressed"
Massage <i>Pain</i>	1	"Some can find massage sore or have history of pain"
Hand grazing	1	"I feel anything delivered by a human builds [positive relationships], but grazing/ hand grazing doesn't seem to gain the same response e.g. most horses vocalise when seeing a human with a hay net/ feed bucket but wouldn't know that it is being led to pasture so not direct enough cause [and] effect".

Some overarching comments included:

"Where I have answered no or don't know this is because I feel whether these scenarios would generate positive affective states would depend on context and the individual horse - and we don't have sufficient evidence to know for certain in these aspects what the response is"

And

"These all look good and as you say you have considered the variables that will affect their valence. Also possible that some of these could induce frustration/over excitement too"

And

"Each is dependent upon the individual animal and the efficacy of handler".

R3 Inclusion and exclusion of behavioural methods to measure positive affect in equines

Participants agreed consensus on thirteen behavioural methods, developed across rounds one and two, to measure positive affect in equines. Results can be viewed in Table 3.14.

Methods one to thirteen all achieved consensus agreement above the 70% consensus threshold.

Table 3.14 Agreement on behavioural methods that could be used to measure positive affect in equines.

Behavioural Method to measure positive affect	Rank	Include % (n)	Exclude % (n)	Don't know % (n)
Body language scoring (ethograms) including muscle tone	1	93.85 (61)	1.54 (1)	4.62 (3)
Scoring of social behaviour/interactions (scoring both affiliative and agonistic interactions with conspecifics)	2	90.77 (59)	3.08 (2)	6.15 (4)
Facial action scoring (scoring of facial action units such as orbital tightening, ear position, cheek muscle tension etc on an ordinal scale)	3	89.23 (58)	3.08 (2)	7.69 (5)
Human-animal relationship tests (such as avoidance distance and voluntary approach tests)	3	89.23 (58)	3.08 (2)	7.69 (5)
Qualitative Behaviour Assessment (via one of two methods: observers are free to choose qualitative descriptors for an animal or a fixed terms approach using	5	83.08 (54)	7.69 (5)	9.32 (6)

predefined descriptors for observers to select between)				
Scoring of play behaviours	6	81.54 (53)	7.69 (5)	10.77 (7)
Time budget analysis (quantifying the behaviours exhibited by an animal during a specified period of time)	7	80.00 (52)	6.15 (4)	13.85 (9)
Cognitive bias testing [presenting the animal with cues that predict positive events and those that predict negative events as well as presenting ambiguous cues; then testing the animal's response to the ambiguous cues as to whether it is positive (optimistic) or negative (pessimistic)]	8	78.46 (51)	10.77 (7)	10.77 (7)
Low facial grimace score (using published Equine Facial Grimace Scoring technique to identify absence of pain)	9	78.46 (51)	12.31 (8)	9.23 (6)
Assessment of vocalisations (identifying the frequency (Hz), patterns and time spend vocalising)	10	76.92 (50)	7.69 (5)	15.38 (10)
Attention bias testing (measuring the animal's	11	73.85 (48)	9.23 (6)	16.92 (11)

attention between positive and negative stimuli)				
Novel object tests (measuring the animal's response to the presentation of a novel object)	12	70.77 (46)	15.38 (10)	13.85 (9)
Absence of frustration (identifying whether there are behaviours exhibited that could indicate frustration such as door kicking, pawing at the ground, negative attention to barriers)	13	70.77 (46)	20.00 (13)	9.23 (6)
Absence of repetitive behaviour (stereotypy) performance (including crib biting, wind sucking, weaving, box walking, fence walking, headshaking)	14	67.69 (44)	26.15 (17)	6.15 (4)
Time spent in REM sleep (identified through rapid eye moved under closed eyes, loss of reflexes and loose muscle tone)	15	61.54 (40)	13.85 (9)	24.62 (16)
Spontaneous blink rate analysis (scoring of blink rate patterns and palpebral aperture features)	16	56.92 (37)	6.15 (4)	36.92 (24)
Time spent in lateral recumbency	17	56.92 (37)	21.54 (14)	21.54 (14)

Biomechanical analysis (e.g. of step rate & frequency, use of accelerometers and/or head and neck position)	17	56.92 (37)	21.54 (14)	21.54 (14)
EBARQ (Equine Behaviour Assessment and Research Questionnaire)	19	53.85 (35)	18.46 (12)	27.69 (18)
Eye gaze analysis (also termed eye lateralisation, scoring which eye is used to focus on stimuli)	20	52.31 (34)	13.85 (9)	33.85 (22)

Twenty-five participants provided qualitative feedback on this question and a summary of the thematic analysis of these is shown in Table 3.15.

Table 3.15 Round three thematic analysis of feedback relating to potential behavioural methods to measure positive affect in equines.

Themes & <i>subthemes</i>	Responses (count)	Selected quotes from respondents
Sleep <i>Quality</i> <i>Complex factors</i>	4	“Sleep and in particular REM sleep is important, but it is quality of sleep rather than overall duration - sleep disturbance etc. and we are only just starting understand how this might vary in the horse - so yes, REM sleep is important but not just duration”. “I think the sleep quality research will take away focus, due to complex factors affecting it”.

Facial scoring <i>FACS</i> <i>Grimace score (HGS)</i>	2	“Facial action coding (FACS) I would say no, that is too much detail and not yet interpretable”. “I would not use the HGS - (this is an old one and made for young stallions) there are at least 2 more facial pain scales and several CPS's [composite pain scales], better to mention ‘Pain assessment scales’”.
Repetitive behaviours <i>Changes over time</i>	2	“...it should rather be a frequency and its variation through time (increase will be "bad" and decrease will be a good sign regarding emotional state)”. “Some stereotypies are protective -like cribbing so would not think that horse had a negative emotion if he were cribbing”.
Qualitative behaviour assessment <i>Free choice measures</i>	2	“...prefer pre-defined characteristics...so everyone on the same page”.
Human-animal relationship tests <i>Standardise</i>	1	“Have some reservations about human-animal relationship tests, could be useful but wonder about standardising how they are applied”.
Novel object tests <i>Standardise</i>	1	“I'm not convinced verbal praise adds much to affect, other than through its possible association with more salient primary reinforcers”
Eye assessment <i>Spontaneous blink rate</i> <i>Eye gaze analysis</i>	1	“I think we are lacking knowledge about their meaning”

E-BARQ <i>Limited evidence/knowledge</i>	1	“this is a relatively unvalidated metric”
Play behaviours <i>Public misunderstanding</i>	1	“Re social behaviours - I do think this could be valuable but I would feel worried about the public over-worrying. I already see a lot of horses kept in isolation because owners interpret play behaviours as aggression and worry about injury. I don’t want this issue getting any worse by people separating more horses on the grounds that they sometimes show some negative interactions”.

Two participants described how the methods would be reliant on the skills and experience of the testers and posited whether human bias could be a confounding factor in subjective tests. Another commented that they disagree with the use of subjective and scale-scored methods and a fourth highlighted their experience of good descriptive tools having higher reliability and validity than behavioural tests or observations.

Seven respondents highlighted that they felt the absence of a negative does not indicate the presence of positive affective state and therefore these measures may be limited, and several comments were provided highlighting the importance of using a variety of different methods within the testing protocol.

R3 Inclusion and exclusion of physiological methods to measure positive affect in equines

Participants achieved consensus on three physiological methods to measure positive affect in equines from the options developed during rounds one and two. Results can be viewed in Table 3.16.

Table 3.16 Agreement on physiological methods that could be used to measure positive affect in equines.

Physiological Method to measure positive affect	Rank	Include % (n)	Exclude % (n)	Don’t know % (n)

Heart rate variability (assessing the balance between the sympathetic and parasympathetic nervous system via a heart rate monitor)	1	92.19 (59)	1.56 (1)	6.25 (4)
Salivary cortisol (testing cortisol levels using saliva collected and run through a calibrated test system such as ELISA)	2	76.56 (49)	7.81 (5)	15.63 (10)
Respiratory rate (via counting the rise and fall of the flanks)	3	75.00 (48)	7.81 (5)	17.19 (11)
Heart rate (via a heart rate monitor)	4	65.63 (42)	17.19 (11)	17.19 (11)
Electroencephalogram (utilising resting or ambulatory EEG technology to assess brain activity)	5	59.38 (38)	4.69 (3)	35.94 (23)
Serotonin levels (testing serotonin levels using blood collected and run through a calibrated test system such as ELISA)	6	57.81 (37)	10.94 (7)	31.25 (20)
Gut microflora assessment (via faecal samples)	7	54.69 (35)	10.94 (7)	34.38 (22)
Urinary cortisol (testing cortisol levels using urine samples collected and run through a calibrated test system such as ELISA)	8	53.13 (34)	15.63 (10)	31.25 (20)
Faecal cortisol (testing cortisol levels using faecal samples collected and run through a calibrated test system such as ELISA)	9	53.13 (34)	17.19 (11)	29.69 (19)
Blood oxytocin (via blood samples)	10	50.00 (32)	4.69 (3)	45.31 (29)

Blood cortisol (testing cortisol levels using blood samples collected and run through a calibrated test system such as ELISA)	11	50.00 (32)	29.69 (19)	20.31 (13)
Skin conductivity/sweat rate (measured using electrodes and sweat rate sensors on the skin)	12	48.44 (31)	7.81 (5)	43.75 (28)
Hair cortisol (testing long term cortisol levels using hair collected and run through a calibrated test system such as ELISA)	13	48.44 (31)	15.63 (10)	35.94 (23)
Infra-red thermography (e.g. of the eye or ear using a thermographic camera)	14	46.88 (30)	15.63 (10)	37.50 (24)
Catecholamine levels (via blood tests for adrenalin and noradrenalin levels)	15	46.88 (30)	20.31 (13)	32.81 (21)
Eye temperature (via thermography)	16	42.19 (27)	9.38 (6)	48.44 (31)
Corticotrophin releasing factor challenge (assessing physiological response, via ACTH and cortisol levels, to intravenous infusion with corticotrophin-releasing hormone)	17	42.19 (27)	17.19 (11)	40.63 (26)
Body temperature (rectal temperature)	17	42.19 (27)	21.88 (14)	35.94 (23)
Immune function tests (testing for immunoglobulin deficiencies/disorders)	19	40.63 (26)	14.06 (9)	45.31 (29)
Opioids and their metabolites (via testing of plasma samples by ELISA)	20	39.06 (25)	7.81 (5)	53.13 (34)

MRI scans (Magnetic Resonance Imaging)	21	34.38 (22)	15.63 (10)	50.00 (32)
Tryptophan levels (measured in plasma via mass spectrometry)	22	32.81 (21)	12.50 (8)	54.69 (35)
Tympanic membrane temperature (measurement of the temperature of the tympanic membrane within the ear using a specialised device)	23	28.13 (18)	10.94 (7)	60.94 (39)
Computed tomography - CT (a series of x-ray images taken at different angles and processed to create cross sectional images)	24	23.44 (15)	25.00 (16)	51.56 (33)

Twenty-nine respondents provided qualitative feedback which was thematically analysed and can be viewed in Table 3.17.

Table 3.17 Round three thematic analysis of feedback relating to potential physiological methods to measure positive affect in equines.

Themes & subthemes	Responses (count)	Selected quotes from respondents
Invasive measures <i>Avoid on welfare grounds</i> <i>Induce negative affect</i>	11	“How do we measure happiness? Would taking blood samples affect that? Intrusive sampling is stressful, and I don’t know if it is worth it”. “I would avoid any invasive measures - both in terms of equine welfare and accessibility”
Cortisol <i>Faecal</i> <i>Hair</i> <i>Time lag</i> <i>Blunt measure</i>	3	“Faecal and hair cortisol provide reasonable longer-term measures but interpretation of cortisol level / response not straightforward - we don't have conclusive measures at the moment”

		“Some of the cortisol tests will show time lags so in certain conditions are less useful” “Testing hair is the most reliable”
Electroencephalogram <i>Feasibility</i>	1	“EEG in horses currently impractical”.
Catecholamines <i>Natural fluctuations</i>	1	“Catecholamines are not so useful unless over standardised periods as these fluctuate normally”
Thermography <i>Feasibility</i> <i>Specificity</i>	1	“Thermography also practically challenging, and my understanding is that peripheral areas cool when blood moves to the core, but only under acute stress. Not sure that there is a physiological reason for +ve [positive] experiences changing blood flow?”.
Serotonin <i>Reliability</i> <i>Specificity</i>	1	“Blood levels of serotonin do not reflect brain levels”.
Heart rate variability <i>Reliability</i>	1	“HRV is very complex and can only be undertaken with an ECG (no heart rate monitors) but may be a possibility”.
MRI <i>Functional MRI</i>	1	“Normal MRI won't tell anything, functional MRI almost certainly would but no machine in the world that can do this with horses”.
CT <i>Radiation</i>	1	“My only reservation re a CT scan is I understand it gives a strong radiation dose? So should not be taken lightly?”.

Twelve participants commented that they felt that there was not enough evidence surrounding the use of many of the methods to measure positive affect in equines and five more commented that it is difficult to use many of the methods to differentiate between positive and negative valence. Two participants commented that the methods would need to be used in conjunction with other physiological and behavioural methods.

3.3.6 Overall result synopsis

Inducing positive affect in equines

When ranks from rounds one, two and three were analysed together the overall ranking from the three rounds identified that the highest ranked methods were high value food and known affiliative conspecific provision and the lowest ranked methods were aromatherapy and forehead rubbing (Figure 3.8).

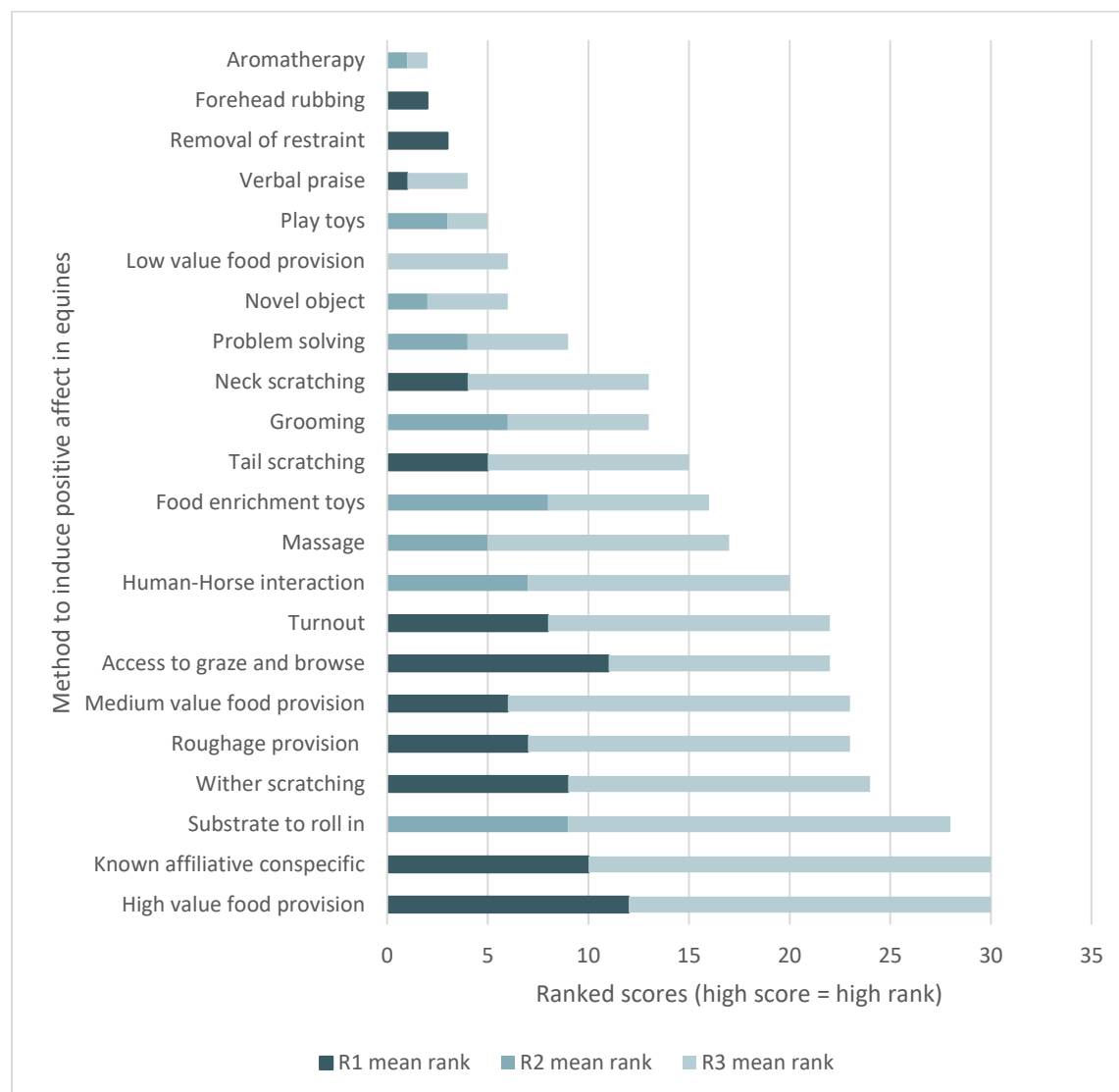


Figure 3.8 Aggregated ranking of the induction methods for positive affect in equines across each round. Rank scores were reversed to allow for clear visualisation of the most popular methods suggested by participants to induce positive affect in equines.

Behavioural measures of positive affect in equines

When ranks from rounds one, two and three were analysed together the ranked methods from each round were extracted for comparison (Figure 3.9). The highest ranked methods were conspecific interactions, body language assessment and facial action scoring. The lowest rank methods were E-BARQ and biomechanical analysis.

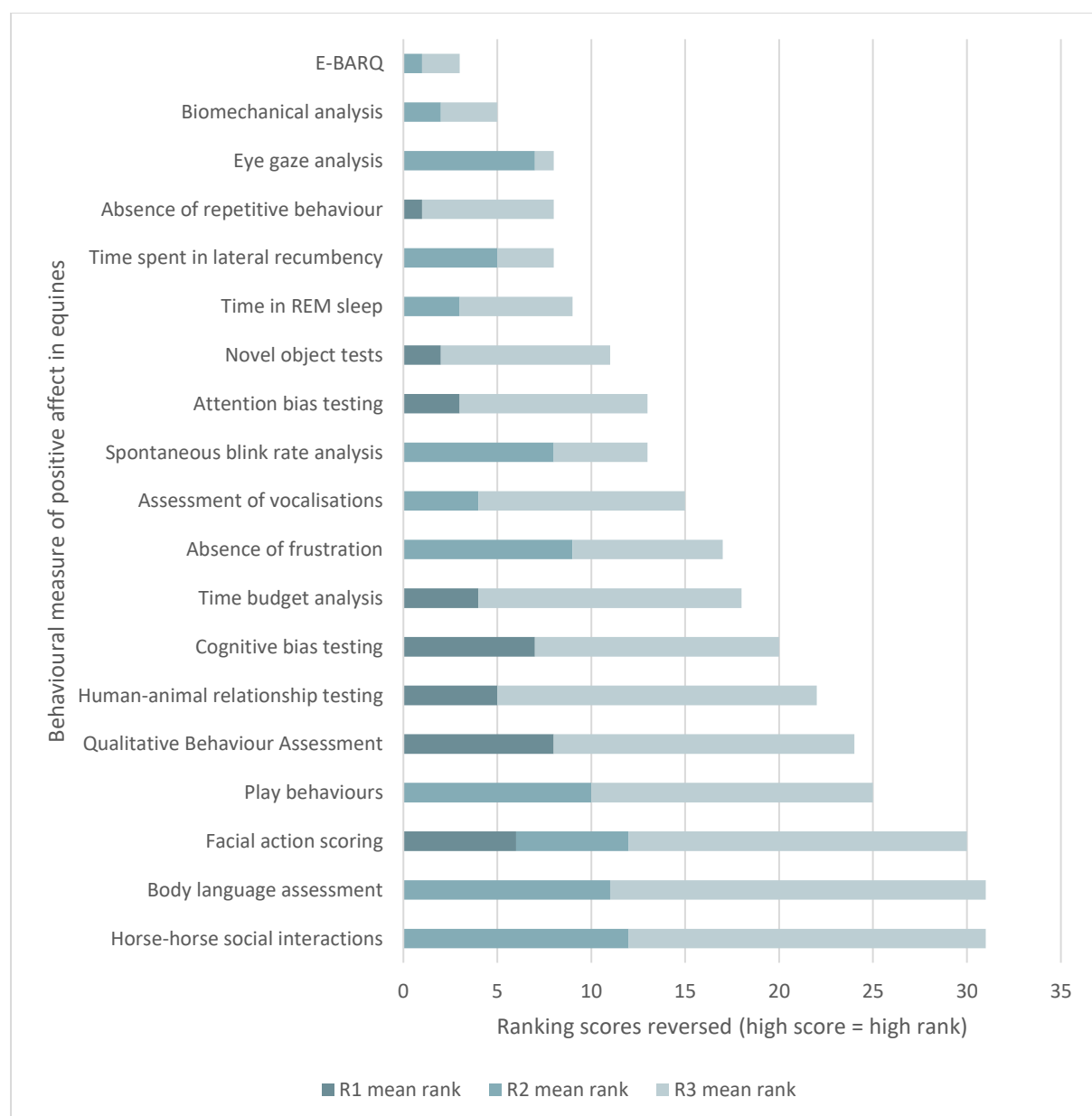


Figure 3.9 Aggregated ranking of the behavioural methods for positive affect in equines across each round. Rank scores were reversed to allow for clear visualisation of the most popular methods suggested by participants to behaviourally measure positive affect in equines.

Physiological measures of positive affect in equines

When ranks from rounds one, two and three were analysed together the ranked methods from each round were extracted for comparison (Figure 3.10). The highest ranked methods were heart rate variability, salivary cortisol and heart rate assessment. The lowest ranked methods were computed tomography and MRI scans.

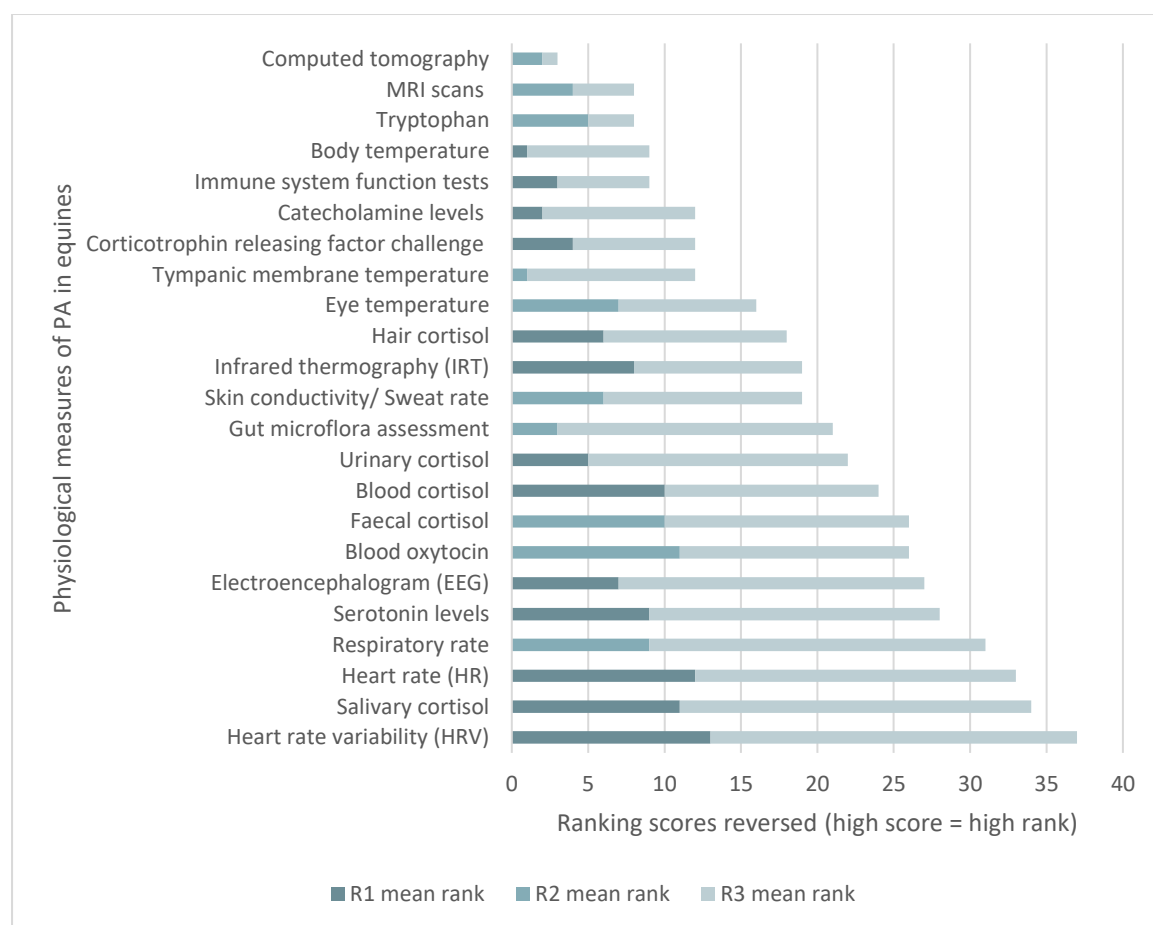


Figure 3.10 Aggregated ranking of the physiological methods to measure positive affect in equines across each round. Rank scores were reversed to allow for clear visualisation of the most popular methods suggested by participants to physiologically measure positive affect in equines.

3.4 Discussion

This Delphi consultation engaged ninety-three equine experts in discussion around methods to induce and measure positive affective state. Achieving consensus amongst a diverse group of stakeholders from across the equine industry is difficult to achieve. This has been identified in other fields such as medicine where Delphi consultations are commonly utilised (Thangaratinam and Redman, 2005). However, this Delphi consultation reached consensus agreement at a minimum of 70% threshold on thirteen statements and three ranking exercises during the three rounds and whilst there was evidence of variations in response distribution, no bi-modal patterns were identified within the results indicating that the data appear reliable in terms of achieving the aim which was to seek consensus through obtaining normative views from a range of experts (Powell, 2003; Mullen, 2003).

3.4.1 Demographic factors

The response rate (55%) to invitation to participate in this Delphi consultation was comparable with that of other studies in the animal welfare sphere (Pearson *et al.*, 2021) and equalled response rates seen in the field of human healthcare where these practices are more commonly undertaken (Baruch and Holtom, 2008). The overall retention rate (71% retained from initial acceptance, $n = 103$, to final completion, $n = 73$) was high especially considering the 30-month duration and compared with other recent equine-related Delphi's with lower retention rates of 56% (DuBois *et al.*, 2018) and 42% (Williams *et al.*, 2023; Pearson *et al.*, 2021). A retention rate of $\geq 70\%$ is considered rigorous and robust (Hasson *et al.*, 2000).

A wide spectrum of the equine industry was represented by the range of experts contributing to the study; however, the representative spread of the number of participants from each area of the sector was imbalanced and skewed towards vets, academics and clinical animal behaviourists. Whilst this factor limited contributions from trainers, those in the competition area, the charity sector and rehabilitation, many participants had expertise in more than one area; for example, many participants were qualified vets who also worked in the charity or competition sectors or behaviourists with training qualifications and experience. Many academics also held other industry based practical or professional qualifications which strengthened the level of both academic and practical knowledge within the survey responses, a benefit described in other equine Delphi's (Collins *et al.*, 2009).

Within this study males were underrepresented at 11.4% of the study population which may have had some effect on the results obtained, however, male equestrian numbers are generally lower than females internationally as reports from the US (22.6% male) and UK (10-33% male) indicate (British Equestrian, 2022) with a decline in male equestrians seen in the US from 41.44% in 2010 to 24.61% in 2019 (Zippia, 2022). However, DEFRA (2004) indicates that the gender profile of those interested in, but not necessarily participating in, equine-related activities in the UK is more equally divided (49% male, 51% female) however the percentage of these who are also equine behavioural experts or scientists is likely to be much lower.

Participants contributing to this study were recruited from an international cohort covering the United Kingdom (62%), North America (12%), Australia (3%), and Europe (23%). Whilst the sample was predominantly European it included thirteen countries, however, more than 50% of the sample was comprised of participants from the UK with no representation of those from Africa, Asia, South America or Russia which may have led to cultural biases in the results obtained (Baumgartner and Weijters, 2015). In this study 45.28% of participants held a level 7 and/or level 8 qualification which is a much larger proportion than is found naturally within equestrianism however this reflects the requirements of the inclusion criteria for participation as recommended by Junger *et al.* (2017) to facilitate expert derived consensus and is therefore a strength of this study.

3.4.2 Expert consensus on emotions in equines

The high level of consensus (97%) regarding the importance of emotional welfare (the ability to produce positive emotions and adapt when confronted with stressful situations) within more holistic equine welfare seen within this study lends strong support to the importance of determining methods to induce and measure positive affective state in equines. This facet is strongly supported by the literature where researchers have highlighted the significance of the affective experience in a wide range of situations; from affective state during exercise (Mellor and Beausoleil, 2017; Dyson *et al.*, 2021), husbandry and veterinary care (McGreevy *et al.*, 2018^a; Pearson *et al.*, 2021), the effect of the equine environment (Horseman *et al.*, 2016), caregiver interactions (Torcivia and McDonnell, 2020), equine assisted therapy interventions (De Santis *et al.*, 2017) and pain (Torcivia and McDonnell, 2021; Lencioni *et al.*, 2021; Anderson *et al.*, 2021). Within this consultation only two participants felt that emotional

welfare was not an important issue in equine welfare, both respondents were veterinary surgeons and ranked this metric low (20% and 38% importance) which may indicate an unconscious bias associated with their work-related interactions with equines being focussed on physiological function and dysfunction rather than psychological aspects.

Valence, arousal and surgency

High levels of consensus agreement surrounding the importance of assessing both valence (89.33%) and arousal (81.33%) in relation to affective state were identified within this consultation. Mendl *et al.* (2010) represents core affect within a two-dimensional framework, adapted from the work of Russell and Barrett (1999) and Burgdorf and Panksepp (2006), which is well known to researchers in the field and may lend explanation to the clear consensus on valence and arousal within this consultation. This finding is supportive of considering an experimental methodology that measures variables that may be indicative of both valence and arousal when attempting to measure positive affective state.

In round one participants did not agree (56.76% agreement) that surgency (the intensity of emotional state experienced) was always important in affective state, however in round two the question was rephrased to ask participants if they felt it was important, unimportant or neither important nor unimportant and participants agreed (95.45%) that it was an important factor. The alteration of the phrasing of the question may have led to this consensus as the removal of the 'always' association provided more scope for consideration of different contexts where surgency may or may not be considered. In addition, the concept of surgency is less well reflected in the literature (returning only 1390 search results across both the human and non-human literature) with arousal and valence being the predominant considerations in affective state assessment to date (returning 203,000 search results). Surgency or power is stated to be more difficult to accurately measure in an animal behavioural context (Martin and Bateson, 2007) and is predominantly cited within the human literature (15,800 search results) as it is concerned with the intensity of the feeling which, unlike in humans, cannot be verbally communicated by animals (Panksepp, 2011). However, surgency could be inferred in animals via assessment of certain components of the behavioural act; for example via calculating the *local rate*, where the frequency of a behaviour is divided by the unit of time to give a rate/time value (Cooper *et al.*, 2020) albeit evidence

within the literature for reliable methods to measure surgency in non-human animals are sparse (Burgdorf and Panksepp, 2006).

When considering the optimal utilisation of the results from the Delphi consultation in later experimental design it is important to consider an appropriate framework to base further studies within. Mendl *et al.* (2010) brings together the evidence relating to discrete *Pankseppian* (separate emotions such as fear, happiness, excitement, anger) and dimensional *Barrettonian* (placement within the arousal/valence core affective space) approaches to conceptualising emotions in animals. They offer up a framework which considers both approaches, which they suggest can:

- (i) offer a structure for integrating different discrete emotions that provides a functional perspective on the adaptive value of emotional states,
- (ii) suggest how long-term mood states arise from short-term discrete emotions, how they also influence these discrete emotions through a bi-directional relationship and how they may function to guide decision-making, and
- (iii) generate novel hypothesis-driven measures of animal emotion and mood.

Mendl's two-dimensional model focuses on shorter term emotional events, as is the focus of this thesis, and does not account for surgency (also termed power or intensity) nor time (how long the individual is spending in a particular state of valence and arousal). Therefore, in further experimental design within this thesis the Mendl model is taken forward within analysis and interpretation of data. Alternative three (surgency) or even four (time) dimensional approaches (Figure 3.11) may best encapsulate the affective state of an individual in longitudinal studies to measure emotional welfare states as proposed by Cabanac (2002) and introduced in other spheres as early as 1754 when Jean le Rond d'Alembert published 'Dimensions', a concept later formalised by Bernhard Riemann in 1854 (Cajori, 1926; Bell, 1965).

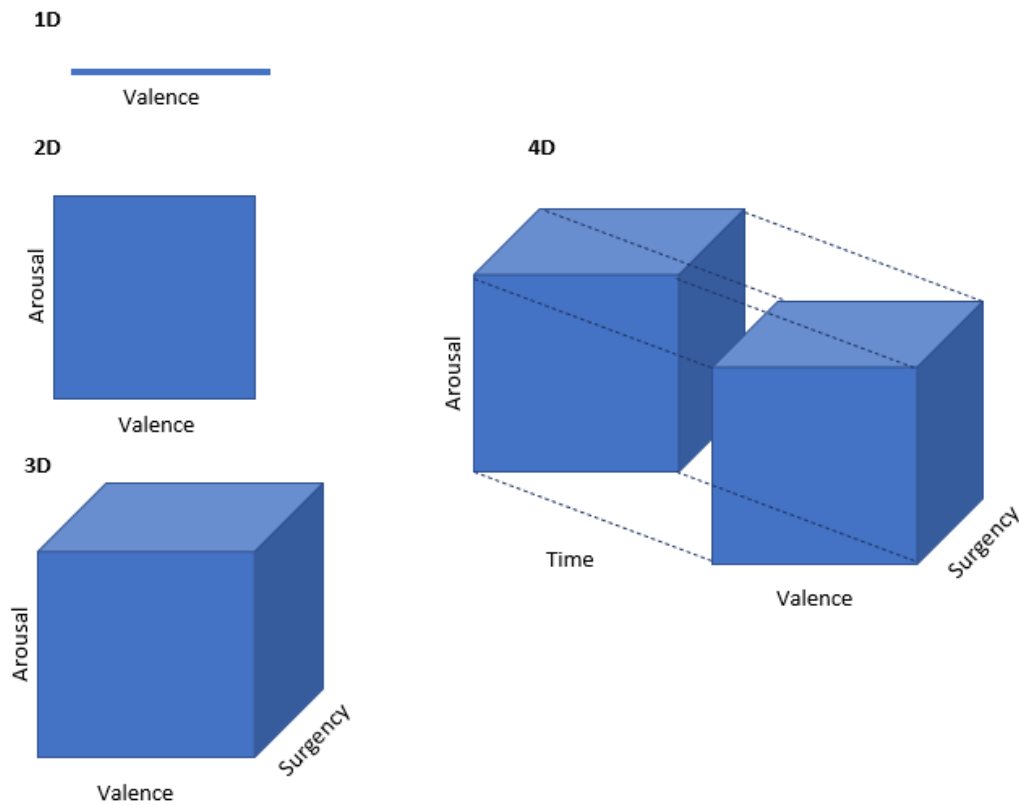


Figure 3.11 Visualising the fourth dimension, a conceptualisation of measures of emotional welfare over time. Authors own diagram adapted from Radiya-Dixit (2017).

Emotional processing

Participants agreed (100%) that emotional processes include motor-expressive, sensory-perceptual, autonomic-hormonal, cognitive-attentional and affective-feeling aspects. They also agreed (100%) that it can be difficult to distinguish between the cognitive (conscious mental activities) and affective (relating to moods and feelings) aspects in field-based research. Indeed, Mendl *et al* (2010) highlight the interplay between facets of emotional processing that give rise to an affective feeling. This interplay can be summarised in Figure 3.12. It is therefore important to evaluate a range of information regarding the individual animal when attempting to induce or measure affective state; this includes the individual's personality (Letzring and Adamcik, 2015; Sih *et al.*, 2015), experiential learning (Boissy *et al.*, 2007a; Christensen *et al.*, 2012; Cooper, 1998) and preferences (Lansade *et al.*, 2022; Platzer and Feuerbacher, 2022; Yarnell *et al.*, 2015) as will be undertaken in this thesis.

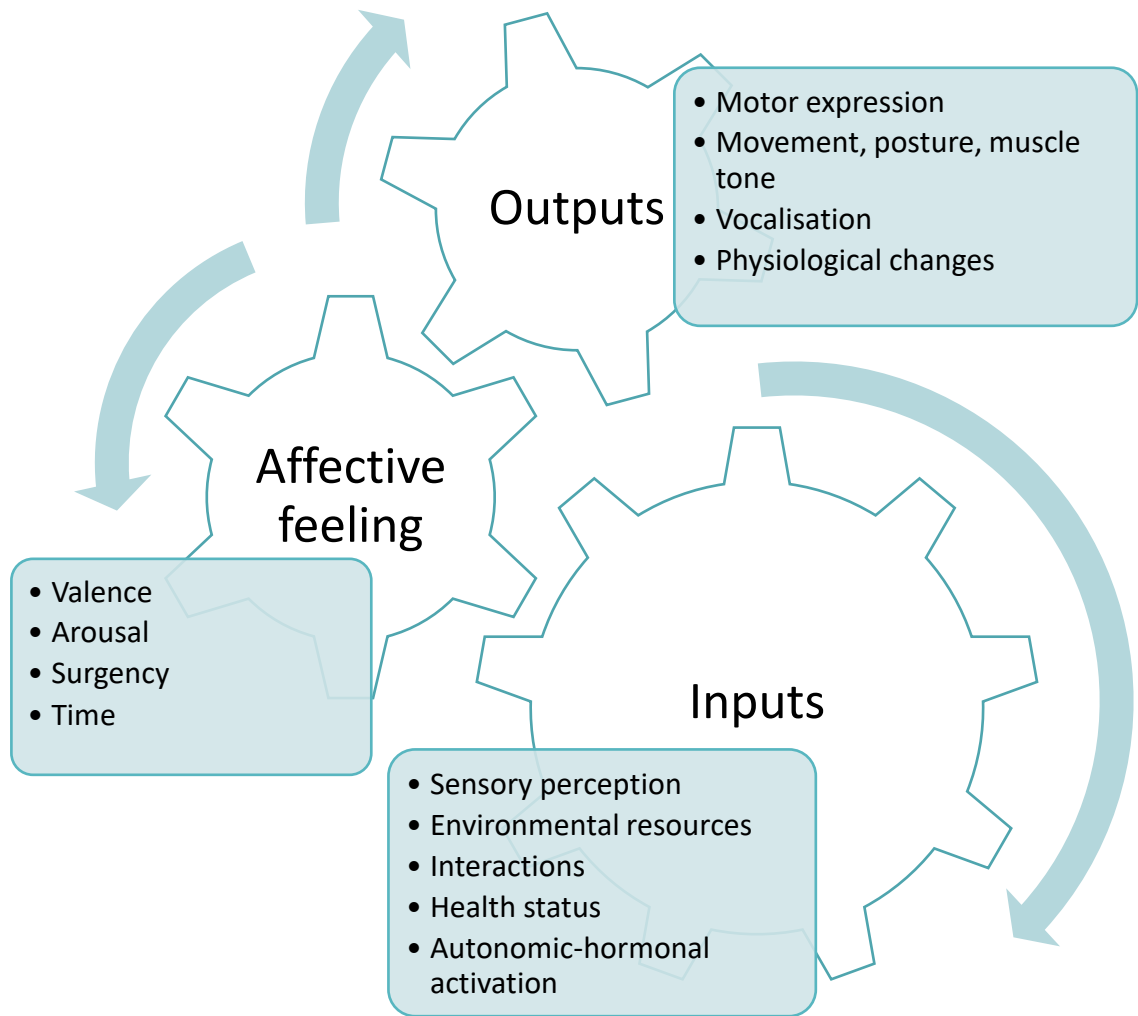


Figure 3.12 Representation of the interplay of emotional processing components (authors own).

The complex interplay between all components in the sensory, environmental, physiological and affective domains have varying inhibitory and disinhibitory effects on the physiology and behavioural outputs of the animal (Cabanac, 2002) and need to be considered when planning methodologies to measure affective state (Waran and Randle, 2017). Individual perception of the environment, sensory status of stimuli, acute and chronic health status and individual physiological factors should be considered within experimental design in order to maximise the potential to achieve positive affective states (Baldwin and Baldwin, 2001).

Positive, negative, and neutral affect

Whilst participants agreed that equines could experience positive (91.66% agreement), negative (93.06% agreement) and neutral (80.55% agreement) affective states the consensus for neutral affective state was ~10% lower than for positive and negative affective states. This difference may be due to a lack of evidence within the non-human animal literature for the presence of neutral affective states in animals. In their review of human neutral affect Gasper *et al.* (2019) define it as “*feeling indifferent, nothing in particular and a lack of preference one way or another*”. The researchers propose a distinct difference between feeling ‘nothing in particular’ (i.e. neutral affect) and the absence of affect (a non-affective condition). Several research groups have studied neutral affect in humans (Storm and Storm, 1987; Zelenski and Larsen, 2000; Tay, 2011; Gasper and Hackenbracht, 2015; Gasper *et al.*, 2021; Samson *et al.*, 2016; Gallegos and Gasper, 2018) with findings suggesting that humans experience neutral states with similar frequency to their experience of positive affective states. Gasper and Hackenbracht (2015) suggest that, similarly to positive and negative affect, neutral affect occupies working memory capacity and therefore neutral affect feelings would be reduced if working memory capacity was reduced via the person being engaged in a cognitively demanding task. Mendl *et al.* (2010) does not consider neutral affect as an affective state from a non-human animal perspective and this may have influenced the lack of research in this area in equines and other domestic animals. Within analysis of experimental data later in this thesis particular attention will be paid to any evidence of the experience of neutral affective state when analysing measures of valence and arousal.

Positive and negative affect are terms much more widely utilised within the equine sphere and have been studied in relation to improving horse-human interactions (Hausberger *et al.*, 2008; Trösch *et al.*, 2020), indeed Trösch *et al.* (2020) identified that horses felt emotion when watching video projections of positive and negative human-horse interactions demonstrated by the differing behaviours they displayed towards the positive and negative human experimenters when presented with them in a live context. Further work has studied the use of facial expressions and oxytocin levels as potential markers of positive affect in equines with data suggesting differences in facial cues, neck position and oxytocin levels, however in this study blood cortisol, heart rate and heart rate variability did not differ between pre and post intervention measurements (Lansade *et al.*, 2018).

Briefer *et al.* (2017b) assessed emotional valence in horse whinnies finding that horses exhibited higher behavioural and physiological arousal measures when presented with unfamiliar whinnies compared with familiar whinnies and showed lower arousal to positively valenced whinnies. The results of this study suggested correlations between behavioural and physiological variables indicative of arousal and affect lending support to experimental designs which consider both paradigms within their methodologies.

Sankey *et al.* (2010b) investigated the use of reinforcement as a mediator of perception in horses, finding that the use of positive reinforcement (food reward) was associated with a greater level of approach during a motionless person test, lower latency to correctly perform the test (respond to a 'back' cue), increased performance of the task (stepping backwards only rather than sideways), increased instances of ears forwards (versus backwards) and lower instances of head raising and neck tension than those being trained with negative reinforcement. This study supports the notion that a food reinforcer appears to generate positive associations and thus potentially positive affective state, as also agreed by the Delphi experts. Similarly, Löckener *et al.* (2016) identified positive cognitive bias in horses pastured in herds (social interactions) using a spatial discrimination test, which may support the notion of intraspecific interaction as an inducer of positive affect; albeit the sample size of 13 was small and some conditions changed during the trial which lasted for over one year.

3.4.3 Expert consensus on inducing positive affect in equines

Within this consultation 22 induction methods for positive affective state were presented to (or suggested by) participants. Six food-related methods, six human-provided tactile methods, six non-food enrichment activities, two methods involving interaction with conspecifics and two human-interaction methods were selected.

Expert consensus identified the top-ranking induction methods as:

1. High value food provision
2. Addition of a known affiliative conspecific
3. Access to a substrate to roll in
4. Wither scratching
5. Roughage provision
6. Medium value food provision

7. Access to graze and browse
8. Turnout
9. Human-horse interaction through training

Each induction method has unique advantages and disadvantages within an experimental paradigm attempting to induce positive affect. Specific factors relating to certain methods reduced their likely inclusion within the final selection. An example of this is the suggested use of rolling behaviour as an induction method, whilst highly likely, in the right situation, to be positive affect inducing, the context is likely very specific to the individual equine and the method likely to be difficult to induce in a practical context on demand. It is noted throughout this chapter that preferences for positive reinforcers are often variable on an individual level to an extent, however the highly specific situation leading to rolling (such as post-exercise or when shedding the coat or marking territory) makes the measure much less generalisable within study parameters; on reflection this type of consideration could have been specified to participants in advance of them putting forward suggestions, conversely however it may have stunted the generation of novel ideas, all of which are valid suggestions within this type of study.

Recommendations for positive affect induction methods to be taken forward to experimental trial.

Following careful evaluation of the literature and the data generated from participants (both quantitative and qualitative) the methods discussed within the consultation were considered in relation to a) their rank (higher rank equals higher chance of inclusion), b) their feasibility within the context and timeframe of further field trials, c) their reliability between instances of testing, d) their cost and e) the availability of the equipment required to facilitate them.

Therefore, the following induction methods have been identified for inclusion in the field study:

1. Provision of food (ranked 1), preference testing of feed prior to the trial required
2. Addition of a known affiliative conspecific (ranked 2)
3. Wither scratching (ranked 4)
4. A simple training task (ranked 9),

Support for the suggested induction methods for PA in equines

High value food provision

High value food provision was cumulatively the highest ranked induction method for PA, being the highest ranked in R1 (90.28%) and third highest rank in R3 (96.88%). This method involves the provision of a foodstuff that is of high value to the individual animal and usually includes foodstuffs that are high in sugars and highly appetitive such as a sweet mix, horse and pony nuts/cubes, fresh fruit, and vegetables and highly flavoured commercially sold training treats. The selection of this induction method by the experts concurs with the literature which indicates that equines find high value food reinforcing and consequently suggesting that this is likely to be a useful method for inducing positive affective state. Indeed, Kieson *et al.* (2020) identified that horses demonstrated a preference for food treats over human interaction regardless of the horse's history and whether the human was a familiar or unfamiliar person. However, all horses in the study displayed increased arousal during non-delivery of the food item (i.e. when a patting or scratching reward was offered instead of a food reward) displayed through tense musculature, high neck, pawing at the ground and in one case biting the target indicative of frustration related behaviours (Pannewitz and Loftus, 2023).

Whilst the literature suggests that all foodstuffs can be effective in inducing positive affect in equines (Platzer and Feuerbacher, 2022), research has identified that response rates to training double when a grain reinforcer is used over a forage-based reinforcer (Ninomiya *et al.*, 2004). All foodstuffs have relative value dependant on the animal's normal diet (Elia *et al.*, 2010) however it is important to establish the relative value of foodstuffs to the individual and also to identify the potential for frustration related to the provision of different types of foodstuffs to the individual equine, through prior preference testing. For this reason, a selection of foodstuffs should be presented to each equine to assess relative value (which type the animal selects to eat) and to assess for signs of frustration and anticipatory behaviours (Peters *et al.*, 2012) using the expert agreed behavioural signs published by Pannewitz and Loftus (2023).

Addition of an affiliative conspecific

As a social prey species, equines are known to live in herd for safety and reproductive purposes (Cameron *et al.*, 2003; Ringhofer *et al.*, 2017). The addition of a known affiliative conspecific companion was cumulatively ranked as the second highest induction method overall, ranked 3rd (80.55% agreement) in round one and 1st (100% agreement) in round 3 and may function as a method of positive affect induction in equines providing that careful assessment of the relationship between the two animals is undertaken prior to utilisation of the method. Particular attention should be paid to interaction type (such as play, investigatory behaviour, resting in proximity) and frequency as well as instances of mutual grooming initiation and response and signs of aggression in order to assess the stability of the relationship and therefore the likelihood of the conspecific presence inducing positive affect (Heitor *et al.*, 2006). Research suggests that horses spending time within one-body length proximity to each other are likely to have stronger affiliative relationships with each other than those spending time 5 or more body lengths from each other. Higher levels of active affiliative interactions such as approaches and mutual grooming also indicate a stronger affiliative bond, however this strength of bond does not reduce likelihood of aggression towards each other (Heitor *et al.*, 2006). Interestingly, Kimura (1998) identified that horses changed their preference for social partners over time and that the time of year significantly affected preferred associate relationships suggesting that ongoing review of preferred conspecifics should occur in order to maintain an accurate representation of preferred social partners in a group of horses. Therefore, within future studies using this methodology real-time and repeated assessment of conspecific relationships should be utilised close to the proposed study period.

Placing of two equines in adjoining areas with the ability to see and touch each other (and with the absence of food which could present as a potential resource guarding trigger) may lead to positive social interactions, or merely social companionship, which could induce positive affect. It is important that the additional equine should be in good health and be generally relaxed to avoid emotional contagion (a phenomenon of an automatic adoption of another's emotional state) of any negative affect (Spinka, 2012).

Wither scratching

Wither scratching was the fourth highest ranked induction method overall and was ranked 3rd (84.72% agreement) in Round one and 6th in Round three (93.75% agreement). Wither scratching is proposed to act as a form of tactile communication between equines to enhance the relationship via social bonding and comfort (McDonnell, 2003; Waring, 1983), which is in turn used by equestrians to imitate allogrooming (McLean and McGreevy, 2010). Indeed, horses have been observed displaying allogrooming towards heterospecifics (Fraser, 2010; McGreevy, 2012). The literature shows that grooming or massaging around the wither area lowers heart rate and enhances relaxed behavioural postures (Feh and de Mazières, 1993; McBride *et al.*, 2004; VanDierendonck and Spruijt, 2012; Lansade *et al.*, 2018; Lansade *et al.*, 2022) and therefore wither scratching may induce a positive affective state similar to that seen in allogrooming. Many studies have assessed the value of wither scratching to the horse with most focussing on latency to learn a new task as an assessment methodology where it is hypothesised that the greater the value of the reinforcer the lower the latency to perform the new task. Wither scratching has been demonstrated to appear more valuable to equines than scratching of other areas and neck patting (Thorbergson *et al.*, 2016) and has resulted in faster learning of new tasks (Ellis and Greening, 2016) and reduced avoidance or stress responses (Watson and McDonnell, 2018). As with food reinforcement it is suggested that individual preference tests should be conducted to establish preferred scratching sites for the individual (Janczarek *et al.*, 2018; Nishimuta *et al.*, 2022) with a conjugate approach to scratching used, whereby continual scratching is offered to the horse who may choose which body areas are scratched, and at what level of pressure. Overall, the literature signifies that wither scratching may be appetitive to equines providing a) it is undertaken in a style that is pleasant to the individual (not too hard nor soft for example), b) the individual enjoys tactile contact with humans and c) that it is perceived as valuable in the specific context (over other interactions for example). Careful and accurate evaluation and monitoring of equine behavioural responses to the stimulation are required to ascertain the appetitive value of the stimulus to each individual as will be considered during the experimental design of the field study.

Human-horse interaction through positive reinforcement training

Human-horse interactions via positive training was ranked 9th overall and was ranked 3rd (65.07% agreement) in Round two and 8th (90.63% agreement) in Round three. The lower

consensus agreement for this measure may be derived from the consideration that it may only be appetitive for equines who deem human interaction to be pleasurable (Hausberger *et al.*, 2019) and providing that the trainer is highly skilled in training the animal and that the task set is not too difficult to risk inducing frustration and other negative affective states (Pannewitz and Loftus, 2023). Using positive reinforcement training may help to reduce negative affect (Freymond *et al.*, 2014) and achieve the desired behaviour or completion of the task with associated higher levels of motivation (Innes and McBride, 2008). It is noted however that the addition of environmental stressors during training markedly decreases performance during both positive and negative reinforcement training (Valenchon *et al.*, 2017) and therefore human trainers should ensure that they set up the training environment to reduce the presence of environmental stressors in order to maximise positive affect during training.

The literature available highlights the importance of an animal's history with humans, their prior training experiences and the body language of the human in relation to how positively the equine may view human interaction. As such each equine must be assessed as an individual to ascertain levels of valence and arousal in response to human contact and interactions during training. Indeed, d'Ingeo *et al.* (2019) demonstrated associative learning related to human cues and Hausberger *et al.* (2019) highlighted how horses build a multisensory and valenced representation of familiar humans based on prior experiences through training and interaction, meaning that regular exposure to well performed positive reinforcement training could form positive associations with humans and facilitate a positive horse-human relationship leading to positive affect. Baragli *et al.* (2015) highlight the importance of the skill of the trainer in horses' affective responses to training tasks; the ability of the trainer to effectively communicate and appropriately reinforce the horse at the correct time is imperative to reduce confusion and frustration (McCall *et al.*, 1993; Pannewitz and Loftus, 2023) and therefore the use of positive reinforcement target training within experimental studies should be undertaken by an expert to increase the likelihood of a positive affective response to training. Positive reinforcement training has been proposed as a measure to optimise welfare (Hausberger *et al.*, 2019) and has been demonstrated to induce positive affect in dogs and wolves (Vasconcellos *et al.*, 2016). Indeed, proxy behaviours that may indicate positive affect in horses, such as increased exploration, increased affiliative

behaviours, reduced defence behaviours and improved memory, have been demonstrated in horses trained using positive reinforcement training as summarised by Hausberger *et al.* (2019).

Utilising positive reinforcement training requires use of a reinforcer that is salient to the animal (predominantly food) and a secondary reinforcer such as a clicker or other marker sound (such as a verbal 'yes' or 'good') thus it may be difficult to distinguish which aspect of the experience is most salient and lends most input to any induction of positive affect, an aspect that would require careful interpretation alongside cross referencing to the results from the food stimulus alone, when considering the results of an experimental trial.

3.4.4 Expert consensus on behavioural measures of positive affect in equines

Within this consultation 19 behavioural measures for positive affective state were presented to (or suggested by) participants. Only seven of the measures achieved a cumulative ranking of ≥ 20 and these were conspecific (horse-horse) social interactions, body language assessment, facial action scoring, play behaviour analysis, qualitative behaviour assessment, human-animal relationship testing and cognitive bias testing. Many of the methods ranked with a score lower than 20 were identified as potentially useful markers but the experts felt that they required further validation before they could support their reliability and validity, measures of note included the absence of negative markers such as frustration and repetitive behaviours and the use of eye gaze or lateralisation analysis.

Recommendations for behavioural measures of positive affect to be taken forwards to experimental trial

Utilising the feedback from the Delphi consultation and evidence from the literature a series of measures were selected for inclusion as behavioural measures of positive affect in equines. Feedback from the experts indicated that they were satisfied that body language and facial action scoring and qualitative behaviour assessment were reliable methods to assess emotional state in equines however many participants commented on the lack of validation of newer methods (eye gaze / lateralisation) and there was disagreement regarding whether the absence of a negative (i.e. absence of frustration and repetitive behaviours) indicated the present of positive affect. As such it was decided to utilise the well validated metrics alongside those with less evidence to investigate the potential reliability of the less validated measures. Assessments that were not feasible within the proposed experimental study, due to the

timescale required for their collection or the resources required were excluded from selection regardless of rank. The following behavioural measures were selected:

1. Body language scoring including facial action scoring (ranked 2nd and 3rd)
2. Qualitative Behaviour Assessment (ranked 5th)
3. Absence of frustration (ranked 9th) and absence of repetitive behaviours (ranked 16th)
4. Eye gaze (lateralisation) (ranked 17th)

Assessment of body language and facial expression

Facial action scoring was ranked third (59.72% agreement) in Round one, Round three (89.23% agreement) and overall. Body language assessment was ranked second overall, second (72.73% agreement) in Round two and first (93.85% agreement) in Round three. Over the last twenty years a plethora of equine ethograms have been tested and validated to varying degrees including the equine discomfort ethogram (Torcivia and McDonnell, 2021), Facial Expressions in Ridden Equines (FEReq) (Mullard *et al.*, 2017), ridden horse pain ethogram (Dyson, 2022), the equid ethogram (McDonnell, 2003) and the equid play ethogram (McDonnell and Poulin, 2002). Many of these have been reviewed in relation to their use in equitation science, equine welfare and behavioural assessment (Hall and Heleski, 2017; Dyson and Van Dijk, 2020; Pierard *et al.*, 2019; Dyson *et al.*, 2020) with Hall and Heleski concluding that ethograms of behaviour are useful primary means of evaluating both the performance and welfare of horses alongside paired physiological measures. This was supported by Pierard *et al.* (2019) who found high reliability between raters (both researchers and practitioners) when scoring behaviour via video clips using an equine ethogram and Dyson *et al.* (2020) who found no significant difference between real-time and video observations of lame horses, nor between volunteers scores and the expert experimenter score, using the ridden horse ethogram.

Whilst overt body language in the form of behavioural ethograms is important in understanding the emotional state of an animal through its behaviour, more subtle signs of affect are displayed by many equines either prior to or in place of more overt body language (Draaisma, 2017). In previous studies looking at more nuanced markers of negative affect in horses, facial actions have been reported to be useful measures. McDuffee *et al.* (2022) studied behaviours indicative of stress in ridden therapy horses and identified thirteen subtle

and nuanced signs of stress in these horses including changes in head postures and facial features, tail movement changes and mouth related behaviours and several authors have identified the importance of facial actions in painful horses (Gleerup *et al.*, 2015; Gleerup and Lindegaard, 2016; Dalla Costa *et al.*, 2014b). These studies lend support to the notion that facial actions change with alteration in affective state and therefore may also be important when assessing positive affective states in equines. Indeed, Lansade *et al.* (2018) investigated facial expression and oxytocin as potential markers of positive affect in horses and identified that there appears to be distinct differences in facial features between horses deemed to be in positive affect (via gentle grooming and scratching) versus those deemed to be in negative affect (through standard grooming). Some features such as neck position (high in negative affect, lower in positive affect) and eye position (open wide in negative affect and half closed in positive affect) agree with previous understanding of equine behavioural repertoires however others such as ear position (backwards in positive affect and forwards in negative affect) appeared to contradict received wisdom that an equine with ears back is an unhappy equine; a finding echoed by Bornmann *et al.* (2021). Practically there is much nuance in many of the body posture signals emitted by equines which demonstrates the need to measure such indices alongside others to be able to attribute meaning to them. Taking the ears as an example, an equine with ears placed flat back and rotated down (likely accompanied by tense lips, cheek muscles and other body movements) is often a clear indication of negative affect and/or pain however many equines rest with their ears rotated backwards (or to the side) in a relaxed manner indicating possible positive affect (McGreevy, 2012). Conversely equines with ears pricked and forwards have historically been deemed as happy however for many equines this posture is accompanied by tense muscle tone, raised head and neck and wide eyes indicating nervousness, apprehension or over-arousal (Draaisma, 2017). These findings are potentially very useful in the study of positive affect in equines, especially as the findings correlated with physiological measures (such as oxytocin levels) that may indicate positive affect in equines and are supported by work undertaken by Trösch *et al.* (2020). Interestingly blood cortisol, heart rate and heart rate variability measures did not vary between groups nor over time within the Trösch *et al.* (2020) study, potentially because neither stimulus was salient enough to provoke a significant reaction; further supported by the fact that no overt behavioural signs were demonstrated in the test groups. In a further study Lansade *et al.* (2022) developed their framework for assessing positive affect using facial actions whilst

utilising automatic brushes in group-housed horses. This study supported the findings of the previous study (Trösch *et al.*, 2020) which is important not only to strengthen the evidence for this measure but also because these horses were at liberty and actively sought the grooming via the automatic brushes rather than it being enforced on them by a person. Interestingly in this second study horses showing a positive affect facial expression were six times more likely to begin allogrooming with another horse, a previously defined measure of a state of relaxation and low threat, supporting the notion that these horses were in positive affect.

The evidence from prior research and the opinions of the experts supports the inclusion of these behavioural measures of positive affect in equines.

Qualitative Behaviour Assessment

Qualitative Behaviour Assessment (QBA) was ranked 1st (77.78% agreement) in Round one and 5th (83.08% agreement) in Round three of the Delphi consultation and 5th overall. It is a popular methodology within equine scientific enquiry with initial literature searches yielding over 40, 000 results. It is cited as a holistic and dynamic tool which can capture expressive qualities of animal behaviour addressing the whole animal and scoring them, often using visual analogue scales (Minero *et al.*, 2018).

QBA methods have been used widely to assess horses' responses to a range of situations from endurance rides (Fleming *et al.*, 2013), short term positive and negative emotional treatments (Hintze *et al.*, 2017) and the presence of humans (Minero *et al.*, 2018), to ill-health (Hausberger *et al.*, 2016), welfare of tethered horses (Mullan *et al.*, 2014) and responses to training (Ruet *et al.*, 2022). Some methodologies utilise observer generated descriptors (termed free choice profiling) whilst most on-farm methods utilise fixed-list methods where descriptors are selected beforehand by the researchers following a review of the related literature (Wemelsfelder *et al.*, 2001). Hall *et al.* (2018) postulate that the use of QBA is a promising measure for positive affect in horses but requires further validation alongside physiological measures to fully access accurate assessment of positive affect. In their review of equine welfare indicators, Dalla Costa *et al.* (2014c) identified the use of QBA as a valid test for the principle of 'good behaviour' (behaviour indicative of good welfare) with high inter-observer repeatability and on-farm feasibility.

The expert Delphi group recommended the use of QBA with fixed-list methods and as such this was the method taken forward to the experimental study via the use of the fixed QBA term list from the AWIN welfare assessment protocol for horses (AWIN, 2015) which was further validated by Minero *et al.* (2018).

Absence of frustration and repetitive behaviour

The absence of frustration was ranked 9th overall, 4th (50% agreement) in Round two and 13th (70.77% agreement) in Round three. Absence of repetitive behaviours was ranked 16th overall, 8th (27.78% agreement) in Round one and 14th (67.69% agreement) in round three. Frustration is defined as the emotion that predominates during the thwarting of impulses or actions that prevents individuals from obtaining something they have been led to expect based on past experience (Amsel, 1992; Dollard *et al.*, 2013). In equines, frustration is thought to indicate a negative affective state with chronic frustration linked to poor welfare (Lesimple, 2020; Mellor *et al.*, 2020), therefore an absence of frustration could be an indicator of, at least, the absence of negative affective state but when combined with other variables, potentially also an indicator of the presence of positive emotional state.

A recent study (Pannewitz and Loftus, 2023) surveyed expert opinion through a Delphi consultation and reached consensus on seventeen causes of frustration in horses and thirteen behavioural signs of frustration (divided into acute and chronic frustration) specific to equines. Highlighted in this study was the presentation of repetitive behaviours as an indicator of frustration but also the current difficulty in differentiating affective states caused by frustration from those induced via other causes (such as pain or anxiety, facets that may be addressed further using biomechanical analysis suggested by Delphi participants in the current study). As a measure to be taken forward to the experimental study an important consideration is the potential for transient alteration in affective state, for example during goal directed behaviours, where initial positive affect linked to goal directed behaviours may switch to negative affect if the goal is not achieved and frustration is experienced (Hintze *et al.*, 2017), a factor that should be evaluated during induction of affective state. In addition, frustration may be short lasting (the inability to access a food reward for example) or longer lasting (for example extended social isolation due to injury related box rest) and is likely to be complexly intertwined within the network relationships in brain physiology and associated neurochemistries (Barrett, 2018) meaning that it may be difficult to disentangle whether

frustration observed relates directly to the stimulus in question or whether it relates more to the overall environment and/or prior experience and learning which could confound the interpretation process.

Frustration in horses has previously been studied in relation to feeding methods (Ninomiya *et al.*, 2004), repetitive behaviours (Broom and Kennedy, 1993), gene expression (Ninomiya *et al.*, 2013) and has also recently been studied in mules (da Luz *et al.*, 2022). It is hypothesised that frustration may lead to a negative feedback loop which may continue to affect behaviour over time and that frustration is accompanied by higher mean and maximum heart rates in the individual compared to baseline measures (Rørvang *et al.*, 2021).

Cooper and McGreevy (2007) linked stereotypy (repetitive behaviour) to frustration whereby environmental deficits may lead to boredom, deprivation or frustration emanating in the development of repetitive behaviours, a theory also presented by McAfee *et al.* (2002) who indicate that the presentation of repetitive behaviours may both cause and intimate poor welfare (both physical and psychological) and thereby their absence may indicate higher levels of emotional and physiological welfare. In addition, the prevention of repetitive behaviour performance can increase stress and frustration, evidenced by higher heart rates and elevated corticosteroid levels, and further compromise welfare (McBride and Cuddeford, 2001; Lebelt *et al.*, 1998; McGreevy and Nicol, 1998), a facet that should be identified during assessments of emotional welfare state. Therefore, it appears logical that the absence of both frustration and repetitive behaviours should be measured when attempting to conclude that an equine is in a positive emotional state.

Eye gaze analysis

Eye gaze analysis or eye lateralisation was ranked 17th overall; 6th (22.73% agreement) in Round one and 20th (52.31% agreement) in Round two; however, experts reported that it could be a useful measure but currently does not have enough evidence to support its reliability to measure affective state. It is a relatively new area of study in equines however studies on lateralised behaviours, including eye gaze, can be found within the cross-species literature dating back to the early 2000's. Studies in other mammals and birds have focussed on aspects such as lateralisation of mother-infant interactions (Karenina and Giljov, 2022; Karenina *et al.*, 2018), asymmetry of behavioural responses to human approach (Sankey *et al.*, 2011), laterality of sucking in foals (Komárková and Bartošová, 2013) and lateralisation of

agonistic and vigilance responses in horses (Austin and Rogers, 2012; Austin and Rogers, 2014).

Eye gaze analysis is a methodology requiring minimal equipment nor training and as such is one that could easily be employed by horse caretakers in an applied environment if proving reliable to address the measurement of positive affect in equines. Early studies investigating eye gaze analysis or lateralisation studies in equines developed upon research undertaken in other mammalian species, birds, reptiles, amphibians and fish (for a full review see Rogers *et al.* (2013)) identifying that hemispheric specialisation exists across taxa with aspects of behaviour expressed asymmetrically, especially in animals with laterally placed eyes, such as equines. Studies across species have identified the use of the left eye (right hemisphere) in unfamiliar situations and the right eye (left hemisphere) in familiar situations (Larose *et al.*, 2006).

A review of the literature by Leliveld *et al.* (2013) identified an evidence base for the processing of negative emotional stimuli in the right hemisphere and positive emotional stimuli in the left hemisphere however within their review they identified contradictory evidence across taxa with 43 papers suggesting right hemispheric dominance in situation of negative emotional valence and 20 papers suggesting the opposite. Their research highlights a strong evidence base for positive emotional processing in the left hemisphere with 21 studies identifying this phenomenon versus 6 studies suggesting the opposite. The authors hasten to point out that there is likely an individual influence on these characteristics and highlight a small number of papers suggesting this occurrence. The use of this measure with the later empirical study may add valuable literature to the current evidence base, especially where the aspects of lateralisation may be paired with other behavioural and physiological data.

3.4.5 Expert consensus on physiological measures of positive affect in equines

Within this consultation 23 physiological measures for positive affective state were presented to (or suggested by) participants. Eleven of the measures achieved a cumulative ranking of ≥ 20 and these were heart rate and heart rate variability, salivary cortisol, respiratory rate, serotonin levels, electroencephalogram, blood oxytocin, faecal, blood and urinary cortisol and gut microflora assessment. Many of the methods ranked with a score lower than 20 were identified as potentially useful markers but the experts felt that they required further

validation before they could support their reliability and validity, measures of note included infrared thermography of the eye and ear, skin conductivity/sweat rate and hair cortisol measurement.

Recommendations for physiological measures of positive affect to be taken forwards to experimental trial

Utilising the feedback from the Delphi consultation and evidence from the literature, a series of measures were selected for inclusion as physiological measures of positive affect in equines. Feedback from the experts indicated that they were satisfied that measures 1-3 (heart rate variability and heart rate, salivary cortisol and respiratory rate) were reliable methods to assess emotional state in equines however many participants commented on the lack of validation of newer methods (infra-red thermography of the eye and ear). As such it was decided to utilise the more highly validated metrics alongside those with less evidence to investigate the potential reliability of the less validated methods. Methods that were not feasible within the proposed experimental study, due to the timescale required for their collection, the invasiveness of the method or the resources required were excluded from selection regardless of rank. The following physiological measures were selected:

1. Heart rate variability and heart rate (ranked 1st and 3rd)
2. Salivary cortisol (ranked 2nd)
3. Respiratory rate (ranked 4th)
4. Infrared thermography of the eye and ear (ranked 13th)

Heart rate and heart rate variability

Heart rate was ranked 3rd overall, 2nd (70.31% agreement) in Round one and 4th (65.36% agreement) in Round three. Heart rate variability (HRV) was ranked 1st overall, 1st (87.5% agreement) in Rounds one and three (92.19% agreement). The literature on the use of heart rate and heart rate variability (HRV) analysis in horses is plentiful and varied in context (such as within psychological stress, exercise, novel situations) to an extent whereby there are numerous review articles supporting its function, validity and reliability (Von Borell *et al.*, 2007; Stucke *et al.*, 2015; McDuffee *et al.*, 2019).

HRV measurement is a non-invasive technique used to investigate the functioning of the autonomic nervous system (Pham *et al.*, 2021). Originally used in human research, since the

early 1990's its use has expanded into the area of non-human animals. The measure analyses changes in sympathovagal balance, effectively the respective dominance of the sympathetic and parasympathetic (or vagal) nervous systems, and the complex functioning and interplay of vagal and sympathetic activity; a process which cannot be fully understood through measuring heart rate alone (Tulppo *et al.*, 1998). In humans the measure has been used to assess autonomic function associated with various diseases processes including mental states such as depression (Carney *et al.*, 2005) with numerous pieces of research using HRV to assess emotional states in humans as outlined by Von Borell *et al.* (2007) in relation to animals. Due to the close similarities in physiology across mammals, especially the endogenous impact of emotions transmitted via the parasympathetic nervous system, it is widely agreed that HRV is as useful a measure in non-human animals as it is in humans (Porges, 1995).

Whilst heart rate and HRV have predominantly been used to assess stress responses in animals, with studies demonstrating a shift to sympathetic dominance in situations of physical or psychological stress (Borstel *et al.*, 2017; Rietmann *et al.*, 2004; McDuffee *et al.*, 2019; Munsters *et al.*, 2012), it has also been asserted that heart rate and HRV can conversely be used to measure relaxation and appetitive behavioural states in animals (Boissy *et al.*, 2007a; Scopa *et al.*, 2020; Kay, 2012; Feh and de Mazières, 1993).

From a methodological perspective it is important that HRV used to assess emotional state includes inter-beat-interval (IBI) data lasting at least five minutes consecutively measured during stationary conditions to negate the effect of movement on cardiac function as well as pre-behavioural physiological anticipatory responses and persisting post-behavioural cardiac responses (Malik *et al.*, 1996). In order to ensure accuracy of ambulatory EEG devices used in equine research a number of studies have compared results of ambulatory devices (such as Polar Equine heart rate monitors) with static EEG devices with high levels of agreement identified at rest (Ille *et al.*, 2014; van Oldruitenborgh-Oosterbaan *et al.*, 1988; Evans and Rose, 1986), however, 'during exercise' results show less agreement across studies with Lenoir *et al.* (2017) and Parker *et al.* (2009) finding low agreement and Fripiat *et al.* (2021) and Ter Woort *et al.* (2021) reporting high levels of agreement between ambulatory and static devices, potentially due to the methods utilised within the different studies and the types of exercise tested. During data analysis inter beat interval (IBI) data should be edited to identify and correct IBI errors as identified by Scopa *et al.* (2020) and correct frequency bands for

species specific analysis should be utilised, which for horses encompasses very low frequency (0.0-0.01Hz), low frequency (0.01-0.07Hz) and high frequency (0.07-0.6Hz) (Stucke *et al.*, 2015).

In their review Von Borell *et al.* (2007) conclude, supported by McDuffee *et al.* (2019), that HRV analysis in horses appears to be a sensitive measure of both physical and emotional responses providing the methodological design and cleaning and analysis of data generated are carefully undertaken. Indeed, Scopa *et al.* (2020) identified vagal (parasympathetic) dominance in horses exposed to a familiar affiliative caretaker in a well-designed study of 23 horses and Loftus *et al.* (2016) used HRV to determine presence or absence of psychological stress in horses in training successfully correlating these measures with behavioural data, as seen also in Visser *et al.* (2003). Therefore, the evidence from the literature, and the strong consensus of the expert Delphi group supports the use of heart rate and heart rate variability within the experimental design of the field trial; both to ascertain affective state and to validate the more novel methods discussed within this chapter.

Cortisol

Salivary cortisol was ranked 2nd overall, 3rd (37.5% agreement) in Round one and 2nd (76.56% agreement) in Round three. The hypothalamic-pituitary-adrenal (HPA) axis is known to be activated during times of both physical and psychological stress (Von Borell, 1995) with cortisol levels changing during these times (Mormède *et al.*, 2007). The literature has demonstrated increases in cortisol levels associated with acute stress however the effect of chronic stress on cortisol levels is less consistent across the literature with it reported to increase, remain the same or be depressed across different studies (Pawluski *et al.*, 2017). Mazzola *et al.* (2021) assessed hair cortisol levels when studying chronic stress responses finding fluctuations over time correlated with seasons and management practices. Hair cortisol is generally used as a longer term measure of stress via cortisol secretion as it passively diffuses into hair follicles from blood vessels during hair growth to give retrospective fingerprints of cortisol levels at points in time (Keckeis *et al.*, 2012) whereas salivary and blood cortisol are useful measures of acute stress over minutes to a few hours (Schmidt *et al.*, 2010) and urinary and faecal cortisol methods give a measure over a 24 hour period (Aurich *et al.*, 2015).

The use of salivary cortisol testing in horses has predominantly focussed on establishing parameters for normal levels and reporting natural fluctuations due to circadian influence as

well as quantifying stress responses to exercise and stimuli in the environment. The measure has been validated against blood serum cortisol levels with high agreement identified (Peeters *et al.*, 2011). Interestingly in their comparison of salivary cortisol tests analysed via liquid chromatography-tandem-mass spectrometry (LC-MS) and enzyme-linked immunoassay (ELISA) Sauer *et al.* (2020) identified that results correlated in terms of their relative increase or decrease however ELISA results were consistently higher than LC-MS results (but remained correlated); a concept that should be further validated in the subsequent experimental study to enhance the reliability of results obtained. To further enhance reliability and validity of results it is recommended that each horse act as their own control especially as differences have been identified across sex, time of year, age and housing parameters when measuring salivary cortisol in horses (Aurich *et al.*, 2015; Contreras-Aguilar *et al.*, 2020). Sauer *et al.* (2019) studied the effect of breed, management and personality on cortisol levels in sports horses following an adrenocorticotrophic hormone (ACTH) challenge and identified that management practices rather than horse use seemed to correlate with cortisol differences whereby stabled horses had an increased response to the challenge as did those housed in groups compared with horses housed outside. They found no effect of personality nor level of competition on cortisol response; a finding that may be pertinent in the later experimental chapter where personality and exercise variables are considered as factors in the responses to the presentation of positive stimuli.

Whilst all the available literature using cortisol measures is focussed on negative physical and emotional welfare states, the method has potential for use as a measure of good welfare if used to review the responses of individual equines to different stimuli and paring these data with other measures of affective state. The literature is more plentiful on the use of salivary cortisol compared with other sample types (blood, urine, faeces, hair), with reference ranges available, and it is demonstrated that rapid pen-side tests on salivary cortisol levels are valid and reliable making this a feasible measure for inclusion within an on-farm study (König v. Borstel *et al.*, 2017; Mazzola *et al.*, 2021; Pawluski *et al.*, 2017; Sauer *et al.*, 2020; Shirtcliff *et al.*, 2015).

Respiratory rate

Respiratory rate was ranked 4th overall, 3rd (56.06% agreement) in Rounds two and three (75.00% agreement). Respiratory rate is known to decrease during rest and relaxation and

increase during arousal however the changes in respiratory rate related to positive valence are still unclear (Fraser, 2009). Hall *et al.* (2018) discuss respiratory rate within their review of assessing affective state in equines but with limited evidence cited for its use as a measure aside from its elevation in arousal. In dogs however, respiratory rate has been cited as a potentially useful measure of emotional valence (Csoltova and Mehinagic, 2020) and in sheep lower respiratory rates have been linked with positive valence (Reefmann *et al.*, 2009) as is reported in the respiratory literature across species where slow and deep respiration is linked to excited positive affect and slow and shallow respiration in calm, relaxed positive affect (Boiten, 1998; Boiten *et al.*, 1994). Neethirajan (2022) highlights the use of respiratory rate analysis in affective state assessment in farm animal species and suggests that this measure is best utilised alongside other measures including heart rate, body temperature and behavioural data to produce a holistic and more reliable perspective on valence and arousal; facets that will be included in the later experimental study.

In horses lower respiratory rate has been associated with positive affect for example, Briefer *et al.* (2017b) identified a lower respiratory rate in horses listening to affiliative reunion whinnies compared with affiliative departure whinnies and all unfamiliar whinnies indicating a physiological respiratory response to a presumed positive stimulus. Supporting this Menchetti *et al.* (2021) identified lower respiratory rates in horses with higher scores for calmness, no avoidance and affiliative responses indicating that this may be a valid metric for assessing positive valence and low arousal.

Whilst there is limited literature to date on the feasibility of respiratory rate in measuring positive affect in horses, the literature that is available suggests that this could be a feasible, simple and on-farm friendly measure if validated alongside other more widely evidenced measures.

Thermography of anatomical features

Thermography was ranked 13th overall, 6th (42.20% agreement) in Round one and 14th (46.88% agreement) in Round three. Whilst thermography has gained popularity as a measure of emotion in equines due, in part, to the increase in available technology to measure this effect both using specialist cameras and add-ons to smartphone or other smart devices; the robustness of this as a measure was clearly debated by experts within the Delphi consultation. In a review of robust indicators of welfare in equitation science Waran and Randle (2017)

indicate that eye temperature may be a suitable metric for assessing affective state in equines with further validation. Temperature of the inner corner of the eye is deemed to indicate core body temperature and as such could be a measure of this metric in the horse however data on this facet is very limited in the literature for horses but increases in core temperature have been associated with anxiety and stress responses in other species (Hall *et al.*, 2018).

In a PhD thesis exploring positive affect in horses Stratton (2022) identified no difference in eye temperature between preferred, intermediate and aversive treatments however the sample size was small (n=13) and data were missing for some treatments. However Cook *et al.* (2001) identified a correlation between elevated salivary and plasma cortisol levels and increased eye temperature measured via infrared thermography potentially signifying an association between elevated eye temperature and activation of the HPA axis, a finding supported by the results of Borstel *et al.* (2017) but with no differences identified in other studies such as Marsh *et al.* (2019) which could be related to a variety of confounding factors including the experimental design, magnitude of the response variables assessed or camera setting utilised. The importance of using multi-modal assessments to measure affect was highlighted by the results of Travain *et al.* (2016) who suggested that eye temperature may be a better measure of arousal level as opposed to valence in their study of positive emotion in dogs, as it was not easy to distinguish differences between dogs in positive affect during food delivery versus those feeling frustrated or conflicted during food delivery using eye thermography as a sole metric. When combined with other metrics validating positive affect the authors identified that changes in temperature appeared to be linked to the arousal element of positive affect rather than the valence itself.

In horses studied post-operatively, ear temperature has been seen to decrease linked to decreased peripheral perfusion related to cardio-pulmonary and neuro-endocrine alterations indicative of hypotension related to stress or shock (Rawlings *et al.*, 1982), a facet that could be interesting from an affective perspective. In dogs, decreases in ear temperature have been linked to separation distress (Riemer *et al.*, 2016) and in dairy calves' eye and ear temperature asymmetries have been identified during a negative experience (disbudding) indicating right hemisphere activation thought to indicate negative affect (Ramirez Montes De Oca, 2021). Ramirez Montes De Oca (2021) identified lateralised increased nostril, eye and ear temperatures indicative of negative affective state during non-rewarded novel object trials

and indicative of positive affective state in rewarded trials in dairy calves and in rewarded and non-rewarded trials in horses, a finding also supported by Kappel (2023) in their PhD thesis examining cognitive and physiological indicators of affective state in horses. These findings are based on the grounding that there are ipsilateral relationships between nostril and ear temperatures and hemispherical brain activation and contralateral relationships between eye and ear temperatures and hemispherical brain activation, a phenomenon requiring further investigation in equines. With these findings in mind thermography of facial landmarks may be a useful measure within a framework to assess positive affective state. However, evidence in horses is very limited within the literature currently and requires further validation against previously validated metrics as will be undertaken during the experimental chapter.

3.5 Limitations of the study

The participants

There are several limitations evident within this study which may have impacted the conclusions drawn. Whilst the range of experts included within the study was broad in scope and well justified, the study did exclude participants who utilised punitive training methods for equines and those working as behaviourists who were not listed as members of professional bodies representing accredited behaviourists. This exclusion criteria may have led to missing data in terms of the views of these individuals which could have contributed to a broader perspective on the themes discussed within the consultation. Additionally, the demographics of the participants may have influenced their responses especially where potential induction or measurement methods may have greater or lesser feasibility or popularity across different countries, professional specialisms or years of experience. Cultural bias, the tendency to interpret or judge phenomena in terms of the society or community to which one belongs (APA, 2024), may also have had an impact on the responses given by participants, particularly surrounding affective states and the equine experience of these. Whilst inclusion and exclusion criteria were carefully selected, set a priori and adhered to, some literature suggests that undertaking a value web approach to the selection of Delphi experts can help to ensure selection of participants with the breadth, depth and objectivity required for the consultation (Devaney and Henchion, 2018), a process that was not undertaken in this study.

Initially ordinal logistic regression was undertaken to assess the impact of participant demographics on responses given, however the imbalanced sample across many demographic features significantly limited the validity of these analyses and they were therefore omitted. In future studies it may be beneficial to plan balanced participant recruitment into the methodological design to aid further statistical analysis, however this could have further negative repercussions if potential participants who could give experienced and valid responses are omitted on a demographic basis only.

The Delphi process

Whilst a three-round Delphi is considered robust in terms of the iterative data collection process it proved difficult to achieve consensus on the proposed methods to induce and measure positive affect using a numerical ranking-based system. As such the selection process was adapted for the final round to ask participants to state whether they felt measures should be included in later experimental design, rather than asking them to rank their method preferences. This adjustment greatly facilitated the development of consensus agreement regarding these measures but was based upon the percentage of people selecting measures to be included rather than there being a numerical ranking of each method by each rater. An alternative third round, or additional fourth round could have been undertaken as a focus group approach whereby a smaller cohort of selected experts engage in a face-to-face or online video-format discussion to come to agreement around the ranking of each set of measures. Unfortunately, within this study there was limited further time available for the completion of a fourth round or the organisation of a focus group, incidentally, this format of final round is not often seen within the equine Delphi literature but is more widely used in other sectors such as healthcare and environmental studies, where it may also be used as a first-round process to define the research question and parameters (Williams *et al.*, 2017a; Canessa *et al.*, 2022).

It is possible that some of the responses given by participants may have been biased by their particular area of expertise, with participants potentially choosing to answer questions for which they had limited practical or theoretical knowledge to make judgements on. This may have occurred because every participant was invited to answer every question, whilst some participants opted not to answer questions and commented where they felt they did not possess the knowledge to add feedback in the specific area, others may have answered

regardless, and this could have led to some opinion-based responses with limited evidence to support them. However it is shown within the literature that expert bias and confidence challenges are generally well overcome through the iterative feedback and social learning that a Delphi consultation enables, which may have reduced the impact of any of these occurrences (Devaney and Henchion, 2018).

Alternative platforms for the hosting of the Delphi consultation may have enabled anonymous, real-time discussion between experts on the questions posed which may have been advantageous in terms of generating data representing multidirectional dialogues and capturing further qualitative information, a process which was undertaken using the Blackboard platform by Rioja-Lang *et al.* (2020). However a disadvantage of using this methodology is that some participants may be inhibited from responding if they felt uncomfortable expressing alternative viewpoints, or as often occurs in focus group approaches, if a small number of participants dominate the discussion (Smithson, 2008). Participants may also change their responses based upon the responses of others which can lead to reduced variation in responses and the potential to overlook important aspects of the questions being posed (Nyumba *et al.*, 2018).

Interpretation of responses

Asking participants to rate and rank their agreement or disagreement with statements may lead to a decrease in the nuance of their selections. Feedback from this Delphi study indicated that some participants struggled to ascribe ordinal ranking to items which they felt were multifactorial. This could mean that the interpretation of the data may have flaws where rankings were decided upon without understanding the contextual nuances of each participant's selections. This aspect was somewhat mitigated by the availability of free-text response boxes after each ranking question within the consultation rounds and many participants utilised these opportunities to provide useful and valid feedback which was analysed to shape further rounds. It is feasible that some participants did not choose to provide this additional feedback to outline their decision-making process.

Qualitative data were numerous throughout this Delphi and, whilst thorough and well considered, the thematic analysis is unavoidably subjective in terms of the placement of feedback items into the thematic groupings identified. It is possible that some qualitative

feedback was misinterpreted within this process potentially leading to erroneous conclusions. To counterbalance this potential the results from each round were presented to all participants for them to indicate their agreement (or disagreement) as a secondary confirmation process to enhance reliability and validity of the results reported.

3.6 Recommendations

Whilst this Delphi consultation included a large and experienced cohort of experts, further studies of this type could look to recruit as balanced a participant sample as possible in terms of subject specialisms, experience, and demographic characteristics. This would enhance the possibility of undertaking valid statistical analysis of the data generated.

Additionally, it may be beneficial to undertake initial pre-Delphi focus groups to identify key questions and themes to ensure that important considerations are not omitted from the consultation process. The addition of a final discussion-based focus group round with a smaller, but balanced, selection of experts may also be useful within studies where it has proven difficult to reach consensus on specific themes or questions, this may strengthen the validity of the consensus reported at the end of the consultation.

3.7 Conclusion

This Delphi consultation highlighted several important considerations surrounding the induction and measurement of positive affective state in equines, including the expert-identified importance of considering the individual's personality within any measures of affective state. Additionally, the importance of preference testing induction methods within individuals was a factor highlighted by many participants, especially in relation to food and touch preferences. Establishing individual preference should therefore be a prerequisite of field studies in this area.

Further, the consideration of the animal's environment and lifestyle was highlighted qualitatively by many participants and as such this data should be collected as part of any study attempting to measure affect. The study highlighted a wide range of methods to induce and measure positive affect, some methods being well evidenced in the literature whilst others more novel and nuanced. The study highlights the differences in viewpoints between specialists in the field, dependant on their specific role and experience, and placed a spotlight on the significant differences in opinion that exist where experts approach a particular issue

from their own perspective, a factor which can make achieving consensus challenging, as evidenced throughout this study.

Several valid invasive measures of affect were suggested throughout the Delphi consultation. These included assessment of serotonin levels, blood oxytocin, blood cortisol, catecholamine levels, tryptophan levels and the use of corticotrophin releasing factor challenge. Whilst a number of these measures are reported within the literature, with many validated relatively widely, it is beyond the remit of this work to undertake invasive procedures in relation to collecting data to measure positive affective state and as such these measures must be excluded from further study design. Additionally, it is feasible that the act of taking these measures could influence affective state, for example discomfort during blood draw could induce negative affect which may alter responses to a subsequent positive stimulus thereby confounding results obtained.

Utilising a range of methods to induce and measure positive affect should allow for individual preferences to be accounted for as well as facilitating the validation of more novel measurements alongside those more commonly cited in the literature. Overall, this study has generated a significant amount of expert data relating to positive affect which can be utilised to better the emotional, and physical, welfare of equines.

Chapter 4 Development of the Equine Personality Eight-Factor Model (EPEM)

4.1 Introduction

This chapter describes the development of the Equine Personality Eight-Factor model (EPEM) and evaluates the evidence collected for the reliability of the model as a potential measure of personality in equines. At the time of the development of this model prior research into assessing equine personality had focussed on utilising the Five Factor Model (Tupes and Christal, 1961; Costa and McCrae, 1989) or the development of novel models based upon analysis of temperament and character traits (Suwała *et al.*, 2016) or specific tests for personality in distinct equine populations such as police horses (Pierard *et al.*, 2017) and sports horses (Hausberger *et al.*, 2011). The aim of this chapter was to consider the cross-application of a previously validated hominoid personality questionnaire (Robinson *et al.*, 2018b; Robinson *et al.*, 2021) for equine use via an iterative process of development and to consider its relation to the extensively used Five Factor Model which has previously been used in equine personality studies (Morris *et al.*, 2002; McGrogan *et al.*, 2008).

A mixed-methods approach was taken to this study whereby the original hominoid personality questionnaire (Robinson *et al.*, 2018b; Weiss *et al.*, 2009) was adapted semantically for equine use and tested on a group of equines to provide data regarding potential correlation of traits. Expert consensus was then sought via a Delphi consultation (Chapter 2 section 2) to iteratively develop traits for inclusion in the final questionnaire. The consensus-developed traits were then tested for validity across a larger cohort of equines. The aim being to utilise this personality assessment within the final experimental study into inducing and measuring positive affective state in equines (Chapter 6) to allow for the effect of personality to be assessed in relation to affective state.

4.1.1 Prior studies on equine personality

Morris *et al.* (2002) analysed the agreement between raters judgements of horse personality focussing in the main on the Five Factor Model of Personality (FFM) originally identified by Tupes and Christal (1961) and developed and championed by McCrae and Costa (1987) (see Chapter 1 section 1.6). This model is presented as NEO-PI-R (NEO personality inventory revised) and NEO-FFI (NEO five factor inventory) (Costa and McCrae, 1989) and covers the global personality traits of Neuroticism, Extraversion, Openness to experience,

Agreeableness, Conscientiousness. Within their study they called upon the prior work of several eminent theorists in human psychology including Funder (1999), Buss (1997), Eysenck (1963), Zuckerman (1991) and Gray (1981) in order to design a study which could reliably evaluate rater agreement on horse personalities. The initial framework for the study was based on Funder (1995)'s work on personality judgement issues which presents the following domains as important considerations when studying personality and led to the development of the Realistic Accuracy Model (RAM):

- a) Good trait - are some traits more easily judged than others?
- b) Good judge - are some people better judges than others?
- c) Good target - are some individuals more easily judged than others?

As with much empirical research there are two obstacles to personality assessment reliability, namely low sample size and lack of standardised taxonomy (Gosling and Bonnenburg, 1998). To attempt to alleviate these issues researchers have often utilised the FFM to ensure standard taxonomy is applied and enable comparisons to be drawn within and between species. Sample size is a factor unique to each piece of research and one which can only be alleviated through inclusion of a large enough cohort to ensure a powerful experiment. In the Morris *et al.* (2002) study the authors utilised standard taxonomy to enable parallels to be drawn when discussing their results however their sample size (n=10, 5 mares and 5 geldings) was low (unlike the cross-species study of Gosling and Bonnenburg (1998), reported later in this section, where n=1640) which may have affected the generalisability of results obtained.

Unlike other studies Morris *et al.* (2002) utilised naive, untrained raters who had a good knowledge of the horses involved in the study but no experience of undertaking personality studies. The rationale for this methodology was to utilise detailed knowledge of the test subjects and identify whether specialised psychological profiling methods could be used by the general population of equine caregivers to assess their horses' personalities. In this study raters were provided with 12 items (descriptive words or phrases) per trait (each of the 'Big Five': Neuroticism, Extraversion, Openness to experience, Agreeableness, Conscientiousness) and were asked to score each item on an analogue scale from low (strongly disagree that the trait is applicable to the individual horse) to high (strongly agree that the trait is applicable to the individual horse). Analysis was undertaken to assess the correlation between individual item scores within each trait per individual and were found to be significantly correlated thus

indicating that individual items were contributing to overall trait scores. Across traits results identified high agreement between raters indicating they can reliably rate all the horses on all the 'Big Five' traits. Interestingly openness and conscientiousness were most difficult to apply in an equine context whereas neuroticism, extraversion and agreeableness were shown to be more highly correlated. The results of this study also highlighted that the level of agreement overall between judges was good, some horses were more difficult to rate for some traits (openness and conscientiousness) and that the individual traits assessed were independent. This study concluded that the NEO-FFI is reliable for use in equines with results comparable to the similarly conducted human studies by Costa and McCrae (1995) and John and Robins (1993) using the same framework. The authors suggest that cues to equine personality are readily available even to naive raters following on from the studies into the Realistic Accuracy Model (RAM) by Funder (1995) and Letzring *et al.* (2020) and that the use of anthropomorphic personality descriptors in non-human species can generate reliable results as supported by Hebb and Williams (1946) and Lloyd *et al.* (2007) however other authors are more questioning of the use of anthropomorphic wording especially when used by naive raters as they feel it may challenge the accuracy of reporting (Breland and Breland, 1961; Kennedy, 1992; Wemelsfelder *et al.*, 2000).

A 2008 study by McGrogan *et al.* (2008) utilised a list of 36 descriptive words formed by 30 horse owners/carers to rate 100 horses for each trait on a 7-point Likert scale, 84 raters rated the horses. They removed two adjectives (stupid and wise) within the analysis due to low reliability and ran a principal component analysis on the remaining 34 items. They used factor analysis to link the adjectives to three factors of the NEO-PI-R Agreeableness, Extraversion and Neuroticism. High internal consistency was also reported for factors 'agreeableness', 'extraversion' and 'neuroticism' and no differences found related to age or sex of the horses in the study. Many of the adjective correlations reported in this study are supported by correlations utilised in previous human studies using the same methodology (McCrae and Costa, 1989; McCrae and Costa, 1992).

Within the equine literature Duberstein and Gilkeson (2010) investigated personality and trainability differences between male and female yearlings (an animal over a year old but under two years old) and found that mares were found to be more skittish, aggressive, tense, and excitable compared to geldings (castrated male horses) who were found to learn quicker

especially when undergoing desensitisation training. This study is limited in that the female horses were entire however the males were not and as such the hormonal influences on behaviour may influence behaviour depending on the stage of the oestrus cycle, and therefore on the results of this study. An improvement to the study would be to include a cohort of entire males alongside the females and castrated males to give a rounded view of the effect of sex on personality and trainability. Pierard *et al.* (2017) designed and conducted a behavioural personality test for Belgian police horses utilising 10 Likert scores and yielding 115 continuous variables. The test assessed responses to novel stimuli including objects in an arena, sudden loud noises, suddenly appearing objects, interactions with another horse and handler interactions. Of the 125 variables assessed only 6 showed a significant difference between the horses categorised as 'well-suited' police horses and those deemed to be 'less well-suited'. The design of the study appears to be very comprehensive and the authors stress that the aim of the initial study was to provide a reference range for currently serving police horses in order that newly recruited horses could be assessed using the test in future to enable appropriate selection of horses for mounted police work. The authors reported that the limited significant differences in behavioural responses identified between the 'well-suited' and 'less well-suited' horses in this cohort are likely to be because these horses have already been selected, trained, and working as police horses and as such their reactivity is likely to be lower than that found in the general population of equines. Further testing of this methodology on other cohorts of trained and naive horses is required to validate the protocol. A similar study by Lansade *et al.* (2016) investigated the development of personality tests linked to performance for show jumping horses. Their study assessed personality using several metrics and measured correlations over time following retesting of the animals. The study also considered the effect of location/situation and as such assessed the horses at events as well as in a known environment. Results showed that the personality measurements were highly correlated across both time and location meaning that they appear to be sensitive and reliable in assessing personality. The sample size for this study was large (n=650) and the repeated measures over 3 years suggest a strong study design and reliable results.

Hausberger *et al.* (2011) undertook what appears to be the first study of its type investigating the effect of work type on horse personality. Their supposition was that, as in humans, job characteristics may, over time, augment personality traits due to placement in conditions

which may suppress emotions or introduce interpersonal conflicts. Indeed, research has shown that different types of work can lead to chronic behavioural disorders such as stereotypy formation outside of the work situation (Hausberger *et al.*, 2009). This study led on from a large cohort study (n=700) by Hausberger *et al.* (2004) which identified that several genetic and environmental factors affected personality including paternal origin, breed, and environmental housing conditions. Hausberger *et al.* (2011) identified that whilst overall display of emotion varied slightly between horses trained for different types of work significant differences were found between the groups interest in a novel object or scenario. Dressage and high school horses were found to be most emotionally reactive, flighty, and anxious compared with show jumping and voltige (acrobatics on horseback) horses. In discussion the authors propose that the higher levels of emotional reactivity in these highly schooled horses may be due to the strict control experienced during their working lives. This could present conflict and prevent expression of emotion which is known to be a source of stress in humans (Richards and Gross, 1999; Yoon *et al.*, 2017; Zapf, 2002). In addition, dressage training develops response sensitivity to ensure rapid response to rider cues which may lead to nervousness and, over time, affect overall personality (Rankins and Wickens, 2020).

Mills (1998b) found that inter-assessor correlation for empirically defined personality traits in horses was low when the traits used were those traditionally prescribed to horses (such as flighty, sharp, honest) and suggested that further delineation of traits, such as those found in human personality assessment, with less emphasis on traditional equestrian terms could be more reliable. Following on from this Suwała *et al.* (2016) proposed a new horse personality model based on assessment of temperament and character with four dimensions identified for both factors. Temperament factors identified were energy (tEN), fearfulness (tFR), sensitivity (tSN) and adaptability (tAD). For character the traits identified were submissiveness (cSB), aggression (cAG), human contact seeking (cCS) and self-reliance (cSR). Of the 801 participants surveyed on the potential of this technique there was strong agreement that all traits were important in personality assessment and that overall selection of equines with high tEN, tSN, tAD, cSB, cCS and cSR traits and low tFR and cAG traits was desirable for trainable equines. To date there is no literature on the effectiveness of this model as an assessment of equine personality. More recently Jolivald (2021) considered physiological

correlates with equine personality traits based on the FFM within their PhD thesis and identified that the equine personality traits used did not correlate with physiological measures highlighting a concerning disparity between what is observed behaviourally using this trait assessment and what is evidenced physiologically. However they did find that scores for personality groups agreeableness, neuroticism, extroversion, and gregariousness were internally consistent and reliable between observers (Jolivald *et al.*, 2022). Their follow-up work assessing chronic stress physiology in relation to the personality trait agreeableness (suggested to be indicative of habitual compliance where horses will comply with instructions given even if they perceive the instruction as negative from an affective perspective) identified a relationship between these factors which is an important consideration from a welfare perspective (Jolivald *et al.*, 2023).

4.1.2 Aims and objectives

This current study aimed to investigate and validate a novel equine personality assessment via the use of an owner-completed questionnaire and the consensus opinion of a group of Delphi experts. The objectives of the study were to a) assess the level of agreement between scorers when scoring fifty-five equine personality traits, b) analyse personality trait data within and between equines to identify personality type clusters, c) present the identified clusters and traits to a group of experts to achieve consensus agreement on the inclusion and grouping of traits, d) re-test the inter-observer reliability of these traits, trait cluster agreement and internal consistency via a second questionnaire and test intra-rater reliability for the final personality traits and groupings and e) develop a short form of the questionnaire to enable rapid but reliable personality assessment.

4.2 Methodology

The methodology encompassed a four-part iterative design to allow for development of an equine personality questionnaire that considered the views of the Delphi experts (methodology in Chapter 2) as well as the objective data collected as described in sections 4.2.1 to 4.2.4 below. An overview of the methodology is given in Figure 4.1 below.

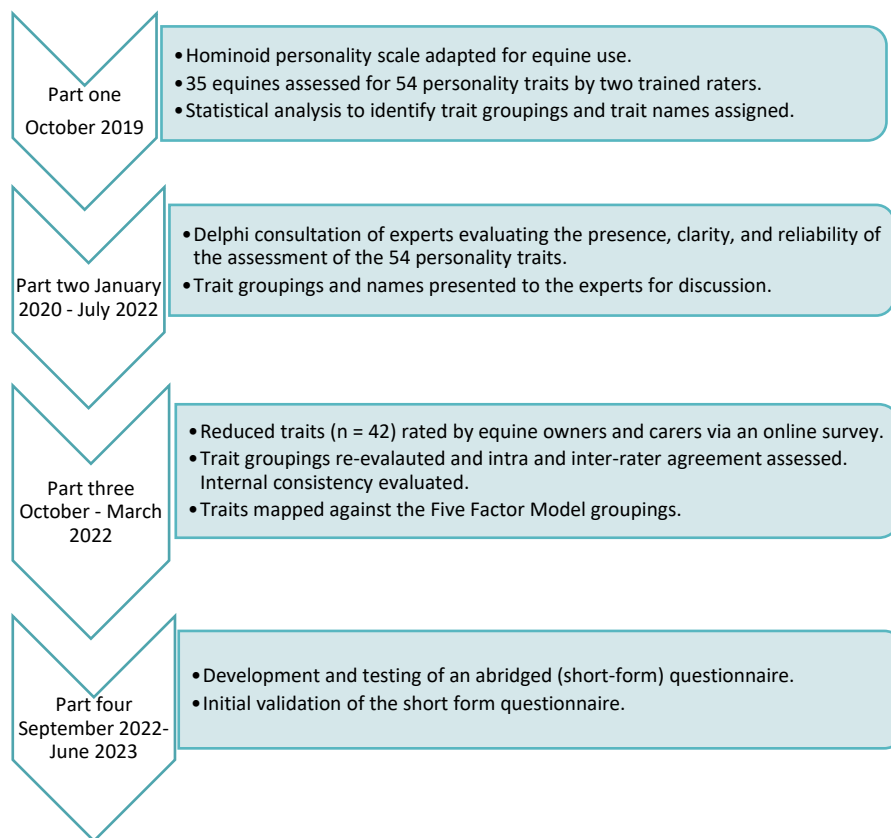


Figure 4.1 An overview of the study methods, the full Delphi methodology can be found in Chapter 2.

4.2.1 Part one

A hominoid personality scale (Weiss *et al.*, 2011) was adapted for equine use (full original scale in Appendix C) as previously reported in other species (Robinson *et al.*, 2016). This study was conducted between 2019 and 2023 via an iterative process to carefully develop the personality questionnaire with the input of Delphi experts and equine owners.

The scale presented 54 personality traits and asked participants to rate the animal on a 1-7 Likert scale for each trait, example in Table 4.1.

Table 4.1 Examples of the same question within the hominoid and equine questionnaires and the Likert scale definitions given to participants.

Likert scale definitions	1. Displays either total absence or negligible amounts of the trait. 2. Displays small amounts of the trait on infrequent occasions. 3. Displays somewhat less than average amounts of the trait. 4. Displays about average amounts of the trait.

	5. Displays somewhat greater than average amounts of the trait. 6. Displays considerable amounts of the trait on frequent occasions. 7. Displays extremely large amounts of the trait.							
Example hominoid question	FEARFUL: Subject reacts excessively to real or imagined threats by displaying behaviors such as screaming, grimacing, running away or other signs of anxiety or distress. leastmost 1 2 3 4 5 6 7 ☐ ☐ ☐ ☐ ☐ ☐ ☐							
Example equine question	FEARFUL: Subject reacts excessively to real or imagined threats by displaying behaviours such as snorting, shying, moving away or other signs of anxiety or distress. LeastMost <table><tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr></table>	1	2	3	4	5	6	7
1	2	3	4	5	6	7		

Study one commenced in November 2019 with 35 healthy equines, known to the researcher, and their owners recruited for the study. All participants were required to have known the animal for at least 6 months prior to participation (mean ownership length 73 months, SD = 54.07). Demographic data were collected on the equines (Table 4.2). Following giving informed consent, participants rated their equines via a Google Form questionnaire on the 54 personality traits from the adapted hominoid personality questionnaire (Weiss *et al.*, 2011), the equines were also independently scored by the researcher.

Table 4.2 Demographic information for the 35 test equines. Definitions: Mare – entire female, Gelding – castrated male, Stallion – entire male 4 years old or over, Colt – entire male less than 4 years old or Filly – entire female less than 4 years old. Turned out – spends 24 hours a day in the field, Stabled – is not turned out in a field, Both – spends time in both the stable and a field in each 24 hour period. Light work – up to an hour a day a few times a week, Medium – daily exercise which includes predominant work in trot or canter with some jumping or schooling, Hard work – daily exercise with predominance of faster work with additional competitions or None – not ridden.

Horse Number	Age (Years)	Breed	Sex	Turnout	Hours of turnout	Type of Work
1	4	Arabian	Gelding	Turned out 24/7	24	Light
2	6	Thoroughbred	Gelding	Stabled	0	Hard
3	8	Thoroughbred	Mare	Stabled	0	Hard
4	5	Thoroughbred	Mare	Stabled	0	Medium
5	6	Hanoverian	Gelding	Both	7	Light
6	6	Hanoverian	Gelding	Both	7	Medium
7	5	Arabian	Mare	Both	6	Light
8	6	Arabian	Mare	Turned out 24/7	24	Hard
9	12	Welsh D	Gelding	Turned out 24/7	24	Light
10	10	British Riding Pony	Gelding	Turned out 24/7	24	Medium
11	19	Anglo Arab	Gelding	Both	4	Light
12	6	Thoroughbred	Mare	Stabled	0	Hard
13	9	Warmblood	Gelding	Both	6	Medium
14	12	Warmblood	Mare	Both	6	Light
15	9	Warmblood	Stallion	Stabled	0	Light
16	14	Arabian	Stallion	Turned out 24/7	24	Light
17	5	Arabian	Stallion	Turned out 24/7	24	Light
18	11	Arabian	Mare	Stabled	0	Hard
19	6	Cleveland Bay	Mare	Both	9	Light
20	15	Cleveland Bay	Mare	Both	9	Light
21	21	Shetland Pony	Mare	Turned out 24/7	24	None
22	13	Welsh A	Mare	Turned out 24/7	24	Light
23	12	Thoroughbred	Mare	Stabled	0	Hard
24	12	Thoroughbred	Gelding	Stabled	0	Hard
25	16	Warmblood	Gelding	Both	8	Light
26	26	Appaloosa	Mare	Turned out 24/7	24	Light

27	15	Shire warmblood	x Gelding	Turned out 24/7	24	Light
28	7	Appaloosa	Mare	Turned out 24/7	24	Light
29	13	Warmblood	Mare	Both	5	Medium
30	11	Hanoverian warmblood	Gelding	Both	7	Hard
31	16	Hanoverian warmblood	Gelding	Both	7	Light
32	4	Thoroughbred	Gelding	Stabled	0	Hard
33	22 months	Thoroughbred	Filly	Turned out 24/7	24	None
34	18 months	Warmblood	Filly	Turned out 24/7	24	None
35	10 months	Warmblood	Filly	Turned out 24/7	24	None

Inter-observer agreement across all responses and within individual traits were tested via intraclass correlation (ICC (2,1), two-way mixed model, for absolute agreement, average measures). Principal component analysis (PCA) was undertaken to investigate relationships between trait scores across the cohort (Greenacre *et al.*, 2022) and to collate overarching personality trait groupings which would be fed back to the Delphi participants in part two of the study. The data were checked for cross-loading with no significant cross loaders identified (Muthen and Asparouhov, 2012) and identified clusters were examined for internal consistency using Cronbach's alpha analysis (Hajjar, 2018). All analyses were undertaken in Minitab 18 (Minitab, 2021) and SPSS 27 (IBM, 2020).

4.2.2 Part two

The data on participants views of equine personality and the presence, clarity and reliability of each trait in equines, collected during round one of the Delphi consultation (method described Chapter 2, section 2.2) were analysed thematically using CATMA (Guis et al., 2024) and descriptively using Minitab (Minitab, 2021).

Quantitative data on the clarity, presence and reliability for each personality trait in equines were not all normally distributed (Anderson Darling test, Clarity $A^2 = 0.24$, $p = 0.766$; Presence $A^2 = 0.84$, $p = 0.029$; Reliability $A^2 = 0.94$, $p = 0.016$) and did not all have equal variance (Clarity : Presence Levene's test $P < 0.001$; Clarity : Reliable Levene's test $P < 0.001$; Presence : Reliable Levene's test $P = 0.254$) and therefore non-parametric statistical analysis was utilised due to the relatively small sample size. Kruskal-Wallis with post hoc Mann-Whitney analysis with Bonferroni correction for multiple comparisons was undertaken to investigate the differences in scores for clarity, presence and reliability awarded across each personality trait. Spearman rank correlation was utilised to investigate the relationship between the scores for the presence of traits in part one of the study and the expert rated 'presence' of each trait in equines.

Round two qualitative data on the trait groupings and group names were analysed thematically and the expert suggested alterations to group names, trait names and traits included within each group were undertaken. Backwards and forwards stepwise analyses were undertaken following each change to trait groupings, with internal consistency measured via Cronbach's alpha. Following this the revised trait grouping (eight groups) and names (42 trait names) were presented to the experts in round three of the Delphi consultation to confirm consensus agreement on all trait groupings and names.

4.2.3 Part three

The initial 54 personality traits now reduced to 42 expert selected traits (with their associated agreed descriptors) were placed within a Google Forms survey instrument; each alongside a seven-point Likert scale, with additional demographic questions including a unique person identifier number for later intra-rater reliability analysis, the type of yard (do-it-yourself livery, part-livery or full-livery), equine name, length of time caring for the equine, how well the participant knew the equine, and the age, sex and breed of the equine; to gather data on the equine being rated. Owners and carers of equines were invited to participate in the study through an online social media recruitment post with a link to the survey placed on the equine pages/forums of Facebook, LinkedIn, Reddit and Instagram. Participants completed the survey between October and November 2022 and were invited to complete it again between February and March 2023 to gather intra-rater reliability data via a test-retest methodology (Berchtold, 2016; Hendrickson *et al.*, 1993). Participants were asked to recruit others who

knew their equines well (other carers or riders of their equines who had known them for over six months) to also complete the survey for the individual equine(s) to gather inter-rater reliability data.

Following survey closure, data were cleaned and organised for analysis. Intraclass correlation coefficient was used to assess inter-rater (ICC (1,k) one-way random effects model for absolute agreement, average measures) and intra-rater agreement (two-way mixed effects model for absolute agreement, average measures). Following this, principal component analysis was undertaken to identify whether the expert identified personality trait groupings were valid when applied to real-life data. Item analysis to identify internal consistency was undertaken via Cronbach's alpha. All analyses were undertaken in Minitab 18 (Minitab, 2021) and SPSS 27 (IBM, 2020).

The final personality groups (n= 8) and their component traits were then mapped against the trait facets and groupings of the Big Five factor model (McCrae and Costa, 1987) to ensure that all model factors were incorporated within this proposed Equine personality Eight-factor Model (EPEM).

4.2.4 Part four

Following finalisation of the EPEM, testing of an abridged (short-form) questionnaire was undertaken using only the overarching personality group names. The aim of this section was to evaluate whether an abridged version of the questionnaire would yield the same results as the full questionnaire and therefore whether the abridged form may be reliable as a shorter method of evaluating personality in equines. As per the method of Smith *et al.* (2000), *a priori* examination of the reliability and validity of the short form in use was conducted on a small subset of the sample. The short form was released to 4 raters across 3 horses initially, these horses had previously been scored using the full questionnaire. Responses for the short-form were compared with the long-form (mean trait scores within personality groups were calculated for comparison) using intraclass correlation coefficient analysis (ICC (2,k) two-way mixed effects model for absolute agreement, average measures) and were then assessed per personality group using intraclass correlation coefficient analysis (ICC (2,k) two-way mixed effects model for absolute agreement, average measures). Initial evaluation suggested that it may be beneficial to add in trait names to the abridged short-form questionnaire groupings, without the full definitions provided in the long form, to aid reliability of scoring.

Following this, twenty equines recruited for a later experimental study (Chapter 6) were scored by their owners using the abridged model with trait names included and the full model and the results were compared between the two sets of scores using intraclass correlation coefficient analysis (ICC (2,k) two-way mixed effects model for absolute agreement, average measures). Construct validity was tested on the final trait groupings using Spearman's rank correlations. All analyses were undertaken in SPSS 27 (IBM, 2020).

4.3 Results

The results of this study are presented in chronological order to demonstrate the developmental process for the design and evaluation of the proposed equine personality assessment.

4.3.1 Part one – October 2019

This considers the initial rating of 35 equines using all the 54 adapted personality traits from the Hominoid personality scale (Weiss *et al.*, 2011) and the evaluation of the scores for each trait to identify personality trait groupings.

Overall rater agreement

Inter-rater agreement overall, across 54 traits and 35 horses was excellent (ICC (2,1) = .893, $p < 0.001$) with substantial rates of absolute agreement (Figure 4.2)

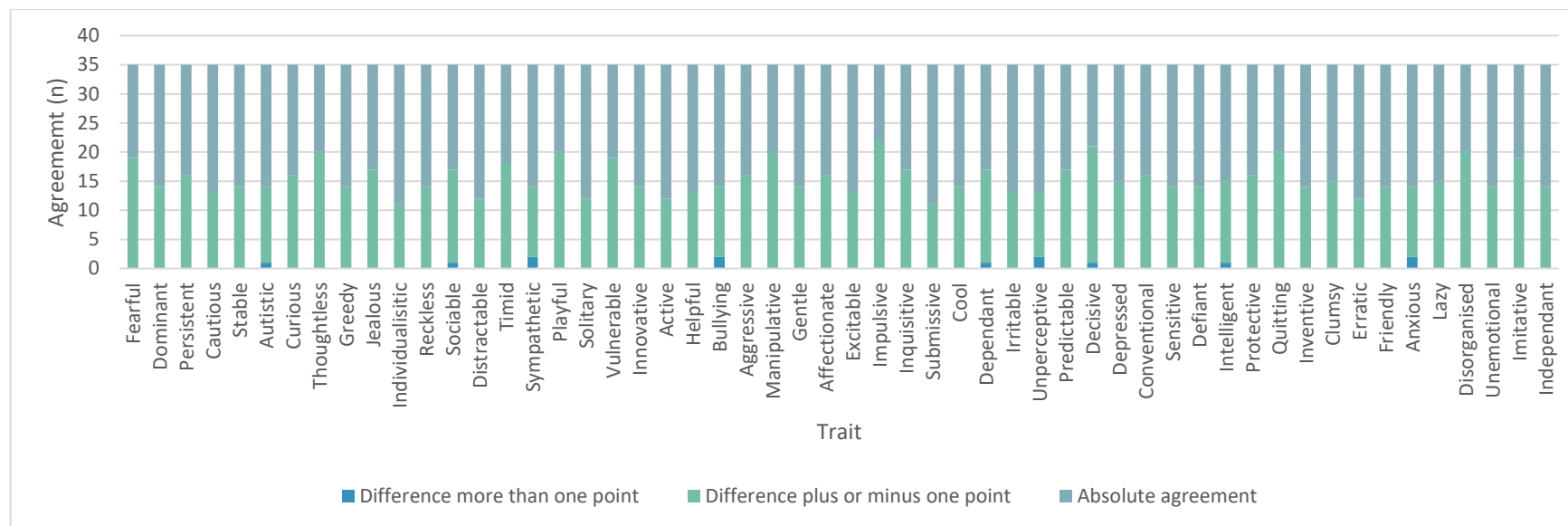


Figure 4.2 Inter-rater agreement per trait (count data). Absolute agreement shown in the lightest shade (top of each bar), differences in scores of plus or minus one point shown in mid shade (middle of each bar) and differences of more than one point shown in dark shade (bottom of relevant bars, only 9 instances across the 54 traits).

Agreement within individual personality traits

Inter-rater agreement varied within traits (two-way mixed effects model for absolute agreement ICC (2,1)) with only four showing required levels of inter-rater agreement as shown in Table 4.3. Agreement level was specified as ≤ 0 – no agreement, 0.01-0.20 slight agreement, 0.21-0.40 fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 good agreement and 0.81-1.00 excellent agreement as per (McHugh, 2012). However, comparison of the means from each rater for each trait were generally similar and only eight traits had instances where disagreement was more than one point different.

Table 4.3 –Inter-rater agreement of individual personality traits. Fair to moderate significant agreement in bold, fair and above non-significant agreement in italics, differences between observer means of less than 0.5 underlined. nd>1 = number of observations per trait with disagreement greater than one point. nAA = number of pairs of ratings with absolute agreement per trait, out of 54 total ratings per trait.

Trait	Agreement level (nAA/54)	Mean observer 1	Mean observer 2	Difference in means	n d >1	ICC (2,1)	<i>p</i>
Fearful	<i>Fair (16)</i>	2.34	2.34	0	0	0.217	0.245
Dominant	Slight (21)	2.31	3.11	0.8	0	0.126	0.257
Persistent	Slight (19)	3.71	3.77	<u>0.06</u>	0	0.102	0.381
Cautious	Slight (22)	4.09	4.23	<u>0.14</u>	0	0.025	0.472
Stable	None (21)	3.94	3.83	<u>0.11</u>	0	-0.488	0.869
Autistic	Slight (21)	1.83	2.14	<u>0.31</u>	1	0.215	0.232
Curious	None (19)	2.83	3.51	0.68	0	-0.367	0.857
Thoughtless	<i>Fair (15)</i>	3.34	3.91	0.57	0	0.383	0.039
Greedy	Slight (21)	3.63	4.26	0.63	0	0.027	0.457
Jealous	None (18)	3.89	4.43	0.54	0	-0.909	0.989
Individualistic	<i>Fair (24)</i>	3.60	3.74	<u>0.14</u>	0	0.254	0.199
Reckless	None (21)	3.51	4.03	0.52	0	-0.030	0.538
Sociable	<i>Fair (18)</i>	4.74	4.34	<u>0.40</u>	1	0.252	0.183
Distractible	Moderate (23)	3.14	4.11	0.97	0	0.498	<0.001
Timid	Slight (17)	4.23	3.89	<u>0.34</u>	0	0.150	0.314

Sympathetic	None (21)	3.94	4.00	<u>0.06</u>	2	-0.522	0.879
Playful	None (15)	4.20	4.71	0.51	0	-0.325	0.816
Solitary	<i>Fair (23)</i>	<i>3.77</i>	<i>3.20</i>	<i>0.57</i>	<i>0</i>	<i>0.224</i>	<i>0.207</i>
Vulnerable	Slight (16)	2.83	3.77	0.94	0	0.048	0.429
Innovative	Moderate (22)	3.34	4.29	0.95	0	0.496	0.001
Active	None (23)	4.11	4.23	<u>0.12</u>	0	-0.075	0.581
Helpful	None (22)	3.91	3.83	<u>0.08</u>	0	-0.361	0.807
Bullying	None (21)	3.94	4.00	<u>0.06</u>	2	-0.662	0.921
Aggressive	Slight (19)	2.34	3.09	0.75	0	0.074	0.365
Manipulative	None (15)	4.14	4.71	0.57	0	-0.272	0.785
Gentle	Slight (21)	2.31	3.11	0.8	0	0.126	0.257
Affectionate	Slight (19)	3.71	3.77	<u>0.06</u>	0	0.102	0.381
Excitable	Slight (22)	4.09	4.23	<u>0.14</u>	0	0.025	0.472
Impulsive	Slight (13)	2.31	2.40	<u>0.09</u>	0	0.185	0.281
Inquisitive	None (18)	3.89	4.43	0.54	0	-0.909	0.989
Submissive	<i>Fair (24)</i>	<i>3.60</i>	<i>3.74</i>	<i>0.14</i>	<i>0</i>	<i>0.254</i>	<i>0.199</i>
Cool	None (21)	3.51	4.03	0.52	0	-0.030	0.538
Dependant	<i>Fair (18)</i>	<i>4.74</i>	<i>4.34</i>	<i>0.40</i>	<i>1</i>	<i>0.252</i>	<i>0.183</i>
Irritable	None (22)	3.80	3.83	<u>0.03</u>	0	-0.145	0.648
Unperceptive	None (22)	3.86	4.00	<u>0.14</u>	0	-0.406	0.833
Predictable	None (18)	2.37	3.09	0.72	0	-0.009	0.516
Decisive	<i>Fair (14)</i>	<i>2.34</i>	<i>2.34</i>	<i>0</i>	<i>2</i>	<i>0.217</i>	<i>0.245</i>
Depressed	Slight (20)	2.34	3.11	0.77	0	0.048	0.408
Conventional	Slight (19)	3.71	3.77	<u>0.06</u>	0	0.102	0.381
Sensitive	Slight (21)	4.09	4.23	<u>0.14</u>	0	0.025	0.472
Defiant	None (21)	3.94	3.83	<u>0.11</u>	0	-0.488	0.869
Intelligent	Slight (20)	1.86	2.14	<u>0.28</u>	1	0.133	0.335
Protective	None (19)	2.83	3.51	0.68	0	-0.367	0.857
Quitting	Fair (15)	3.34	3.91	0.57	0	0.383	0.039
Inventive	Slight (22)	3.63	4.26	0.63	0	0.027	0.457
Clumsy	Moderate (20)	3.34	4.29	0.95	0	0.496	<0.001

Erratic	None (23)	4.11	4.23	<u>0.12</u>	0	-0.075	0.581
Friendly	None (21)	3.89	3.83	<u>0.06</u>	0	-0.382	0.818
Anxious	None (21)	3.94	4.00	<u>0.06</u>	2	-0.662	0.921
Lazy	Slight (20)	2.34	3.09	0.75	0	0.074	0.365
Disorganised	None (15)	4.14	4.71	0.57	0	-0.272	0.785
Unemotional	Slight (21)	2.31	3.11	0.8	0	0.126	0.257
Imitative	<i>Fair</i> (16)	2.34	2.34	<u>0</u>	0	0.217	0.245
Independent	Slight (21)	2.31	3.11	0.8	0	0.126	0.257

Agreement within individual equines

Agreement within individual horses, between observers, was excellent (ICC (2, 1) = 0.853, $p < 0.001$), (Figure 4.3).

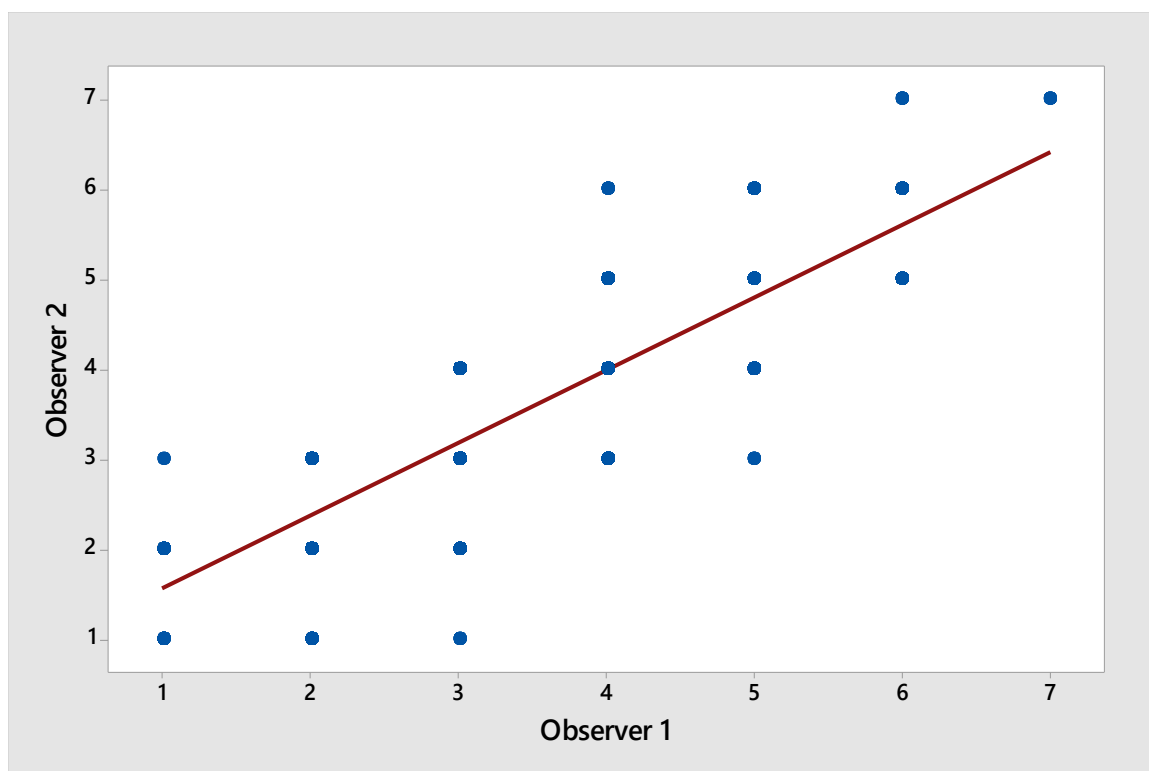


Figure 4.3 Mean scores correlation between observer one and observer two.

Relationship between trait scores

Initial Principal Component Analysis (PCA) identified that the 54 personality traits could be grouped into 10 principal components with Eigenvalues greater than 1 which cumulatively accounted for 86.7% of the variation in the data. Analysis of the scree plot identified the first

two components (with Eigenvalues above 4) to be taken forward for further analysis (Figure 4.4).

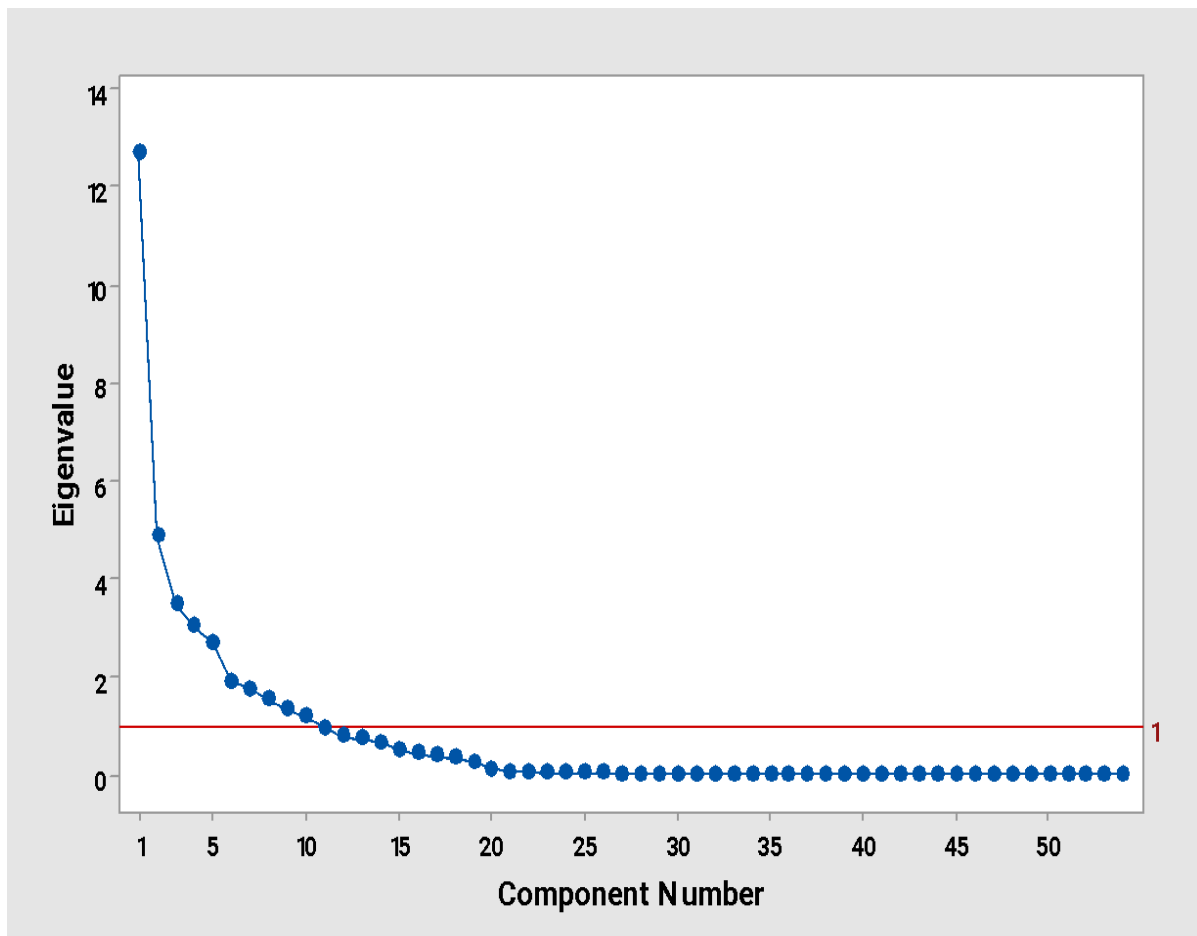


Figure 4.4 Principal components scree plot identifying 10 PCs with eigenvalues greater than 1 (1 signified by the red line).

The data were checked for cross-loading with no significant cross loaders identified. Following this, 3 large clusters and 5 smaller clusters on the loading plot (Figure 4.5) were selected for further analysis beginning with examination of internal consistency via Cronbach's alpha.

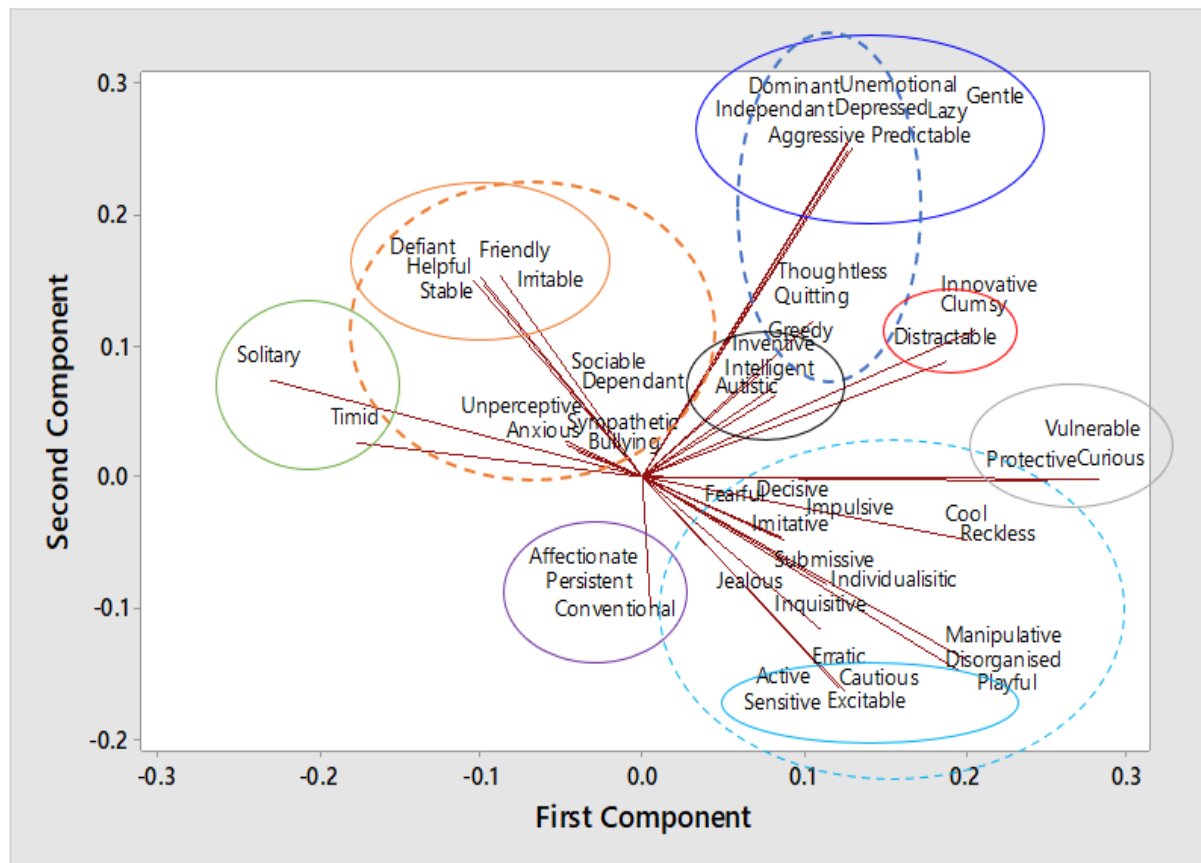


Figure 4.5 Loading plot for principal components 1 and 2. Dashed lines represent three distinct large clusters and continuous lines represent the subdivided three larger and five smaller clusters.

Cluster 1 (dark blue, Figure 4.5) showed very high internal consistency between variables (Cronbach's alpha = 0.9978). When variables 'dominant', 'independent' and 'aggressive' were removed from the analysis, internal consistency remained high at 0.9957 (Cronbach's alpha). When variables 'gentle' and 'predictable', which did not appear to align semantically with 'unemotional', 'depressed', and 'lazy', were removed from the analysis, internal consistency between traits 'unemotional', 'depressed', and 'lazy' remained high (Cronbach's alpha = 0.9920). When 'quitting', which appeared semantically aligned, was added to 'lazy', 'unemotional', and 'depressed' these were also highly associated (Cronbach's alpha = 0.8779) and these four traits formed the 'DEMOTIVATED' grouping. The removed traits were further evaluated, and their internal consistency was assessed. 'Gentle' and 'predictable' were analysed together showing high internal consistency (Cronbach's alpha = 0.9819) for a 'STEADY' grouping. Similarly, 'Dominant', 'independent', and 'aggressive' were analysed

together again showing high internal consistency (Cronbach's alpha = 0.9947). 'Thoughtless' and 'greedy' were analysed with 'dominant', 'independent', and 'aggressive' (cluster 1 expanded, blue dashed cluster) showing good internal consistency (Cronbach's alpha = 0.8251) with 'thoughtless' and 'greedy' sharing the same loading on PC1 as 'dominant', 'independent', and 'aggressive' but lower loadings on PC 2. This new grouping was termed 'CONTROLLING'.

Cluster 2 (orange) showed high internal consistency between the traits 'friendly', 'irritable', 'stable', 'helpful', and 'defiant' (Cronbach's alpha = 0.9921). When traits 'irritable' and 'defiant', which did not appear to align semantically with the other traits, were removed internal consistency increased (Cronbach's alpha = 0.9936). When traits 'sociable' and 'sympathetic' (dashed orange), which appeared semantically aligned, were added to 'friendly', 'stable', and 'helpful' internal consistency decreased but remained at an acceptable level (Cronbach's alpha = 0.6619) and formed the AFFILIATIVE grouping. When traits 'unperceptive' and 'bullying' were added to 'irritable' and 'defiant', due to their perceived semantic alignment, internal consistency lowered but remained at an acceptable level (Cronbach's alpha = 0.6750) and these formed the ANTAGONISTIC grouping.

Cluster 3 (Green) showed acceptable internal consistency between traits 'timid' and 'solitary' (Cronbach's alpha= 0.6523) and was termed the LONER grouping.

Cluster 4 (light blue) showed high levels of internal consistency (Cronbach's alpha = 0.9988) between all traits in the cluster. This cluster of traits included highly aroused attributes with varying valences to produce 'cautious/sensitive- erratic' (Cronbach's alpha = 0.9979) and 'excitable/active- erratic' (Cronbach's alpha = 0.9979) subgroups. When 'playful' and 'inquisitive' were added to the whole grouping internal consistency remained high (Cronbach's alpha = 0.8290) and allowed for a further high and positive arousal subgroup 'PLAYFUL' to be identified including traits 'excitable', 'playful', 'active', 'inquisitive, and 'erratic' with high internal consistency between traits (Cronbach's alpha = 0.8691). When 'submissive' was added to the whole group internal consistency remained high (Cronbach's alpha = 0.9486) however 'submissive' appeared semantically more aligned with 'cautious' and 'sensitive' and when added to the 'cautious-sensitive-erratic' group to form a 'NERVOUS' group; internal consistency was also high (Cronbach's alpha = 0.8687). When fearful and imitative were added to the NERVOUS group internal consistency remained above acceptable

levels (Cronbach's alpha = 0.7680). Analysis of further traits (dashed light blue) showed high internal consistency amongst all eight traits (individualistic, manipulative, cool, jealous, decisive, reckless, impulsive, disorganised) (Cronbach's alpha = 0.8551) and the grouping was termed 'COERCIVE'.

Cluster 5 (purple) showed the highest level of internal consistency between traits 'affectionate', 'conventional', and 'persistent' (Cronbach's alpha = 1). This group of traits was termed 'SUPPORTIVE'. Cluster 6 (grey) showed high internal consistency between traits (Cronbach's alpha, 0.8653) of 'vulnerable', 'protective', and 'curious' – culminating in the grouping 'HESITANT'. Cluster 7 (red) showed acceptable internal consistency between traits 'clumsy' and 'distractible' to form a subgroup of 'GOOFY' (Cronbach's alpha = 0.6888). Cluster 8 (black) also showed acceptable internal consistency between traits 'intelligent', 'innovative', 'autistic', and 'inventive' (Cronbach's alpha = 0.6614) depicting an 'INTELLECTUAL' grouping.

Figure 4.6 depicts the internal consistency between traits within the groupings outlined in Figure 4.5 above and explains how traits were further evaluated via Cronbach's alpha analysis and either included or excluded from groupings.

From these identified clusters new personality groupings (Table 4.4) were suggested based on the tested internal consistency characteristics denoted above. These were then presented to participants via the Delphi consultation (part 2).

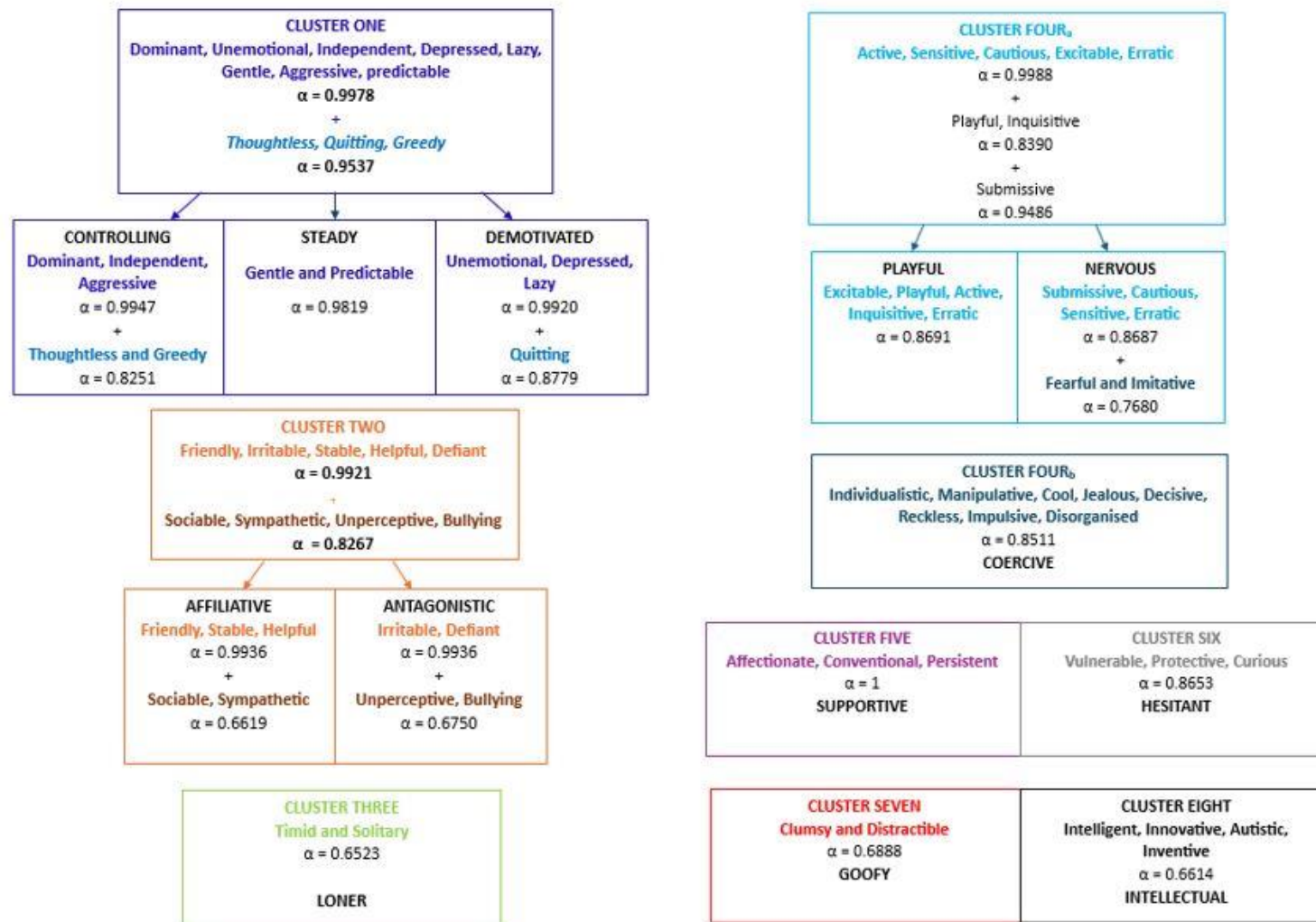


Figure 4.6 Schema depicting the internal consistency levels of each original group and how each large group was divided into small groups with the new group internal consistency ascribed.

Table 4.4 New personality groupings arising from PCA and internal consistency analysis.

New personality group name	Internal consistency (α)	Traits included
DEMOTIVATED	0.8779	Lazy, Depressed, Unemotional, Quitting
STEADY	0.9819	Gentle and Predictable
CONTROLLING	0.8251	Dominant, Aggressive, Independent, Thoughtless, Greedy.
AFFILIATIVE	0.6619	Friendly, Helpful, Stable, Sociable, Sympathetic.
ANTAGONISTIC	0.6750	Irritable, Defiant, Unperceptive, Bullying.
LONER	0.6523	Timid and Solitary.
NERVOUS	0.7680	Cautious, Sensitive, Submissive, Fearful, Imitative, Erratic.
PLAYFUL	0.8691	Excitable, Playful, Active, Erratic, Inquisitive.
COERCIVE	0.8364	Individualistic, Manipulative, Cool, Jealous, Decisive, Reckless, Impulsive, Disorganised.
SUPPORTIVE	1	Conventional, Persistent, Affectionate.
HESITANT	0.8653	Vulnerable, Protective, Curious.
GOOFY	0.6888	Clumsy and Distractible.
INTELLECTUAL	0.6614	Intelligent, Innovative, Autistic, Inventive.

4.3.2 Part two – January 2020 – July 2022

Delphi experts' views on equine personality

Within the first round of the Delphi consultation participants were asked if they believed that equines possess individual personality traits, 95.6% stated 'yes' (n = 87) with the other 4.4% (n = 4) stating 'probably yes'. Of the respondents 47.25% (n = 43) reported that they actively followed research into equine personality, 45.05% (n = 41) reported that they were aware of current research into equine personality but did not follow new research avidly and 7.69% (n = 7) reported that they were aware of it as a subject area but did not feel they had an up-to-date knowledge regarding the published literature on the topic.

Participants mean agreement with the statement provided to them that 'Personality is the individuals general behaviour compared with those found in the general population' was 66.14% (SD 30.47), whilst mean agreement for the statement provided to them that 'Behavioural syndromes describe variation within populations, not individuals. They describe the distribution of personalities within a population, for example horses which crib bite' was lower at 57.83% (SD32.75) and even lower for the statement provided which stated that 'Temperament is the animal's behavioural response in a single context measured on some form of scale, for example flight distance' at 43.21% (SD34.51). Therefore consensus (set at 70%) was not reached for any of these statements in round one.

Reaching consensus, 72% (n = 54) of participants agreed that an individual equine's personality would influence the intensities of different emotional states experienced by the individual, 28% (n = 21) were unsure.

Clarity, presence, and reliability of individual personality traits

Ninety-one Delphi responses were recorded for each personality trait and each of the three characteristics tested (Clarity of the trait, Presence of the trait in equines, Reliability of identifying the trait in equines); totalling 14,742 response items. Data were not normally distributed and did not have equal variance. Kruskal-Wallis with post hoc Mann-Whitney analysis with Bonferroni correction for multiple comparisons (Table 4.5) identified that there was a significant difference in scores awarded for Clarity, Reliability, and Presence (higher score = higher clarity, reliability and presence of the trait in equines) (Kruskal-Wallis, $H_2=22.06$, $p < 0.001$) with participants identifying that they felt Reliability of measurement of the traits was lower than their Clarity and Presence in equines.

Table 4.5 *Post hoc* Mann-Whitney analysis to identify differences in expert scores for the clarity, presence and reliability of the traits in equines (significant pairings highlighted in bold) with the median for Reliable significantly lower than the medians for Clarity and Presence.

Characteristics	Mann-Whitney result
Clarity (Median 38)/Present (Median 40)	W=2866.50, n=54,54, $p=0.641$
Clarity (Median 38)/ Reliable (Median 24)	W=3654.00, n=54,54, $p<0.001$
Present (Median 40)/Reliable (Median 24)	W=3553.50, n=54,54, $p<0.001$

Overall ‘sociable’ was the highest ranked trait when the expert scores for Clarity, Presence and Reliability characteristics were analysed together, followed by ‘fearful’ and ‘playful’. Twenty-four traits received high scores when the individual scores for ‘Clarity’, ‘Present’ and ‘Reliable’ were aggregated, with the remaining thirty traits scoring lower.

For the characteristic ‘Clarity’ trait ‘fearful’ was ranked second, ‘excitable’ third, ‘aggressive’ fourth and ‘playful’ fifth. For characteristic ‘Present’ trait ‘playful’ was ranked second, ‘inquisitive’ third, ‘fearful’ fourth and ‘aggressive’ fifth. For characteristic ‘Reliable’ trait ‘fearful’ was ranked second, ‘aggressive’ third, ‘playful’ fourth and ‘stable’ fifth. Figure 4.7 displays the total scores for ‘Clarity’, ‘Presence’ and ‘Reliability’ of each trait in a clustered bar chart with the overall score (scores for ‘Clarity’, ‘Presence’ and ‘Reliability’ summed and displayed via a total score line.

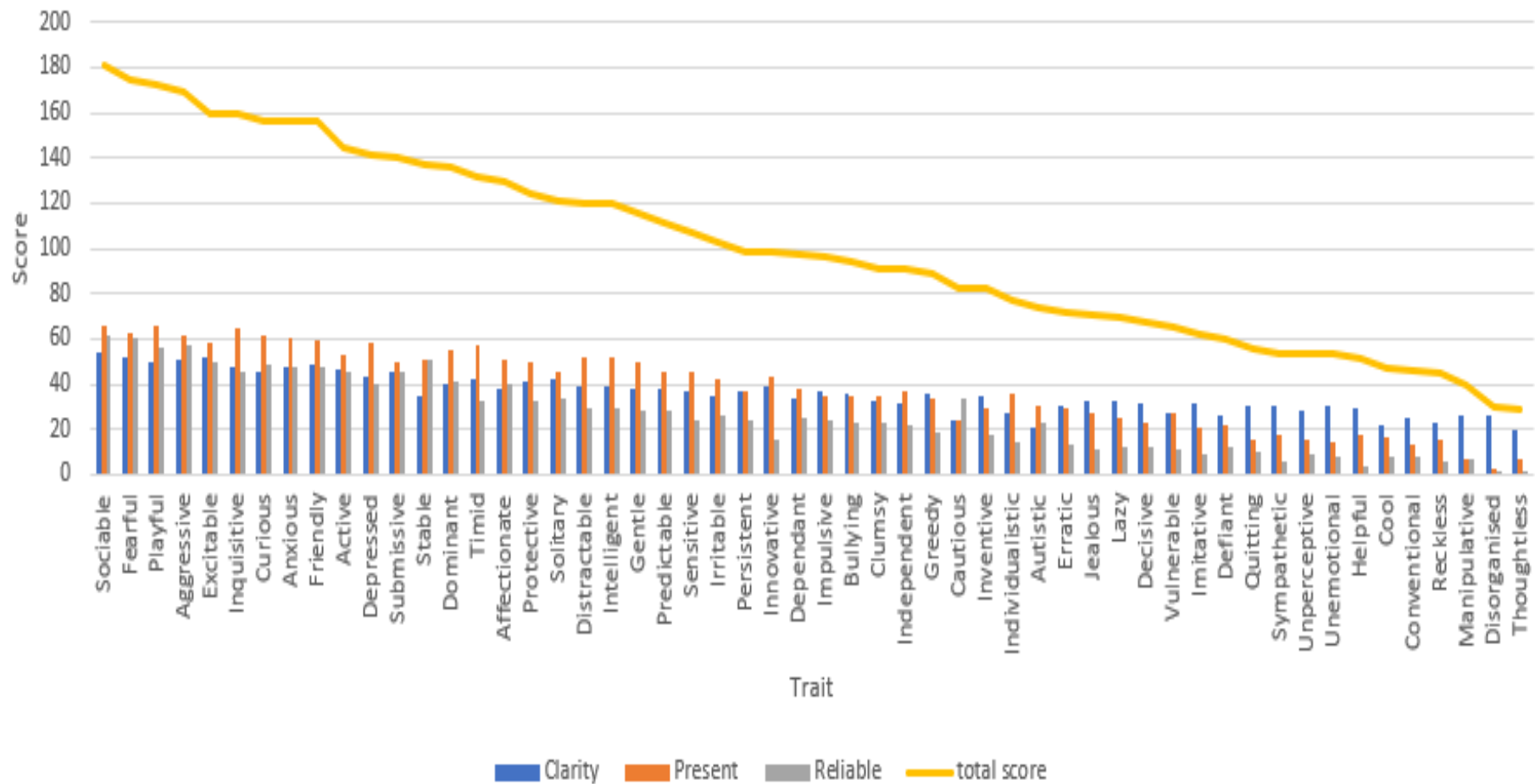


Figure 4.7 Clustered bar chart displaying expert rated scores for each trait across each characteristic with a total score line displayed in yellow.

Personality groupings and group names

Within the second round of the Delphi consultation participants were asked to provide feedback on trait groupings and group names presented to them following the analysis of the results of part one of this study.

Delphi results from a consultation of experts resulted in a wide range of agreement level and feedback with expert suggested alterations to the proposed groupings as discussed thematically below. Alterations to group names and trait inclusion was only made when at least two experts made the same suggestion.

Playful

Includes *Excitable, Playful, Active, Inquisitive, and Erratic*

Expert consensus was 77.28% (SD24.38) for this grouping. The experts suggested alteration of the group name 'Playful' (which was designated as a tautology) to 'Boisterous' or 'Exuberant'. Feedback suggested that the term *Erratic* is not symbiotic with the other three terms within this category but would be required to name the category 'Boisterous'; as such this category name was changed to 'Exuberant' and *Erratic* was removed. The experts suggested that inquisitive was similar to curious and should therefore be removed from the trait list altogether.

Nervous

Includes *Cautious, Erratic, Fearful, Submissive, Imitative, Sensitive*

Expert consensus was 73.74% (SD25.68) for this grouping. Respondents objected to use of the term *Submissive* and suggested it be changed to *Acquiescent*. *Imitative* was highlighted by respondents as unsuitable for this group, as such this was removed from the group.

Demotivated

Includes *Lazy, Depressed, Unemotional, Quitting*

Expert consensus was 69.19% (SD29.54) for this grouping. Respondents suggested alteration of the name of this group to 'Unresponsive'. They requested that the trait *Lazy* be removed.

Steady

Includes *Predictable, Gentle*

Expert consensus was 78.44% (SD25.05) for this grouping. Some respondents felt that *Gentle* was a poor fit for this grouping. They also suggested that the trait *Stable* be included into this group.

Controlling

Includes *Dominant, Aggressive, Independent, Thoughtless, Greedy*

Expert consensus was 57.96% (SD31.85) for this grouping. Respondents did not feel that the term 'Controlling' was a suitable term for equine use. They suggested renaming the grouping to 'Bold'. It was suggested to rename the term *Dominant* to *Forceful, Bullish* or *Assertive*. To avoid attributing negative connotations to equines within this category the term *Assertive* was selected. Participants also felt that the terms *Thoughtless* and *Greedy* were not applicable to this grouping, these were therefore removed. *Aggressive* was renamed *Confrontational* following feedback from participants.

Affiliative

Includes *Friendly, Helpful, Stable, Sociable, Sympathetic*

Expert consensus was 80.04% (SD24.60) for this grouping. Respondents asked for *Stable* to be moved to the 'Steady' group.

Antagonistic

Includes *Irritable, Defiant, Unperceptive, Bullying*

Expert consensus was 64.42% (SD32.05) for this grouping. Participants suggested the use of the group name 'Hostile' to replace 'Antagonistic'.

Loner

Includes *Timid, Solitary*

Expert consensus was 68.58% (SD29.38) for this grouping. Participants suggested replacing the group name with the term 'Shy'.

Coercive

Includes *Individualistic, Cool, Decisive, Impulsive, Manipulative, Jealous, Reckless, Disorganised*

Expert consensus was 52.29% (SD34.47) for this grouping. Participants were not in agreement with many of the characteristics of this grouping. They reported that the terms *Manipulative*, *Jealous*, *Disorganised* and *Reckless* were too highly cognitive to be applied to equines and should be removed.

Terms such as 'Standoffish' and 'Confident' were suggested by some respondents however these terms don't necessarily encompass all the trait characteristics included within the grouping.

Following the feedback of the Delphi experts it was decided to split this group into two more nuanced groups to allow for the differences in emotional state within the terms included. The first new grouping was termed 'Determined' and included *Decisive* and *Impulsive*. The next new grouping was termed 'Aloof' and included *Individualistic* and *Cool*.

Supportive

Includes *Conventional*, *Persistent*, *Affectionate*

Expert consensus was 61.26% (SD31.82) for this grouping. Participants felt that this group was superfluous and that terms within it could be placed into other existing groupings. As such *Conventional* was moved to the 'Steady' group, *Persistent* was moved to the 'Determined' group and *Affectionate* was moved to the 'Affiliative' group.

Hesitant

Includes *Vulnerable*, *Protective* and *Curious*

Expert consensus was 56.13% (SD31.60) for this grouping. Participants felt that this category should be renamed 'Careful' or 'Self-protective'. It was suggested that this group could be removed entirely, and the remaining traits be reallocated to other groups. *Vulnerable* was moved to the 'Nervous' group and *Protective* was moved to the 'Affiliative' group. *Curious* was reallocated to the new 'Investigatory' group.

Goofy

Includes *Clumsy* and *Distractible*.

Expert consensus was 58.10% (SD36.21) for this grouping. Participants felt that 'Goofy' was not an appropriate term and the alternative terms 'Uncoordinated', 'awkward', 'inelegant' and 'inattentive' were suggested. 'Inattentive' was selected as the new group name.

Intellectual

Includes *Intelligent, Innovative, Autistic and Inventive*.

Expert consensus was 60.86% (SD32.43) for this grouping. Participants felt that the group should be renamed either 'Investigatory' or 'Experimental' and that *Curious* should be moved from the 'Hesitant' group into this group and *Autistic* should be removed altogether. 'Investigatory' was selected as the new group name.

Forty-two personality traits across 12 groups were carried forward to part three (Table 4.6).

Table 4.6 Version two personality groupings utilised in the owner survey and presented to experts in Round three of the Delphi.

Group name	Traits included
EXUBERANT	<i>Excitable, Playful, Active</i>
NERVOUS	<i>Cautious, Erratic, Fearful, Acquiescent, Vulnerable, Sensitive</i>
UNRESPONSIVE	<i>Depressed, Unemotional, Quitting</i>
STEADY	<i>Predictable, Stable, Conventional</i>
BOLD	<i>Assertive, Confrontational, Independent</i>
AFFILIATIVE	<i>Friendly, Helpful, Sociable, Sympathetic, Gentle, Affectionate, Protective</i>
HOSTILE	<i>Irritable, Defiant, Unperceptive, Bullying</i>
SHY	<i>Timid, Solitary</i>
DETERMINED	<i>Decisive, Impulsive, Persistent</i>
ALOOF	<i>Individualistic, Cool</i>
INATTENTIVE	<i>Clumsy, Distractible</i>
INVESTIGATORY	<i>Intelligent, Inventive, Innovative, Curious</i>

4.3.3 Part three – October – March 2022

To establish if the new trait groupings were still highly correlated, that is grouped together within PCA analysis, they were placed into an owner questionnaire and the equine-owning public were invited to complete the survey for each of their equines. The new trait groupings were also provided to experts within Round three of the Delphi for them to rate their level of agreement and provide qualitative feedback.

Demographic data

The survey recruited responses for 117 equines by 136 independent raters during the first distribution, 11 responses were removed due to being incomplete leaving 125 raters and 106 equines in the sample which consisted of 55 geldings, 48 mares, 2 stallions and 1 colt. The horses ranged in age from 1 – 30 years, with a median age of 13 years. The sample therefore consisted of 4452 Likert scale values for the first round, with an additional 672 data points contributed by 16 participants who re-rated their horses 3 months later in Round 2 contributing 5124 data points in total.

Human participants demographically represented a wide cross-section of the UK and kept or cared for equines across a wide spectrum of yard types from competition yards, livery yards, breeding establishments and private yards.

Inter-rater reliability

Fourteen horses had all 42 traits scored by two or more raters. Inter-rater reliability varied across horses and raters but was generally good with significant levels of agreement (Table 4.7).

Table 4.7 One-way random effects Intraclass Correlation Coefficient (ICC (1, 1) for inter-rater agreement across 14 horses and 2, 3 or 4 raters, significant agreement highlighted in bold and good or higher agreement in italics.

Horse number	ICC (1, <i>k</i>) and <i>p</i> value	Number of raters
78	<i>0.684, <0.001</i>	2
94	<i>0.630, <0.001</i>	3
95	0.362, 0.04	3
96	<i>0.766, <0.001</i>	2
98	<i>0.486, 0.017</i>	2
99	0.249, .179	2
100	0.355, 0.08	2
101	<i>0.794, <0.001</i>	4
102	<i>0.687, <0.001</i>	2
103	<i>0.461, 0.025</i>	2
104	<i>0.777, <0.001</i>	3
105	<i>0.420, 0.04</i>	2
106	<i>0.688, <0.001</i>	2
107	<i>0.694, <0.001</i>	2

Intra-rater reliability

Intra-rater reliability across all traits, scores and raters was excellent (two-way mixed effects model for absolute agreement, average measures, 0.906, $p = <0.001$) for the horses ($n=16$) who were scored twice (3 months apart) by their raters.

Trait groupings

Principal component analysis identified 28 components with eigenvalues above 1 using the Kaiser criterion, that together explained 94.2% of the variation, the first twelve components made up 72% of the variation in the data. (Figure 4.8). The first two components had eigenvalues with values (28.224 and 20.707 respectively) that were more than double the next two components (13.062 and 10.301 respectively); therefore, the first two components were taken forward for further evaluation.

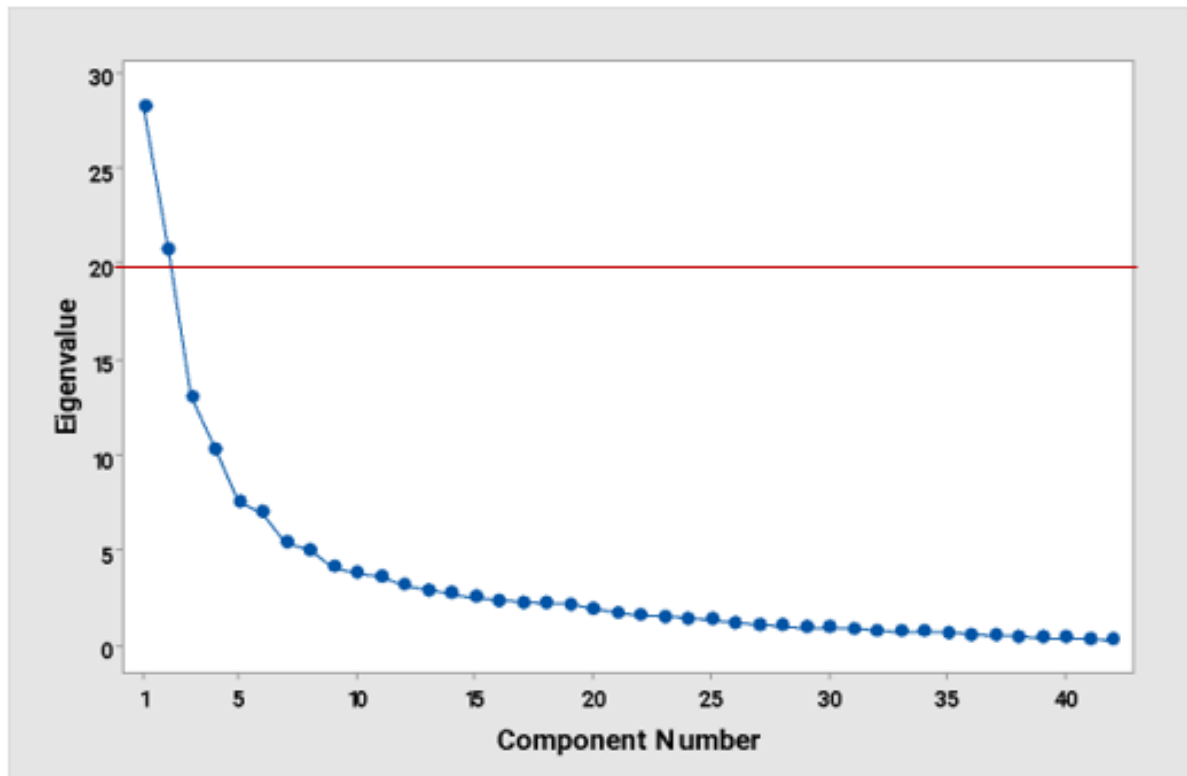


Figure 4.8 Principal components scree plot identifying 2 PCs with eigenvalues greater than 20 (20 signified by the red line).

The loading plot identified that traits within the 'Exuberant' group had strong influence on the first component and weak influence on the second component; the traits within group 'Nervous' had weak influences on the first component and stronger influence (with some negative loadings) on the second component. The group 'Unresponsive' had neutral influence on both the first and second components. Group 'Steady' had stronger influence on the second component whilst the groups 'Bold', 'Inattentive' and 'Aloof' had mild to moderate positive influence on the first component and neutral to mild negative influence on the second component. Group 'Affiliative' had significant positive influence on both the first and second components while group 'Unfriendly' had weak influence on the first component and strong influence on the second component. Group 'Determined' had variable positive influence on the first and variable negative influence on the second components. Group 'Shy' had neutral loadings on the second component and negative loadings on the first component and Group 'Investigatory' had positive influence on the first component and neutral or negative influence on the second component (Figure 4.9). This indicates that groups containing traits which are considered low arousal and positive valence generally share

positive influence across components 1 and 2. Groups containing traits considered higher in arousal and more negative in valence generally had negative loadings on the second component and mild to moderate loadings on the first component. Groups containing traits with lower arousal and negative valence had more neutral loadings on each component.

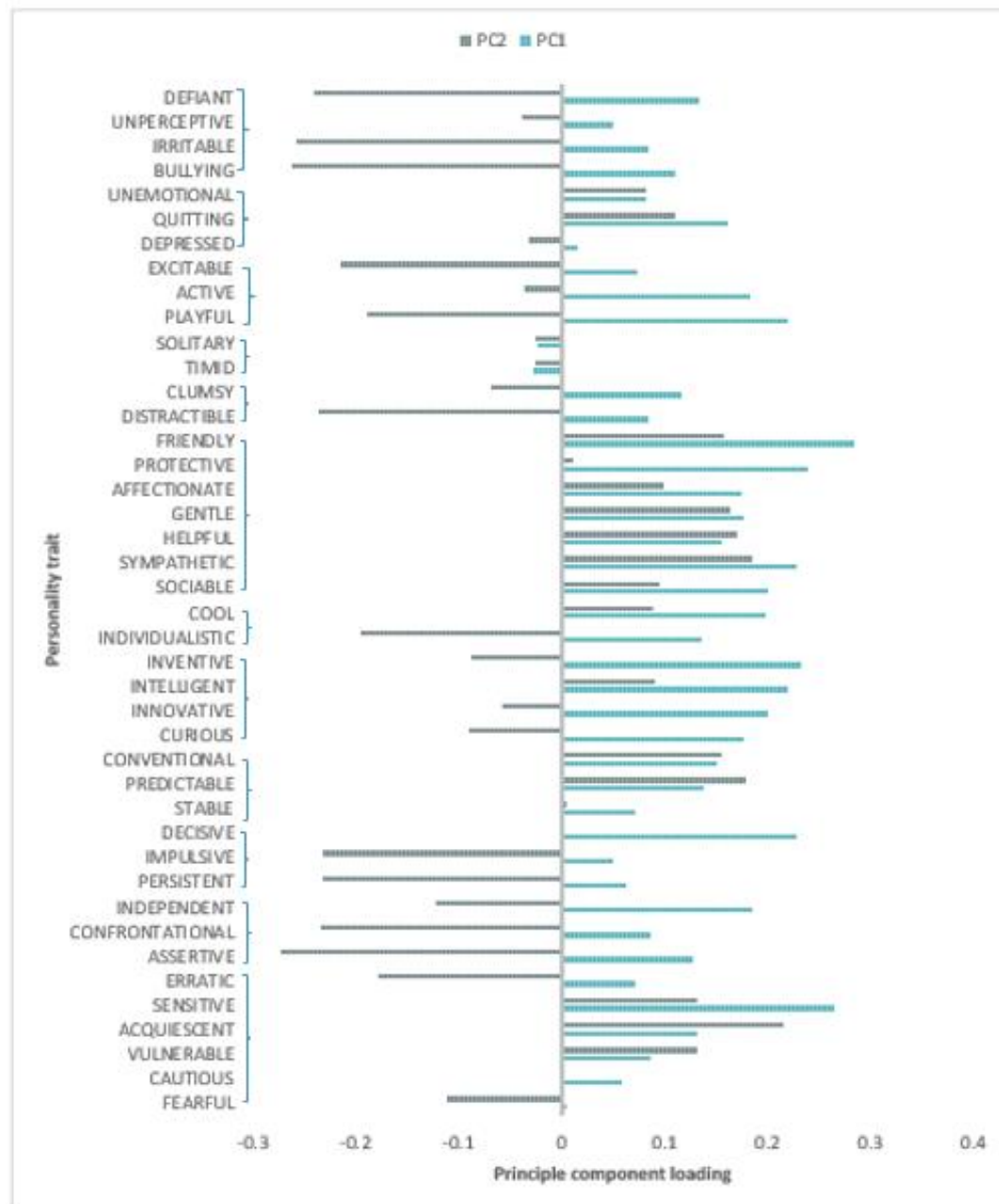


Figure 4.9 Trait influences on the first and second components. Groupings denoted by '}' are the original groupings identified in Part two.

Item analysis to identify internal consistency was performed on the survey data beginning with the personality groupings identified by the expert Delphi group. Backwards and forward

stepwise analysis was then undertaken to identify the most highly correlated groupings. Both the initial internal consistency and consistency after stepwise analysis can be seen in Table 4.8.

Table 4.8 Expert derived group names and traits from the Delphi consultation with traits and groups removed following backwards stepwise analysis (strikethrough) and traits added (bold) after forward stepwise comparison analysis of the owner-rater generated data. Expert agreement level from Round three of the Delphi (consensus level reached highlighted in bold). Group level internal consistency is reported via Cronbach's Alpha value (excellent agreement in bold, very good agreement in italics and moderate agreement in standard text).

Group name	Second round expert derived traits	Internal consistency (Cronbach's Alpha)	Third round Delphi agreement level (%)	Final traits following stepwise comparison	Internal consistency (Cronbach's Alpha)
EXUBERANT	<i>Excitable, Playful, Active.</i>	<i>0.5906</i>	92.31	<i>Excitable, Playful, Active, Distractible, Impulsive.</i>	0.7260
NERVOUS	<i>Cautious, Erratic, Fearful, Acquiescent, Vulnerable, Sensitive.</i>	<i>0.5982</i>	76.92	<i>Cautious, ~~Erratic,~~ Fearful, Acquiescent, Vulnerable, Sensitive, Timid.</i>	<i>0.6900</i>
UNRESPONSIVE	<i>Depressed, Unemotional, Quitting.</i>	<i>0.2856</i>	64.62	<i>Depressed, Unemotional, ~~Quitting,~~ Solitary, Unperceptive.</i>	<i>0.5421</i>
STEADY	<i>Predictable, Stable, Conventional.</i>	<i>0.4124</i>	92.31	<i>Predictable, ~~Stable,~~ Conventional, Cool.</i>	<i>0.6729</i>

BOLD	<i>Assertive, Confrontational, Independent.</i>	0.5997	76.92	<i>Assertive, ~~Confrontational,~~ Independent, Decisive, Persistent, Individualistic.</i>	0.6746
AFFILIATIVE	<i>Friendly, Helpful, Sociable, Sympathetic, Gentle, Affectionate, Protective.</i>	0.8400	89.23	<i>Friendly, Helpful, Sociable, Sympathetic, Gentle, Affectionate, ~~Protective.~~</i>	0.8464
UNFRIENDLY	<i>Irritable, Defiant, Unperceptive, Bullying.</i>	0.6617	76.92	<i>Irritable, Defiant, ~~Unperceptive,~~ Bullying, Confrontational.</i>	0.8117
SHY	<i>Timid, Solitary.</i>	-0.1494	86.15	~~<i>Timid, Solitary</i>~~	Group removed
DETERMINED	<i>Decisive, Impulsive, Persistent.</i>	0.1940	78.46	~~<i>Decisive, Impulsive, Persistent</i>~~	Group removed
ALOOF	<i>Individualistic, Cool.</i>	0.1504	83.08	~~<i>Individualistic and Cool</i>~~	Group removed
INATTENTIVE	<i>Clumsy, Distractible.</i>	0.4797	87.69	~~<i>Clumsy, Distractible</i>~~	Group removed
INVESTIGATORY	<i>Intelligent, Inventive, Innovative, Curious.</i>	0.7047	81.54	<i>Intelligent, Inventive, Innovative, Curious.</i>	0.7047

Expert qualitative feedback on the new groupings and subsequent final item analysis

In Round three of the Delphi consultation the experts gave feedback on any items that they felt required adjustment. The feedback was analysed thematically and is provided in Table 4.9 below.

Table 4.9 Expert feedback on the second-round traits and trait groupings.

Trait (<i>italics</i>) or group (capitalised)	Number of expert comments	Feedback theme (n similar comments)
NERVOUS	9	Remove acquiescent (7). Remove erratic. Change group name to sensitive or cautious (3).
INVESTIGATORY	2	Remove intelligent. Add playful.
BOLD	11	Remove confrontational (9), bullying defiant and unperceptive. Change confrontational to confident (2). Could describe a curious, investigative horse.
HOSTILE	3	Add confrontational. Remove unperceptive. Change group name (2).
ALOOF	4	Remove bullying. Add solitary. Change group name. How do we define cool?
DETERMINED	11	Remove impulsive (11).
AFFILIATIVE	3	Remove helpful. Remove sympathetic, affectionate, gentle and protective (2).
<i>Clumsy</i>	3	Inappropriate trait name, likely veterinary in origin.
<i>Distractible</i>	2	Remove trait.

Item analysis for Nervous

Further item analysis applying the expert comments confirmed that when *Acquiescent* was removed from the group 'Nervous' the Cronbach's alpha value dropped from $\alpha = 0.6900$ to $\alpha = 0.6473$, therefore this trait remained. The group name was not altered to avoid tautology as highlighted by experts in Round two; both suggested alternative names (cautious and sensitive) were already traits within the grouping. *Erratic* had been previously removed following the item analysis using the survey results.

Item analysis for Investigatory

When *Intelligent* was removed and *Playful* added, the Cronbach's alpha value rose from $\alpha = 0.7047$ to $\alpha = 0.7998$. However, when *Playful* was moved from group 'Exuberant' the Cronbach's alpha value dropped from $\alpha = 0.7260$ to $\alpha = 0.6697$ therefore *Playful* remained in the 'Exuberant' group. The final grouping of *Inventive*, *Innovative*, and *Curious* had a Cronbach's alpha score of $\alpha = 0.7827$.

Item analysis for Bold

Confrontational had previously been removed from this group following item analysis using the survey results and therefore no further analysis of this group was required.

Item analysis for Hostile

Confrontational had previously been added and *Unperceptive* removed from this group following item analysis using the survey results and therefore no further analysis of this group was required. The group name was changed to 'Unfriendly'.

Item analysis for Affiliative

Protective had previously been removed from this group following item analysis using the survey results and therefore no further analysis of this group was required.

Aloof, Determined, and Inattentive

The 'Aloof', 'Inattentive' and 'Determined' groups had previously been removed following item analysis using the survey results therefore no further changes were required.

Comparison of final traits and groups with the Five Factor Model

These final eight overarching personality traits and their associated new trait facets, named the Equine Personality Eight-factor Model (EPEM), is proposed as a potentially reliable

method to assess equine personality, and can be considered in relation to the Big Five (NEO PI-R) model proposed by Costa and McCrae (1989) (Table 4.10). Within our proposed model all facets of the NEO PI-R are accounted for with the trait facets of our model spanning several different NEO PI-R categories.

Table 4.10 Comparison of the final proposed Equine personality eight-factor model (EPEM) in this study with the Big Five (NEO PI-R) model

EPEM group trait	Trait facets	Big Five factor model	Trait facets (aligned traits highlighted in bold)
EXUBERANT	<i>Excitable, Playful, Active,</i>	EXTROVERSION	Warmth, Gregariousness, Assertiveness, Activity, Excitement-seeking, Positive emotions
	<i>Distractible, Impulsive</i>	NEUROTICISM	<i>Anxiety, Hostility, Depression, Self-consciousness, Impulsiveness, Vulnerability</i>
NERVOUS	<i>Cautious, Fearful, Vulnerable,</i>	NEUROTICISM	Anxiety, Hostility, Depression, Self-consciousness, Impulsiveness, Vulnerability
	<i>Acquiescent, Sensitive, Timid</i>	AGREEABLENESS	<i>Trust, Straightforwardness, Altruism, Compliance, Modesty, Tendermindedness</i>
UNRESPONSIVE	<i>Depressed, Unemotional, Solitary, Unperceptive</i>	NEUROTICISM	<i>Anxiety, Hostility, Depression, Self-consciousness, Impulsiveness, Vulnerability</i>
STEADY	<i>Conventional, Cool,</i>	AGREEABLENESS	Trust, Straightforwardness, Altruism, Compliance, Modesty, Tendermindedness
	<i>Predictable</i>	CONSCIENTIOUSNESS	<i>Competence, Order, Dutifulness, Achievement striving, Self-discipline, Deliberation</i>
BOLD	<i>Assertive, Independent,</i>	EXTROVERSION	<i>Warmth, Gregariousness, Assertiveness, Activity, Excitement-seeking, Positive emotions</i>
	<i>Decisive, Persistent, Individualistic</i>	CONSCIENTIOUSNESS	Competence, Order, Dutifulness, Achievement striving, Self-discipline, Deliberation
AFFILIATIVE	<i>Friendly, Helpful, Sociable,</i>	EXTROVERSION	Warmth, Gregariousness, Assertiveness, Activity, Excitement-seeking, Positive emotions

	<i>Sympathetic, Gentle, Affectionate</i>	AGREEABLENESS	<i>Trust, Straightforwardness, Altruism, Compliance, Modesty, Tendermindedness</i>
		OPENNESS TO EXPERIENCE	<i>Fantasy, Aesthetics, Feelings, Actions, Ideas, Values</i>
UNFRIENDLY	<i>Irritable, Defiant, Bullying, Confrontational</i>	NEUROTICISM	<i>Anxiety, Hostility, Depression, Self-consciousness, Impulsiveness, Vulnerability</i>
INVESTIGATORY	<i>Inventive, Innovative, Curious</i>	OPENNESS TO EXPERIENCE	<i>Fantasy, Aesthetics, Feelings, Actions, Ideas, Values</i>
		EXTROVERSION	<i>Warmth, Gregariousness, Assertiveness, Activity, Excitement-seeking, Positive emotions</i>
		CONSCIENTIOUSNESS	<i>Competence, Order, Dutifulness, Achievement striving, Self-discipline, Deliberation</i>

4.3.4 Part four – September 2022 – June 2023

Development of a short-form (abridged) questionnaire

As per the method of Smith *et al.* (2000) *a priori* examination of the reliability and validity of the short form in use was conducted on a small subset of the sample. The short form was released to 4 raters across 3 horses with moderate to good agreement between raters (ICC (2,k) = 0.742, $p < 0.001$).

Agreement between scores using the full questionnaire and the short form (abridged questionnaire) was good overall (ICC (2,k) = 0.697, $p < 0.01$) with some trait groups showing greater levels of agreement than others (Table 4.11).

Table 4.11 Group level agreement between full questionnaire scores and short form (abridged scores). Significant associations in bold.

Group	ICC (2,k)	P
EXUBERANT	0.784	0.01
NERVOUS	0.949	<0.01
UNRESPONSIVE	0.802	0.017
STEADY	0.516	0.190
BOLD	0.0	1.0
AFFILIATIVE	0.847	0.08
UNFRIENDLY	0.487	0.221
INVESTIGATORY	0.359	0.383

As such it was decided to further test the abridged form with the inclusion of trait names within groups for short screening and the full questionnaire for greater depth of analysis, comparing each method. Twenty further equines were scored using the new abridged questionnaire (with trait names included) and the full questionnaire with results showing excellent agreement between methods across all personality groups (Table 4.12).

Table 4.12 Agreement between abridged and unabridged questionnaire scores in twenty horses.

Group	ICC (2,k)	<i>p</i>
EXUBERANT	0.978	<0.001
NERVOUS	0.959	<0.001
UNRESPONSIVE	0.983	<0.001
STEADY	0.984	<0.001
BOLD	0.952	<0.001
AFFILIATIVE	0.966	<0.001
UNFRIENDLY	0.948	<0.001
INVESTIGATORY	0.981	<0.001

Construct validity

Divergent construct validity was supported with data demonstrating negative relationships between groupings with theoretically opposing adjectives as shown in Table 4.13 below, indicating good initial construct validity.

Table 4.13 Construct validity tested via Spearman's rank between presumed opposing personality type groupings.

Groupings tested	Spearman's rank correlation	P value
Exuberant vs Unresponsive	-0.796	<0.001
Exuberant vs Steady	-0.191	0.003
Nervous vs Bold	-0.305	<0.001
Affiliative vs Unfriendly	-0.322	<0.001
Nervous vs Investigatory	-0.148	0.020
Unresponsive vs Investigatory	-0.123	0.050

4.4 Discussion

This study aimed to develop a reliable full and abridged personality trait assessment for equines building on the previously validated hominoid personality questionnaire (Weiss *et al.*, 2011). The results have demonstrated an apparently reliable set of questions for inclusion within the equine personality questionnaire that can be reduced to an abridged eight-factor scale, both of which encompass all the facets of the established Big Five Factor personality model (Tupes and Christal, 1961; Costa and McCrae, 1989).

This multi-staged approach to the development of the proposed Equine Personality Eight-factor Model (EPEM) facilitated the inclusion of expert consensus opinion within an empirical framework to successively refine and undertake initial validation of a reliable model with which to assess equine personality. Beginning with a previously validated Hominoid personality questionnaire (Robinson *et al.*, 2018b; Weiss *et al.*, 2011; Weiss *et al.*, 2009) afforded a well-substantiated starting point for equine application which is considered valid due to the evidence for the similarities in brain structure, cognitive abilities, and lifespan between equine and hominoid species (Johnson *et al.*, 2019; Matsuzawa, 2017; Matsuzawa, 2010).

Our approach followed the methodology of Robinson *et al.* (2018b) who validated the use of a full 54-item Hominoid Personality Questionnaire (HPQ) and an abridged expert-consensus version to expediate assessment processes. To this end our research has produced a full 37-item EPEM questionnaire with an abridged eight-dimension version (EPEM-R) (Appendix D) assessed for reliability and attenuation through statistical testing following the model of Smith *et al.* (2000).

4.4.1 The assessment instrument

Part one

Whilst the original instrument, consisting of 54 items, showed a high level of agreement between the two raters across all equines it did not show good levels of agreement between raters within individual traits. Three traits showed no agreement, 13 showed slight agreement, 25 showed fair agreement and 13 showed moderate agreement with no items showing substantial agreement meaning that in its original format the instrument did not appear reliable for equine use. However, agreement between raters within horses was higher (16 of 35 horses showing moderate or substantial agreement between raters) than within

individual traits, indicating that some traits may have been more reliable than others when measured, as identified by Mirkó *et al.* (2013) in their evaluation of subjective personality ratings in dogs.

Factors which may affect rating reliability include variation in the subjective assessment of the raters (Mirkó *et al.*, 2013) and expression of personality traits within individual equines, with some horses appearing easier to assess than others (Dutton, 2008; Bolig *et al.*, 1992). In addition, rater interpretation of personality trait meanings and the level of experience each rater has with the individual horses across a range of situations where traits may be expressed likely alter ratings (Martau *et al.*, 1985). Prior research has demonstrated that traits which express more subtle social or cognitive abilities are less reliably assessed than others (Dutton, 2008), a factor that was highlighted in this study by the Delphi experts and indicated by the variability in agreement between raters on specific traits. Prior experiences and the current environment of the animal may also affect expression of personality traits (Martin, 2005) as a result of their individual levels of valence and arousal at the time they are tested.

The groupings identified via PCA in Part One were carefully examined for internal consistency prior to presentation to the expert Delphi group in part two of the study. This ensured that statistically valid groups were presented for consideration and refinement as per the methods of Peterson *et al.* (2007) whereby a Cronbach α coefficient of 0.70 or greater was deemed acceptable.

Part two

Participants agreed on the definition of personality but identified that whilst the traits presented were clear and largely present in equines, they were not unanimous in their view regarding the reliability of measuring each trait in equines. As such it was important to establish the individual reliability of identification of each trait in a cohort of equines across numerous observers, a facet that was studied in part three of the study.

Within Part two participants gave significant levels of feedback in relation to the traits included within personality groupings as well as the trait and group names suggested, with many common themes identified across their responses. This allowed for re-examination of the groupings and terminology used, prior to further exploration of the reliability of the questionnaire in part three. This process also allowed for trimming of the number of traits

and groupings within the questionnaire which aided the development of greater brevity which is shown to increase completion of such measures by participants (Liu and Wronski, 2018).

Traits rated highly for the clarity of the trait, the presence of the trait in equines, and the reliability of identifying the trait in equines were typically those which appear to have relative ease of identification in equines and appear in other equine personality assessments such as that presented by Ijichi *et al.* (2014), further tested by Jolivald *et al.* (2022) and included traits such as sociable, fearful, playful, aggressive, excitable and inquisitive. More nuanced traits, or those deemed by the expert group to be more anthropomorphic such as cool, conventional, reckless, manipulative, disorganised and thoughtless were rated much lower in terms of their presence, reliability, and clarity.

This process of gaining expert agreement prior to final questionnaire testing significantly enhanced the preparation of the questionnaire and added important experiential consideration to the groupings presented to later lay-participants.

Part three

Within this study temporal ratings of individuals were stable over the short-term, with excellent agreement in intra-rater test-retest scores (ICC two-way mixed effects, absolute agreement 0.906, $p = <0.001$), as also reported by Uher *et al.* (2008); but were limited to a 3-month test-retest protocol which could be usefully extended to retest a further 6 and 12 months after initial testing. However previous studies in humans (Roberts *et al.*, 2006) and non-human animals (Weiss *et al.*, 2007; King *et al.*, 2005) have identified that trait ratings can change as animals mature with higher ratings for exuberance, playfulness and curiosity often seen in younger animals, meaning that age or sex-related social changes may reduce temporal reliability of trait assessment in these cohorts.

Real world application of the personality questionnaire across 106 equines demonstrated relatively good levels of agreement between raters even in horses where there were 3 or 4 raters for the individual, with 8 of the 14 horses rated by multiple raters showing good or excellent agreement and only 3 of the remaining 6 horses showing agreement below moderate. Agreement between raters may have been affected by individual equine or human differences during rating related to the behavioural repertoire of the horse-human dyad

and/or rater understanding of equine behaviour specific to the individuals, rather than being related to the construct validity of the questionnaire. This is likely to be the case because this study employed converging forms of empirical evidence to mitigate against elements of rater bias by testing for (a) inter-rater reliability, (b) inter-item consistency, (c) cross-situational consistency and (d) cross-temporal stability as recommended by Figueredo *et al.* (1995). Further, the study design facilitated critical design features recommended by Kenrick and Funder (1988) for personality assessment including (a) using raters who are thoroughly familiar with the individuals being rated (set in our inclusion criteria), (b) rating based on multiple behavioural observations over time (set in our inclusion criteria), (c) rating by multiple observers (study design), (d) rating dimensions that are publicly observable (agreed through expert consensus) and (e) rating behaviours relevant to the dimension in question (agreed through expert consensus).

The collected equine personality data enabled further revision of the personality questionnaire following expert consensus through its application to a large mixed cohort of equines to allow for greater validity in the generalisability of the questionnaire. Whilst inter-item analysis (internal consistency) was good for the personality groupings selected by the experts when applied to real world data, subtle nuances in the distribution of traits within individuals became apparent in the data collected and enabled further analysis of internal consistency between traits to develop reliable final grouping of traits.

Evidence for the strength of the instrument and the expertise of the expert group was highlighted by the agreement between the inter-item analysis of the data generated from the 106 equines in part three of the study and views of the expert group in round three of the consultation where items suggested by experts to be removed or added to specific trait groups were also identified in the independent statistical analysis.

Initial mapping of the EPEM groups and traits to the FFM (Costa and McCrae, 1989), a well-established personality model, has indicated good coverage of the FFM traits and trait facets within our personality model although further modelling utilising additional data generated for equines scored using both the FFM and EPEM models would help to validate this aspect further. The EPEM, containing a greater number of factors (8 versus 5) than the FFM, could be shown to be more effective in summarising individual differences in personality; a limitation specified by dissenters of the FFM is that five categories are too few to meaningfully

allow for this differentiation and allow only for global characterisation (McCrae and John, 1992).

Part four

Evidence within the literature shows that the use of short-form or abridged questionnaires can provide valid outcomes providing that sufficient information is provided to the scorer to make valid judgements on trait items and that prior analysis of the independence of scored items has been established (Pelle *et al.*, 2013). Within this study short-form development was carefully scrutinised at each stage to determine the validity and reliability of scores awarded using the short-form compared with the full questionnaire. Within the short-form up to six personality traits were assessed via one questionnaire item, which, when evaluating the agreement between raters scoring their equines using both the short-form and full questionnaire, appeared to reduce reliability, indicating that the assessment of many traits within one questionnaire item may be less valid. In order to improve validity, the short-form was altered after the first analysis of data to give scorers further information on trait inclusion within trait groups in order to improve the validity of judgements. The result of this iterative design process demonstrated very high levels of agreement between short and long form responses albeit the final cohort tested was relatively small ($n = 20$). Our findings are supported by other studies which have evaluated the validity of the use of short forms compared with long forms including in subjective assessment such as quality of life in dogs (Reid *et al.*, 2013), attitudes to animals (Bradley *et al.*, 2020) and pain in horses (Howard *et al.*, 2024). Further validation of the use of the short form is warranted to establish its stability and reliability across raters, horses, and temporal scales.

4.4.2 The sample

The equine sample

Within this study 164 equines were evaluated during the iterative phases of the questionnaire design. The horses included in the sample were representative of the types of horses found within the general population in terms of the range of ages, breeds, sexes, backgrounds, and types of work (leisure, unaffiliated competition, elite competition) undertaken. The sample size, whilst small in the context of the number of equines in the UK, is higher than that utilised by Jolivald *et al.* (2022), Lansade *et al.* (2016) and Lloyd *et al.* (2007) and similar to the sample utilised by Hausberger *et al.* (2011) in similar personality-based equine studies undertaken in

recent years; but is still likely underpowered with a retrospective sample size calculation suggesting that a sample size of 384 equines would be more representative of the UK population. Further studies will address this sample size weakness through the collection of additional personality data from a greater number of equines.

The human sample

The human population within this study can be divided into two components a) the Delphi experts and b) the owner-raters of equines.

Ninety-three Delphi experts contributed to the discussion and consensus agreement on the equine personality traits and trait groupings. Nearly 50% of these experts self-reported to be either actively engaged in personality research or studied and monitored new research into animal personality on a regular basis suggesting that these experts could be considered reliable to give valid opinions surrounding equine personality. This supposition is supported by the independent agreement, within Part three of the study, between the item analysis results and the Delphi participants' views on traits which should be included and excluded from groupings.

In total 162 raters scored personality traits for equines in their care, this large cohort of raters allowed for a good level of inter-rater analysis throughout the study which strengthens the results obtained. In addition, our raters were, overall, reliable from an intra-rater perspective over temporal scales indicating that they provided valid subjective opinions on the personality traits in their animals as has been demonstrated in prior reviews of observer ratings of animal characteristics (Meagher, 2009; Gmel *et al.*, 2020; Randle, 2015). All raters scored the horses independently of each other throughout the study which ensured that raters were not influenced by the ratings of others. Inter-rater reliability within specific personality traits was seen to more variable than for overall personality groupings, this is likely to be caused by humans finding some traits more difficult to identify and score than others, or having a preference for some traits over others, factors which have been quite widely discussed within the personality literature (Funder and Dobroth, 1987; Ready *et al.*, 2000; Waiyavutti *et al.*, 2012; Iwanicki and Lehmann, 2015).

4.4.3 EPEM in the context of other personality models

Previous research has identified the relative alignment between human and non-human animal personality traits and the temporal stability of these traits over time (Chamove *et al.*, 1972). Whilst early non-human animal personality models tended to be based around the Emotions Profile Index (EPI) (Kellerman and Plutchik, 1968; Plutchik and Kellerman, 1974) with forced choices between 66 pairs of personality trait combinations, later research developed personality assessments based on Likert scales with behaviourally defined lists of adjectives, as used within this study, to allow for a more nuanced personality assessment compared with the use of the blunter forced-choice methodology (Stevenson-Hinde and Zunz, 1978). Developments in animal personality assessment have led to an interweaving of psychological and ethological approaches to assessment to allow for reliable rating of cross-situational consistency of measures (Stevenson-Hinde and Hinde, 2011).

Previously validated statistical methodologies to condense questionnaire items were rigorously followed within this study. The use of PCA to condense items is well documented (Stevenson-Hinde and Hinde, 2011) and allows for reliable rapid assessment of personality traits across a range of taxa. Further consideration could be given to the relative loading of items on the main components, specifically in relation to their relative positive and negative loadings across components to develop a model of the intersection of personality traits. Considering the literature surrounding the relatively arbitrary unidirectional categorisation of personality used within the Five factor NEO PI-R model and the spanning of multiple categories of this by our groupings on the proposed EPEM it may be considered that there would be a greater level of content validity to our model when following the more bi-dimensional circumplex approach of Peabody and Goldberg (1989) to consider the intersections of personality traits within an individual, following the work of Hofstee *et al.* (1992) and Johnson and Ostendorf (1993). To satisfy this further analysis and refinement could be undertaken in subsequent studies to consider the intersections of traits facets, plotted with reference to the Big Five NEO PI-R factors with positive and negative signalment for high and low scores respectively following the methodology of Peabody and Goldberg (1989) where Roman numerals and plus and minus signs are attributed to two intersecting factors in a two-way table to depict all of the possible nuanced trait combinations and strength of trait presence across individuals.

Final testing of both the EPEM and EPEM-R was conducted on a relatively small sample of twenty equines and therefore the questionnaires would benefit from further validity and reliability testing across a greater sample size and using a wider number of raters who also rate the animals over longer time periods of three, six and 12 months to further validate the temporal validity of the traits and their groupings.

4.5 Limitations

This study is limited by the sample size of 164 equines which suggests that a cautious approach should be taken to the application of the EPEM prior to further validation being undertaken. Additionally, whilst statistical evaluation of the robustness of the short-form compared with the full questionnaire yielded excellent agreement, the sample size was very low at just 20 horses indicating that this aspect of the survey development requires much greater validation across both a larger sample size and over extended time periods.

Further evaluation of the temporal stability of the EPEM is required to assess its validity over longer timeframes of six and twelve months at a minimum. This should be undertaken on mature but not aged equines who are likely to be relatively stable in their personalities (Bleidorn *et al.*, 2022; Gosling and John, 1999) to provide a useful frame of reference for further studies which could then look to investigate temporal stability of equine personality in younger horses as they mature and in mature horses as they age. However it is important to consider the impact of life events on the differential stability of personality traits, a factor that has been studied extensively in humans (Vaidya *et al.*, 2002) but is not as well reported in equines (Hausberger *et al.*, 2004).

Whilst a robust methodology was employed within this study, using both Delphi expert opinions and statistical analysis of real-world data, it is feasible that human bias towards individual personality traits may have affected the grouping of traits to some degree. The potential for fundamental attribution error, where people wrongly attribute personality traits when asked to make inferences regarding an animal's underlying disposition (Haselton *et al.*, 2015), is a potential cognitive bias that could also have impaired the judgements of the animal raters and subsequently influenced the statistical analysis.

Brief consideration has been given to the relatedness of the EPEM and FFM in this study through the mapping of traits from the EPEM to the FFM. However, further empirical testing

of this mapping is required to validate the assumptions made within this study through statistical analysis of the results of screening equine personality using both models.

4.6 Conclusion

This study has utilised the strengths of two types of methodological design (Delphi consultation and survey-based instruments) to iteratively evaluate and test the development of a previously validated hominoid personality questionnaire for use in equines. The advantages of using a previously validated animal model allowed for this study to focus on fine-tuning the questionnaire for equine use which was contributed to equally usefully by the expert group consensus and owner-rater assessments.

The development of both long and short form versions of the questionnaire allows for future use in a wide variety of situations where each form may be judged to be most applicable and is suggested to enable rapid evaluation of equine personality alongside more detailed analysis using the long form.

Whilst this study has demonstrated very good levels of inter and intra-rater reliability as well as short-term temporal stability, both the short and long forms require further validation across a greater sample size and timescale. Initial data suggests that the EPEM may be a valid and reliable model for measuring equine personality however it should also be considered within the context of other newly formed equine personality assessments which were not available at the time the study commenced.

The importance of the effect of personality on affective state is widely recognised within the literature and therefore this personality model provides a means of assessing this aspect within wider assessment of equine emotional and physiological welfare with the aim of enabling more nuanced analysis of welfare assessments and the tailoring of management strategies to meet the needs of the individual equine.

Chapter 5 Pilot study investigating the effect of positive stimuli application on behavioural and physiological measurement parameters in equines

5.1 Introduction

Emotional welfare in equines has become an integral facet of overall welfare over the last twenty years, with the mental state now considered one of the essential five domains of welfare (Mellor *et al.*, 2020). Specifically, the presence and measurement of positive affective state (good emotional welfare) is an important area of research that ensures high levels of emotional welfare can be provided and measured within individuals and groups (Mellor, 2012). When considering emotional welfare both valence and arousal are important factors (Starling *et al.*, 2013), a concept represented by Mendl *et al.* (2010) following on from the work of Russell and Barrett (1999) and Burgdorf and Panksepp (2006). In addition to this, Cabanac (2002) suggests that both longer-term and shorter-term affective states (i.e. moods) and the surgency (or power) of these are important in emotional welfare and the ability to recognise and measure these allows us a greater understanding the level of emotional welfare of the individual both in the moment and over time.

This chapter reports the results of a pilot study that was undertaken to trial and refine the methodology for the final study (Chapter 6) and to evaluate the implementation and analysis of some of the expert proposed methods (Chapter 3 - Delphi study) to induce and measure positive affective state in equines.

5.1.1 Emotional processing and affective state

Emotional processes include a combination of motor-expressive, sensory-perceptual, autonomic-hormonal, cognitive-attentional and affective-feeling components with the interplay between these leading to the varying inhibitory and disinhibitory effects on physiology, behavioural outputs of the animal and aligned affective feelings (Mendl *et al.*, 2010; Cabanac, 2002). Processing of internal and external stimuli may lead to negative, positive (Hausberger *et al.*, 2019; Trösch *et al.*, 2020) or even neutral affective state (Gasper *et al.*, 2019; Samson *et al.*, 2016). Positive and negative affect are the most widely used of the three terms in the equine field and studies have evaluated these aspects in relation to responses human-animal interactions (Hausberger *et al.*, 2019; Trösch *et al.*, 2020), responses to grooming (Lansade *et al.*, 2018) and conspecific separation and reunion (Briefer *et al.*,

2017a). Emotional processing and affective state are further differentiated within individuals with unique personality traits modulating emotional valence, arousal and intensity, as well as the effect of these aspects on cognitive processing and performance (Fortin *et al.*, 2018; Goschke and Bolte, 2014; Ijichi *et al.*, 2014; Lansade *et al.*, 2017).

5.1.2 Positive stimuli for equines

Food stimuli

Food stimuli have been demonstrated to be highly appetitive for equines and preferable over human interactions (Kieson *et al.*, 2020) and conspecific interactions (Lee *et al.*, 2021) with experts (Delphi consultation - Chapter3) ranking provision of food as the most reliable method of inducing positive affect in equines. However equines often show signs of heightened arousal during food delivery, especially if delivery is delayed or the animal has previously experienced threats to food resources, which can cause changes in valence from positive to negative (frustration) in some individuals (Pannewitz and Loftus, 2023). To ameliorate for this a lower value type of food stimulus can be utilised with Platzer and Feuerbacher (2022) identifying that a range of feedstuffs can be utilised as a positive stimulus for equines (such as concentrate food such as coarse mix, pony nuts, succulents such as carrots and herbal treat cobs) however some (such as hay) may not be considered a high value reinforcer by some individuals (Ninomiya *et al.*, 2007) depending on individual diet and environment (Elia *et al.*, 2010).

Affiliative conspecifics

The presence of affiliative conspecifics is ranked by experts (Chapter 3) as a highly rated method of inducing positive affect in equines relating to their natural ethology as a herd-living prey species (Ringhofer *et al.*, 2017; Cameron *et al.*, 2003). Equines are capable of cross-modal individual recognition (matching of sensory cues to identity using stored information about the individual from other modalities) to identify conspecifics (Proops *et al.*, 2009). They are also adept at recognising facial expressions in conspecifics (Proops *et al.*, 2018; Wathan *et al.*, 2016); as such it is important to accurately distinguish between affiliative and non-affiliative conspecifics when attempting to induce positive affective state to avoid mistakenly introducing a non-affiliative conspecific which could lead to the production of negative affective state. The literature suggests that this can be undertaken through assessment of proximity to conspecifics, duration of proximity, and the performance of active affiliative

behavioural expressions; such as approaches and mutual grooming (Heitor *et al.*, 2006; Sigurjónsdóttir *et al.*, 2003). However preferences for social partners often change over time, including during seasons of the year, and as such it is important that continual assessments of preferred social partner are undertaken to avoid misattribution (Kimura, 1998). It is also important to assess the behaviour of the preferred social partner when designing and implementing studies utilising this methodology as emotional contagion (such as the transference of heightened arousal or fear responses) from one partner to the other could confound the physiological and behavioural data collected (Spinka, 2012).

Verbal praise

Verbal praise is a secondary reinforcer utilised commonly in equine training and handling (Hockenhull and Creighton, 2013). Prior to conditioning (via classical conditioning methods) verbal praise is a neutral stimulus for equines and as such it is important that verbal praise phrases are positively conditioned before use by pairing them with a primary reinforcer such as food (Mills, 1998a; McGreevy, 2012). Experts (Chapter 3) rated verbal praise low as a positive stimulus for equines and as such it was included within this study to examine the potential differences in response between this method and the highly ranked positive stimuli (food and conspecific companionship).

5.1.3 Measurements of affective state

Body language and facial expression

Assessment of body language via the use of ethograms has been undertaken for many years and has been reviewed extensively (Hall and Heleski, 2017). Facial expression is a newer aspect of this overarching behavioural assessment modality and was initially studied from a pain and negative affect perspective (Dalla Costa *et al.*, 2014a; Gleerup *et al.*, 2015) but has more recently been used as a measure of positive affect (Lansade *et al.*, 2018; Trösch *et al.*, 2020; Lansade *et al.*, 2022). Markers such as ear and eye position have been used in recent studies to determine positive affective state and as such may prove reliable indicators amongst other metrics (Lansade *et al.*, 2022; Lansade *et al.*, 2018) as they were additionally highly ranked by experts (Chapter 3). In humans, facial expression is known to correlate with specific emotions and subsequent research in animals has identified changes in facial markers during pain in rodents (Miller *et al.*, 2016), sheep (McLennan *et al.*, 2016) and horses (Gleerup *et al.*, 2015; Dalla Costa *et al.*, 2014a; Dalla Costa *et al.*, 2016b). More recently facial

expressions have been used to identify positive affective states in rats being tickled (Finlayson *et al.*, 2016), sclera prominence in cows (Lambert and Carder, 2017) and eye position in horses (Hintze *et al.*, 2016). This preliminary evidence suggests that facial actions may be useful measures for characterising affective state in horses, who are predominantly visual animals, so postural and visual communication is an important part of the behavioural and communicative repertoire suggesting that changes in these features are likely to occur with changes in attentional and affective state (Wathan *et al.*, 2016; McDonnell, 2003). Additionally, equines have a complex facial musculature structure, similar to primates and humans, enabling them to mobilise and alter many of their facial features during emotional expression (Wathan *et al.*, 2015).

Eye gaze analysis (lateralisation)

Changes in attention biases, such as eye gaze, indicate alterations in cognitive processing that may be used to infer affective state (Paul *et al.*, 2005b; Mendl *et al.*, 2009; Mendl *et al.*, 2010). Attentional state markers can be quantified behaviourally by measuring frequency and/or duration of eye gaze (Winters *et al.*, 2015). Eye gaze analysis or lateralisation is a relatively new focus of study in equines however some literature on lateralised behaviours can be found dating back to the early 2000's (Austin and Rogers, 2012) with hemispheric specialisation identified across taxa including equines (Rogers and Andrew, 2002).

The emotional valence lateralisation hypothesis suggests that negative emotions are associated with increased activity in the right hemisphere while positive emotions are correlated with increased left hemispherical dominance, a theory proposed by Levy *et al.* (1983). Currently the literature with regard to right and left hemispheric lateralisation is contradictory with some evidence suggested right hemispheric activation indicates negative emotional arousal and others suggesting the opposite, for a review see Leliveld *et al.* (2013). Some studies have suggested that both hemispheres are important for processing of positive stimuli (Des Roches *et al.*, 2008) while others have found positive processing with right hemispheric (left eye) dominance (Farmer *et al.*, 2010; Bourne, 2010). Furthermore, it has been suggested that social learning may modulate the lateralisation of emotional valence (Shamay-Tsoory *et al.*, 2008) and that there are differences in lateralisation at an individual level, meaning that it may be important for individual preferences to be assessed prior to use of this measure (Baragli *et al.*, 2021).

Whilst the two halves of the brain are believed to play different roles in emotional processing, the specific contributions of each hemisphere is widely debated (Killgore and Yurgelun-Todd, 2007). It may be the case that instead of the two popular theories (the valence specific hypothesis and the right-hemisphere hypothesis) being in direct opposition to each other, they may in fact reflect different facets of a complex emotional processing system (Killgore and Yurgelun-Todd, 2007). This theory was recently supported by Ross (2021) who put forward an emotion-type hypothesis suggesting that primary emotions and their related behaviours are modulated by the right hemisphere and social emotions and related behaviours are modulated by the left hemisphere, a proposition which echoes the prior findings of Shamay-Tsoory *et al.* (2008). Due to the conflict within the literature this measure was specifically included within this study to add further evidence to the current argument.

Heart rate

Support for the use of heart rate analysis in equines is plentiful, and varied in context, with the literature identifying decreases in heart rate as a marker of positive emotional state relating to affiliative behaviours (Aureli *et al.*, 1999; Boccia *et al.*, 1989; Laister *et al.*, 2011). However, some positive stimuli may increase arousal which is likely to lead to increases in heart rate due to activation of the sympathetic nervous system and as such this measure may not be sufficient alone to identify valence of the emotional state. Heart rate was utilised within this study to attempt to draw conclusions as to its reliability as a measure of affective state rather than simply emotional arousal and to trial the equipment which would be utilised in the later trial (Chapter 6).

Heart rate is an indicator for shifts in the net interaction of the parasympathetic and sympathetic nervous systems (PNS and SNS respectively) also known as sympathovagal balance, within the autonomic nervous system (ANS). The ANS is important within emotion regulation due to its role in monitoring internal body state and making short-term adjustments to address changes in both internal and external demands (Levenson, 2014). The PNS is predominantly responsible for conservation and restorative functions (rest and digest) whereas the SNS influences preparation for stressful experiences (flight and fight responses) (Kreibig, 2010). Chronic stress is often characterised by autonomic imbalance in the form of decreased PNS activity or increased SNS activity whereas shorter-term negative emotional states are characterised by more transient episodes of SNS dominance (Von Borell *et al.*,

2007). The PNS controls the sinoatrial (SA) node which is the pacemaker of the heart and dampens heart rate whereas the SNS has a stimulating effect on heart rate, indicating that at rest (where heart rate is not being influenced by physical exertion) heart rate may be indicative of PNS or SNS activity where values are seen to decline or increase respectively (Von Borell *et al.*, 2007) as identified by Scopa *et al.* (2020) in horses, where decreases in heart rate were associated with relaxation. Within this pilot study heart rate variability was not utilised, however it was included within the final trial (Chapter 6).

Infrared thermography

Infrared thermography (IRT) measures changes in cutaneous temperature of body surfaces via detection of infrared radiation emitted, as reviewed by Travain and Valsecchi (2021). The use of IRT of the eye has been posited as a potentially useful metric for measuring changes in stress response in equines (Waran and Randle, 2017) with increases in eye temperature shown to correlate with elevated salivary and plasma cortisol levels (Cook *et al.*, 2001). Other studies have found associations between elevated eye temperature and activation of the hypothalamic-pituitary-adrenal axis (von Borstel *et al.*, 2017) and some suggest that alone, it may be more of a useful measure of arousal rather than valence *per se*, however when used in conjunction with other measures it could be a reliable measure of the valence component of positive affective state (Travain *et al.*, 2016). Indeed Ramirez Montes De Oca (2021) examined the use of IRT in assessing negative emotional state in dairy calves and equines and identified that there were lateralised increases in nostril, eye and ear temperatures indicative of negative emotional state during trials. These findings are based on the grounding that there are ipsilateral relationships between nostril and ear temperatures and hemispheric brain activation and contralateral relationships between ear and eye temperatures and hemispheric brain activation.

Similarly to heart rate, core and peripheral body temperature is controlled by the ANS through the action of the SNS which alters resistance in the vascular system resulting in changes to blood pressure and flow which affects heat dissipation and therefore changes in both the core and peripheral temperatures of the body (Travain and Valsecchi, 2021). Physical or psychological stress leads to a physiological response whereby blood flow to the internal organs is prioritised over that to the periphery leading to an increase in core body temperature and cutaneous vasoconstriction leading to a decrease in peripheral

temperature. This stress-induced hyperthermia (SIH) is seen across endothermic species (Mota-Rojas *et al.*, 2021) and has been demonstrated in dogs (Riemer *et al.*, 2016), mice (Bouwknicht *et al.*, 2007) and rats (Dallmann *et al.*, 2006).

The advantages of using IRT include the non-invasive nature and real-time appraisal ability as well as the increasing availability of handheld devices or smartphone adapters making this technology more accessible for use. Cumulatively these measures may provide insight into the emotional processing and resultant affective state in equines provided with different stimuli aiming to induce positive affective state. The interplay between measures and differences in measurements within subjects may elucidate individual preferences for stimuli and provide support for the inclusion of induction measures within the final trial. In addition, prior testing of methodologies allows for refinement in the research process and potential reduction in trial time and repetition which aids in the production of reliable data and enhances welfare through consideration of the 3Rs (replacement, reduction and refinement) in animal research (Kirk, 2018; DeGrazia and Beauchamp, 2019).

5.1.4 Aims and Objectives

This pilot study aimed to trial several of the proposed induction and measurement methods for positive affect suggested by the expert Delphi participants (Chapter 3) to evaluate and refine the proposed methodology for the final study which was designed to induce positive affect in equines.

The objectives were to:

1. Evaluate different methods of inducing positive affective state through the provision of positive stimuli (verbal praise, addition of an affiliative conspecific and provision of high value food),
2. Apply behavioural (facial actions, eye gaze) and physiological methods (heart rate and infrared thermography of the eye) to measure the presence of positive affective state,
3. Evaluate the induction and measurement methods employed and investigate relationships between the different methods.

It was hypothesised that individual equines would respond differently to the three induction stimuli due to variances in individual personality and prior experience and that verbal praise

(low rated by Delphi experts) may be less appetitive than the addition of food or an affiliative conspecific.

Secondly it was hypothesised that there may be relationships between heart rate and eye temperature linked to changes in valence and arousal during affect induction and that these physiological measures may also relate to changes in eye and ear position (facial actions, eye gaze) linked to attention and affective state.

5.2 Methodology

5.2.1 Ethical approval

The protocol and methods of this study were approved by the Animal Welfare Ethics Review Board (AWERB) of Newcastle University (Project ID: 1025) and the Ethical Review Committee of University Centre Askham Bryan (Approval number: PGP532019) prior to study commencement. Permission was also granted by Queen's Veterinary School Hospital, University of Cambridge for data collection at their site.

5.2.2 Pilot study focus and design overview

This pilot study was undertaken to trial the proposed methodology for the final study which would utilise the expert Delphi derived induction methods and behavioural and physiological measures to attempt to induce and measure positive affective state in a group of equines. A pilot study should ask the same questions as a feasibility study; namely whether a project should be undertaken and if so how it should be implemented but it also has a design element where the planned future study is conducted on a smaller scale (Eldridge *et al.*, 2016). This pilot study aimed to trial the following features of the planned experimental design for the final study reported in Chapter 6, the focal aspects were to:

1. Identify optimal methods to introduce a range of positive stimuli
2. Explore camera placement to capture behavioural data (eye and ear position as well as Qualitative Behaviour Assessment and indicators of frustration) and physiological data (respiratory rate)
3. Compare equine behaviours when cross-tied versus when handler held during the experimental period and analyse the quality of data generated.
4. Test heart rate monitor placement and settings
5. Trial the settings for the infrared camera

6. Design and evaluate scoring systems for behavioural data
7. Undertake initial statistical analyses of the measures recorded

The pilot study took place in from January to March 2021 in two locations. An overview of the methods is detailed below (Figure 5.1).

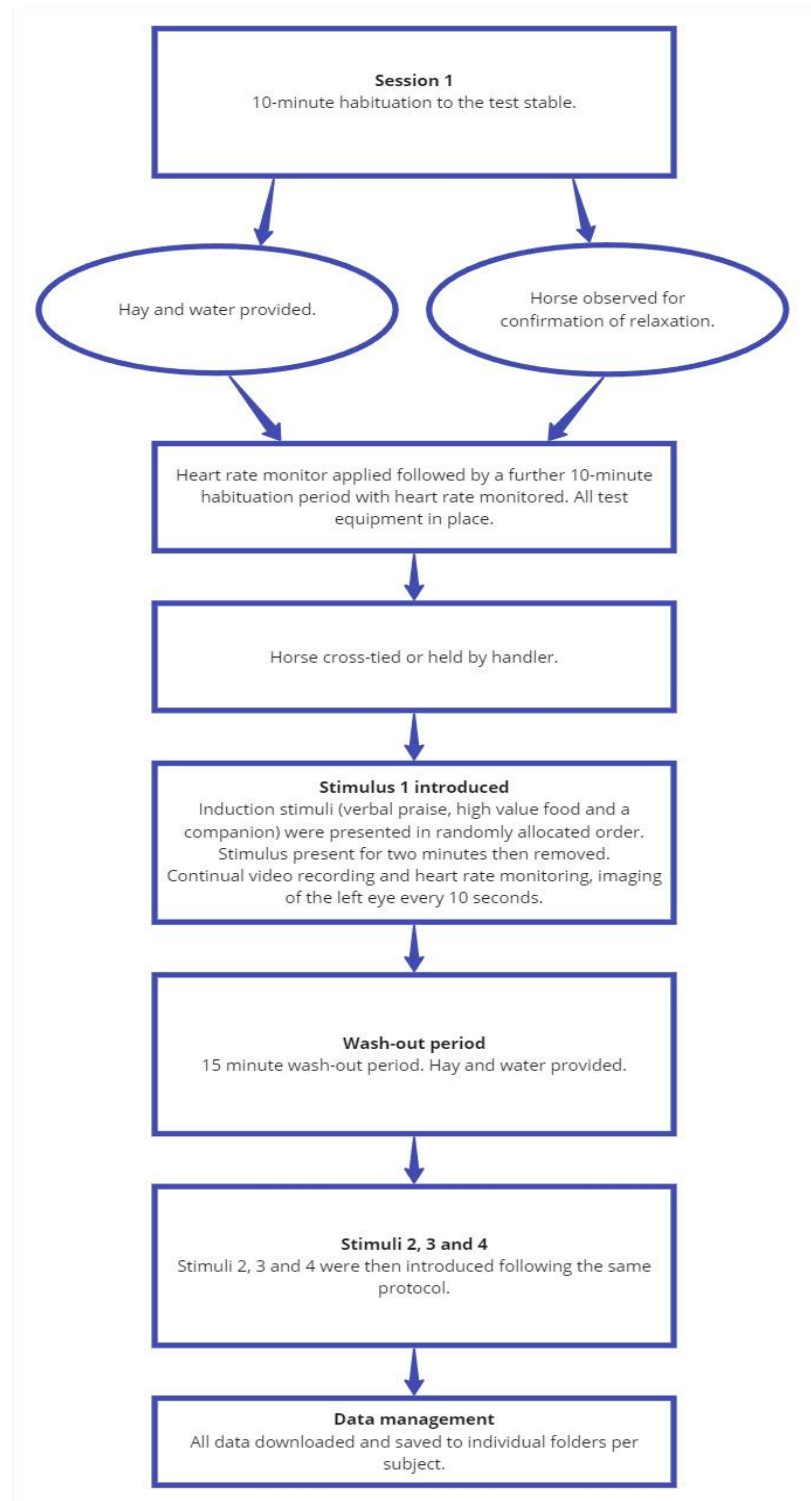


Figure 5.1 Overview of the pilot study methodology.

5.2.3 Subjects

The 23 equines included within this study were aged 6-19 years of age, mares (n=10), geldings (n=13), and represented a range of breeds and crossbreeds. They ranged in height from 13hh to 16.3hh.

All equines were utilised within university teaching facilities at two university campuses (nine at Queen's Veterinary School Hospital, University of Cambridge (site A) and 14 at University Centre Askham Bryan (site B)). All had been housed in their relative facilities for a minimum of six months and were previously habituated to the areas used during data collection on at least five occasions (range 5-11 occasions) prior to the study as described in the literature (Christensen *et al.*, 2006; Christensen *et al.*, 2008b). Both populations of equines were housed in American barn style stabling with small group turnout daily and fed high fibre diets with *ad lib* hay. Measurements at site A were taken over a 3-day period and measurements at site B were taken over four days split over a two-week period. All measurements were undertaken at the same time of year and were undertaken in the morning after feeding time to reduce the effects of food anticipation and to consider the effects of circadian rhythm (Fenner *et al.*, 2016).

5.2.4 Habituation to the study area and equipment

All equines were habituated to the stable used for data collection prior to the trials by being placed in the stables regularly for 10–30-minute periods with hay and water. They were considered habituated when they quickly (within 1 minute) showed postural (relaxed muscle tone, low head carriage) and behavioural (standing and eating hay or resting/dozing) signs of relaxation on entering the stable. Horses that did not habituate to the stable area were excluded from the trial (n = 2). At the beginning of each trial all equines were given a ten-minute acclimatisation period prior to the application of equipment for the trials. All equines were also habituated to the wearing of the Polar equine heart rate monitor on three occasions prior to the sessions, and all were used to wearing items around the girth as they were all ridden equines. Habituation was deemed to have occurred when heart rate did not increase more than 20 beats per minute during the heart rate monitor application (Leiner and Fendt, 2011). During habituation to the test stable all equines were also presented with the video camera set up in terms of positioning of the front and rear stable cameras and handheld thermography camera. All equines were previously habituated to cross-tying as this method

of restraint was utilised frequently in their teaching scenarios. A cross-tying method was identified to allow for suitable positioning of the head for thermography video collection whilst allowing for relatively natural movement of the head in both dorsoventral and mediolateral planes to facilitate behavioural measurements and to reduce frustration related responses (Pannewitz and Loftus, 2023) as shown in Figure 5.2. Horses that did not habituate to cross-tying (habituation classed as no rise in heart rate over 20 beats per minute and an absence of avoidance, escape or frustration behaviours) over three trial periods were held by a handler on the same length lead rope during the trials to avoid potential stress responses confounding results (Stull and Rodiek, 2002). The handler did not interact with the horse during the trial and stood outside the stable and to the right of the closed stable door to reduce handler effect (Birke *et al.*, 2011; Górecka *et al.*, 2007).



Figure 5.2 Cross-tied subject with heart rate monitor fitted under Polypad and surcingle.

5.2.5 Test procedure

All measurements were recorded by the same two experimenters throughout the data collection with additional handlers at each site who were known to the individual equines.

The test protocol was identical for each equine. The subject was led to the research stable by a known handler and released for a 10-minute habituation period during which they had *ad lib* access to hay and water (Lambert and Carder, 2019). A familiar affiliative equine was placed in a stable on the same side of the barn approximately 15 metres away during the entire trial to avoid isolation stress (Lansade *et al.*, 2018; Zhao *et al.*, 2020). After 10-minutes the subject was fitted with the Polar equine heart rate monitor. The subject was left to habituate freely for a further ten minutes after which the subject was cross-tied, and the hay and water removed. The heart rate monitor and the two cameras started recording video data with the time on the heart rate monitor and cameras synchronised for later analysis. The subject was presented with the three stimuli chosen to induce a positive affective state as identified in the first round of the Delphi consultation described in Chapters two and three. The stimuli were presented to all equines in a randomly allocated order using online randomisation software (Urbaniak and Plous, 2013). Heart rate was measured continuously once the subject was cross-tied along with thermal imaging of the left and right eye and video recording of the head and body for scoring of eye and ear positions and body language. A period of two minutes was timed from the beginning of data collection to obtain baseline readings ('before' data). This process was repeated for each stimulus trial in each horse.

Stimulus observations commenced at the beginning of the approach of each stimulus with each stimulus presented from ten metres by walking in a straight line towards the subject and stopping in front of the subject. The subject was free to interact with the stimulus up to the end of a two-minute period when the stimulus was removed in a straight line away from the subject. Once the stimulus moved out of the subjects view the recording of behavioural and physiological parameters continued for a further two minutes and were then stopped, following the method of Des Roches *et al.* (2008). A fifteen-minute wash-out period was used between treatments (Robinson *et al.*, 2018a) to allow for heart rate, respiratory rate and behavioural responses to return to baseline. Whilst within the behavioural literature washout periods are reported which vary between a few hours (Walker *et al.*, 2006), 12 hours (Greening *et al.*, 2021), 24 hours (Williams *et al.*, 2017b) to several days (Barnes *et al.*, 2012) or even weeks (Ramos *et al.*, 2006) depending on the context of the research and the duration of the research stimuli however within this pilot study limitations arose in terms of the

availability of all of the horses for all three trials, necessitating a shorter than optimal wash-out period.

The Delphi aligned stimuli presented in random order were:

1. High value food (F)– 1 Stubbs™ food scoop (~1.5kg) (Stubbs, 2024b) of concentrate mix presented in a familiar vessel which had been previously associated, by the subjects, with receiving feedstuff on a daily basis; presented by the caregiver in an overdoor feeder (Stubbs, 2024a).
2. Care giver verbal praise (P) – offered every five seconds using habituated phrases (previously paired with a primary reinforcer of food for at least 20 repetitions by the caregiver) for the individual subject such as ‘Good, ‘Yes’, ‘Super’ depending on the individuals prior habituation (Martin and Friedman, 2011; Dorey and Cox, 2018; Pfaller-Sadovsky *et al.*, 2020).
3. Addition of a conspecific companion (C) – an affiliative equine known to the subject that they have previously been noted to spend time with and allogroom with during turnout; handled by the familiar caregiver. The companion was introduced in the same manner as the other stimuli and was stood in front of the stabled horse with full flexibility to engage in sniffing, touching and allogrooming throughout the two-minute period.

5.2.6 Behavioural and physiological measures

Behavioural measures

Behavioural observations were recorded using two GoPro cameras affixed to the rear left corner of the stable at a height of 2 metres and on a tripod in front of the subject at 3 metres distance. These cameras facilitated continual recording of the subjects’ behavioural responses from two perspectives and the approach and retreat of the three stimuli meaning that behavioural data timepoints could be paired with heart rate and infrared thermography data in later analyses. Data were stored on a removable SD card (Sandisk™ Extreme Pro SDXC) and uploaded to secure cloud storage after each trial.

Each subject had the following behavioural data scored by the experimenters using video footage after the trials. Durations (seconds) were calculated for each behavioural output across each horse and trial.

1. Ear position – Forwards, Asymmetrical, Backwards (Figure 5.3)
2. Eye position – White, Half, Open; left eye, right eye and both eye bias (Figure 5.3)
3. Notes were also taken regarding the animal's general responses to the stimuli such as signs of frustration, over-arousal and avoidance.



Left eye bias



Sclera visible



Half open eye



Ears forwards



Ears backwards



Ears asymmetrical

Figure 5.3 Eye and ear positions recorded.

Physiological measures

All subjects were fitted with a Polar Equine Set – H2 Wearlink Transmitter and surcingle enabling the attachment of a Polar Equine Heart Rate Monitor M400 (Polar Electro Oy, Proffessorintie 5, 90440 Kempele, Finland) for continuous recording of heart rate. All equines were either clipped or had shed their winter coats therefore the use of ultrasonic transmission gel alongside dampening the electrode application area facilitated reliable recording of readings. The electrodes were held in place with an elasticated surcingle and polypad. Data were transmitted to Polar FlowSync (Polar Electro Oy, 2022) and saved as subject number files.

An infrared camera (ThermaCam T604, FLIR Systems AB, Danderyd, Sweden) was used to gather thermographic data of each subjects left eye, recorded from the left-hand side at 1

metre distance. The camera was calibrated to the observation distance, temperature and relative humidity and recalibrated for each subject. Continual recording of the medial posterior palpebral border of the lower eyelid and lacrimal caruncle (Figure 5.4) was utilised to avoid missing data highlighted in previous studies (Fenner *et al.*, 2016; Yarnell *et al.*, 2015).

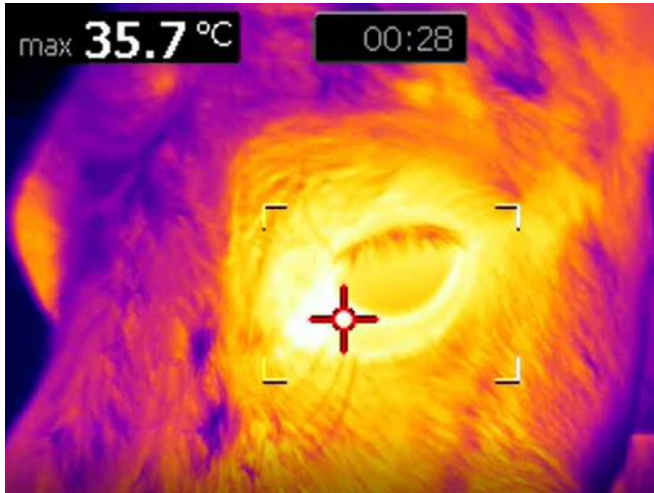


Figure 5.4 Anatomical recording position (medial posterior palpebral border of the lower eyelid and lacrimal caruncle) for eye temperature.

5.2.7 Statistical analysis

Statistical analyses were undertaken in Minitab 18 (Minitab, 2021) and SPSS 27 (IBM, 2020). Data were time matched for each subject to allowing pairing of variables for analysis. An exploratory rather than explanatory approach was taken to the analysis with the aim of highlighting measures that may be of interest in further studies. The independent variables were the induction methods, and the dependant variables were the physiological and behavioural measures.

Eye bias data were not normally distributed (Anderson Darling test, all $p < 0.005$) and did not have equal variance (Levene's test 54.58, $p < 0.001$ corrected for multiple comparisons). Eye position data were not normally distributed (Anderson Darling test, all $p < 0.005$) and did not have equal variance (Levene's test 37.52, $p < 0.001$ corrected for multiple comparisons). Eye temperature data were not normally distributed (Anderson Darling test, all $p < 0.005$) but did have equal variance (Levene's test 1.14, $p = 0.859$ corrected for multiple comparisons). Ear position data were not normally distributed (Anderson Darling test, all $p < 0.005$) apart from 'After Companion Forwards' (Levene's test $A^2 = 0.23$, $p = 0.487$) and did not have equal variance (Levene's test 9.36, $p < 0.001$ corrected for multiple comparisons). Heart rate data

were not normally distributed (Anderson Darling test, all $p < 0.005$) and did not have equal variance (Levene's test 61.93, $p < 0.001$ corrected for multiple comparisons).

Friedman tests were utilised as the non-parametric alternative to the analysis of variance (ANOVA) for repeated measures, followed by Wilcoxon signed rank tests for any significant Friedman results to identify within which groups the difference(s) were found (Riffenburgh, 2012). Spearman rank correlations were used to investigate potential relationships between continuous variables within this non-parametric dataset. Relationships between duration of eye bias and eye temperature, eye bias and heart rate, eye bias and eye position, and ear position and eye position were undertaken in an exploratory manner to allow consideration of which measures may be taken forwards to the final experiment.

5.3 Results

To evaluate changes in measurement variable values between induction methods over the different trials and phases, each measurement variable was analysed alone to discover trends; and following this, initial exploration of the potential relationships between variables were analysed to consider whether variables appeared related to each other.

5.3.1 Behavioural measures to evaluate positive affect in equines

Eye bias

There were no statistically significant differences in left eye bias across trials and phases ($\chi^2(2) = 0.67, p = 0.717$), nor in right eye bias ($\chi^2(2) = 0.67, p = 0.717$), or both eye bias ($\chi^2(2) = 0.00, p = 1.000$), however trends relating to mean eye bias expression were identified when data were graphed (Figure 5.5) with both eye bias being higher during the food stimulus compared with the other induction stimuli and phases. Left eye bias appears to increase during the companion induction compared with the food and praise inductions and right eye bias appears to increase during the companion induction and after the food induction compared with the other phases and treatments.

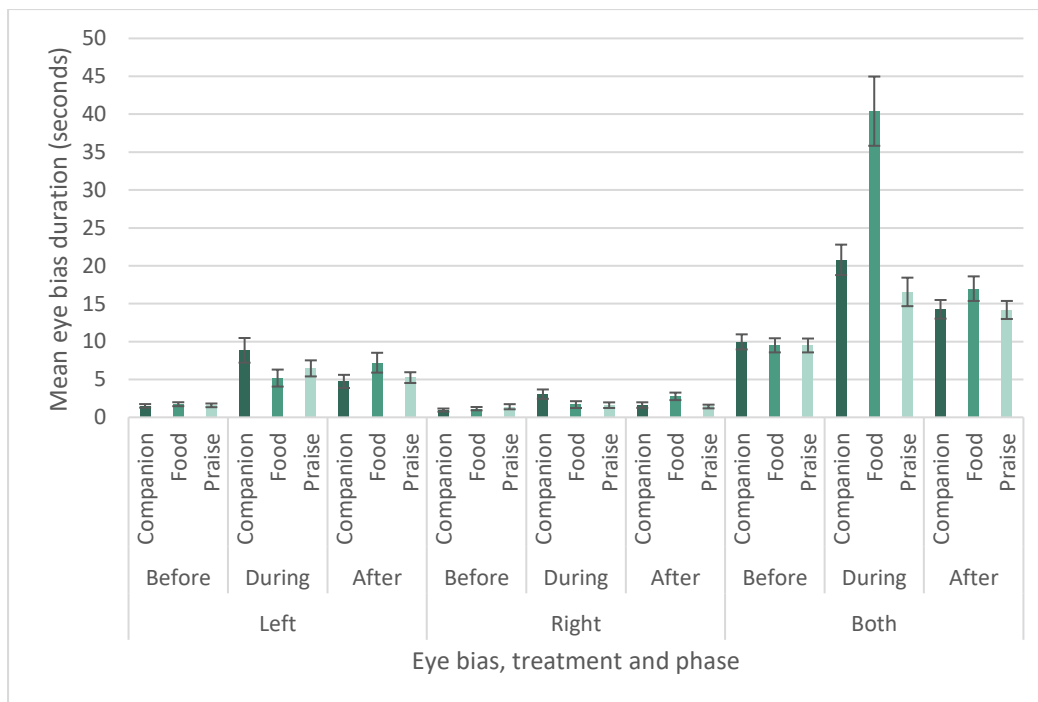


Figure 5.5 Mean duration (seconds) with SEM, spent in each eye bias position across treatments and phases.

However, there was a statistically significant difference between left, right and both eye bias across all treatments and phases ($\chi^2(2) = 18.00, p < 0.001$) with all eye positions significantly different from each other (all $Z = -2.666, p = 0.008$) (Figure 5.6).

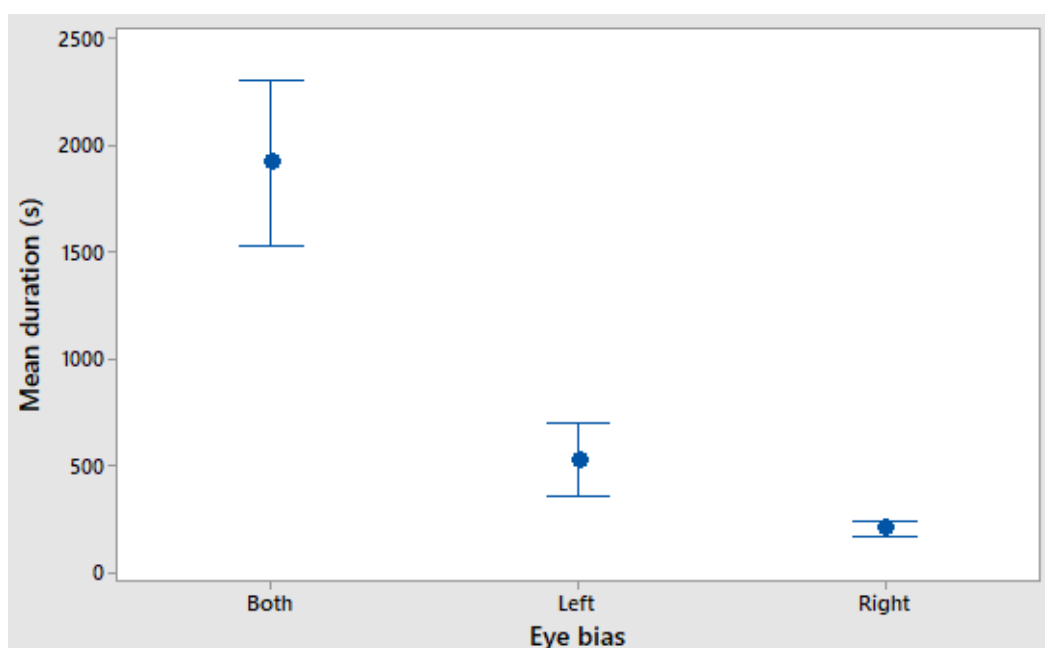


Figure 5.6 Mean duration (seconds) with 95% confidence intervals, spent in eye bias across treatments and phases.

Eye position

There were no statistically significant differences in sclera duration across trials and phases ($\chi^2(2) = 0.67, p = 0.717$), nor in eye open duration ($\chi^2(2) = 4.67, p = 0.097$), or half eye duration ($\chi^2(2) = 0.67, p = 0.717$). Trends were identified suggesting that sclera presence may increase after the induction trial and that half eye position was most seen during an induction trial, particularly during the food trial (Figure 5.7).

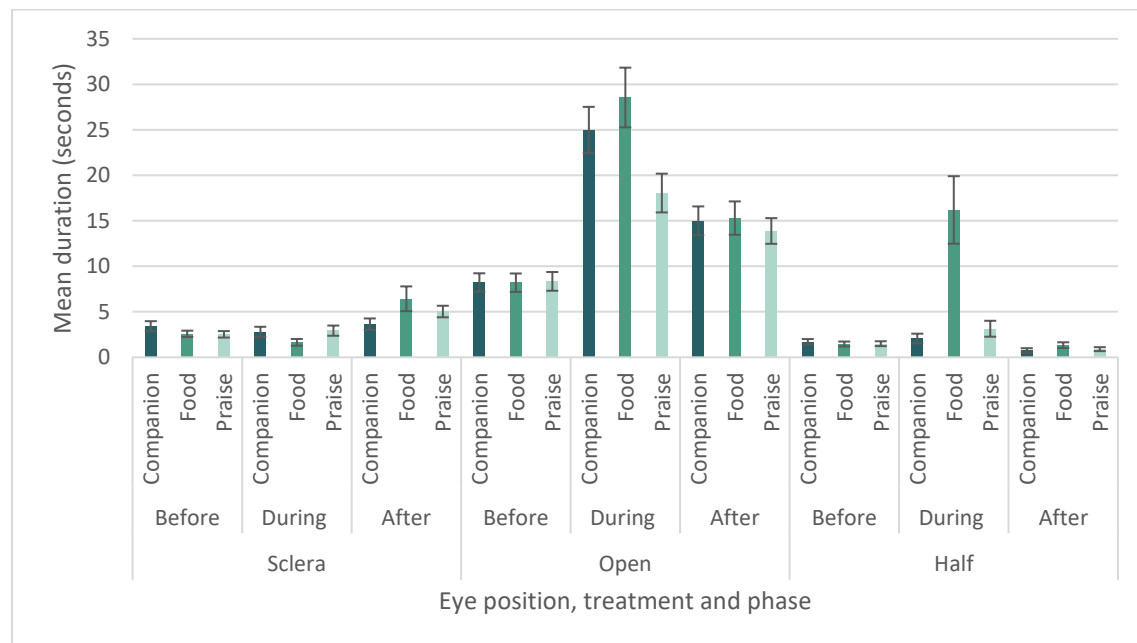


Figure 5.7 Mean duration (seconds) with SEM, spent in each eye position across treatments and phases.

There was also no statistically significant difference between eye sclera, eye open and eye half across treatments and phases ($\chi^2(2) = 4.09, p = 0.894$) (Figure 5.7). However, eye positions half and open were significantly different from each other ($Z = -2.666, p = 0.008$) as were eye positions sclera and open ($Z = -2.521, p = 0.012$) (Figure 5.8).

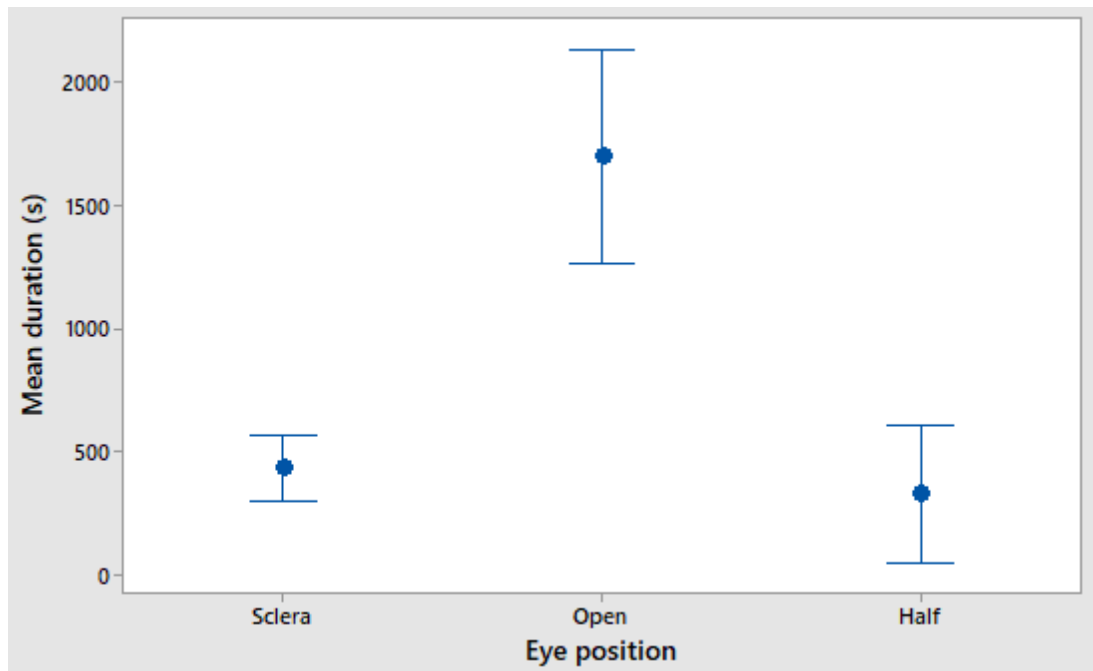


Figure 5.8 Mean time (seconds) with 95% confidence intervals, spent in eye positions across treatments and phases.

Ear position

There were no significant differences identified in the ear positions asymmetrical ($\chi^2(2) = 0.00$, $p = 1.000$), forward ($\chi^2(2) = 1.00$, $p = 0.607$), nor backward ($\chi^2(2) = 0.00$, $p = 0.1.000$) across treatments and phases. Some interesting trends were observed including the reduction in ears asymmetrical during the companion treatment with a substantial increase after companion treatment (Figure 5.9).

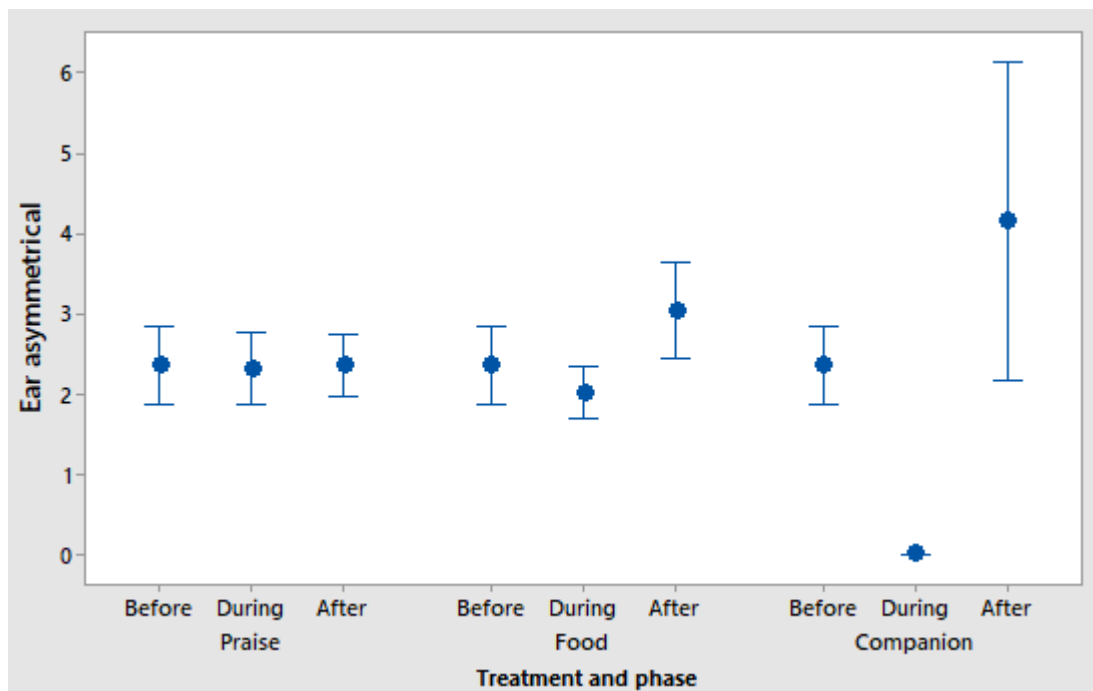


Figure 5.9 Ear asymmetrical (mean duration, seconds) with 95% confidence intervals, during all phases and treatments.

Ear backward was seen to increase substantially during the food treatment (Figure 5.10).

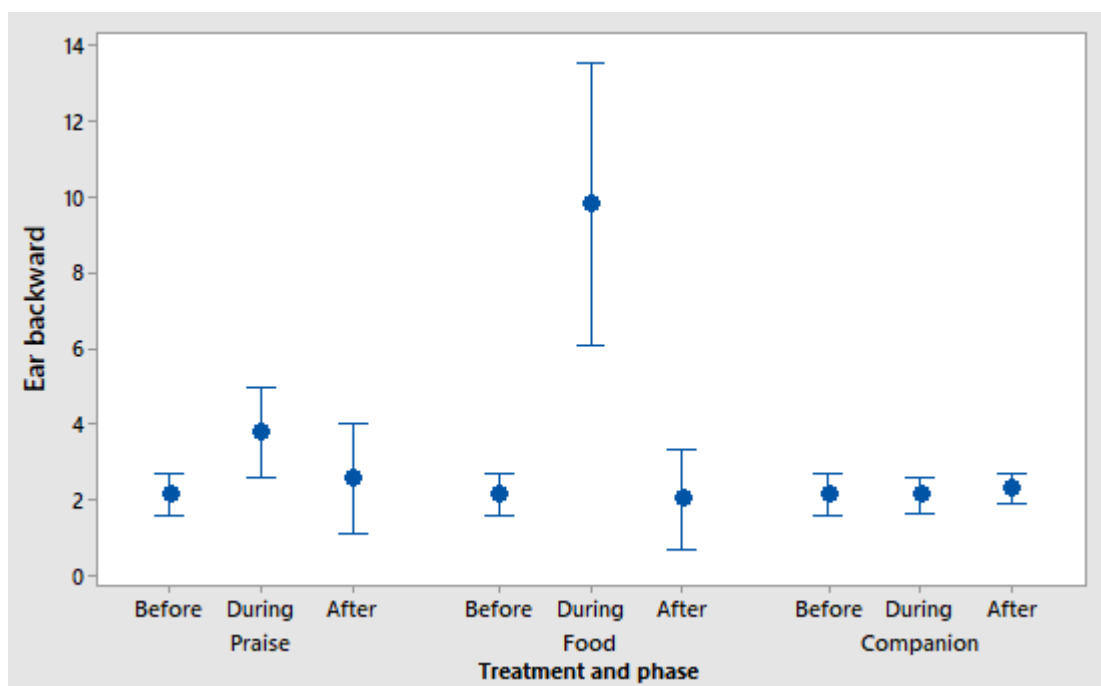


Figure 5.10 Ear backward (mean duration, seconds) with 95% confidence intervals, during all phases and treatments.

Ear forward was substantially lower after the companion treatment; however, all other treatments and phases had overlapping confidence intervals (Figure 5.11).

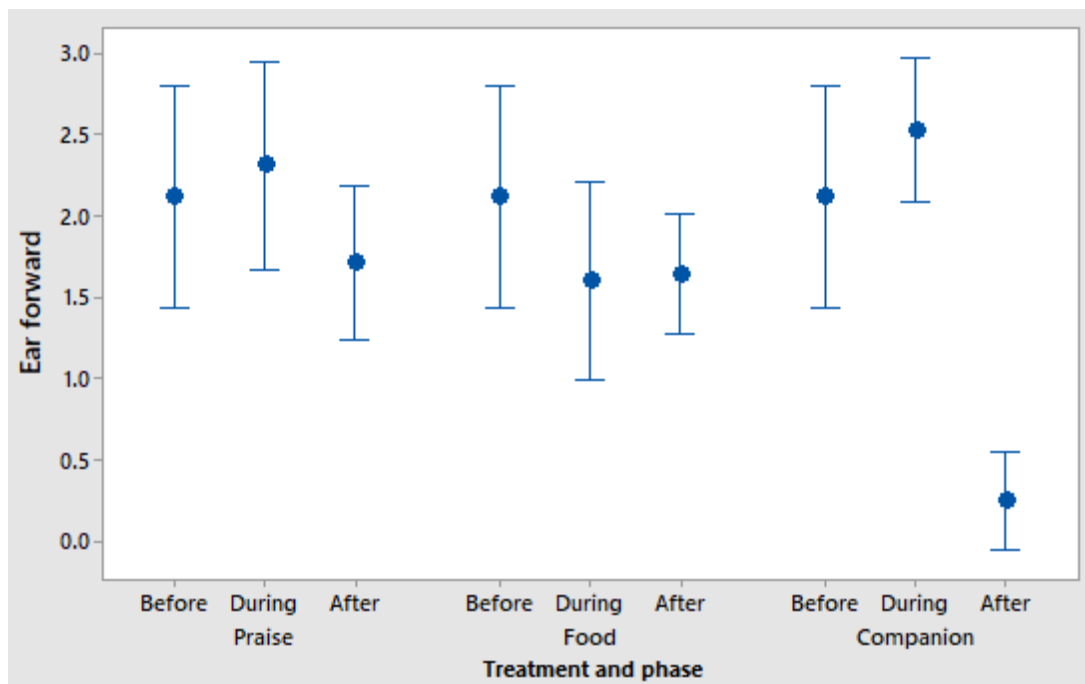


Figure 5.11 Ear forward (mean duration, seconds) with 95% confidence intervals, during all phases and treatments.

5.3.2 Physiological measures to evaluate positive affect in equines

Eye temperature

Treatment and phase did not significantly affect eye temperature overall ($\chi^2(2) = 2.00$, $p = 0.368$) (Figure 5.12).

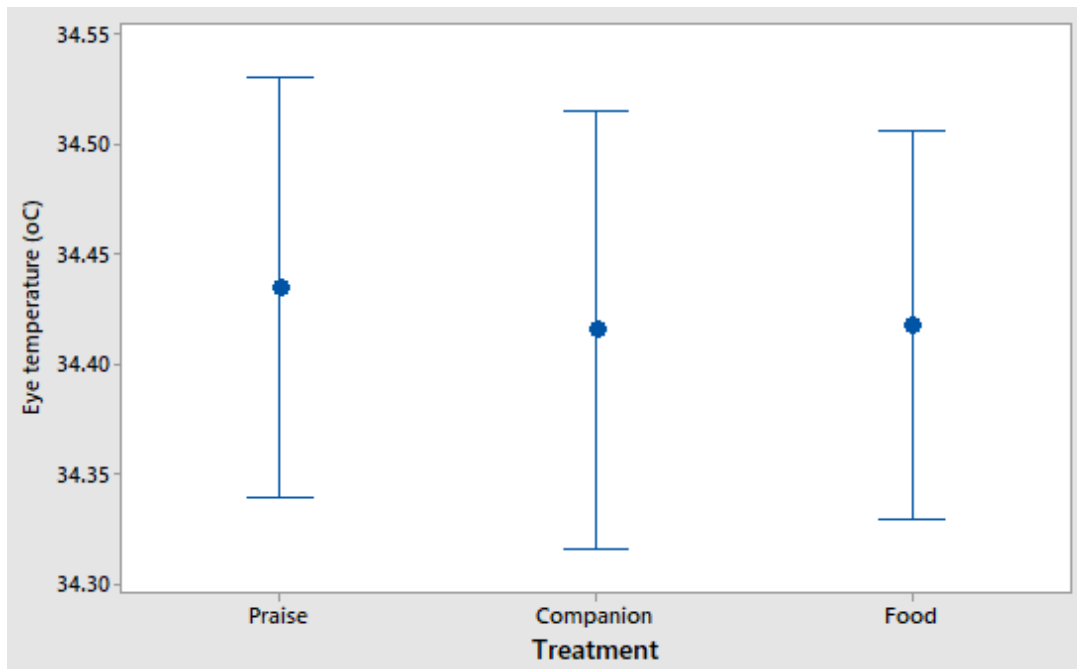


Figure 5.12 Eye temperature (°C) with 95% confidence intervals, across the three treatments.

However, there were some trends of interest identified including a decrease in eye temperature after the praise treatment and an increase in temperature after the companion treatment (Figure 5.13). These trends were further supported by Wilcoxon analysis which identified that there were significant differences in eye temperature within the praise group between before and during ($W = 47256.00$, $Z = 2.386$, $p = 0.017$), during and after ($W = 43756.50$, $Z = 9.672$, $p < 0.001$) and before and after ($W = 47907.50$, $Z = 2.008$, $p = 0.037$). A significant difference was reported in eye temperature in the food group between during and after phases ($W = 50960.50$, $Z = 8.516$, $p < 0.001$) which appears incongruous. In the companion trial a difference was identified between the during and after phases ($W = 44518.50$, $Z = 9.018$, $p < 0.001$). However, there is significant potential for type 1 error across these findings and as such they cannot be deemed reliable.

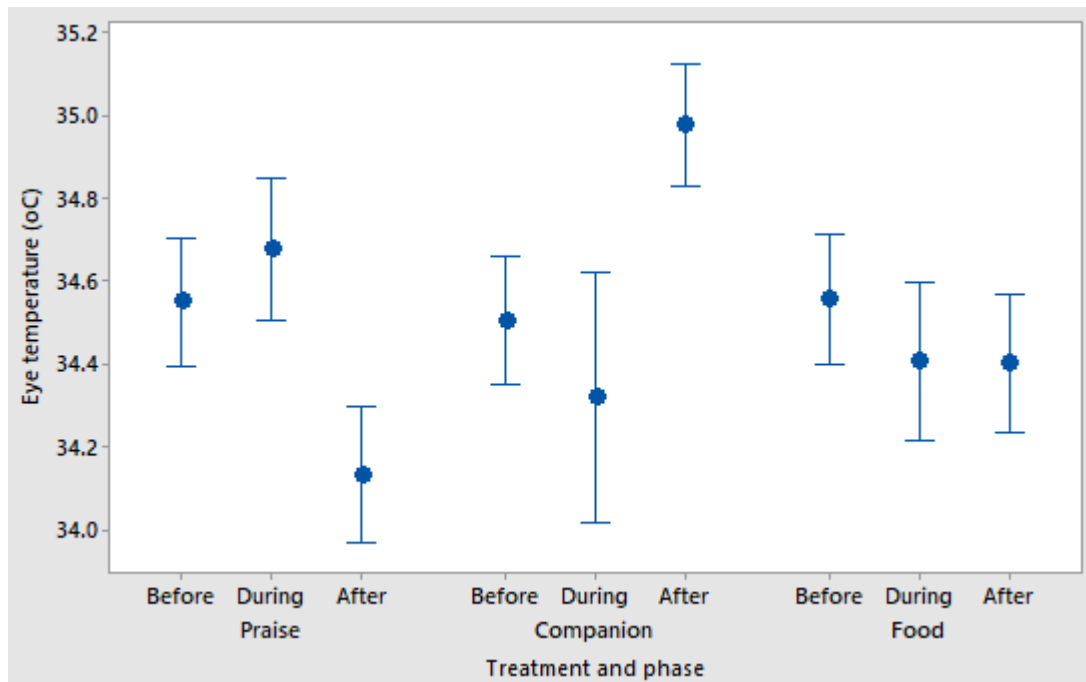


Figure 5.13 Eye temperature (°C) with 95% confidence intervals, during all phases and treatments.

Heart rate

There was no significant difference in heart rate between all treatments and phases when analysed together ($\chi^2(2) = 2.00, p = 0.368$). However, there were interesting trends observed including a continual decrease in heart rate during and after the praise treatment, a continual increase in heart rate during and after the food treatment and decrease in heart rate during the companion treatment followed by a substantial increase after treatment (Figure 5.14).

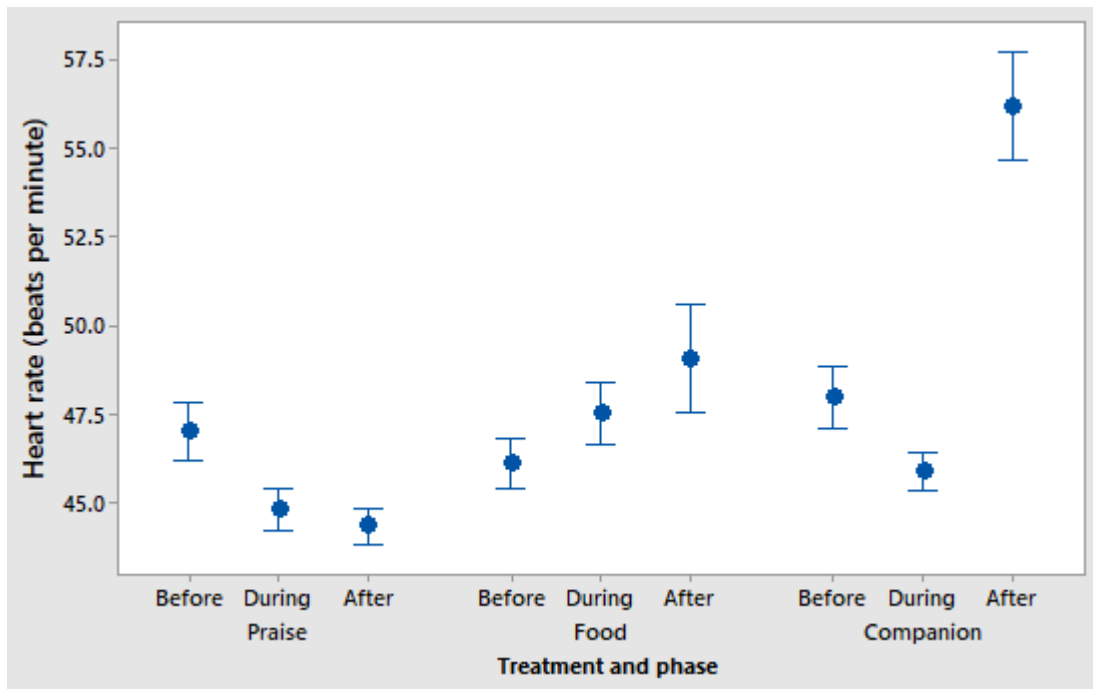


Figure 5.14 Heart rate (beats per minute) with 95% confidence intervals, during all phases and treatments.

5.3.3 Relationship between eye bias, eye position, eye temperature, ear position, and heart rate

Left eye bias and left eye temperature

There was a weak positive correlation between left eye bias and left eye temperature during the Food treatment (Spearman Rank Correlation, $r_s = 0.278$, $N = 124$, $p = 0.014$) (Figure 5.15) and a moderate positive correlation after Companion treatment (Spearman Rank Correlation, $r_s = 0.407$, $N = 103$, $p < 0.001$) (Figure 5.16).

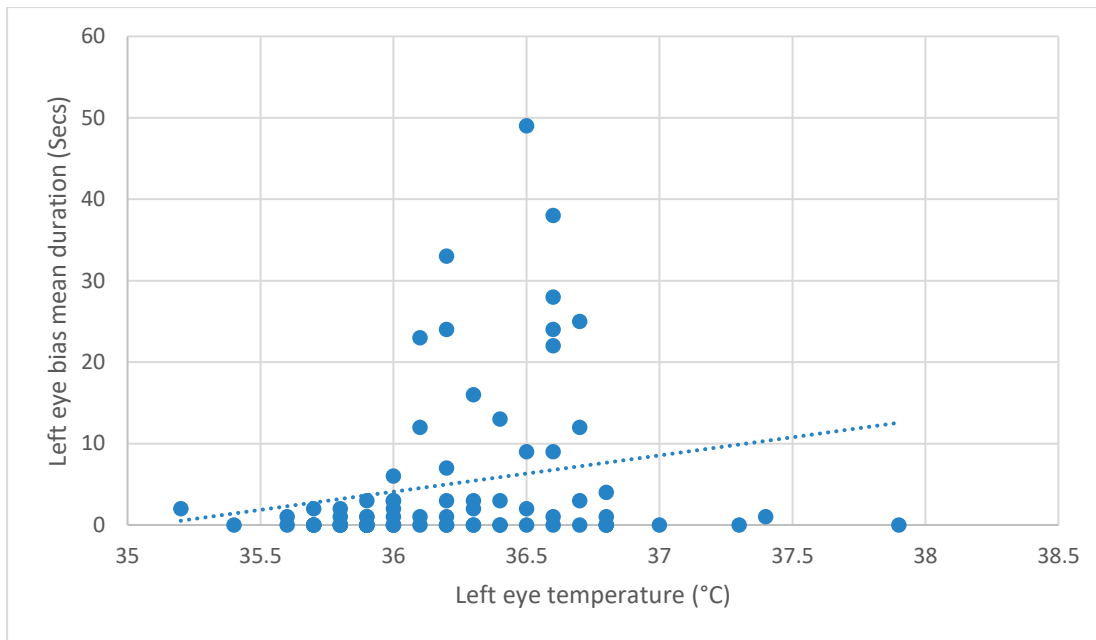


Figure 5.15 Relationship between left eye bias and temperature during Food treatment.

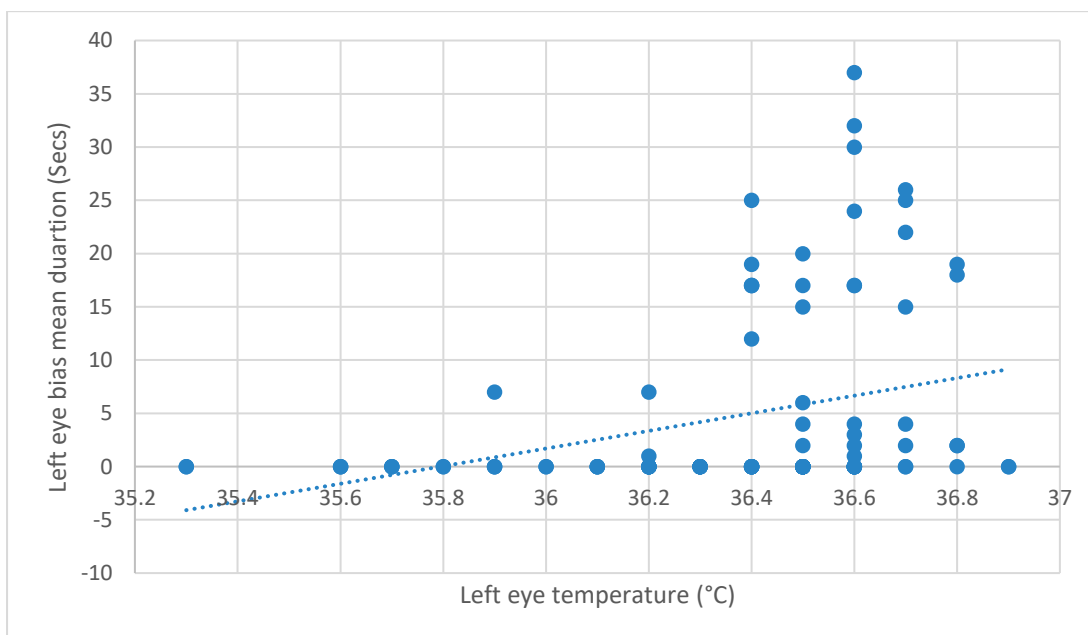


Figure 5.16 Relationship between left eye bias and temperature after Companion treatment.

Left eye bias and heart rate

There were weak to moderate positive correlations between heart rate and left eye bias in the Praise before, Praise during, Praise after, Food before, Food after, and Companion before phases with heart rate increasing with increases in left eye bias. Interestingly groups Food during, Companion during and Companion after reported statistically significant negative

correlations, with heart rate decreasing as left eye bias duration increased; the opposite effect was found for all other treatments (Figure 5.17).

Left eye bias and eye position

There were many significant positive correlations between left eye bias and eye position as detailed in Figure 5.18. The strongest positive correlations were found between left eye bias duration and sclera and eye half durations in the during and after phases across all treatments. The companion group showed the strongest correlations overall.

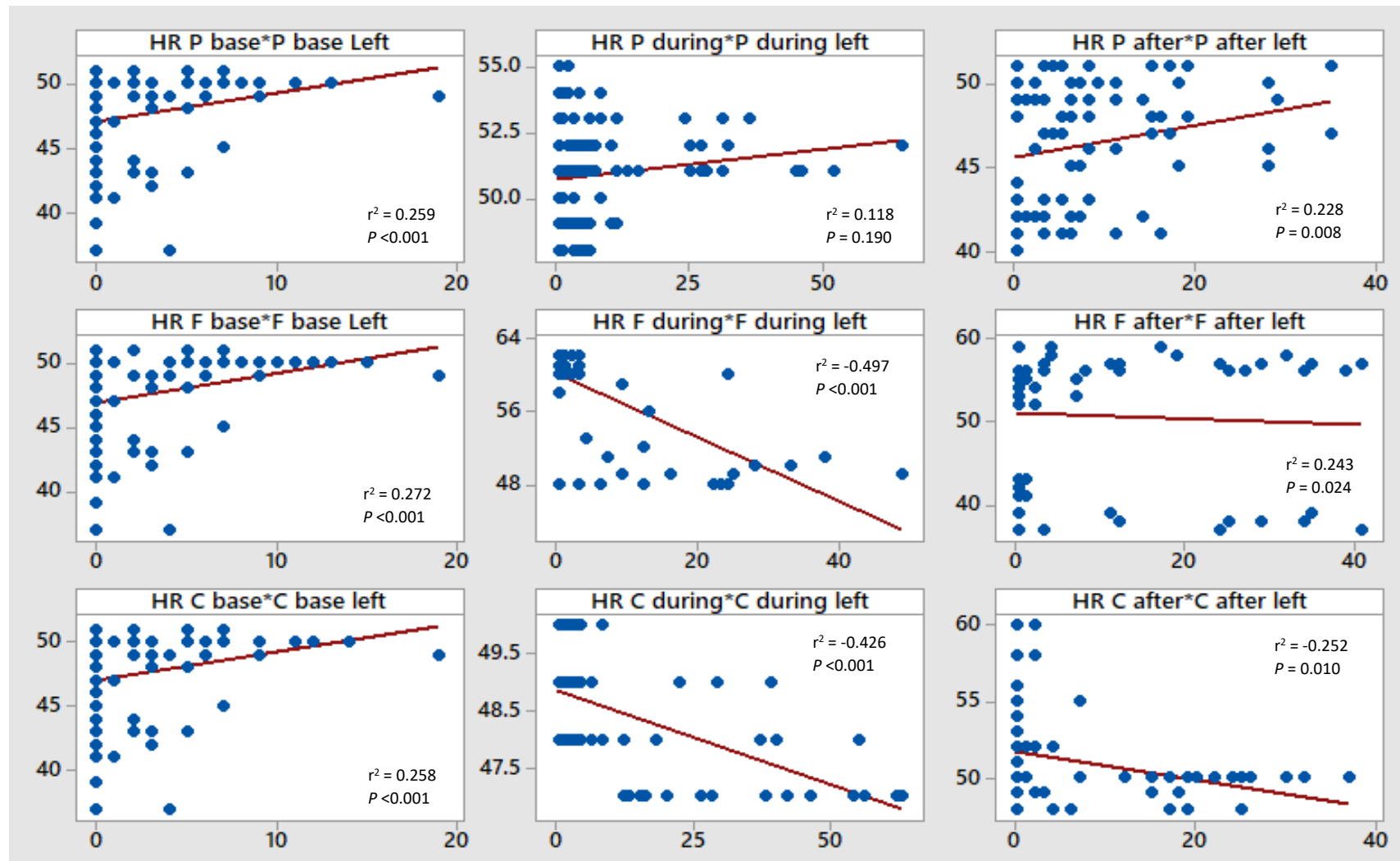


Figure 5.17 Relationships between left eye bias duration and heart rate for all phases and all treatments with Spearman rank correlation results.

C = Companion, F = Food , and P = Praise.

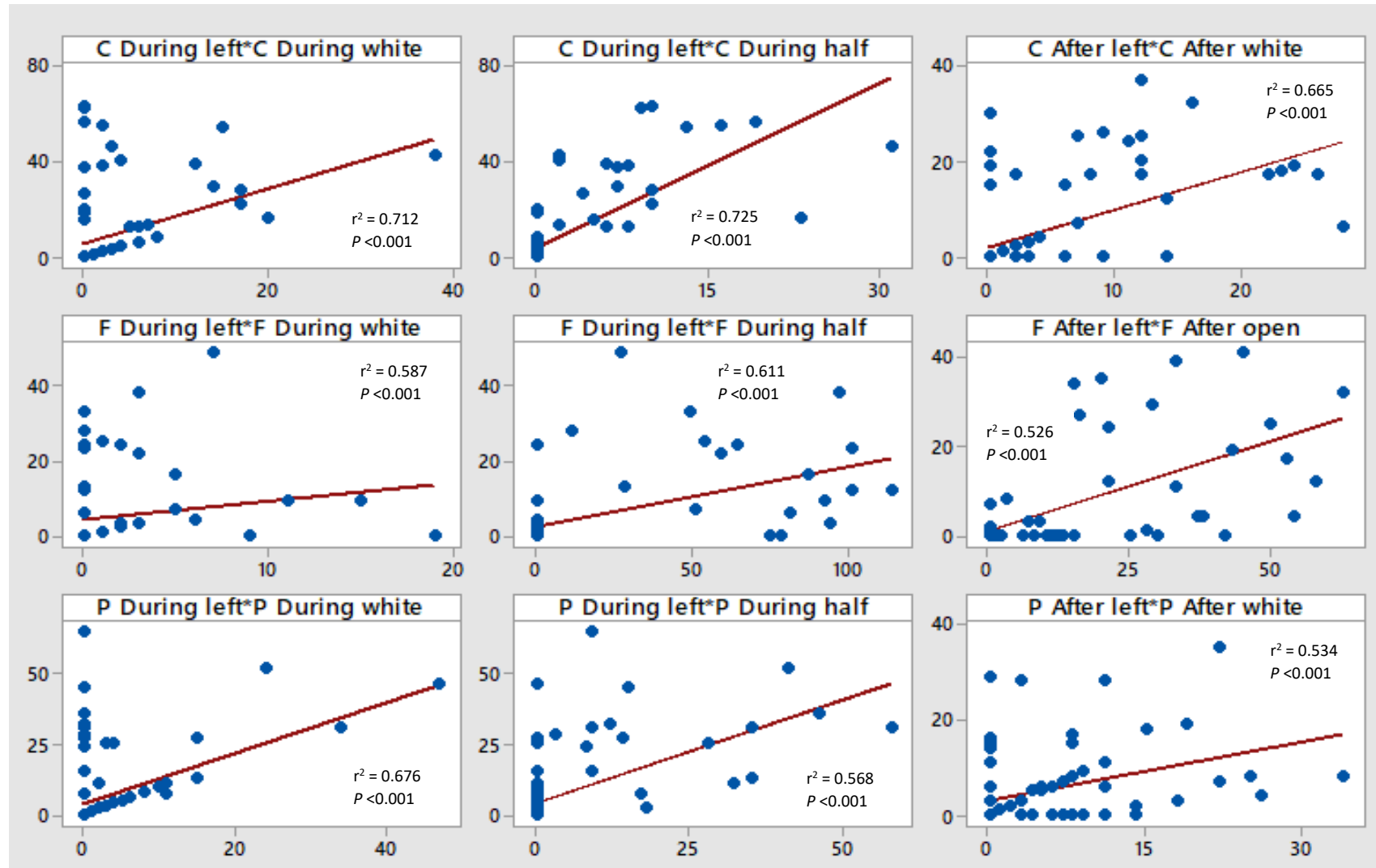


Figure 5.18 Relationships between left eye bias duration and heart rate for all phases and all treatments with Spearman rank correlation results. C = Companion, F = Food , and P = Praise.

Ear position and eye position

There were no significant correlations between 'ears asymmetric' and any eye position nor 'ears forward' and any eye position however there was a strong negative correlation between 'Backwards ears' and 'Sclera' (Spearman Rank Correlation, $r_s = -0.763$, $N = 6416$, $p = 0.017$) and a very strong positive correlation between 'Backwards ears' and 'Eyes open' (Spearman Rank Correlation, $r_s = 0.898$, $N = 18585$, $p = 0.001$) (Figure 5.19).

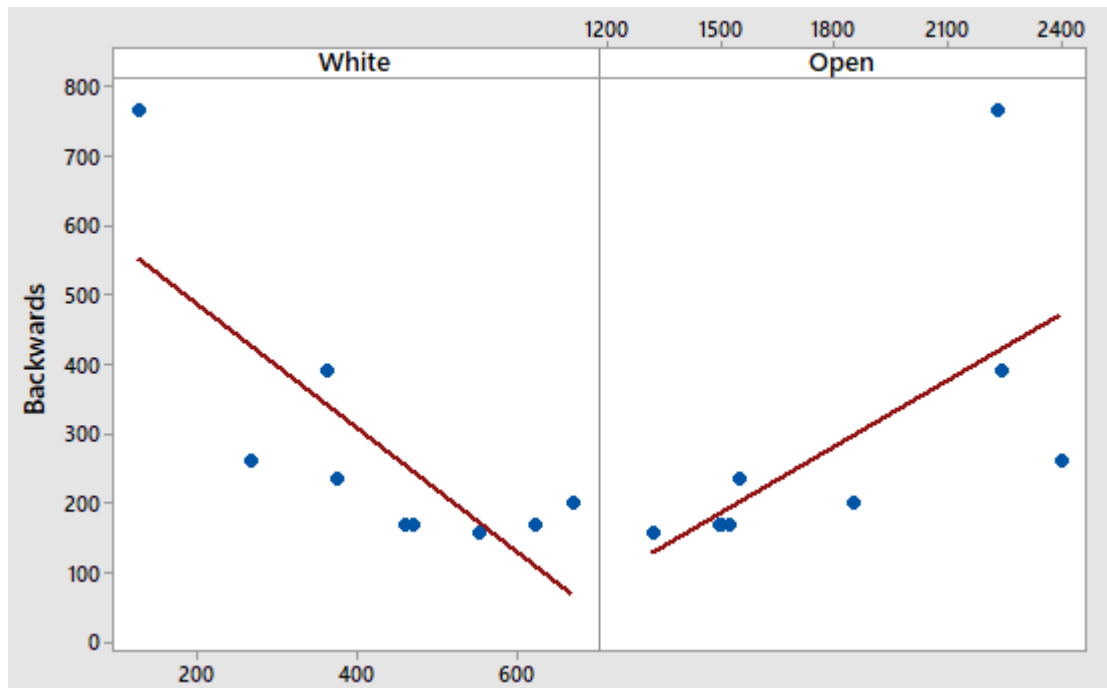


Figure 5.19 Correlations between ear position 'backwards' duration (seconds) and eye position (eye white/sclera on the left and eyes open on the right) duration (seconds).

5.4 Discussion

This study aimed to trial the proposed methodology for the full study (Chapter 6) to allow for refinement and optimisation of the methodological design whilst investigating potential behavioural (eye bias, eye position, and ear position) and physiological (eye temperature and heart rate) markers of interest measured during the attempted induction of positive affective state via three methods (provision of high value food, addition of a conspecific companion, and verbal praise). The results obtained are discussed below with reference to the methodological design, induction methods and behavioural and physiological measures used within this pilot study. Some interesting significant correlations were identified including

positive correlations between ears back and eyes open, left eye bias and left eye temperature, left eye bias and sclera and eye half duration, as well as a negative correlation between ears back duration and sclera duration.

5.4.1 The methodological design

This study utilised a repeated measures experimental design where each equine received each treatment, counterbalanced by treatment order randomisation to reduce any effect of treatment order on the results. The utilisation of this within-subjects design likely reduced the effect of individual differences on the results reported (Verma, 2015). The demographic variance within the subjects was broad in terms of equine breed, age, sex and height which allowed for evaluation of a cross-section of the equine population however the sample size was not sufficient to allow for further blocking within data analysis to evaluate any differences in the results related to these demographic factors. The small sample size significantly limits the generalisability of the results to the equine population however application across the equine population was not an aim of this pilot study.

The processes undertaken to habituate each subject to the study area and equipment, and to exclude equines who did not habituate appeared to be valid; with 92% of equines habituating fully to the procedure. All equines (n= 23) who were deemed habituated went on to complete the trial successfully indicating that the inclusion criteria for continuation into the trial were valid. The identification of two horses who required exclusion from the trial suggests that the exclusion criteria were sensitive to those individuals displaying behaviour indicative of non-habituation to the conditions.

The test procedure was easily replicated across study sites suggesting that providing there was a covered test area at each potential study site, with the capacity to house the study horse and a conspecific close by, the study equipment and process could be replicated across multiple sites, a facet deemed to be important in behavioural trials of this nature (Edmond *et al.*, 2019) and one which would be a requirement in the final trial (Chapter 6). Within the test process itself it was noticed that there may have been carry-over effects from one induction trial to another due to the short wash-out period utilised within this pilot study. A carry-over effect may occur in any situation where an individual's prior experience influences their current performance in a given situation (O'Connor *et al.*, 2014), and in this study the wash-out period of fifteen minutes may have been insufficient to allow for physiological and

behavioural variables to return to base rate, evidenced by continued elevated arousal, from base rate, in some horses in between induction trials. In addition, the univariate analysis undertaken within this study may have led to the results being influenced by any carry-over effects, suggesting that multivariate analyses may be more appropriate in this situation (Diaz-Uriarte, 2001), learning taken forward to the final trial (Chapter 6).

5.4.2 Induction methods for positive affect in equines

The induction methods utilised within this pilot study were taken from the results of Round One of the Delphi consultation due to the extended timescales of the Delphi necessitating initial Round One data to be used for timely undertaking of the pilot study in preparation for the final trial. Three induction stimuli were selected to allow for assessment of stimuli which were deemed to have potentially differing levels of ethological value to the subjects. The food and companion stimuli were highly rated by the Delphi experts in round one and were both deemed to be highly salient to equines due to their relative ethological value in survival and maintenance (Goodwin, 1999; Kieson *et al.*, 2020). However, it is likely that these stimuli have differing levels of significance to individuals depending on their prior experience and current environment, therefore it was hypothesised that responses to these stimuli may differ between subjects (Christensen *et al.*, 2008a; Hartmann *et al.*, 2011; Strand *et al.*, 2002). Furthermore, verbal praise was selected as a stimulus rated low by the Delphi experts with the intention that it would allow for comparison between the high and low rated induction methods, as well as allowing for further investigation of this method, which is commonly used within the practical equine field as a reinforcer (Heleski *et al.*, 2008). On reflection, it may have been useful to include two low ranked induction stimuli to balance the experimental design by utilising two high rated and two low rated methods.

5.4.3 Behavioural measures of positive affect in equines

Eye bias

As anticipated for a prey species with limited binocular vision (55-65°) and laterally placed eyes, frontal approach, application and removal of a stimulus was predominantly viewed with both eyes across all trials and phases. In this study both eye dominance was higher during food application than for companion and praise trials. However, previous research suggests that hemispheric specialisation exists across taxa but especially in species, such as equines, with laterally placed eyes (Rogers and Andrew, 2002). Almost complete decussation of optic

fibres in equines provides the potential for relatively accurate quantification of lateralisation in this species (Saslow, 2002). In the current study, during stimulus application there was an increase in left eye bias compared with right eye bias across all treatments. Levels of right eye bias were highest when the companion stimulus was provided but remained lower than left eye dominance throughout all trials and treatments. These results are interesting as they indicate that presentation of a positive stimulus results in left eye lateralisation and thus right hemispheric dominance which contradicts the findings of some studies but is supported by others.

Studies across species have identified the use of the left eye (right hemisphere) in unfamiliar situations and the right eye (left hemisphere) in familiar situations (Larose *et al.*, 2006). In humans, Bogen (1985) suggested that the right hemisphere specialised in the processing of negative emotional stimuli and the left specialised in processing positive emotional stimuli, a concept followed up by Davidson (1992) who studied hemispheric activity in emotion. However Leliveld *et al.* (2013) identified significant contradictory evidence within the literature with 43 studies suggesting right hemispheric dominance in processing negative affective states and 20 papers reporting the opposite finding providing support for the left eye bias identified in this study to be indicative of positive emotional state.

Larose *et al.* (2006) identified that equines with higher emotionality scores used their left eye more often to gaze at novel objects indicating that eye bias may be related to arousal more generally rather than valence *per se*. Within this study subjects demonstrated signs of higher arousal when treatments were applied, with individual differences between stimuli recorded in the behavioural notes. Some equines displayed signs of mild frustration (Pannewitz and Loftus, 2023) during different phases of treatment, most noticeably after food removal which may account for the maintenance of left eye bias after stimulus presentation. However these factors were not measured within this study but were consequently taken forwards for inclusion consideration within the final study (Chapter 6). Des Roches *et al.* (2008) reported similar findings to (Larose *et al.*, 2006), also identifying that both hemispheres appeared to be important in the processing of emotional stimuli. Farmer *et al.* (2010) reported right hemispheric lateralisation in equines processing affiliative conspecific social interactions, a finding supported by the results of this study, and their findings showed similar eye bias patterns to those reported by Austin and Rogers (2012) during negative interactions; further

supporting the theory that left eye bias may indicate emotional processing as a whole rather than just negative emotional processing. Furthermore eye lateralisation may differ at an individual rather than group level as suggested by (Baragli *et al.*, 2021) who identified clear lateralisation preferences within individuals but not across the group of 77 equines as a whole. This suggests the need for analysis of the individual as well as the whole group when considering lateralisation data, a factor that was not considered within the analysis of this study. From the results of this study, it appears unlikely that eye bias alone is sufficient to reliably identify valence of emotional state, a concept that is supported within the literature (Hall *et al.*, 2018), however it may be useful in conjunction with other behavioural and physiological measures to build a more complete picture of levels of valence and arousal.

Eye position

Open eye was the dominant position during stimulus presentation across all treatments with sclera duration lowest during stimulus application, especially in the companion treatment. Eyes half open was highest across treatments during stimulus application, especially the companion treatment and lowest after stimulus presentation. Sclera and eye half open were also significantly correlated with left eye bias in the during and after phases of the trials across induction methods. Eye positions reported in the current study are comparable with those reported by Lansade *et al.* (2018) where eyes open was the predominant eye position during both gentle grooming and standard grooming suggested to be indicative of positive affective state. In addition eyes half open was seen significantly more often during the presentation of the more positive stimulus (gentle grooming) in the Lansade *et al.* (2018) study and discussed by the authors as being indicative of a greater level of positive affect; a finding replicated in the current study. Sclera was the position identified the least by Lansade *et al.* (2018) contrary to the findings in this study where sclera was significantly correlated with left eye bias during and after most treatments, albeit frequency and duration still remained low. This difference could be attributed to the potential for food and companion stimuli being more arousing than the gentle grooming utilised by Lansade *et al.* (2018) with the potential for elements of frustration (Pannewitz and Loftus, 2023) during treatment in these groups and possible conflicted emotions in the praise treatment, especially as sclera has previously been attributed to negative emotional processing (Bell *et al.*, 2019b; Hall *et al.*, 2018; Siniscalchi *et al.*, 2021). In order to further investigate eye position as a measure of affective state another

variable, blink rate, could be added to the measures to identify the presence or absence of a stress response with increased blink rate reported to be highly correlated with other physiological variables indicative of stress and *vice versa* (Mott *et al.*, 2020).

Ear position

Three ear positions were evaluated during this study as described by Lansade *et al.* (2018). Ear position varied between treatments and phases with higher levels of ears backwards during food and higher levels of ears forwards during companion treatments. Ear asymmetry was highest after companion with increases in ear asymmetry also seen after the food stimulus. It is unclear whether these asymmetries were lateralised as this aspect was not measured within the study but may usefully be in future studies as described by Smith *et al.* (2018) who identified significant differences in equine ear asymmetry between exposure to positive and negative human vocal signals.

Ear asymmetry, a proposed indicator of negative affect (Lansade *et al.*, 2018) was identified significantly more after positive stimulus application and was not seen at all during the companion stimulus; this and the evidence that changes in asymmetry did not feature significantly during the praise treatment phases supports the proposal that ear asymmetry may readily identify a more negative emotional state than ears forwards. Ears forwards duration was seen to rise during the companion phase indicating a potential positive affective state, when supported by the other behavioural and physiological measures associated with this phase. The change in ear position predominance from forwards to asymmetrical and backwards after the companion stimulus indicates that this cohort may have found the removal of the companion negative; similarly, the ear position in the food trial also altered to predominantly asymmetrical after the food trial indicating a potentially more negative and conflicted or frustrated emotional state, this is potentially supported by a rise in heart rate seen in conjunction with asymmetrical ears. These results may indicate that the valence and relative importance of a stimulus may be more easily understood if changes in ear position are evaluated both during and after the stimulus presentation (Hall *et al.*, 2018; Lansade *et al.*, 2018), a facet that should be taken forwards to the final trial.

Interestingly ears backwards was the modal ear position during the food stimulus which, as a standalone marker, could be suggested to indicate a negative affective state. However it is

noted that ears backwards was positively correlated with eyes open and eyes half and negatively correlated with sclera indicating that ears backwards could indicate a more positive valence when considered in association with other variables such as eye position, eye bias and eye temperature as well as heart rate (which showed a lowering trend associated with ears backwards) and overall body posturing, a finding also reported by Lansade *et al.* (2018). Both grooming and eating have previously been associated with endorphin release (Keverne *et al.*, 1989; Dunbar, 2010; Rokade, 2011) which could explain the similar physiological and behavioural response between stimuli however it is also important to consider the discrete emotional experience of the individual to a particular stimulus where some individuals may find certain stimuli delivery and removal more arousing (either positively or negatively) than others (Barrett, 2017b; Barrett *et al.*, 2001). Ears backwards during eating may also be linked anatomically to the physical process of eating and also potentially to a feeling of apprehension relating to the method of food delivery (by a person rather than in a bucket on the floor) which may have been perceived by some equines as a threat to their food security; it is notable however that the equines within this study did not have their ears pinned back in a defensive or offensive manner, rather they were slightly dropped backwards. Indeed ears forwards and backwards were associated with lower levels of left eye bias, with the inverse seen for asymmetrical ears potentially indicating that ears backwards and forwards may be indicative of lower emotional arousal overall compared with ears asymmetrical as reported in the literature (Peters *et al.*, 2012).

5.4.3 Physiological measures of positive affect in equines

Eye temperature

Eye temperature remained relatively constant across treatments and phases but was higher during praise compared with during food and companion. Eye temperature dropped during the food and companion stimuli, but not significantly so, and were markedly raised after companion removal. These results may suggest that the subjects showed a stress response to the praise treatment and also when the companion was removed, as reported by Fenner *et al.* (2016) who suggested that rises in eye temperature are potentially indicative of a mild stress response.

The lower eye temperatures recorded during the food and companion treatment phases may indicate a level of positive arousal during these treatments compared with both praise and

the after phase of the companion treatments (Kappel, 2023). Within the current study continuous data collection was challenging due to subject head movement resulting in fewer than optimal measurements which may have affected the data collected, a methodological issue also experienced by Uddin *et al.* (2019). Furthermore, eye thermography has been shown to be influenced both at an individual level by endogenous factors (sex and breed) and more globally by environmental factors (situation, location, time of year) and as such should be used carefully with these factors controlled for and used alongside other behavioural and physiological measures (Jansson *et al.*, 2021).

Despite the challenges, these results do provide encouraging foundations that eye temperature could be used as a measure of arousal and may, when paired with other data be predictive of positively valenced emotional states. Previous studies have used eye temperature to measure physiological stress in equines (Redaelli *et al.*, 2019) with eye temperature shown to be distinct from body temperature (Jansson *et al.*, 2021) and therefore useful as a measure independent of the effects of physical exertion. Other studies have utilised eye temperature to identify psychological stress in equines (De Mira *et al.*, 2021; Hall *et al.*, 2014) with eye temperature seen to increase alongside other markers of stress, suggesting that a lower eye temperature may indicate lower arousal and potentially a more positive emotional state (König v. Borstel *et al.*, 2017). However other theories suggest that lower eye temperature signifies higher levels of arousal due to an increase in SNS-mediated changes in blood flow to the lacrimal caruncle and medial canthus leading to a reduction in surface temperature in this area (Stewart *et al.*, 2008; Kappel, 2023). The evidence suggests that eye temperature is likely to be a useful indicator of arousal however the consensus regarding its use as a measure of valence is much lower. Some theorise that, in line with the emotional valence lateralization hypothesis (Davidson, 1992; Leliveld *et al.*, 2013), asymmetric temperatures within bilateral facial areas such as the eyes may be linked to valence (Ramirez Montes De Oca, 2021); however there is much debate regarding the mechanisms for this. Regardless of data interpretation methods, it is important that conditions for thermography use are set and maintained throughout trials in order to optimise reliability of results; these conditions include maintaining a constant distance from the subject, presentation of the camera at an angle of 90° to the sagittal plane (Ijichi *et al.*,

2020) and adjustment for ambient temperature, humidity and light conditions (Okada *et al.*, 2013).

Heart rate

Whilst heart rate did not vary significantly amongst treatments and phases, trends were observed with heart rate reducing slightly during praise and increasing during and after food. Most noticeably heart rate decreased during companion but was markedly raised when the companion was removed. Heart rate was selected as a measure due to the absence of physical activity within the trials which may have otherwise confounded the data collected in this study, and has been used in prior work as an indicator of emotional state (Wascher, 2021). Indeed, decreases in heart rate have been identified as markers of positive emotional state relating to affiliative behaviours (Aureli *et al.*, 1999; Boccia *et al.*, 1989; Laister *et al.*, 2011) a finding supported by the results of the current study where heart rate decreased during affiliative conspecific interactions and was strongly associated with a trend for 'ears forwards'. Increases in heart rate (exclusive of increased physical exertion) are noted in situations of negative emotional state such as agonistic interactions (Wascher, 2021) or potentially the removal of a positive stimulus within the current study where heart rate increased above that seen during the stimulus; but also in situations of positive valence but high arousal (Ellenberg *et al.*, 2013); potentially including the delivery of feedstuffs as seen in some of the subjects within this study. Heart rate has also been shown to increase during learning tasks in equines (Mengoli *et al.*, 2014) however the valence during the task was unclear in Mengoli's study.

Alone, heart rate cannot reliably determine affective state due to the identified increases during both positive and negative experiences (Briefer *et al.*, 2015; Reefmann *et al.*, 2009) but alongside other measures and ideally with the inclusion of heart rate variability parameters (Scopa *et al.*, 2018) it could be a useful measure of arousal and aid in the disentanglement of the effects of negative and positive stimuli (Wascher, 2021; Wascher *et al.*, 2018; Trondrud *et al.*, 2021). It is also suggested that the longitudinal assessment of heart rate in an individual may provide further information regarding consistency of phenotype, including personality, leading to a greater knowledge of how behaviour and physiology of an individual affects decision making (Careau *et al.*, 2008; Roche *et al.*, 2016) and repeated measures of heart rate can help to distinguish between adaptive and maladaptive responses to stimuli (Selye, 1973; Proudfoot and Habing, 2015).

5.4.4 Limitations

Within this pilot study several limitations were identified which required further consideration when planning the later trial. Initially it is important to consider that a limited number of measurement indicators were utilised within the study meaning that the measures selected may have not adequately captured reliable data relating to the ability to measure affective state. The decision to utilise a small number of measures was made to facilitate timely review of the planned methodology prior to the commencement of the full trial (Chapter 6) and to allow for initial exploratory analysis of the measures and their viability for inclusion within the later trial. Secondly, a relatively short wash-out period was utilised within this study, again to facilitate timely data collection, which may have resulted in carry-over effects from one trial to subsequent trials which could affect the responses seen. To reduce risk of this in the later trial a much greater wash-out period should be built into the methodological design.

A significant limitation of this study is the small sample size combined with the statistical approach. Within this small cohort it was difficult to identify any significant differences between treatments and trial phases and any differences identified are potentially vulnerable to type one error (Diaz-Uriarte, 2001) due to the use of univariate analyses. Whilst some significant correlations between measurement variables were identified it is unwise to attribute these as reliable and generalisable to a larger population due to the low study sample size. The univariate analyses used, whilst valuable in identifying potential trends, limited the interpretation of the results and for the final study an alternative multivariate analysis approach would be preferable to allow for reduction in the potential for type one error and to allow for a more cohesive evaluation of the relationship between measurement variables and any effects of treatment on the response variables. However, this pilot study identified interesting trends in terms of the differences within individual equine responses to the three induction methods highlighting potential between subject differences occurring because of individual personality, prior life experience and current environment. This finding is supported by the views of the experts who identified that careful preference testing of individuals is required to provide them with stimuli that are most likely to induce positive affective states and avoid negative states and high levels of arousal. Furthermore, this pilot study has identified the difficulty we face when attempting to decipher the valence of

emotional states, rather than just ascribing arousal levels. Previously studied negative states such as pain have been easier to study due to the more overt physiological and behavioural responses seen in many animals who are suffering, whereas indicators of positive emotional state, especially those in low arousal appear to be much more nuanced and subtle, making their reliable identification more difficult.

5.4.5 Recommendations

To ameliorate for the difficulties and limitations outlined above it is suggested that the use of preference testing within individuals is required prior to attempts to induce positive affect. Exploration of methods which appear likely to induce positive affect may determine a hierarchy of methods which are likely to induce positive affect across a broad spectrum of the population however there will always be individual difference to account for, meaning that high subject numbers within studies may mitigate, somewhat, this limitation.

When considering methods to measure positive affect it is unlikely that individual measures can definitively identify positive affect and as such a multi-measure approach is likely to be more informative and reliable. The results of this study suggest that there may be relationships between eye bias and eye temperature with each measure increasing in value together; stronger negative relationships were identified between eye bias and heart rate potentially indicating that these measures may have an inverse relationship. Various relationships were suggested between the different eye and ear positions depending upon trial phase and induction stimulus which require further investigation alongside the inclusion of other measures in a more complex statistical model to elucidate potentially useful combinations of measure for different valences and arousal levels of affective state. These facets are explored within the multivariate explorative analysis in the final trial (Chapter 6).

5.5 Conclusion

The ability to induce and measure positive affective state in equines is an important but, as yet, incomplete area of knowledge and application in the equine science and welfare field. Within this area the use of behavioural lateralisation as a tool to aid in assessing emotional state is very limited to date. This preliminary study has identified some interesting trends in left eye bias during the application of stimuli designed to induce positive affective state in equines which is indicative of right hemispheric dominance in emotional processing in response to these positive stimuli. Links between heart rate and the trial phase were

identified and were different between treatments, potentially highlighting differences in the salience of the different stimuli during different trial phases. In addition, heart rate decreased and left eye bias increased during food and companion stimuli potentially indicating a further link between left eye dominance and a more positive (lower arousal) emotional state. Ear position varied according to stimulus and phases; however, an interesting negative correlation was identified between left eye bias and ears back, potentially indicating a relationship between the reduction in an indicator of negative emotional state (ears back) and an increase in right hemispheric dominance. Whilst some findings appear contradictory, and the sample size is limited in this study, these potential markers of affective state provide interesting aspects for further study to elucidate easily identifiable behavioural signs of positive affect in equines.

This study suggests, through measurement of physiological and behavioural markers, that different positive stimuli may elicit different levels of arousal before, during and after their presentation and that differences may occur within individual equines likely in relation to their current environment and previous experiences of those stimuli. As such it is imperative that preferences for stimuli are tested within individuals prior to trials such as this and that the animal's general responses to the stimuli are recorded to add context to both physiological and behavioural markers. In addition, the individual personality of the equine may influence their responses to a variety of different induction stimuli, a facet that should be considered within further studies via the assessment of individual personality and the accounting for this within analysis of data gathered.

Whilst this study has several limitations, in relation to the type of data collected (no heart rate variability analysis especially) and methods of collection (such as difficulties in capturing reliable eye thermography data), it does provide useful information surrounding potential markers of positive affective state. It has also highlighted the potential propensity for some individuals to perceive intended positive induction methods as negative or very highly arousing according to individual preference or experience. This information should be further tested alongside additional measures with the aim of producing a series of induction and measurement methods that could be utilised by researchers and laypersons alike to begin to further validate measures of affective welfare in equines.

Chapter 6 Inducing and measuring positive affective state in equines: an exploratory study

6.1 Introduction

This chapter is the culmination of work undertaken in previous chapters which shaped the design of this final study. The chapter brings together the learning from all three rounds of three Delphi consultation (Chapter 3) and the pilot study (Chapter 5) to investigate a range of potential induction and measurement methods for positive affective state in equines. In addition, this chapter considers within its study design and analysis the effect of personality (Chapter 4) on individual responses to the utilised induction methods for positive affect.

6.1.1 Outcomes from the pilot study

The pilot study undertaken (Chapter 5) allowed for refinement of the methodology employed in this study including the consideration of study area set-up, wash-out period between treatments and methods utilised for data processing and analysis. Within the pilot study there were limited types of both behavioural (ear position, eye position, and eye bias) and physiological measurement (heart rate and eye temperature) methods utilised and therefore within this present study a greater range of measurement methods were employed, as suggested by Hall *et al.* (2018), to increase the probability of identifying trends in potentially useful measures. The temporal proximity of the trials to each other within the pilot study was identified as a potential confound with a high likelihood of carry-over effect between trials and therefore the wash-out period was significantly increased (from 15 minutes to a minimum of 6 hours) within this current study to match those identified as acceptable within the previous literature (Kang and Lee, 2016). Additionally, the data analysis within the pilot study took the form of exploratory relationship modelling via univariate statistical methods. This allowed for some identification of trends, in terms of increases and decreases in the measured variables, within results generated via the behavioural and physiological measures but did not allow for more complex modelling of relationships between the various factors. Therefore, within this study more complex multivariate analyses are undertaken to explore the data generated. Within the pilot study it appeared that individuals responded variably to the induction measures however their personality profiles were not assessed. Consequently,

within this study these measurements of personality are factored into the statistical models to allow for evaluation of any possible effects of personality on the results obtained.

6.1.2 Positive induction stimuli utilised within this study

To maximise the likelihood of inducing positive affect in the equine cohort within this study the induction stimuli utilised were those rated most highly by the expert group and supported by the literature. The cross-applicability of induction methods was also considered within the selection process whereby methods that may be highly appetitive to equines but only in a specific or limited circumstance (such as access to a substrate to roll in) were excluded (Table 6.1). This process resulted in the selection of the following four induction methods: (i) High value food provision, (ii) the addition of an affiliative conspecific, (iii) wither scratching and (iv) application of positive training via target training.

Table 6.1 Ranked induction methods from the Delphi consultation with exclusion rationale for excluded methods.

Rank	Induction method	Included in this study	Exclusion rationale
1	High value food provision	Yes	-
2	Known affiliative conspecific	Yes	-
3	Substrate to roll in	No	Value of stimulus is likely to be very variable depending on the current environment and physiological state of the animal.
4	Wither scratching	Yes	-
5	Roughage provision	No	May not be of a high enough biological value to individuals.
6	Medium value food provision	No	High value food provision selected instead as likely to be more salient to individuals.
7 & 8	Access to graze and browse and turnout	No	Would require equines to be at liberty therefore reducing the measurement variables that could be employed.
9	Human-horse interaction via training	Yes	-
10	Massage	No	Would require an accredited equine physiotherapist and extended trial length.

11	Food enrichment toys	No	A food related method had already been selected. Limited research on the application of different types of food enrichment toys would have hampered selection of the toy. Potential for increased frustration if unable to access food from the toy.
12-14	Tail scratching, grooming and neck scratching	No	Wither scratching, supported with greater evidence from the literature, already selected.
15-16	Problem solving and novel object provision	No	Potential for frustration and fear responses respectively.
17	Low value food provision	No	May not be of a high enough biological value to individuals.
18	Play toys	No	Would require equines to be at liberty therefore reducing the measurement variables that could be employed.
19-22	Verbal praise, removal of restraint, forehead rubbing, and aromatherapy	No	Very low ranking by experts.

High value food provision

As the highest ranked induction method (Chapter 3) and one well supported by the literature (Kieson *et al.*, 2020; Ninomiya *et al.*, 2004) this method was included within the study. This induction stimulus was of particular interest due to its high value, ethologically, to equines (McGreevy, 2012) but also due to the potential for it to induce alternative affective responses to those aimed for within the study. For example, it is recognised that high value food provision may induce higher arousal levels in some equines (Kieson *et al.*, 2020) which may lead to anticipatory behaviour (Peters *et al.*, 2012) and frustration related responses (Pannewitz and Loftus, 2023) rather than a low arousal positive affective state. Arousal activates the hypothalamic-pituitary-adrenal (HPA) axis stimulating the sympathetic nervous system (Feh and de Mazières, 1993); as such this method was selected for inclusion but not as the ‘gold standard’ measure. Within the pilot study ‘eyes half’ (a potential marker of relaxation) was identified more frequently in the food group and as such this supports its inclusion as a measure within this study.

In addition, the relative value of perceived high value feedstuff within individuals, according to their personality and prior experiences, was deemed likely to result in a range of affective

experiences to induction using this method (Elia *et al.*, 2010), a factor which was identified as potentially useful within later modelling of the measures employed to evaluate affective state achieved.

Addition of a known affiliative conspecific

Ranked second by the experts, and supported within the ethological literature (Heitor *et al.*, 2006; Kimura, 1998; Ransom and Cade, 2009), this measure was included within the study as a potential induction method for positive affect in low arousal in equines. To utilise the addition of a conspecific to induce positive affect in equines an understanding the nature of the relationship between the two parties is critical. This can be achieved through careful examination of their interactions in the period leading up to the trial and to consider carefully both the proximity placement of the added conspecific and the valence and arousal levels of said conspecific to avoid inducing negative affect. Proops *et al.* (2009) identified that horses are capable of cross-modal individual recognition using auditory, visual, and olfactory information. They utilise this to recognise known conspecifics and to identify unfamiliar conspecifics which is important when a horse is added to the social group or proximity of the individual. Horses are also proficient at discriminating between facial expressions of conspecifics and show behavioural and physiological changes in response to differing facial expressions suggesting that they glean important social information about a potential interaction through analysis of the other horse's facial expression (Wathan *et al.*, 2016). Kimura (1998) identified that horses changed their preference for social partners over time and that the time of year significantly affected preferred associate relationships. This means that it was important within this methodological design that recent review of preferred conspecifics had taken place to present a preferred social partner during the trial. In addition, it is acknowledged that emotional contagion between conspecifics can multiply both positive and negative emotions in each animal and effect changes in affective state, thus it is imperative that any introduced individuals are also in a state of positive affect to avoid negative changes in affective state of the subject (Spinka, 2012). Within the pilot study the addition of an affiliative conspecific was associated with significantly lower heart rate and higher instances of 'ears forward' which also leant support to its inclusion as a measure in this trial.

Wither scratching

Whilst ranked fourth by the experts, wither scratching is highly supported within the equine literature as a phenomenon that is broadly experienced as positive by equines. It is thought to contribute to relationship building and comfort (McDonnell, 2003; Waring, 1983) and has been observed between heterospecifics (Fraser, 2010; McGreevy, 2012) with research identifying links between grooming and/or massaging around the wither area and reduction in heart rate combined with relaxed behavioural responses (Feh and de Mazières, 1993; VanDierendonck and Spruijt, 2012; Lansade *et al.*, 2022; Lansade *et al.*, 2018). Therefore, wither scratching may induce a positive affective state similar to that seen in allogrooming. Prior work has demonstrated a preference for scratching of the wither over other areas such as the neck or face (Thorbergson *et al.*, 2016) and as such this was selected as the focal scratching area within this study. However, in line with previous evidence surrounding individual preferences (Lansade *et al.*, 2022; Janczarek *et al.*, 2018), wither scratching method and exact location were tested in all equines within this study prior to its use as an induction method. Due to its prior recognition as a method to induce calmness and relaxation in horses (Rochais *et al.*, 2014; Watson and McDonnell, 2018) this method was selected as one which may induce positive affect with low arousal and due to the strength of evidence within the literature, was positioned as the gold standard method against which other induction methods would be considered within the analysis of the data gathered. Nishimuta *et al.* (2022) identified that tactile reinforcement could effectively be used to shape new behaviours in horses (and mules) but reported that the reinforcement delivery process is likely to be an important factor in ensuring efficacy of the approach. Preference over choice of body area for tactile reinforcement was highlighted as a particularly important facet in training success, a suggestion supported by Feh and de Mazières (1993), Lansade *et al.* (2018), Normando *et al.* (2003) and Watson and McDonnell (2018).

A simple target training task

Engagement with a simple training task (target training) using positive reinforcement training methods was ranked 9th by the experts but was the fourth most highly ranked induction stimulus considered practically viable for use within this trial. Methods ranked more highly than this were excluded due to their relative ambulatory nature (turnout, grazing and browsing) which would have limited data collection; and similarity to other selected measures

(medium value food and roughage provision) or transient value (substrate to roll in). This method was deemed to be a novel method which could have good application for inducing positive affect but one that is not widely researched to date. This technique has been previously proposed as a method to optimise equine welfare (Hausberger *et al.*, 2019) and although it remains unutilised within equine studies attempting to induce positive affect there is evidence of its efficacy in dogs and wolves (Vasconcellos *et al.*, 2016). Furthermore, in pigs enhanced attention and cognitive activity has been shown to induce greater positive emotional states (Puppe *et al.*, 2007), a finding that may be transferable to equines and one considered by Sankey *et al.* (2010b) and Sankey *et al.* (2010a) within their studies looking at reinforcement and positive long-lasting memories in equines, indicating that human attention may be a source of positive emotions for equines. In terms of training, handlers with a more positive attitude have been shown to help horses negotiate courses with greater success and increase horse confidence during human interactions (Chamove *et al.*, 2002; Lesimple and Hausberger, 2014). In addition, humans have been suggested to help horses approach novel objects more confidently (Henry *et al.*, 2017). However, Ijichi *et al.* (2014) found no discernible difference between horses responses to tasks when accompanied by a familiar or unfamiliar person suggesting that it may be the mere presence of a human companion, rather than the specific companion, that promotes security in a horse with no surrounding conspecifics; a factor that would fit well with equine ethology as a social prey species. Using positive reinforcement training may help to reduce negative affect (Freymond *et al.*, 2014) and achieve the desired behaviour or completion of the task with associated higher levels of motivation (Innes and McBride, 2008). To reduce the potential for frustration or other negatively valenced emotions during training (Pannewitz and Loftus, 2023) it was decided that the training would be undertaken by a professional trainer (the experimenter) who was adept at and confident in delivering target training with horses (Chamove *et al.*, 2002; Lesimple and Hausberger, 2014). In addition, a previously conditioned secondary marker (verbal 'yes' or 'good') would be utilised alongside a low value food reinforcer (chaff or other forage) to negate, as much as possible, any crossover in effect between this trial and the high value food trial (Kieson *et al.*, 2020).

6.1.3 Behavioural measures to evaluate positive affect in equines

Following the pilot study (Chapter 5) it was identified that a wider range of behavioural measures should be utilised within this study to expand the potential to identify measures that may be indicative of positive emotional state in equines. Therefore, based on the results of the Delphi expert consensus (Chapter 3) and evidence within the literature, four modes of behavioural assessment were undertaken within this study. These were selected on a ranking basis with exclusion of methods deemed infeasible for inclusion within this study. Excluded methods, alongside their rationale are further detailed in Table 6.2. The measures selected were (i) body language assessment, (ii) facial action scoring, (iii) qualitative behaviour assessment including absence of frustration and repetitive behaviours, and (iv) eye bias analysis.

Table 6.2 Ranked behavioural measures from the Delphi consultation with exclusion rationale for excluded methods.

Rank	Behavioural measure	Included in this study	Exclusion rationale
1	Horse-horse social interactions	No	Would require equines to be at liberty therefore reducing the measurement variables that could be employed.
2	Body language assessment	Yes	-
3	Facial action scoring	Yes	-
4	Play behaviours	No	Would require equines to be at liberty therefore reducing the measurement variables that could be employed.
5	Qualitative behaviour assessment	Yes	-
6-8	Human-animal relationship testing, cognitive bias testing and time budget analysis	No	Measures not suited to the trial duration within the methodological design.
9	Absence of frustration	Yes	-
10	Assessment of vocalisations	No	Minimal vocalisations anticipated within the induction trials based upon evaluation of the pilot study.
11	Spontaneous blink rate analysis	No	Lower ranking by the experts and limited analysis time resulted in omission.

12- 13	Attention bias testing and novel object tests	No	Not feasible within the timeframe of the study nor the methodological design.
14	Time in REM sleep	No	Not feasible within the timeframe of the study nor the methodological design.
15	Absence of repetitive behaviour	Yes	-
16	Eye bias analysis	Yes	-
17	Biomechanical analysis	No	Not feasible within the timeframe of the study nor the methodological design.
18	E-BARQ	No	Very low ranking by experts and limited validation within the literature in relation to this study's purpose.

Body language and facial action scoring

As highly ranked methods (2nd and 3rd respectively) the scoring of body language, focusing on facial actions, was identified as a potentially useful measure of affective valence in equines. Evidence within the literature provides a multitude of ethograms for equine behavioural assessment (Torcivia and McDonnell, 2021; Dyson, 2022; McDonnell, 2003; McDonnell and Poulin, 2002) and inter-rater reliability is often found to be high across their use (Pierard *et al.*, 2019; Dyson and Van Dijk, 2020) making them useful measures in equines. Within this study a focus on ear and eye position was determined based upon prior work by Lansade *et al.* (2018) which indicated that changes in these features may be useful markers of positive affect in equines. Within the pilot study distinction was not made between the presentation of different types of 'ears backwards' which was felt to be a limitation as it could be a significant factor in determining affect. Therefore, varying types of ear position were utilised within this current study including 'ears gently backwards' and 'ears fully backwards' to allow for potential interpretation of the nuance of ear positions. Prior research has also linked ear position with eye bias (Peters *et al.*, 2012), a finding that we hoped to follow-up on within this study. As with ear position, our pilot study results appeared to support those of Lansade *et al.* (2018) with a predominance of 'eye open' displayed. However, the Lansade study identified higher frequency of 'eyes half open' during interventions that were perceived to induce positive affect and as such this appeared a potentially valid indicator worthy of further investigation. Potentially 'eyes half open' and 'eye white' could be at opposite ends of the affective valence scale indicating positive and negative affect respectively and this hypothesis

requires further testing following on from the work of Lansade *et al.* (2018) on positive affect and Siniscalchi *et al.* (2021) and Bell *et al.* (2019b) who looked at negative affect.

Body language was considered more globally within qualitative behaviour assessment discussed below.

Qualitative behaviour assessment (QBA)

The use of QBA protocols in welfare assessment is well documented within the equine literature (Minero *et al.*, 2018; Hausberger *et al.*, 2016; Hintze *et al.*, 2017; Ruet *et al.*, 2022) and the previously validated AWIN indicators (AWIN, 2015) provide a useful and reliable framework for equine QBA studies (Dalla Costa *et al.*, 2016a). Delphi participants suggested that a fixed-terms approach using predefined descriptors would be the most suitable method (Wemelsfelder *et al.*, 2001). The use of QBA to explore emotional state in horses has recently been undertaken by Minero *et al.* (2018), and Estevez *et al.* (2017) discussed the use of the AWIN application across species. As this method has previously been identified as being repeatable and reliable as well as highly feasible 'on farm' (Dalla Costa *et al.*, 2014c; Hintze *et al.*, 2017) but in need of further validation alongside physiological measures (Hall *et al.*, 2018) it was identified that this further validation could be undertaken via this study. As a highly ranked behavioural measure (ranked 5th) this method was utilised by adopting the full AWIN QBA methodology throughout evaluation of the video footage in this study. Due to the strength of support for this measure within the literature and the availability of an already validated assessment system this method was considered the 'gold-standard' behavioural measure against which other measures would be evaluated.

Absence of frustration and repetitive behaviours

Behavioural signs of frustration and repetitive behaviours such as crib-biting, wind-sucking, box-walking and weaving are deemed to be markers of acute or chronic welfare compromise (Pannewitz and Loftus, 2023; Lesimple, 2020) which may be interlinked (Cooper and McGreevy, 2007) and are associated with negative emotional states (McAfee *et al.*, 2002; Rørvang *et al.*, 2021). Absence of frustration was ranked 13th in round three of the consultation and 9th overall as a measure that may indicate positive affect. In equines, frustration is thought to indicate a negative affective state with chronic frustration linked to poor welfare (Lesimple, 2020; Mellor *et al.*, 2020). Therefore, an absence of frustration could

be an indicator of, at least, the absence of negative affective state but when combined with other variables, potentially also an indicator of the presence of positive emotional state. Within this measure the concept of the presence of repetitive behaviours could also be considered utilising the expert derived measures of frustration from a tangential project relating to this study (Pannewitz and Loftus, 2023) whereby a Delphi consultation was undertaken to identify measures of acute and chronic frustration in equines. Experts in the Pannewitz and Loftus (2023) study identified repetitive behaviours as a sign of chronic frustration and within the Delphi consultation contained within this thesis experts reported the absence of repetitive behaviours as one measure that could contribute to excluding negative affect, thereby potentially helping to identify positive affect. Within the Delphi consultation the experts agreed that positive affect could not co-vary with the presence of frustration and repetitive behaviours and as such the presence of either of these indicators would exclude a determination of positive affective state. As such these measures were included as a 'check and balance' measure so that if identified their relationship with other measures could be examined to aid in identification of markers that may indicate multivalence from markers which appear to indicate solely positive valence.

Eye gaze (eye bias/lateralisation) analysis

Whilst ranked 17th within the expert rating of proposed behavioural measures due to the limited evidence for the reliability of the measure, the Delphi experts were interested in the potential use of eye gaze analysis for assessing affective state in equines. Across other species lateralised behaviours have been more widely studied (reviewed by Rogers *et al.* (2013)) and theories regarding lateralised hemispheric processing of positive and negative stimuli have been widely postulated (reviewed by Leliveld *et al.* (2013)). Disagreement within the literature regarding right and left sided processing dominance in positive and negative affective states suggests that further research into this phenomenon is warranted and as such it was deemed important to include this more controversial measure alongside more highly validated and widely accepted measures to attempt to elucidate further information on this variable. Within the pilot study it was identified that left eye bias increased during provision of perceived positive stimuli, a concept supported by some studies but refuted by others (as summarised by Leliveld *et al.* (2013)) and therefore this measure was considered worthy of further investigation. Interestingly Baragli *et al.* (2021) identified significant differences in

lateralised eye tendencies in individuals regardless of the valence of the stimuli presented and as such the assessment of both right and left eye bias appeared logical in this experimental design to measure differences in the presentation of different eye bias directions.

6.1.4 Physiological measures to evaluate positive affective state in equines

Similarly to the behavioural methods, the pilot study (Chapter 5) identified that a greater number of physiological measures warranted inclusion within this study and the Delphi consultation (Chapter 3) provided clear expert consensus on methods that could be further explored in relation to measuring positive affective state in equines. Four physiological methods were selected for inclusion within this study based upon the expert ranking and support for their inclusion within the literature. These methods were: (i) heart rate and heart rate variability analysis, (ii) salivary cortisol measurement, (iii) respiratory rate measurement and (iv) infrared thermography of the eye and ear. Several potential methods were excluded, see Table 6.3 for exclusion rationale. The measures chosen were selected to provide balance between previously well validated measures in affective state recognition in equines and more novel methods that may gain increased validation through inclusion within this study.

Table 6.3 Ranked physiological measures from the Delphi consultation with exclusion rationale for excluded methods.

Rank	Physiological measure	Included in this study	Exclusion rationale
1	Heart rate variability	Yes	-
2	Salivary cortisol	Yes	-
3	Heart rate	Yes	-
4	Respiratory rate	Yes	-
5-7	Serotonin levels, electroencephalogram, blood oxytocin	No	Either invasive measures or equipment availability reduces viability for the method.
8-10	Faecal, blood and urinary cortisol	No	Blood cortisol requires invasive methods and salivary selected over urinary and faecal cortisol due to increased feasibility.
11 & 12	Gut microflora assessment, skin conductivity/sweat rate	No	Equipment and expertise unavailable.

13	Infra-red thermography	Yes	-
14	Hair cortisol	No	A longer-term measure of stress.
15	Eye temperature	Yes	-
16	Tympanic membrane temperature	No	External ear temperature selected instead.
17-21	Corticotrophin releasing factor challenge, catecholamine levels, immune function tests, body temperature and tryptophan levels	No	Invasive measures.
22-23	MRI and CT scans	No	Equipment and expertise unavailable.

Heart rate and heart rate variability analysis

Ranked 1st and 3rd overall respectively heart rate variability and heart rate were deemed by the Delphi experts and through evidence within the literature to be valid measures that have previously been utilised within experimental designs aiming to study affective state and welfare in equines (Von Borell *et al.*, 2007; Stucke *et al.*, 2015; McDuffee *et al.*, 2019). As well as being widely used to assess negative affective states, these measures have prior evidence supporting their use as measures of relaxation and appetitive motivation in animals, including horses (Boissy *et al.*, 2007b; Scopa *et al.*, 2020; Feh and de Mazières, 1993). There is also a wide array of methods papers available which have evaluated the various metrics and methods used within heart rate and HRV analysis in equines and other species (McDuffee *et al.*, 2019; Stucke *et al.*, 2015; Von Borell *et al.*, 2007). It was hoped that this well validated measure may be useful when considered alongside less well validated measures to provide further evidence for the potential use of more novel measures within later experimental designs.

Salivary cortisol analysis

The use of salivary cortisol measurement to quantify responses to stressors, including exercise, is well documented within the equine literature (Fazio *et al.*, 2008; Fazio *et al.*, 2013; Covallesky *et al.*, 1992; Alexander and Irvine, 1998; McGreevy and Nicol, 1998) and is well validated as an accurate measure against blood cortisol values (Peeters *et al.*, 2011). The use of salivary cortisol testing in horses has predominantly focussed on establishing parameters for normal levels and reporting natural fluctuations due to circadian influence as well as quantifying stress responses to exercise and stimuli in the environment. Aurich *et al.* (2015)

investigated the effect of season, age, sex and housing variables on cortisol levels in horses over a six-month period to produce useful equine reference ranges for salivary cortisol and highlighting considerations which should be undertaken when planning experimental design.

Elevations in salivary cortisol are associated with increased activation of the hypothalamic-pituitary-adrenal (HPA) axis (Duran *et al.*, 2017) indicative of a stress response, therefore it follows that reductions in salivary cortisol may be useful as a measure of lower arousal levels and potentially of positive emotional responses in equines (Buchanan *et al.*, 1999). However, interpretation of salivary cortisol levels is not necessarily straightforward and some evidence suggests that low levels may be connected to chronic poor welfare (Pawluski *et al.*, 2017) and conversely high levels may signify positive arousal (Ralph and Tilbrook, 2016). As such salivary cortisol was included, not only due to its high rank (2nd), but also as a measure to enable further consideration of the significance of changes in salivary cortisol levels in relation to differing emotional states in equines.

Respiratory rate

Ranked 4th overall, respiratory rate was a measure applicable within this study with support within the literature suggesting that changes in respiratory rate may be useful in assessment of valence as well as arousal (Reefmann *et al.*, 2009; Csoltova and Mehinagic, 2020; Boiten, 1998; Boiten *et al.*, 1994; Neethirajan, 2022). Whilst there is less empirical research into this metric as a measure of affective state in equines, some studies have linked lower respiratory rate with positive affect, calmness and affiliative responses (Briefer *et al.*, 2017b; Menchetti *et al.*, 2021). Further investigation of this phenomenon within this study may provide insight into its potential use as an affect marker in equines, one which could be easily measured and evaluated by researchers and lay persons due to its relative ease of measurement.

Infrared thermography of the eye and ear

Whilst thermography of the eye and ear was not one of the highest rated measures by the Delphi experts (ranked 13th) an increased interest in this measure within studies of affect in horses (Stratton, 2022; Kappel, 2023; Ramirez Montes De Oca, 2021) suggests that it is a marker of interest for which further validation may be achievable within future empirical trials. Thermography has gained popularity as a measure of emotion in equines due, in part, to the increase in available technology to measure this effect both using specialist cameras

and add-ons to smartphone or other smart devices. In a review of robust indicators of welfare in equitation science, Waran and Randle (2017) indicate that eye temperature may be a suitable metric with further validation. A study by Ramirez Montes De Oca (2021) identified lateralised increased nostril, eye and ear temperatures indicative of negative affective state during non-rewarded novel object trials and differences in lateralised responses indicative of positive affective state in rewarded trials. These findings are based on the grounding that there are ipsilateral relationships between nostril and ear temperatures and hemispherical brain activation and contralateral relationships between eye and ear temperatures and hemispherical brain activation, a phenomenon requiring further investigation in equines. Some studies suggest that thermography data may be a more reliable measure of arousal rather than valence (Travain *et al.*, 2016) due to its reported association with increases in HPA axis activity (Borstel *et al.*, 2017) identified through salivary cortisol measures by Cook *et al.* (2001). Within this study eye and ear thermography data are considered alongside measures with greater evidence to support them to identify potential relationships between more well validated measures and novel measures such as this one.

6.1.5 Concurrent measures of affective state

Evidence from the Delphi experts (Chapter 3) and the literature highlights the importance of considering a combination of measures when attempting to use them as proxies to measure affective state (Hall *et al.*, 2018; Wemelsfelder *et al.*, 2001; Burghardt, 2005; Boissy *et al.*, 2007a; Désiré *et al.*, 2002). Within this study a componential approach is taken which addresses behavioural, autonomic and subjective components to fulfil the aims of inducing and measuring positive affective state in equines (Mortillaro and Mehu, 2015). There is overwhelming consensus that emotional phenomena are multi-componential beginning with appraisal of an event leading to altered motivational state, which in turn causes changes to face and body expression, changes in physiological responses and production of a subjective feeling (Mohammadi *et al.*, 2019) (Figure 6.1).

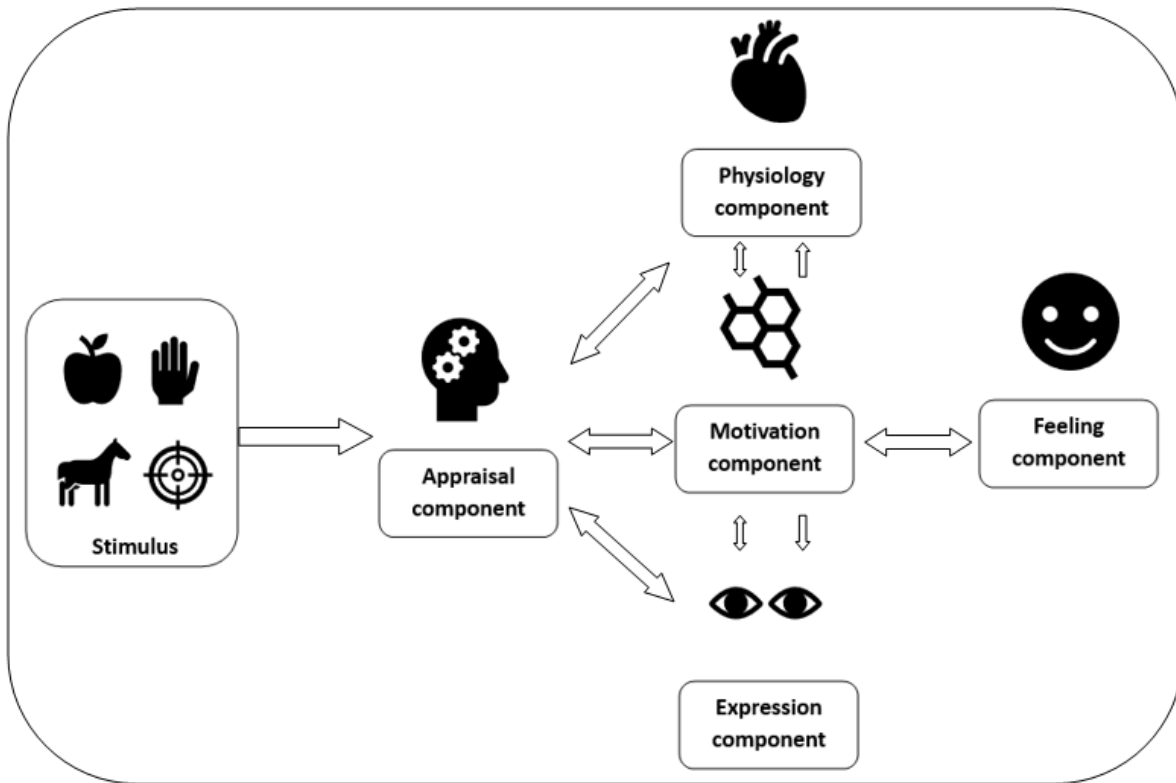


Figure 6.1 Schema depicting the component model of emotion adapted from Scherer (2009) and Mohammadi *et al.* (2019).

The study aims to place componential aspects within a dimensional framework of arousal and valence (Russell, 2003; Mendl *et al.*, 2010), defined further through the more granular affective circumplex (Yik *et al.*, 2011) via the appraisal of the affect induction stimulus, changes in physiology, motivation and expression and the production of the feeling component, the affective experience.

6.1.6 Aims and objectives

The aim of this study was to utilise a range of expert suggested methods (supported by the evidence in the available literature) to attempt to induce positive affect in a mixed cohort of equines and to use the expert suggested behavioural and physiological measures to evaluate affective state induced in a real-world setting. The study also considered assessment of personality within the subject cohort and factored the results of these assessments into the statistical analysis undertaken.

The objectives of the study were to:

1. Evaluate equine responses to four different methods (high value food provision, wither scratching, positive training via target training and the addition of an affiliative conspecific) of inducing positive affective state in the subject cohort of 20 equines,
2. Apply four behavioural (Qualitative Behaviour Assessment, frustration and repetitive behaviour indices, facial action scoring and eye gaze) and four physiological (heart rate and heart rate variability, respiratory rate, salivary cortisol and infrared thermography of the eye and ear) methods to attempt to identify methods which have potential value in measuring positive affective state in equines,
3. Assess personality in the subject cohort and utilise this data within exploratory analyses to identify if there is an effect of personality on equine responses to the four induction stimuli,
4. Undertake exploratory statistical modelling of the results obtained to evaluate the potential usefulness of specific methods to induce and measure positive affect in equines.

It was hypothesised that individual equines would respond differently, potentially partially because of individual personality traits, to the four induction stimuli and that this would be evidenced through differences in the behavioural and physiological measures recorded between induction stimuli, considered alongside personality metrics. It was also hypothesised that there may be identifiable groupings of measures that may indicate differing levels of arousal and types of valence, across the subject cohort.

6.2 Methodology

6.2.1 Trial preparation

Subjects were recruited through the professional networks of the researcher via written invitation. Potential participants were provided with a participant information sheet detailing the requirements for, and planned process of, the study (Appendix E). Inclusion criteria for trial subjects were defined as follows:

Participants must:

- Have at least one equine that they own or care for,
- Have another friendly equine available for use during the trial,
- Confirm that their equine is physically healthy,
- Confirm that their equine will be comfortable wearing a heart rate monitor surcingle band,
- Confirm that their equine will be comfortable having saliva swabs taken,
- Confirm that they have access to a stable or tying up area,
- Be over the 18 years of age and have the capacity to consent to their involvement and the use of the equine for study purposes,
- Complete the pre-trial questionnaire at least 24 hours prior to the allocated study period,
- Be available on two consecutive days for a period in the morning and afternoon of both days to facilitate data collection (each trial lasts around 20-30 minutes per horse and each horse will undertake 4 trials),
- Be comfortable handling their equine.

Following acceptance of the invitation, participants completed a pre-trial information form via Google Forms where they confirmed their eligibility and willingness to contribute to the study and provided details about their individual equine (Appendix F) including scoring their personality using the EPEM abridged scoring system (Chapter 4). Prior to the trial date participants also scored their equines personality using the full EPEM personality questionnaire as part of the personality questionnaire validation discussed in Chapter 4.

Each participant was then given a two-day period where the trial for their equine would take place. Data collection ran from June to August 2023. Prior to the trial days each subject underwent a food preference testing protocol as outlined in the methods section below.

6.2.2 Subjects

Twenty healthy equines (Table 6.4) were recruited to the study between March and May 2023.

Table 6.4 Demographic data for the twenty subjects.

Horse number	Age	Breed	Sex	Use	Time in ownership or care
1	3	Mini Cob	Gelding	Non-ridden	> 6 months < 1 year
2	9	Thoroughbred	Gelding	Unaffiliated competition	> 6 years
3	5	Irish x Thoroughbred x Arab x Trakehner	Mare	Unaffiliated competition	5-6 years
4	16	Warmblood	Gelding	Non-ridden	> 6 months < 1 year
5	11	Native crossbreed	Gelding	Unaffiliated competition	1-2 years
6	25	Highland	Mare	Hacking	> 6 years
7	28	Native crossbreed	Gelding	Non-ridden	> 6 years
8	9	Thoroughbred	Gelding	Elite competition	> 6 years
9	10	Thoroughbred	Gelding	Elite competition	> 6 years
10	4	Thoroughbred	Gelding	Elite competition	1-2 years
11	10	Thoroughbred	Gelding	Elite competition	> 6 years
12	5	Thoroughbred	Gelding	Elite competition	1-2 years
13	26	Cob	Mare	Hacking	> 6 years
14	30	Connemara cross	Mare	Non-ridden	> 6 years
15	18	Native crossbreed	Mare	Non-ridden	> 6 years
16	6	Thoroughbred	Mare	Elite competition	1-2 years
17	6	Thoroughbred	Mare	Elite competition	1-2 years
18	5	Thoroughbred	Mare	Elite competition	> 6 months < 1 year
19	6	Thoroughbred	Gelding	Elite competition	1-2 years
20	7	Thoroughbred	Gelding	Elite competition	3-4 years

6.2.3 Methods

Preference testing

Food: At least 24 hours prior to the trial day (but not more than 72 hours prior) each subject undertook a food preference test to ensure that a high value food reinforcer was selected for their food trial. A method similar to that of Vinassa *et al.* (2020) was followed whereby a three-choice repeated measures preference test was set up. A negative control (0.5kg of the horse's usual hay or haylage) was placed in one red bucket (previously left with the owner of the subject for habituation through provision of food in the bucket on two occasions prior to the preference test). Two additional trial feedstuffs (500 grams of Baileys Horse Fibre Plus Nuggets and one Horslyx Mini Original equine lick tub) were placed separately in two other identical red buckets (RedGorilla, 2024). The buckets were placed 0.5m apart opposite the entrance, in a test space (usually the horses stable) and the equine was released to approach the buckets at will. The feedstuff approached first, time spent eating each foodstuff and eaten weight after a two-minute period were recorded for each horse. This method was repeated twice more with the presentation of buckets in a randomised location order at each trial presentation to reduce order or location effect. The feedstuff that the horse spent longest eating and ate most of on each trial was recorded as the preferred substrate and the modal feedstuff recorded across the trials was selected as the feedstuff to be presented during the study.

Wither scratching: Prior to the trial day each equine was tested on three occasions for a two-minute period to identify their responses to different style (fast or slow scratching), and location of wither scratching. Preferences were recorded for each trial and the modal preference was selected for the study. Behaviours such as leaning into the scratches, head and neck extension, changes in mouth position (Figure 6.2) and attempts to groom the human scratching them or to groom inanimate objects were identified as positive indicators of enjoyment of the scratching style (McGreevy *et al.*, 2009). Avoidance behaviours (Draaisma, 2017) were identified as indications of a negative response to the scratching style. The preferred scratching style was then utilised on the individual equine during their trial.



a



b

Figure 6.2 Examples of positive responses to wither scratching a) head and neck extension with mouth position changes and b) grooming of the human scratcher during wither scratching.

Preference test results for each subject are presented in Table 6.5 below.

Table 6.5 Preference test results for wither scratching area and food preference. Results for each of the three tests per induction method are presented. LF = Left-hand side fast, LS = left-hand side slow, RF = right-hand side fast, RS = right-hand side slow, TF = top of wither fast, TS = top of wither slow. H = Hay, N = Nuggets, L = Lick.

Equine number	Wither preference test result			Method selected	Food preference test result			Method selected
	Test 1	Test 2	Test 3		Test 1	Test 2	Test 3	
1	LF	LF	LF	LF	N	N	N	N
2	TS	TS	TF	TS	L	N	N	N
3	LF	LF	LF	LF	N	N	N	N
4	RS	RS	TS	RS	L	L	L	L
5	LS	LS	LS	LS	L	N	N	N
6	RF	RF	RF	RF	N	N	N	N
7	LS	LS	LS	LS	L	N	N	N
8	LF	LF	LF	LF	N	N	N	N
9	LF	LF	LF	LF	L	L	L	L
10	TS	TS	TF	TS	N	N	N	N
11	RS	RS	RS	RS	N	L	N	N
12	LS	LS	LS	LS	N	N	N	N
13	LS	LS	LS	LS	N	N	N	N
14	TS	TS	TS	TS	L	L	L	L
15	RS	RS	RS	RS	N	N	N	N
16	LF	LF	LF	LF	N	N	N	N
17	LF	LF	LF	LF	N	N	N	N
18	LF	LF	LF	LF	N	N	N	N
19	LS	LS	LS	LS	L	N	N	N
20	TS	TS	TS	TS	N	N	N	N

Trial area set up and habituation

On the first trial day the trial area was set up in an identical manner at each of the five test locations (see Figure 6.3 for examples) with three GoPro Hero 8 cameras (GoPro Inc, USA) placed 3 metres to the front of the equine, 1-2 metres to the side of the equine and 3 metres to the rear of the equine to capture a wide field of view and enable scoring of behavioural and respiratory data.

Location	Front view	Side view	Back view
1			
2			

Figure 6.3 The three views generated by camera placement at two of the five study locations.

A thermography camera FLIR T430SC (FLIR Systems AB, Danderyd, Sweden) was positioned next to the front facing camera to collect thermography data and was time paired with all of the GoPro cameras and the heart rate monitor. All equines were habituated to the area used for data collection prior to the trials by being placed in the area regularly for 10–30-minute periods with hay and water. They were considered habituated when they quickly (within 1

minute) showed postural (relaxed muscle tone, low head carriage) and behavioural (standing and eating hay or resting/dozing) signs of relaxation on entering the area. Horses who did not habituate to the stable area were to be excluded from the trial, however all equines habituated well. Prior to commencement of the trial each equine was given a ten-minute acclimatisation period prior to the application of equipment for the trials. All equines were also habituated to the wearing of the Polar equine heart rate monitor on three occasions prior to the sessions, and all were used to wearing items around the girth as they were all currently or had been previously ridden equines. Habituation was deemed to have occurred when heart rate did not increase more than 20 beats per minute during the heart rate monitor application (Leiner and Fendt, 2011). During habituation to the test stable all equines were also presented with the video camera set up in terms of positioning of the front, side and rear-view cameras and the thermography camera placed on a tripod.

In contrast to the pilot study (Chapter 5) most equines were not accustomed to being cross-tied and therefore all subjects were held on a lead rope for each trial in order to avoid introducing a stressor to the environment. The horses were held by a known handler on the same length lead rope during the trials to avoid potential stress responses confounding results (Stull and Rodiek, 2002). The handler did not interact with the horse during the trial and stood outside the stable and to the right of the closed stable door or through another form of barrier to reduce handler effect (Birke *et al.*, 2011; Górecka *et al.*, 2007).

Test procedure

The test protocol was identical for each equine. The subject was led to the research area by a known handler and released for a 10-minute habituation period during which they had opportunity for free movement. Differently to the pilot study, free access to hay and water were not provided during the habituation period as it was identified in the pilot study that the removal of the items prior to stimulus introduction caused arousal in some subjects as identified previously by Ricci-Bonot and Mills (2023) and Pannewitz and Loftus (2023). A familiar affiliative equine was placed nearby (approximately 10-15 metres away) during the entire trial to avoid isolation stress (Lansade *et al.*, 2018; Zhao *et al.*, 2020). After 10-minutes the subject was fitted with the Polar equine heart rate monitor (Figure 6.4). The Polar band was soaked in water for at least ten minutes prior to each trial and ultrasound gel (Anagel,

Ana Wiz Lt, Surrey) was placed liberally along the electrode strip to maximise conductivity and generate consistent data as per the method of Nath *et al.* (2022).



Figure 6.4 Placement of the Polar Equine Heart rate monitor using a Polar equine band.

The subject was left to habituate freely for a further ten minutes after which the subject was placed on a lead rope. The heart rate monitor, the three GoPro cameras and the FLIR camera started recording video data with the time on the heart rate monitor and cameras synchronised for later analysis.

The first salivary cortisol swab was taken two minutes after the lead rope was attached and prior to stimulus introduction. The first stimulus was presented to the subject in a randomly allocated order (Appendix G) identified through online randomisation software (Urbaniak and Plous, 2013). Following a timed two-minute period of stimulus presentation the stimulus was removed, and a second salivary cortisol swab taken. After a further two minutes the subject was then allowed free movement until a period of fifteen minutes had elapsed whereupon a third salivary cortisol test was taken. The heart rate monitor was then stopped, data saved, and the monitor and band removed. The subject was washed down to remove remaining ultrasound gel and returned to their field or stable. A summary of the trial protocol can be seen in Figure 6.5.

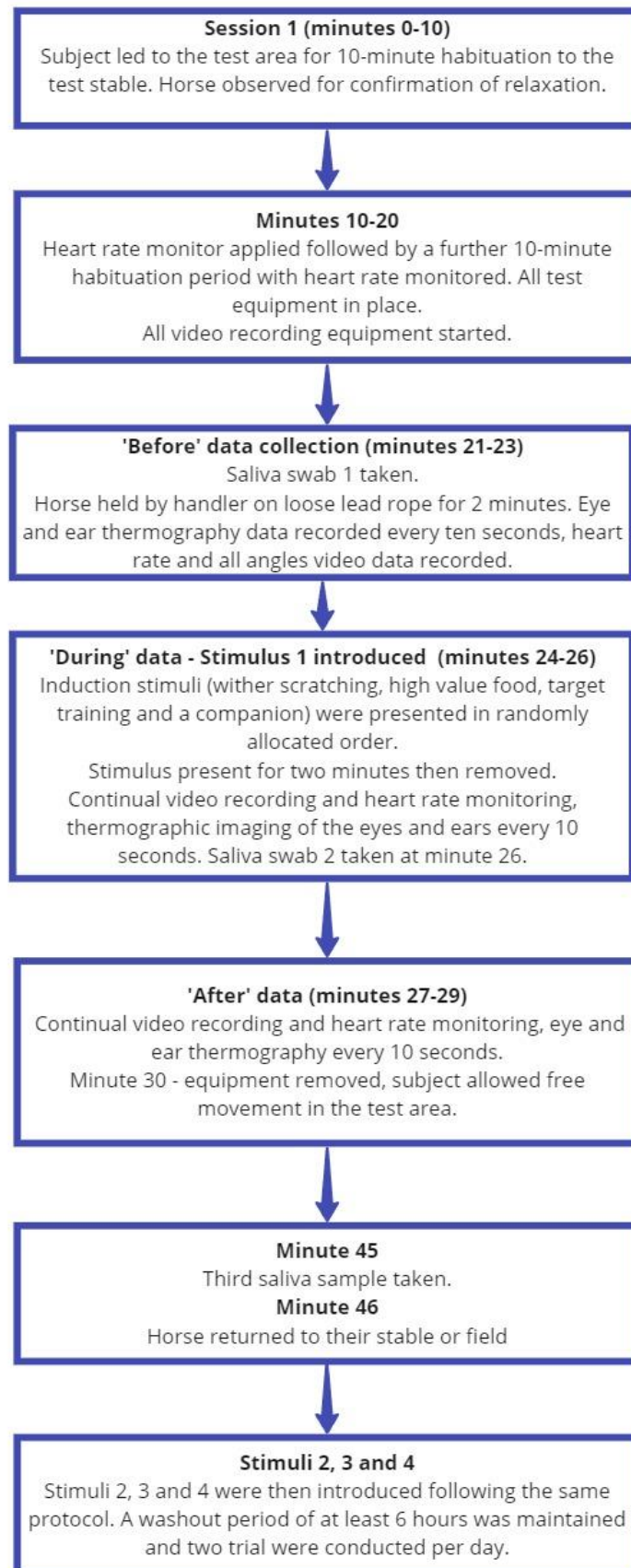


Figure 6.5 Flowchart depicting an outline of the data collection method.

Four trials were conducted over two days (Figure 6.6) to allow a period of at least six hours wash-out period between trials. It was deemed appropriate to conduct two trials per day as per the method of Kang and Lee (2016) who were also collecting predominantly behavioural data with associated physiological measures. Longer washout periods are required for repeated measures, within-subjects designs where bio-active or psycho-active supplements or medications are administered depending on their metabolism rates and half-lives.

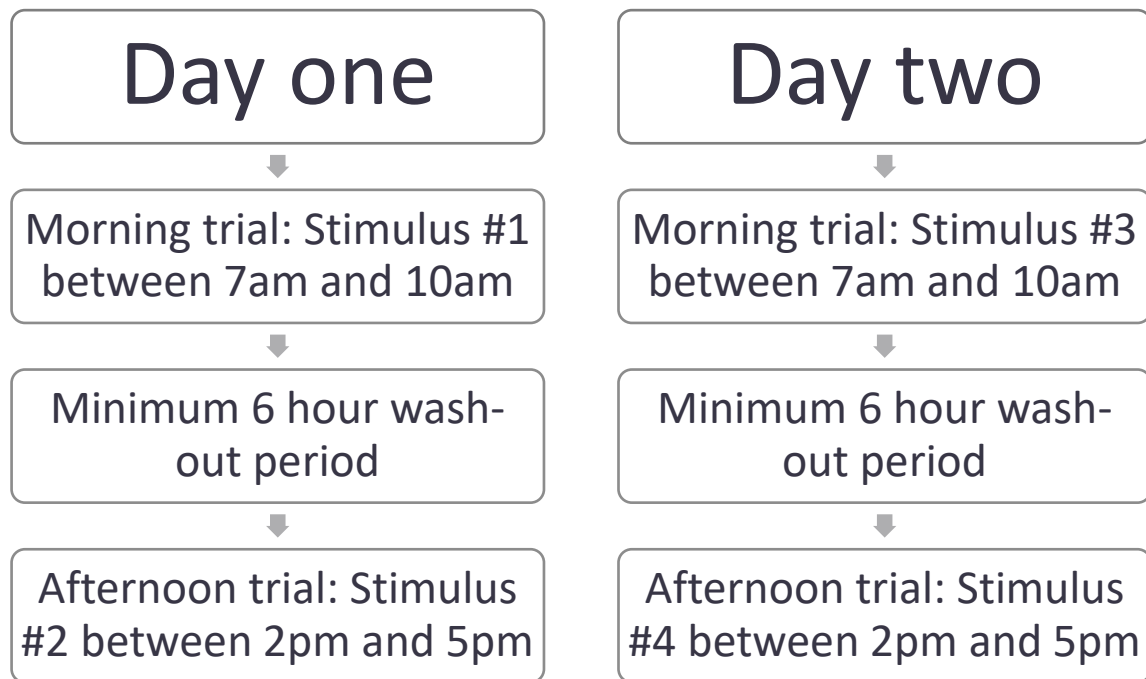


Figure 6.6. Timings of each trial ensuring that there would be at least a six-hour wash-out period between trials.

Video and physiological data storage and handling

Following each day of data collection video data were uploaded to secure cloud storage and each file named with the subject number and trial number (blinded to induction method treatment) as well as the date and time of the trial session, they were stored in folders named 'front', 'back' and 'side' according to the camera angle they were taken from. Thermography data were uploaded to FLIR Ignite (TeledyneFlir, 2024) secure cloud storage after each day of data collection and the files named in the same method as for video files. Cortisol buffer bottles were labelled at the time of collection with the subject number and trial number and the swab immediately placed in the buffer bottle and sealed once saliva collection was complete. The saliva samples were analysed via Lateral Flow Device (LFD) on the day of

collection and then sent in batches for laboratory ELISA analysis. Heart rate and heart rate variability data were uploaded to Polar Flow software (PolarElectro, 2024) at the end of each day of data collection with each dataset labelled with the subject number and trial number and the date and time automatically attributed via the Polar software.

6.2.4 Positive induction stimuli utilised within this study

Addition of an affiliative conspecific

Within the trial the affiliative conspecific was defined as such following observation of at least five occasions of allogrooming and/or affiliative interaction with the subject horse in the month preceding the trial. The conspecific was brought into view and walked up to the subject horse whereupon the two horses were free to choose to interact or not for the two-minute trial duration (Figure 6.7). Following this the conspecific was led away out of view.



a



b

Figure 6.7 Affiliative conspecific interactions during the conspecific trial.

High value feedstuff

Following preference testing described in section 6.2.3, this trial included the presentation of 1kg of the preferred substrate in one of the preconditioned buckets (RedGorilla, 2024) which the horse could eat continuously for the two-minute trial period (Figure 6.8), after which the stimulus was removed.



Figure 6.8. Mid-way through a food trial.

Wither scratching

Within this study wither scratching was trailed with all subjects prior to the trial (section 6.2.3) and the preferred method and exact site (Figure 6.9) was utilised for each individual subject. The subject received wither scratching for the entirety of the two-minute trial period and it then abruptly stopped.

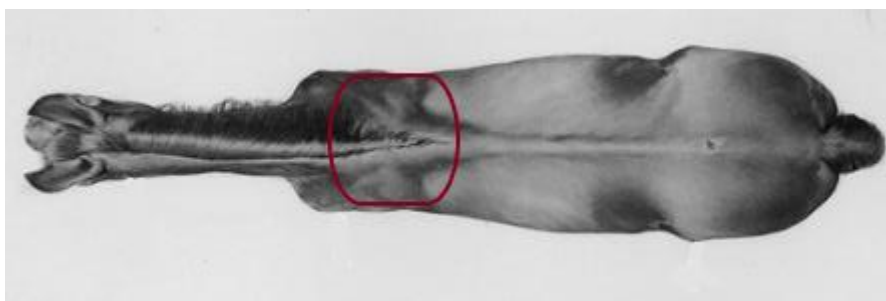


Figure 6.9. The wither area permitted for scratching circled in red, according to individual equine preference, adapted from Dittrich (1898).

Training

Positive training in the form of target training (Figure 6.10) with positive reinforcement was selected as a novel induction method in this study. Each equine undertook target training for a two-minute duration for this induction trial. The equine was presented with the target stick and upon touching the target with the muzzle, the target was withdrawn, and a low value food reinforcer (a pinch of Dengie Hi-Fi chaff) was delivered. Following which the target was presented again to the equine; this process continued for the two-minute trial duration.

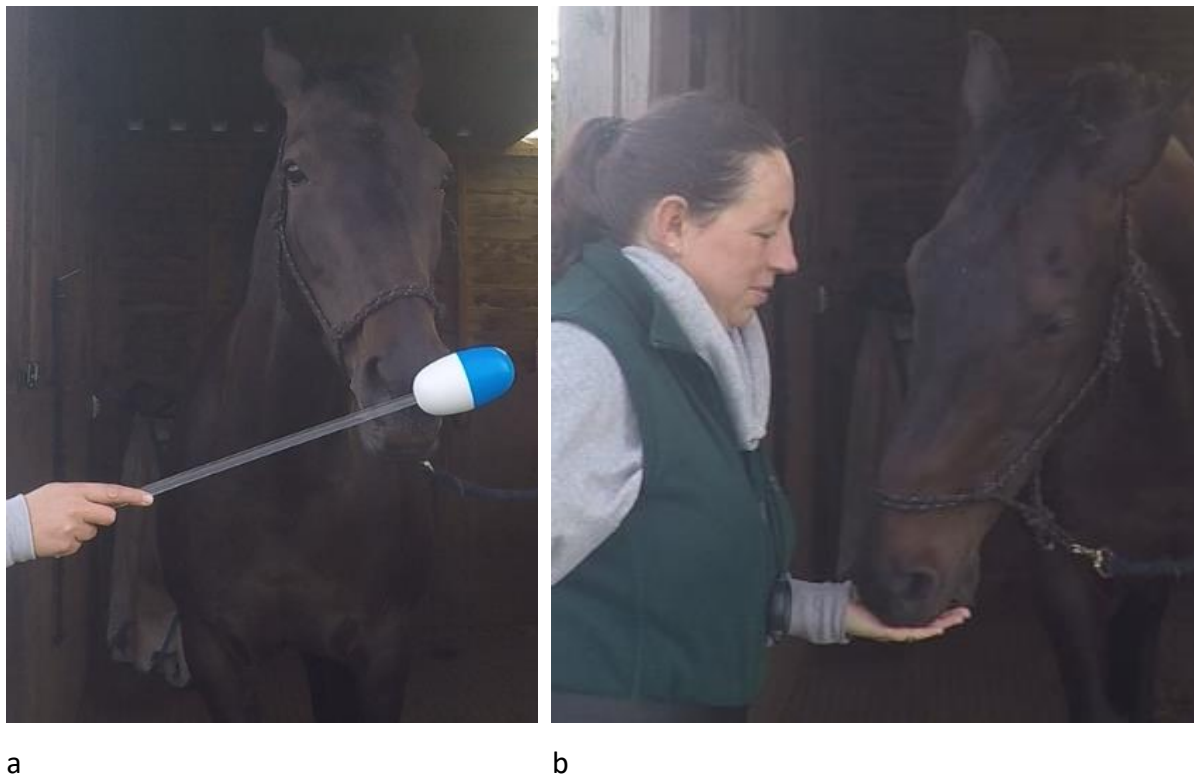


Figure 6.10 Example of target training with a) the horse touching the target with their muzzle and b) receiving a low value food reinforcer (Dengie Hi-Fi chaff).

6.2.5 Behavioural measures of positive affect in equines utilised within this study

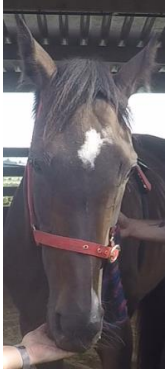
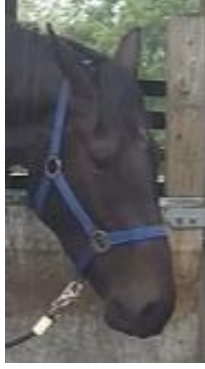
Ear position

Ear position data were scored as described in Table 6.6 below and were scored for both frequency and duration via continuous recording.

Eye position

Within this study eye positions 'open', 'half open' and 'sclera' were scored for frequency and duration via continuous recording (Table 6.6).

Table 6.6 The ear and eye positions with descriptors and images as utilised by the video scorers. Ethogram developed from Lansade *et al.* (2018).

Ears forwards	Ears sideways	Ear asymmetrical	Ear gently back	Ear back	Open	Half	Sclera
							
Both ears facing forwards. The inner ear can be seen.	Boths ears facing to the side and relaxed.	Each ear facing in a different direction, may be forward, sideways or backwards.	Both ears rotated backwards but are relaxed, inner ear not seen.	Both ears rotated fully backwards and flattened.	The eye is fully open with no sclera (eye white) showing.	The eye is half closed with the upper eyelid partially covering the lower eye.	Some white of the eye is showing.

Eye lateralisation (eye bias)

Eye lateralisation or eye bias was scored for left, both and right eye bias, with scoring of both duration and frequency undertaken (Table 6.7).

Table 6.7 Eye bias positions with images and descriptors. Ethogram developed from Des Roches *et al.* (2008) and Austin and Rogers (2007).

Both	Left	Right
		
The horse is focussing on the stimulus with both eyes equally or is looking straight ahead.	The horse angles or turn the head to focus more with the left eye than the right eye.	The horse angles or turn the head to focus more with the right eye than the left eye.

Qualitative Behaviour Assessment (QBA)

QBA assessment was undertaken using the full AWIN QBA descriptor list (AWIN, 2015) with both frequency and duration of any behaviours identified recorded for each equine across all trials (Table 6.8).

Table 6.8 QBA names and descriptors from AWIN (2015).

Indicator name	Descriptor
Aggressive	Hostile, attacking, wants to fight/attack, dominance, defensive aggression, (i.e. may display the following: bite/kick, position of ears flat-back against head, dilated nostrils, turns the hindquarters towards object of aggression, intention to harm, tail-swishing).
Alarmed	Worried/tense, apprehensive, jumpy, nervous, watchful, on guard against a possible threat/danger (i.e. rigid stance, startled reaction to loud noise, looking around/vigilant, moving ears).
Annoyed	Irritated, displeased, bothered by something, disturbed, upset, troubled, exasperated (i.e. may display rapid tail-swishing, stomping).
Apathetic	Having or showing little or no emotion; disinterested, indifferent, isolated, depressed, unresponsive, motionless.
At ease	Calm, carefree, peaceful.
Curious	Inquisitive, desire to investigate (i.e. approach person/object of curiosity, engaged in exploratory behaviour; possibly displaying head and neck extended toward object of curiosity, with ears pricked forward).
Friendly	Affectionate, kind, not hostile, receptive, positive feelings toward people, confident (i.e. the horse approaches the person, may sniff or interact in some way).
Fearful	Afraid, hesitant, timid, not confident, not necessarily linked with something going on in the environment (i.e. you may see the body tremble, flared nostrils, tail clamped).
Happy	Feeling, showing or expressing joy, pleased, lively, playful, satisfied.
Look for contact	Actively looking for interaction, interested, close proximity, eager to approach.
Relaxed	Not tense or rigid, easy-going, tranquil.
Pushy	Assertive or forceful (i.e. not leaving space, head butting out of the way, exhibits dominant behaviour, may be mouthy or nippy).
Uneasy	Afflicted, uncomfortable, unsettled, restless.

Indicators of frustration

This measure was included, utilising the descriptors reported by Pannewitz and Loftus (2023) (Table 6.9) and scored for both frequency and duration of occurrence (Martin and Bateson, 2007).

Table 6.9 Indicators of frustration from Pannewitz and Loftus (2023).

Indicator name	Descriptor
Acute - conflict	The horse shows some resistance to the stimulus or a person including avoiding interaction, trying to move away, being hesitant.
Acute - displacement	The horse engages in another behaviour such as bed eating, increased eating or drinking, active looking away, mouthing items or people, undertaking crib-biting, wind sucking, weaving or repetitive pacing) rather than engaging with the stimulus.
Acute - vocalisation	The horse produces a sound.
Acute – redirected aggression	The horse is aggressive to something that is not the stimulus being presented.
Acute – direct aggression	The horse is aggressive to the stimulus being presented.
Acute – increased locomotion	The horse moved around (this may be anything from one step to an entire movement sequence.
Acute – muscle tension	The horse becomes tense in its posture.
Chronic – repetitive behaviours	The horse engages in repetitive behaviours for a prolonged period of time.
Chronic – learned helplessness	The horse is unresponsive and inappetent for a prolonged period of time.
Chronic – altered cognitive bias	The horse shows a prolonged change in judgement about the stimulus.

6.2.6 Physiological measures of positive affect in equines utilised within this study

Heart rate and Heart rate variability (HRV)

From a methodological perspective it is important that when HRV is used to assess emotional state, inter-beat-interval (IBI) data lasting at least five minutes consecutively is measured during stationary conditions to negate the effect of movement behaviour on cardiac function as well as pre-behavioural physiological anticipatory responses and persisting post-behavioural cardiac responses (Malik *et al.*, 1996). To satisfy these requirements both heart rate and heart rate variability were recorded continually from the start to the end of each trial phase and for a minimum of 2 minutes per phase of each trial giving a minimum recording time of 6 minutes and a maximum of 12 minutes within this study. All data were recorded using the Polar V800 device and Polar H10 Heart rate sensor band. Fitting of heart rate

monitor was undertaken as explained in section 6.2.3 and all equines were previously habituated to the sensation of the band around the girth area.

Respiratory rate

Respiratory rate was scored via video footage where the rise and fall of the flank was counted as 1 unit. A period of 30 seconds was scored within the middle section of each trial phase across each induction stimulus for each equine and scores were multiplied by 2 to give a 'beats per minute' value. Scoring was repeated later, by the same scorer, to assess intra-rater reliability.

Salivary cortisol

Within this study each equine acted as its own control for each trial with salivary samples taken before each trial as well as immediately after and 15 minutes after each trial.

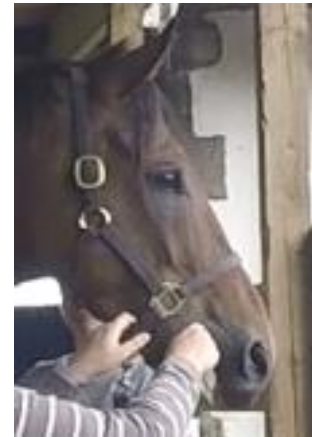
Soma equine salivary cortisol test kits (Soma Bioscience, Wallingford, UK) were utilised to take salivary cortisol samples alongside the Soma Cube reader (Soma Bioscience, Wallingford, UK) which analysed lateral flow device (LFD) samples at penside (see Figure 6.11). Subsequently all samples were laboratory analysed via Enzyme Linked Immunosorbent Assay (ELISA) (Alhajj *et al.*, 2023) by Soma Bioscience to evaluate the reliability of the LFD analysis.

Step one



One oral swab, LFD cartridge and oral fluid collector (OFC) bottle with buffer fluid were used per phase of each trial in each equine.

Step two



The saliva sample is collected using the swab inserted in the diastema between the incisor and molar teeth and maintained in position until the blue indicator line shows on the stem of the swab.

Step three



The swab is placed in an OFC bottle labelled with the subject number, trial number (T1-T4) and trial phase (.1 = before, .2 = immediately after, .3 = 15 minutes after) and gently inverted for 1-2 minutes. The corresponding LFD cartridge is labelled in the same manner as the bottle.

Step four



After a 10-minute incubation period two drops of the saliva-buffer mix are dropped onto the sample window (circled red) of the LFD cartridge and when the test window (green rectangle) turns red in colour the test is ready to run. The cube reader is turned on and the RFID calibration card is placed on the cube reader to verify the test. The test is run by pressing the run button on the top of the cube reader and the result appears on the cube screen within 4 seconds. Results were recorded in an excel sheet.

Figure 6.11 Salivary cortisol collection and LFD analysis process.

Ear and eye temperature

Within this study eye and ear temperature were recorded using a Flir T430SC (FLIR Systems AB, Danderyd, Sweden) camera on a tripod set one metre away from the subject. Data were recorded at ten second intervals, and data files uploaded to Flir Ignite online software for analysis. The camera was calibrated to the observation distance, temperature and relative humidity and recalibrated for each subject. Still images of the medial posterior palpebral border of the lower eyelid and lacrimal caruncle of the left and right eyes were taken every 10 seconds to limit data loss, which was apparent in the pilot study when continual recording was undertaken, and the camera focus did not adjust quickly to head movement from the horse, leading to a high number of invalid images.

6.2.7 Personality assessment

Personality assessment was undertaken for each equine as per the method outlined in Chapter 4. All twenty equines included in this final study were scored by their owners or carers on all the individual personality traits and using the abridged version of the questionnaire as part of the study reported in Chapter 4 and to contribute data to this study via the abridged questionnaire results.

6.2.8 Data processing and statistical analysis

Cortisol data

Cortisol data from the LFD were stored in an excel spreadsheet (Microsoft Excel V16) to which the ELISA results were added. Spearman's rank analysis was undertaken to compare the LFD and ELISA results which showed moderate agreement ($r^2 = 0.4$, $p < 0.001$), therefore, to ensure accuracy the ELISA results were taken forward for analysis.

Heart rate data

Heart rate data were extracted from Polar Flow individually per equine, per stimulus. The data were then uploaded to Kubios version 3.5.0 (Kubios Oy, Kuopio, Finland) (Tarvainen *et al.*, 2014). Data were treated using the method outlined by Scopa *et al.* (2020), initially the R peaks of the QRS complexes were determined using Kubios' inbuilt QRS detector and the Pan-Tompkins algorithm method. Data were passed through a bandpass filter (frequency band of 0.05-32Hz and window width of 150ms) to smooth close-by peaks. Medium correction was utilised to avoid over-correcting and missing normal inter-beat intervals. Three spectral bands

were identified that have been specifically recommended for equine HRV analysis (Scopa *et al.*, 2020); Very Low Frequency (VLF) was set to 0-0.01Hz, Low Frequency (LF) was set to 0.01-0.07Hz and High Frequency was set to 0.07-0.6Hz.

In the time domain Mean RR, SDNN, Mean HR, SDHR, Minimum HR, Maximum HR, RMSSD and pNN50 data were extracted. In the frequency domain AR values were extracted as these are most applicable for data collected over shorter timeframes. These included Peak (Hz), Power (ms^2), Power (%) and Power (n.u.) for the VLF, LF and HF domains; LF-HF ratio was also computed. In the non-linear domain SD1 and SD2 were computed. All extracted data were placed into an excel sheet (Microsoft Excel V16).

Respiratory rate data

Respiratory rate was calculated from video data extracted from a side-facing GoPro camera which was mounted at a height and position to capture the rise and fall of the flanks. Each rise and fall were counted as one respiratory unit. Respiratory rate was counted using a four-digit clicker counter (Sdarming, Dongguan Xunze Trading Company, China) for 30 seconds in the middle section of each phase of each trial and the number was doubled to give a beat per minute value. Intra-observer reliability was tested through rescoring of 20% of the counts 2 weeks following the initial count. All data were inputted into an excel sheet (Microsoft Excel V16) and assessed via ICC (3,1), two-way mixed model, absolute agreement.

Thermography data

Still images from each phase of each stimulus presentation were analysed using Flir Ignite software (Teledyne, 2024). Initially images were inspected for clarity, focus and availability of measurement of each of the four data points. Images that were not in focus or did not allow for the four-point measurement were discarded. Four thermography readings were taken from each still image (see Figure 6.12) within the Flir Ignite software. All data were placed into an excel sheet (Microsoft Excel V16).

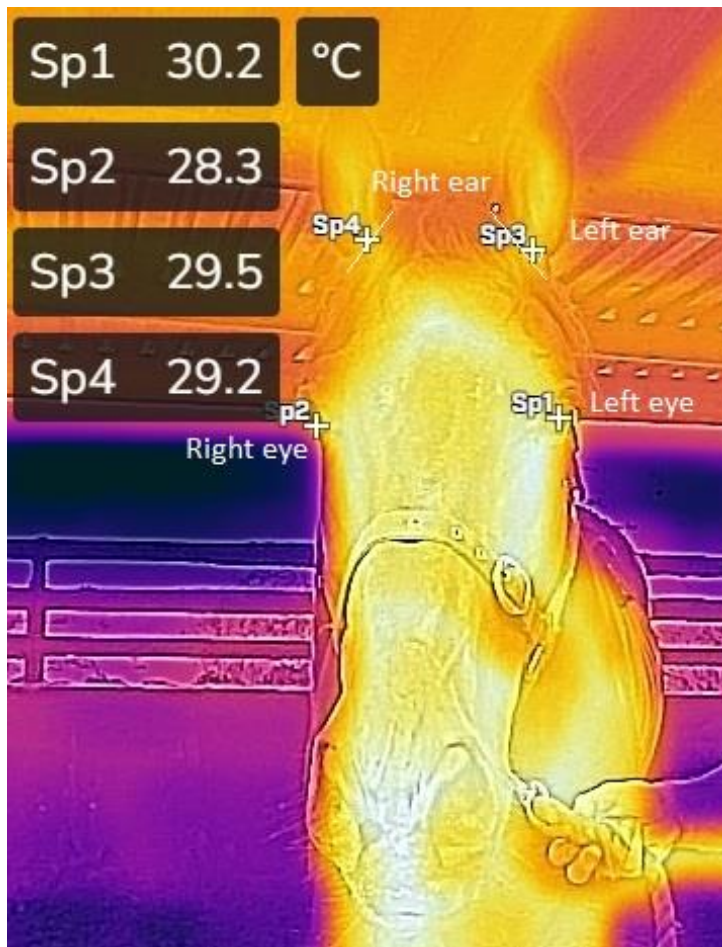


Figure 6.12 Data reading positions for the left eye (Sp1), right eye (Sp2), left ear (Sp3) and right ear (Sp4). A marker line indicated the slope across the base of each ear to aid data spot recording accuracy. Eye temperature was taken from the medial canthus of the eye as suggested by Kim and Cho (2021).

Behavioural data

Behavioural data were scored from video using CowLog™ software (Pastell, 2016) and followed a predetermined method (Figure 6.13 and Appendix H).

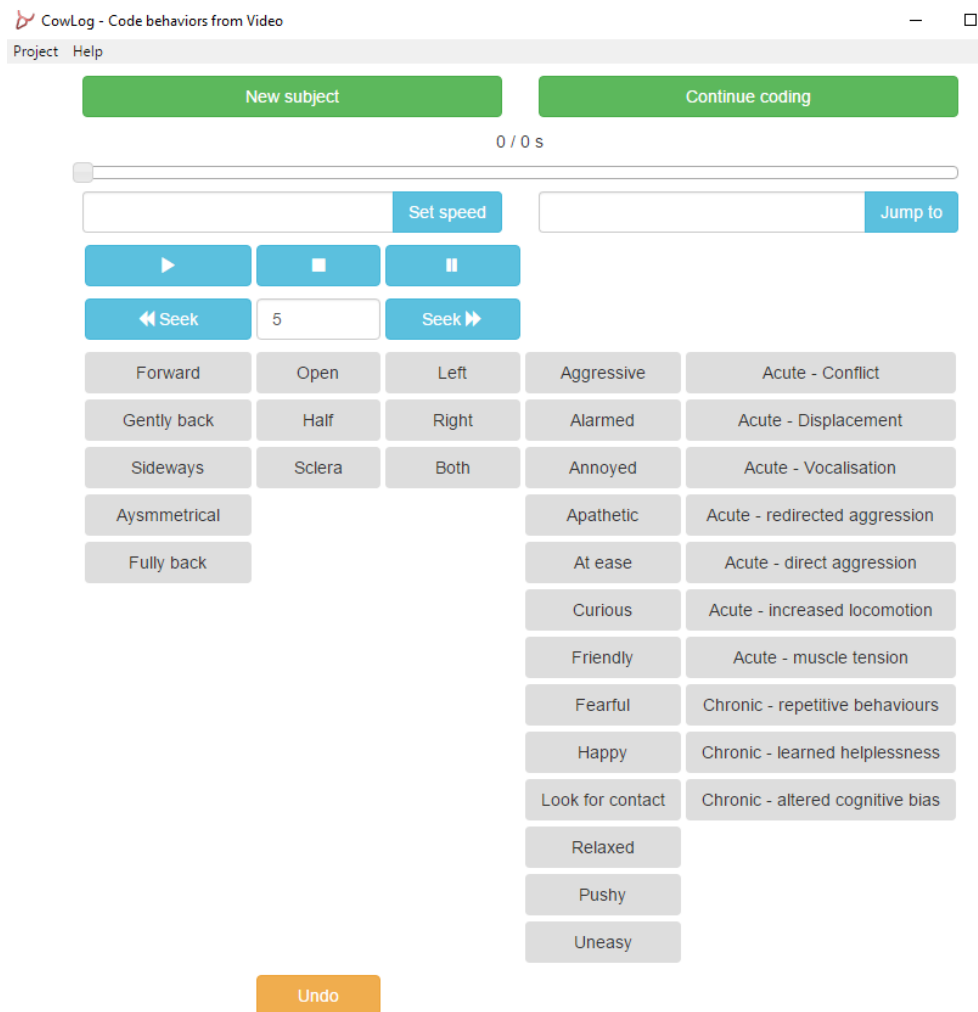


Figure 6.13 CowLog coding set-up.

Once scored the behavioural data were placed into an Excel spreadsheet (Microsoft Excel V16) with pre-set formulas which calculated total duration and frequency of each type of scored behaviour. The total duration percentage of time and the frequency rate per minute were then calculated automatically using a predefined formula. These data points were then transferred to a master data sheet spreadsheet (Microsoft Excel V16) containing the data from each equine for each trial.

Inter-observer reliability was tested utilising intraclass correlation coefficient (ICC 1, k , one-way random model for absolute agreement) analysis on data from 5 additional scorers who between them re-scored 20% of the videos (at least one per equine and three per stimulus). Two scorers were specialists in equine behaviour, two scorers were equine specialists but not specialists in behaviour and the final scorer was neither an equine nor a behaviour specialist. All scorers received the same CowLog training, CowLog file, behaviour ethogram and scoring

instructions and all scored the videos completely independently of each other. Random allocation of videos to the scorers was determined using online randomisation software (www.random.org).

Statistical analysis

An exploratory approach was undertaken throughout the statistical analysis to attempt to identify behavioural and physiological markers of affective state (both positive and negative) and to consider the effect of the treatment (stimulus), personality and the random effect of the horse within analyses. Four *a priori* assumptions were made (i) There would be covariation of measures (Masko *et al.*, 2020; Bennett-Wimbush *et al.*, 2020; Contalbrigo *et al.*, 2021; Cerasoli *et al.*, 2022); (ii) Wither scratching induces positive affect (Feh and de Mazières, 1993; Lansade *et al.*, 2022; Lansade *et al.*, 2018), (iii) Personality would influence behavioural and physiological responses to different induction stimuli (Kremer *et al.*, 2020; Marinier and Alexander, 1994) and (iv) QBA measures are valid indicators due to their prior validation (AWIN, 2015; Dalla Costa *et al.*, 2016a; Dalla Costa *et al.*, 2014c). These assumptions were made following review of the literature which has previously identified covariation between some behavioural and physiological measures of equine affective state (Lansade *et al.*, 2022; Lansade *et al.*, 2018) and identified wither scratching as a method which is likely to induce relaxation and positive emotions in equines (McGreevy *et al.*, 2009; Thorbergson *et al.*, 2016; Lansade *et al.*, 2018; Watson and McDonnell, 2018). Evidence from the literature also identified the likely influence of personality traits on equine behavioural responses (Fortin *et al.*, 2018; Lansade *et al.*, 2017; Fureix *et al.*, 2012).

All analyses were undertaken using R Studio (RCoreTeam, 2023) utilising the packages “car”, “corr”, “corrplot”, “devtools”, “FactoMineR”, “factoextra”, “fans”, “ggcorrplot”, “ggord”, “KlaR”, “lme4”, “MASS” and “Psych”. Initially outliers were removed from the dataset using the 1.5 x IQR rule (Voudouris *et al.*, 2022) and variables with a high number of missing values (>10 per variable) were excluded from analysis. This included several of the QBA and frustration measures and where variables contained low levels of missing values (≤ 10 per variable) the median value for the variable was computed and imputed in place of the missing values (this occurred mainly for thermography and HRV data). Distributions were checked visually via histograms after imputation to ensure relative normality. The statistical methodology is described in full in Figure 6.14.

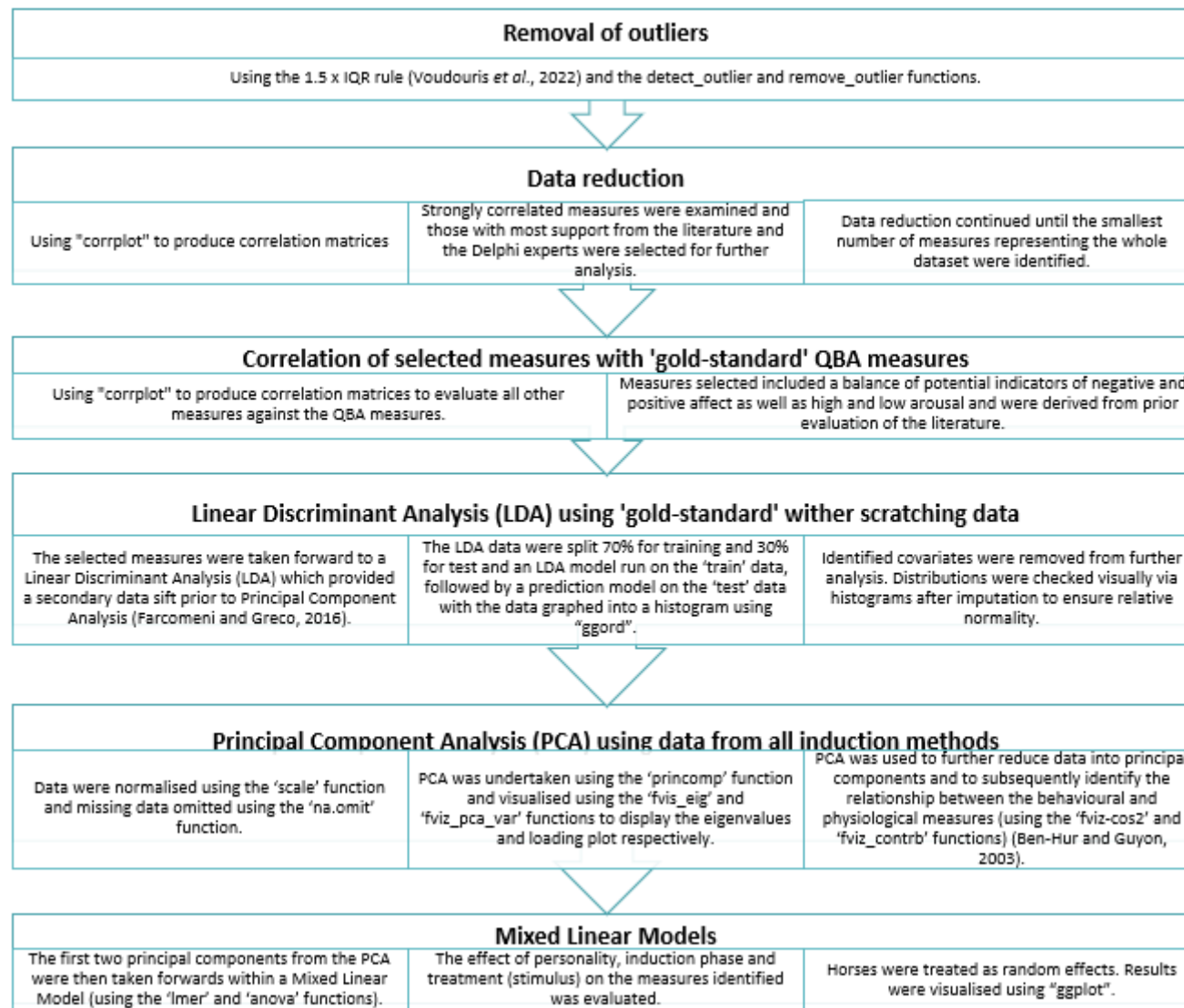


Figure 6.14 Flow chart detailing the statistical analysis methodology undertaken.

After data analysis further LDA analysis was undertaken on the 'Training', 'Companion' and 'Food' induction methods independently of each other to identify differences in the useful measurement variables within each induction method.

6.3 Results

6.3.1 Reliability of measures

Salivary cortisol

Salivary cortisol results measured using the Cube reader from lateral flow assays were compared with laboratory ELISA results via Spearman correlation after assessment of Q-Q plots determined non-normal distribution. There was moderate agreement between methods ($r^2 = 0.40$, $p < 0.01$), possibly due to several outliers (Figure 6.15) therefore, ELISA results were utilised within subsequent analyses for maximum accuracy.

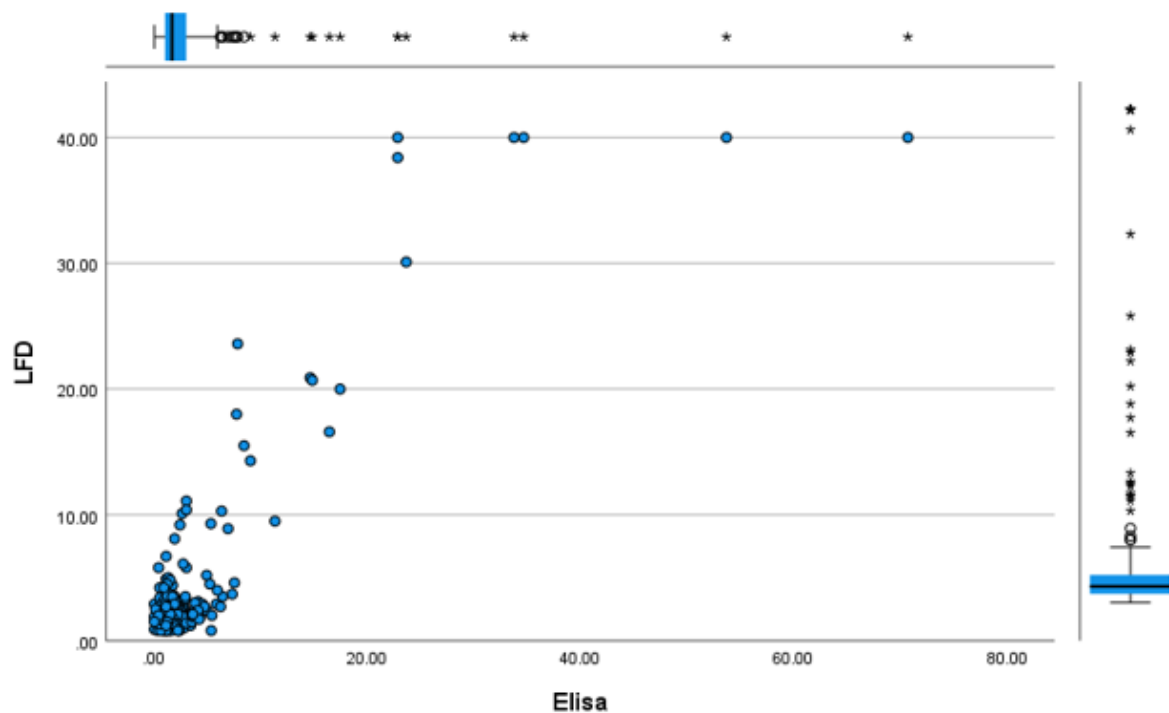


Figure 6.15 Correlation plot of LFD and ELISA salivary cortisol results with box and whisker to depict the mean and range of both data sets. Outliers are depicted by *.

Respiratory rate

Respiratory rate data were scored twice by the same observer and intra-rater agreement was excellent (ICC (3,1), two-way mixed, absolute agreement = 0.995, $p < 0.001$).

Behavioural data

Agreement between the principal rater and five other raters of the behavioural data was excellent as evaluated via ICC, one-way random, absolute agreement. Each equine was scored by the principal rater and by at least one other rater. Agreement levels can be found in Table 6.10.

Table 6.10 ICC results for behavioural measures for each horse.

Horse	Stimulus	n raters	ICC	<i>p</i>
1	Training	2	0.926	<0.001
2	Companion	2	0.793	<0.001
3	Training	4	0.923	<0.001
4	Companion	2	0.800	<0.001
5	Companion	2	0.846	<0.001
6	Training	2	0.837	<0.001
7	Food	2	0.795	<0.001
8	Food	2	0.901	<0.001
9	Training	2	0.936	<0.001
10	Food	2	0.981	<0.001
11	Food	2	0.892	<0.001
12	Food	5	0.904	<0.001
13	Scratch	2	0.967	<0.001
14	Scratch	2	0.901	<0.001
15	Companion	2	0.891	<0.000
16	Training	2	0.927	<0.001
17	Food	2	0.929	<0.001
18	Scratch	2	0.916	<0.001
19	Training	2	0.872	<0.001
20	Scratch	2	0.988	<0.001

6.3.2 Outlier removal and correlation of measures

Following removal of outliers, using the 1.5 x IQR rule (Voudouris *et al.*, 2022), correlation matrices undertaken on each behavioural and physiological measure dataset were examined visually and numerically for covariation.

Left and right eye and ear temperatures did not vary significantly between trial phases and treatments (Appendix I-1) and were all positively correlated (>0.7) apart from left eye and right ear which was 0.7 and as such were removed from later analyses. Ear position data were not generally positively correlated apart from 'ears gently back time in seconds' with 'ears gently back frequency'; 'ears asymmetrical time in seconds' with 'ears asymmetrical frequency' and 'ears fully back time in seconds' with 'ears fully back frequency'. Within eye position and bias data 'eye sclera time in seconds' was positively correlated (0.74) with 'eye sclera frequency' and 'eye open frequency' (0.50). 'Eye bias both frequency' was positively correlated with 'eye bias left time in seconds' (0.60) and 'eye bias left frequency' (0.81) and 'eye bias right time in seconds' (0.49) and 'eye bias right frequency' (0.80). 'Eye bias right time in seconds' was positively correlated with 'eye bias right frequency' (0.71) (Figure 6.16a, b, c).

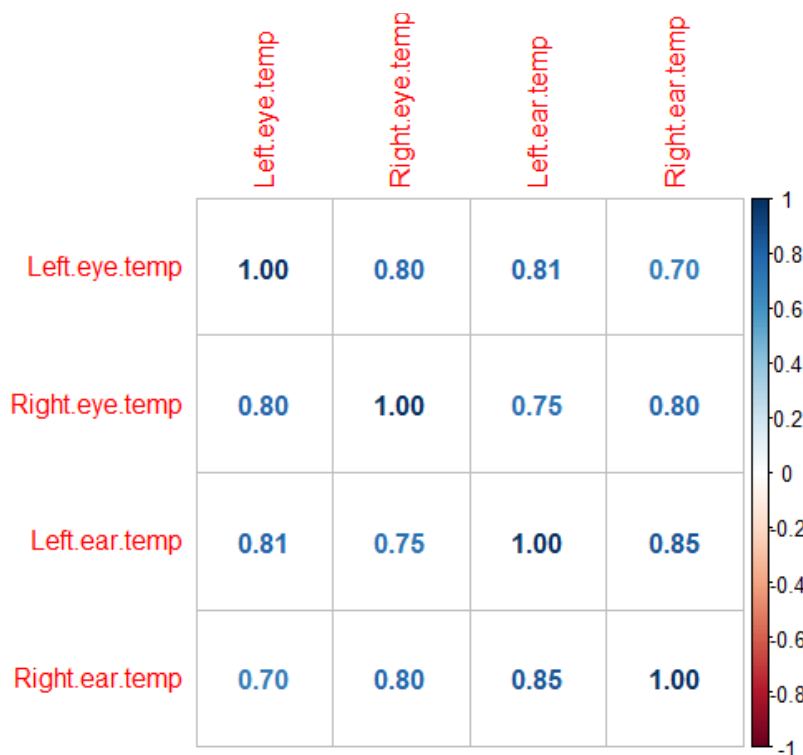


Figure 6.16a Corplot of ear and eye temperature (number method).

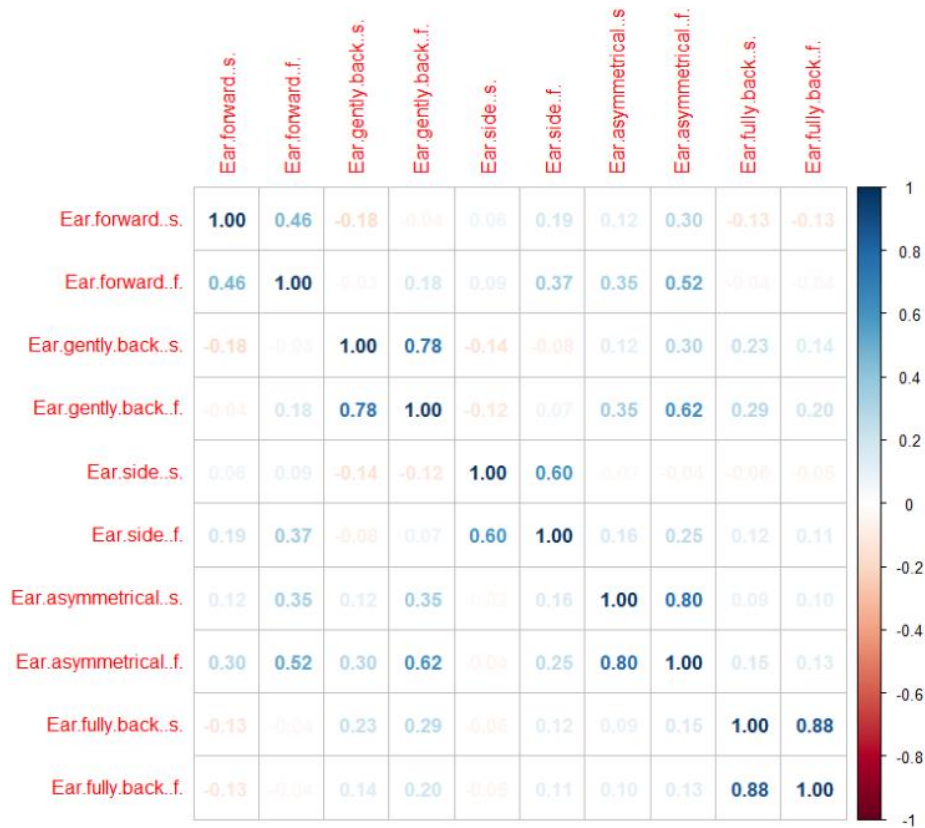


Figure 6.16b Corplot of ear positions (number method).

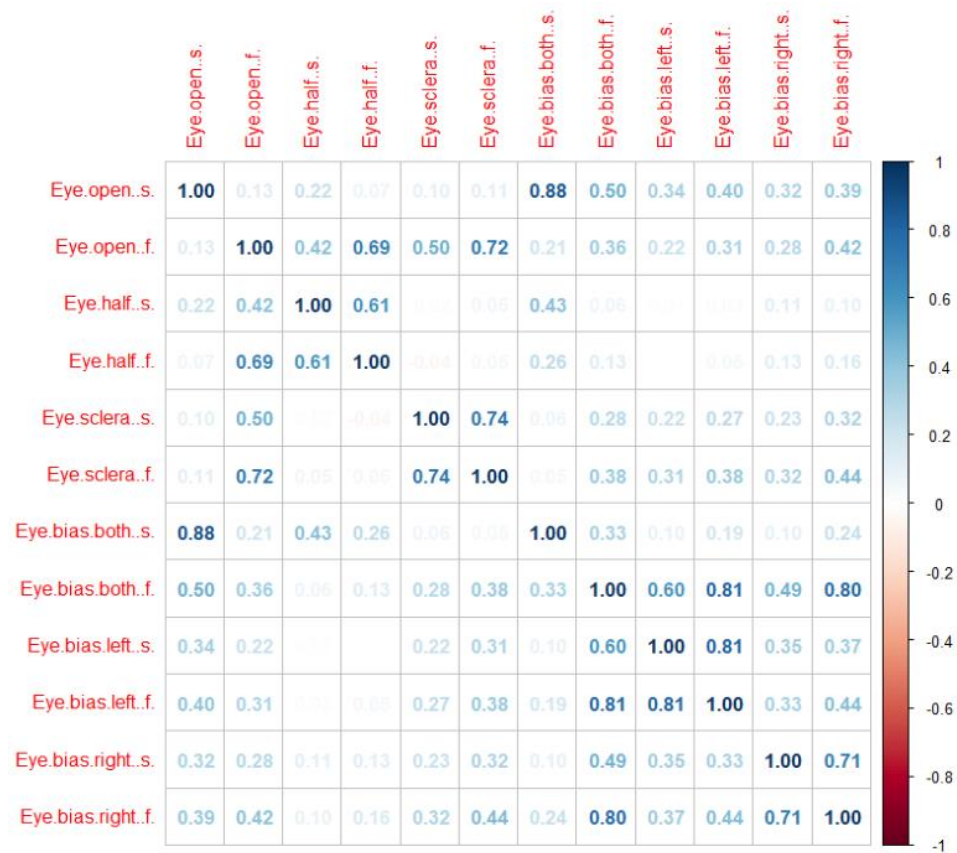


Figure 6.16c Corplot of eye position and eye bias (number method).

Many of the frequency QBA measures correlated with the duration QBA measures, for example 'QBA Alarmed time in seconds' was positively correlated (0.87) with 'QBA Alarmed frequency', therefore the frequency data were taken forward for use in later analyses (Figure 6.17).

Similarly, within the frustration indices time in seconds and frequency data were positively correlated for 'acute conflict (0.72)', 'acute displacement' (0.74). 'Acute increased locomotion' (0.66) and 'acute muscle tension' (0.71). Therefore, the frequency data were taken forward for later analyses (Figure 6.18).

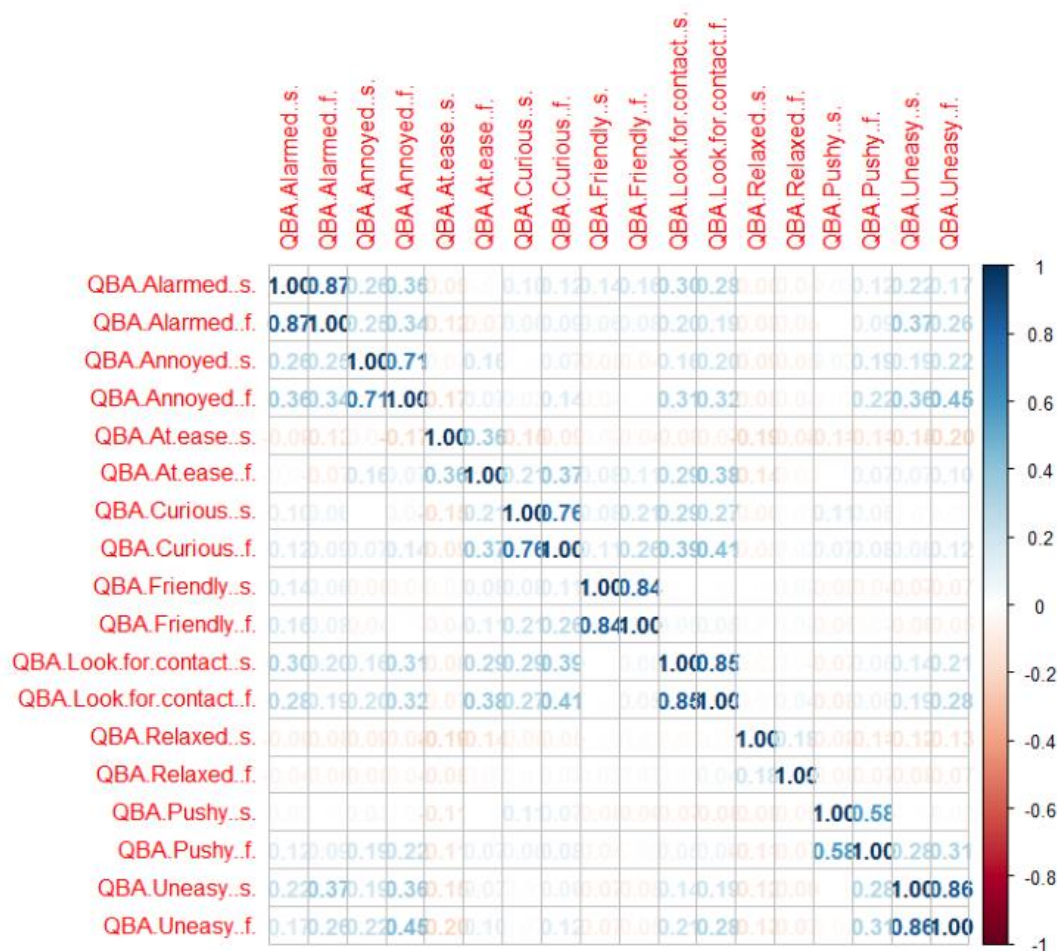


Figure 6.17 Corrplot of QBA measures (number method).

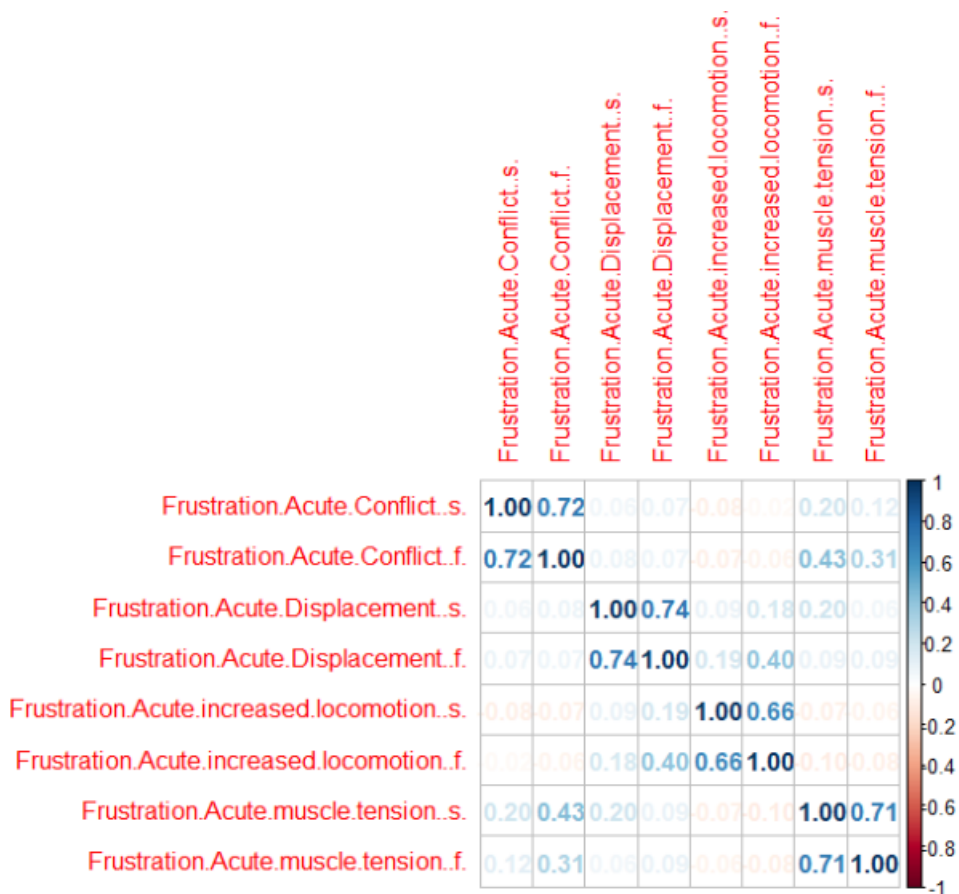


Figure 6.18 Corrplot of Frustration measures (number method).

Many of the heart rate variability indices were correlated (Figure 6.19). In particular, minimum heart rate and mean heart rate (0.92), SDRR and RMSSD (0.80), pNN50 and RMSSD (0.84), SD1 and SDRR (0.80), pNN50 and SD1 (0.84). Low frequency and high frequency power values were negatively correlated as would be expected due to their opposing measures.

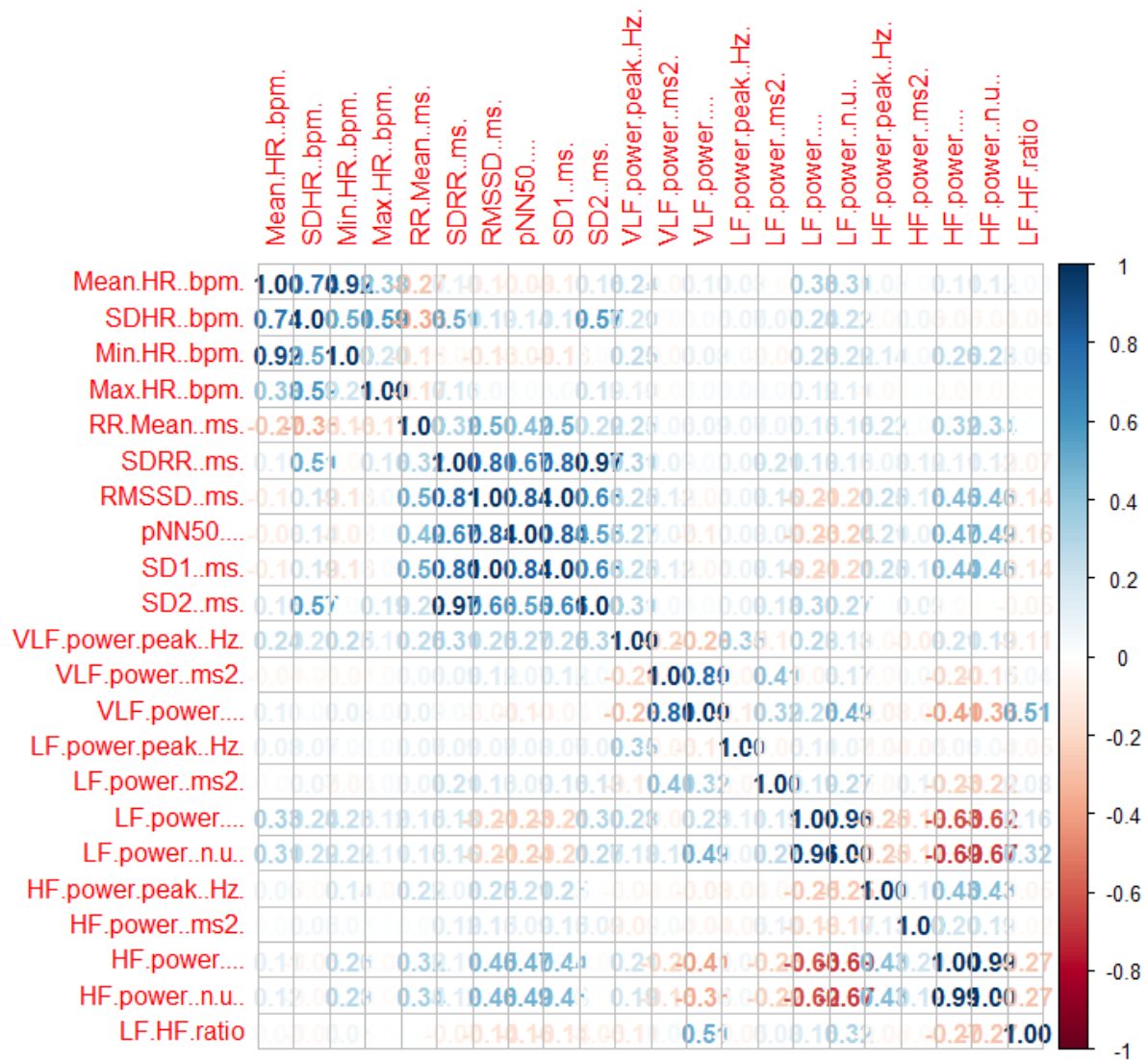


Figure 6.19 Corrplot of heart rate indices (number method).

The QBA data frame was scaled and normalised. The well validated QBA measures (AWIN, 2015; Dalla Costa *et al.*, 2016a; Ruet *et al.*, 2022) were evaluated for correlations (using corrplot number method) with the other measures (ear data, eye data, frustration data and heart rate data). A balance of measures with both negative and positive correlations to the QBA variables were selected to be taken forwards using evidence for their value from the literature (Table 6.11).

Table 6.11 Measures taken forward for further analysis. Rationale for taking each measure forward is presented along with the results of the Pearson's product-moment correlation, the correlation estimates and whether the item positively or negatively correlated with the QBA measures.

Measure	Rationale	t value	df	p value	Correlation	+/-
Ears forward (rate per minute)	Included as a potential positive measure (Tomberg <i>et al.</i> , 2023).	0.779	232	0.4363	0.051	+
Ears fully back (rate per minute)	Included as a potential negative measure (Tomberg <i>et al.</i> , 2023).	-10.442	232	<0.001	-0.565	-
Eye bias both (rate per minute)	Included as one of the more highly negatively correlated measures (Des Roches <i>et al.</i> , 2008).	-3.155	232	<0.01	-0.203	-
Eye bias right (rate per minute)	Included as one of the more highly negatively correlated measures. Inclusion support by the pilot study and Des Roches <i>et al.</i> (2008).	-4.051	232	<0.001	-0.257	-
Frustration: Acute increased	Included as a positively correlated	0.985	232	<0.001	0.058	+

locomotion (rate per minute)	measure and potential negative affect measure (Pannewitz and Loftus, 2023).					
Mean heart rate (beats per minute)	Included as a positively correlated measure supported by Loftus <i>et al.</i> (2016).	3.052	228	0.002	0.198	+
RR mean (ms)	Included as a comparison to mean heart rate (Scopa <i>et al.</i> , 2020).	-1.240	228	0.216	-0.081	-
SDRR (ms)	Included as a negatively correlated measure. Inclusion support by the literature (Scopa <i>et al.</i> , 2020).	-0.038	228	0.969	-0.002	-
RMSSD (ms)	Included as a positively correlated measure. Inclusion support by the literature (Scopa <i>et al.</i> , 2020).	0.042	228	0.965	0.002	+
pNN50 (%)	Included as a positively correlated measure. Inclusion	0.950	228	0.343	0.062	+

	support by the literature (Scopa <i>et al.</i> , 2020).					
Low frequency power (%)	Included as a positively correlated measure. Inclusion support by the literature (Scopa <i>et al.</i> , 2020).	0.494	228	0.621	0.032	+
High frequency power (%)	Included as a positively correlated measure. Inclusion support by the literature (Scopa <i>et al.</i> , 2020).	51	228	0.725	0.023	+
Low frequency/high frequency ratio	Included as a negatively correlated measure. Inclusion support by the literature (Scopa <i>et al.</i> , 2020).	-0.334	228	0.738	-0.022	-

Cortisol and respiratory rate data were also added as physiological variables of interest as highlighted by the literature. There appeared to be some differences between treatments and trial phase within the cortisol data collected (Figure 6.20) and the respiratory data (Figure 6.21).

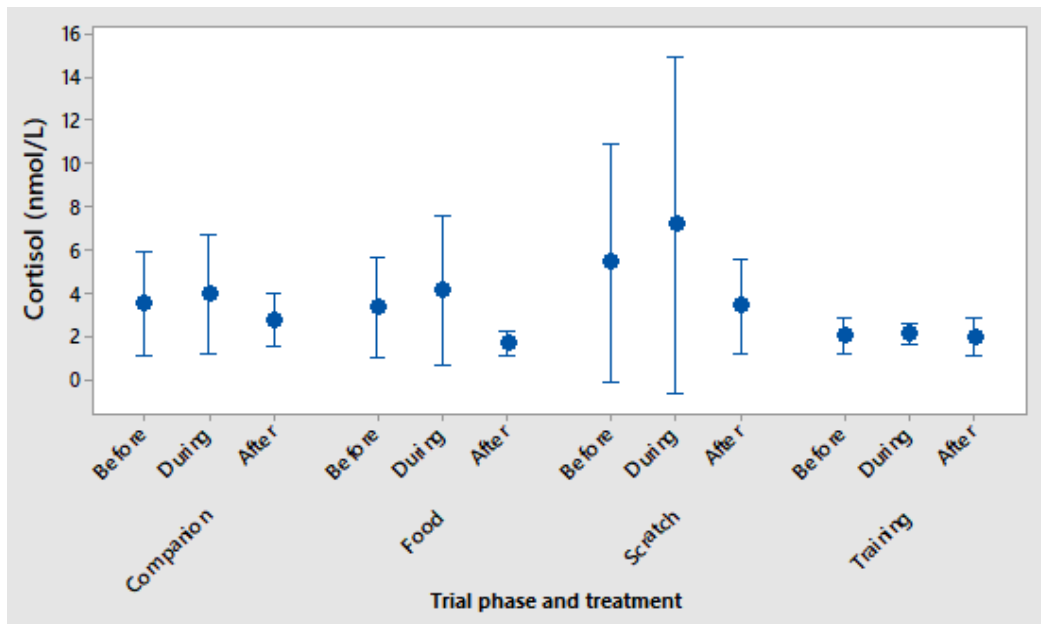


Figure 6.20 Interval plot of cortisol data by trial phase and treatment with confidence intervals.

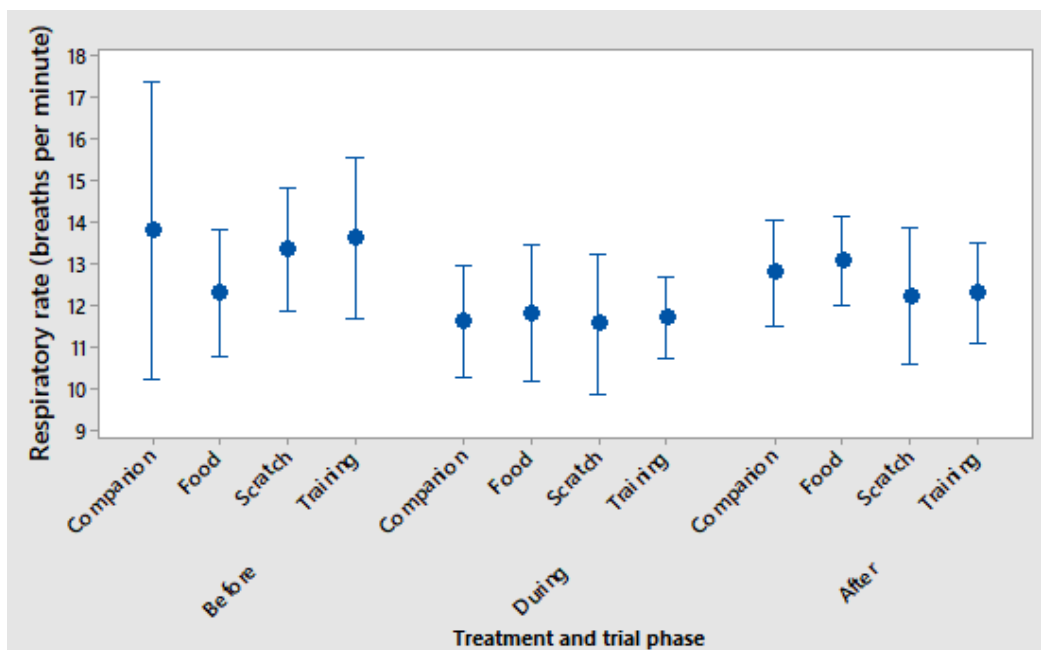


Figure 6.21 Interval plot of respiratory data by trial phase and treatment with confidence intervals.

6.3.3 Linear Discriminant Analysis (LDA)

As an additional data sift an LDA was run using only the wither scratch data with the rationale that wither scratching was the most highly validated induction method for calm, positive affect according to evidence from the literature (Lansade *et al.*, 2022; Lansade *et al.*, 2018;

Thorbergson *et al.*, 2016; Watson and McDonnell, 2018). The LDA was only run on phases 'before' and 'during' to avoid the impact of any carry over effect during phase 'after' which may have skewed the data (Cleophas, 1993). Nice separation between before and after phases was identified within the LDA plot (Figure 6.22) indicating distinct differences between the two phases. As such all variables were taken forward to Principal Component Analysis (PCA).

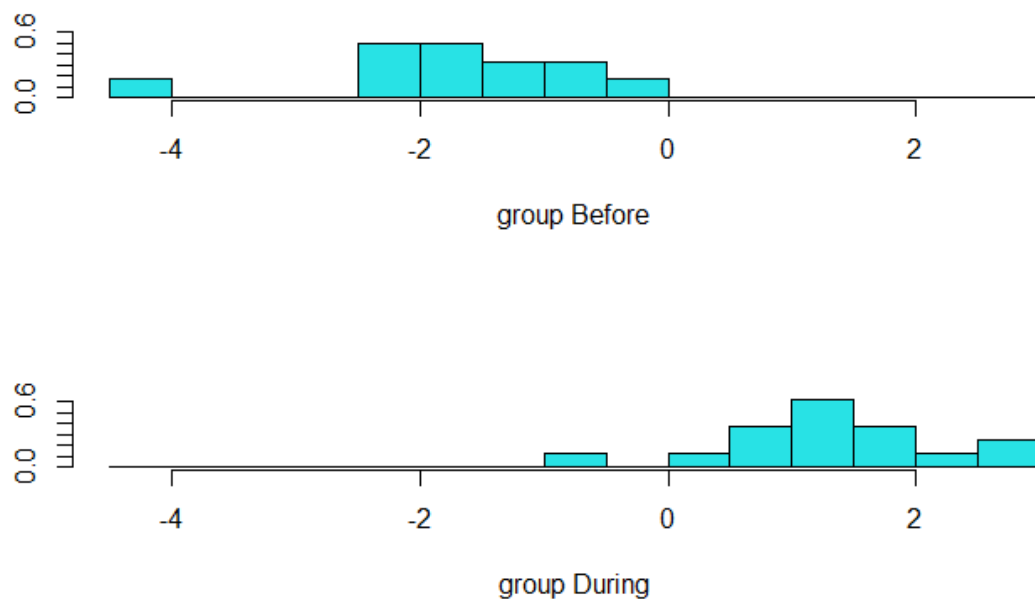


Figure 6.22 Separation of variables between the 'before' and 'during' phases of the scratch trial. With correlations of greater than 0.7 only identified between variables 'Low frequency power' and 'Low frequency/high frequency ratio'.

6.3.4 Principal Component Analysis (PCA)

Principal component analysis was undertaken for all induction methods (with scratching, high value food, companion and training) across all trials (before, during and after) with the variables reported in section 6.3.2. Data were normalised and missing values excluded.

Initial PCA analysis identified 15 principal components with the first four components accounting for 77% of the variation (Figure 6.23).

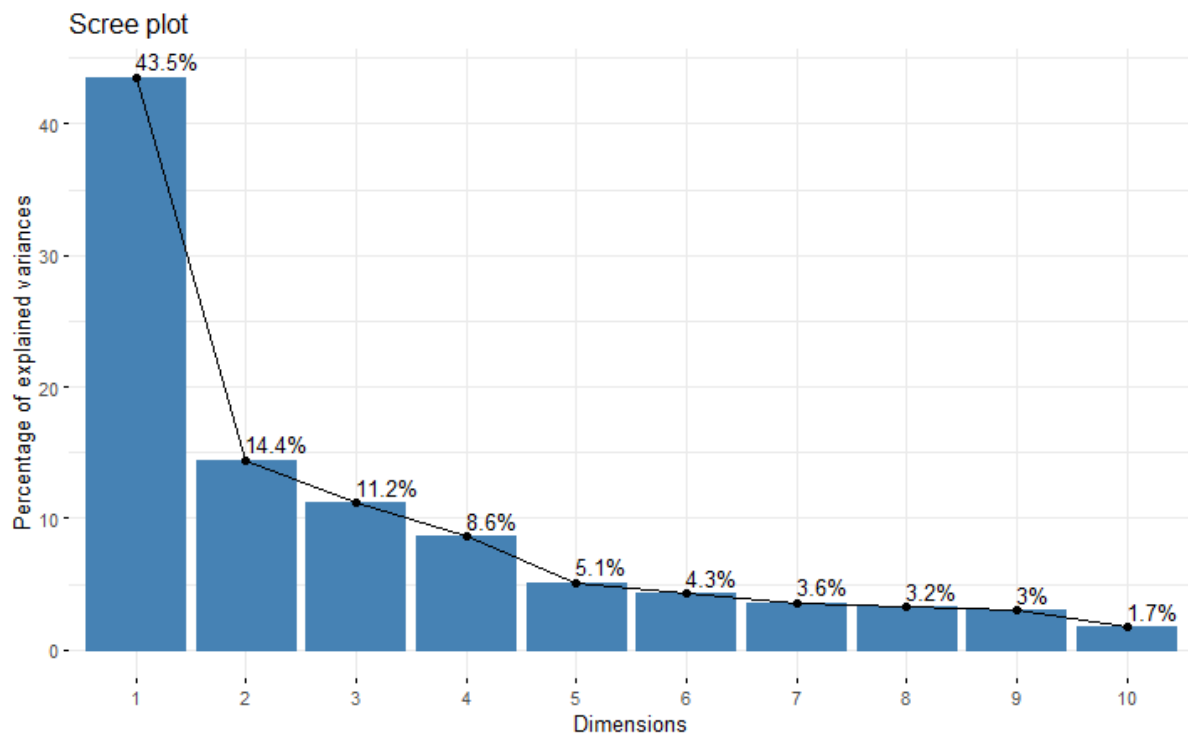


Figure 6.23 Scree plot of the first ten principal components.

The PCA loading plot identified variation in weightings between variables across the first two components (Figure 6.24).

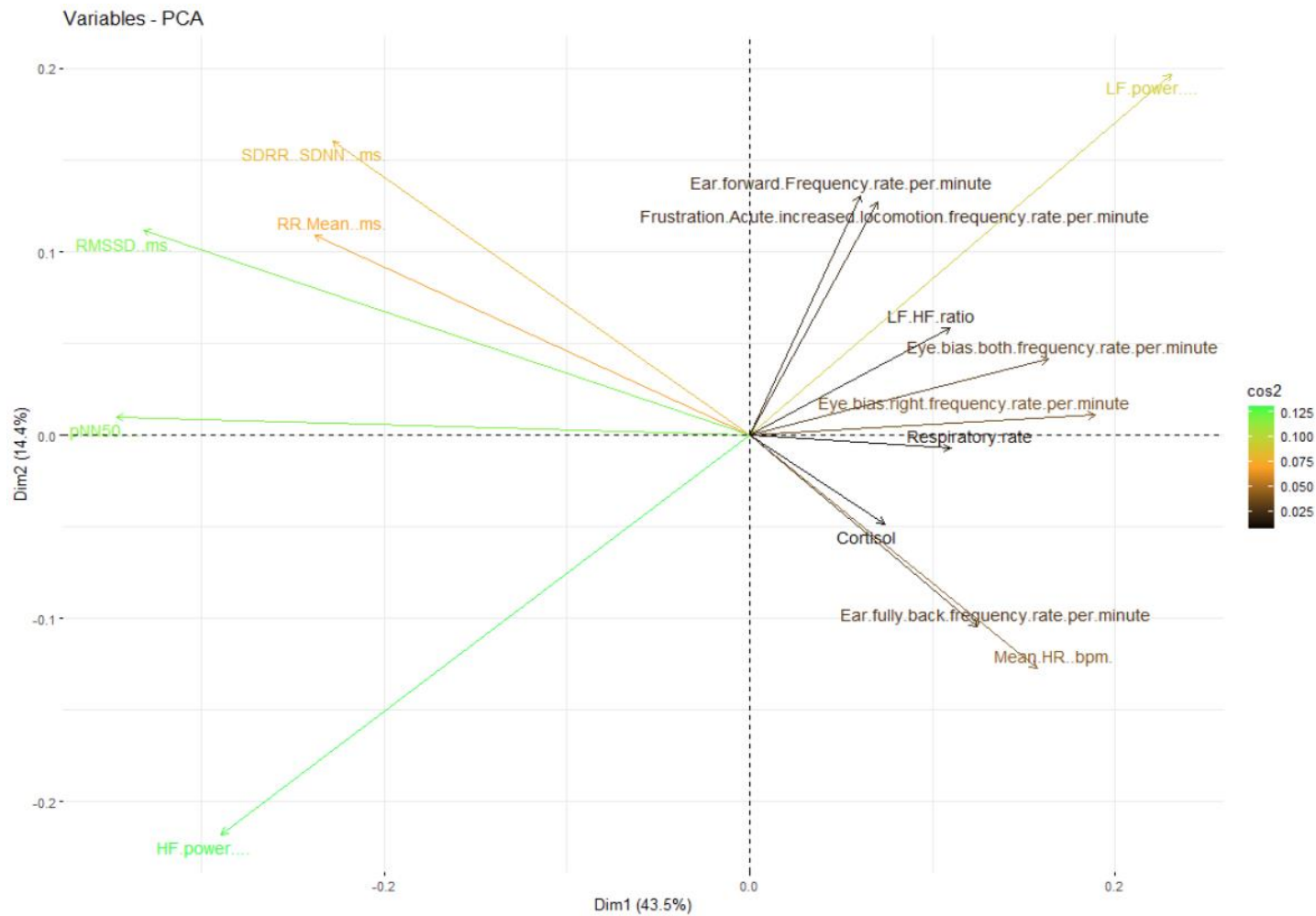
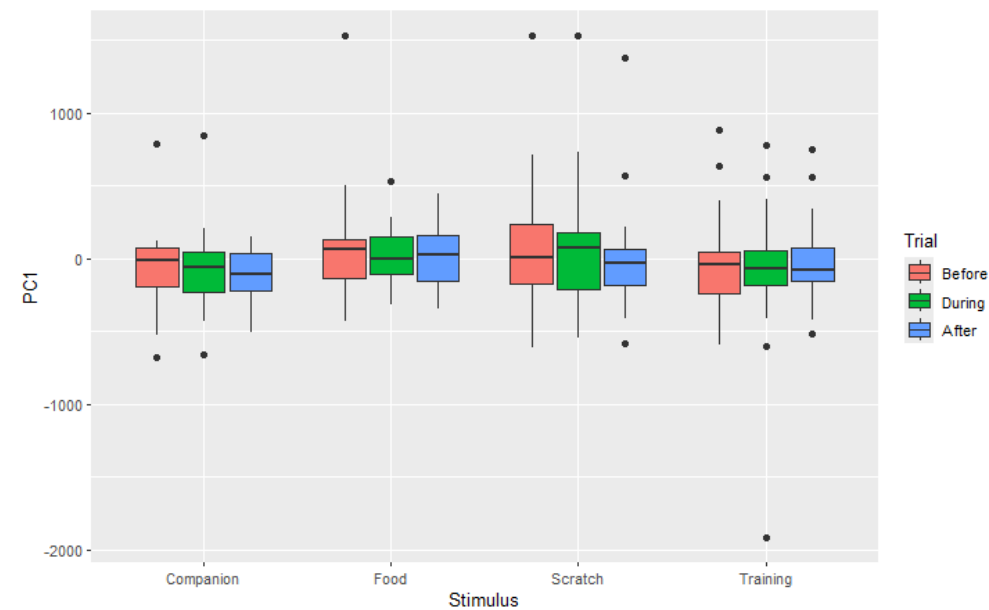


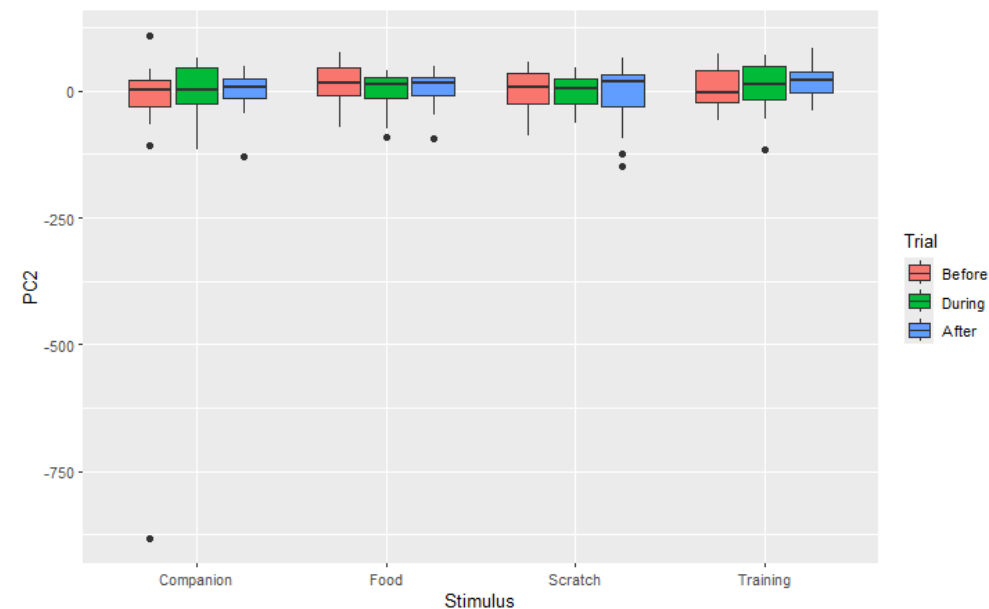
Figure 6.24 Loading plot of variables showing clustering of variables across negative and positive dimensions of principal components 1 and 2. Coloured lines represent the amount of each variable that is represented within each given component with green and yellow lines indicating higher levels of representation than brown and black lines.

6.3.5 Linear modelling of PCA components by stimulus and trial

Linear models were run to identify any effect of trial phase and stimulus type on PCA component data. A small effect of stimulus type was identified within principal component 1 ($F = 2.5609$, $df = 2$, $p = 0.1$; Figure 6.25a) but no effect of trial phase was found. No significant effects were identified for principal component 2 (Figure 6.25b).



a



b

Figure 6.25a Boxplot of trial and stimulus on principal component 1, and 6.25b Boxplot of trial and stimulus on principal component 2.

A mixed model was run to consider the effect of individual horses (inputted as random effects) on principal component 1 with a significant difference between horses across all induction stimuli (Type iii ANOVA with Satterthwaite's method, Stimulus: $F = 4.3784$, $df = 193.92$, $p = 0.005$) and a significant effect of horse found on the stimulus 'with scratching' ($t = 2.702$, $df = 193.554$, $p = 0.007$).

Within principal component 2 there was no significant difference between horses across all trial phases and induction stimuli overall (Type iii ANOVA with Satterthwaite's method, Trial: $F = 0.5811$, $df = 193.44$, $p = 0.56$; Stimulus: $F = 1.3906$, $df = 194.18$, $p = 0.25$). However there was a significant effect of horse on component 2 for 'trial during' ($t = 2.526$, $df = 192.86$, $p = 0.01$), 'trial after' ($t = 2.131$, $df = 193.42$, $p = 0.03$), 'stimulus food' ($t = 2.898$, $df = 193.81$, $p = 0.004$), 'stimulus scratch' ($t = 2.406$, $df = 193.54$, $p = 0.01$), 'stimulus training' ($t = 2.424$, $df = 193.19$, $p = 0.01$), 'trial during stimulus food' ($t = -2.225$, $df = 192.97$, $p = 0.02$), 'trial after stimulus food' ($t = -2.065$, $df = 193.73$, $p = 0.04$) and 'trial during stimulus scratch' ($t = -1.995$, $df = 192.86$, $p = 0.04$).

A linear model was undertaken to consider the effect of personality on equine responses to trial phase and stimulus. Personality group 'Unfriendly' was identified as showing a significant relationship with the with scratching stimulus ($t = 3.740$, $df = 198.29$, $p < 0.001$). None of the other seven personality groups showed significant interactions with stimulus or trial phase variables.

Following review of the relationships between variables it was noted that there did not seem to be sufficient variables included within the PCA to meaningfully distinguish between potential indicators of negative and positive affective state. Therefore, further variables were included within a second PCA analysis. Additionally, there was a predominance of physiological variables ($n = 10$) over behavioural variables ($n = 5$) suggesting that further behavioural variables should be added to the second PCA.

6.3.6 Further LDA and PCA analysis

A further twelve behavioural variables were added to a second round of analysis based on support for their inclusion from the equine and affective state literature (Table 6.12).

Table 6.12 Rationale for adding further measures into principal component analysis.

Measure	Rationale
Ears sideways	Supported as a potential measure of affect by Bornmann <i>et al.</i> (2021) and Ricci-Bonot and Mills (2023). Action defined by Wathan <i>et al.</i> (2015).
Ears gently back	Supported as a potential measure of affect by Bornmann <i>et al.</i> (2021) and Ricci-Bonot and Mills (2023). Action defined by Wathan <i>et al.</i> (2015).
Ears asymmetrical	Supported as a potential measure of affect by Bornmann <i>et al.</i> (2021) and Lansade <i>et al.</i> (2018). Action defined by Wathan <i>et al.</i> (2015).
Eye sclera	Supported as a marker of negative affect by Lansade <i>et al.</i> (2018) and Ricci-Bonot and Mills (2023). Action defined by Wathan <i>et al.</i> (2015).
Eye half	Supported as a marker of positive affect by Lansade <i>et al.</i> (2018) and of changes in affect by Hintze and Schanz (2021). Action defined by Wathan <i>et al.</i> (2015).
Eye bias left	Supported as a potential measure of negative affect by Ricci-Bonot and Mills (2023) and arousal with variable valence by Leliveld <i>et al.</i> (2013) and Evans <i>et al.</i> (2024).
Frustration Acute conflict	Supported as a potential measure of negative affect by Pannewitz and Loftus (2023).
QBA relaxed	Previously validated measures (AWIN, 2015; Dalla Costa <i>et al.</i> , 2016a; Ruet <i>et al.</i> , 2022; Minero <i>et al.</i> , 2018). Potential measure of positive affect.
QBA at ease	Previously validated measures (AWIN, 2015; Dalla Costa <i>et al.</i> , 2016a; Ruet <i>et al.</i> , 2022; Minero <i>et al.</i> , 2018). Potential measure of positive affect.
QBA friendly	Previously validated measures (AWIN, 2015; Dalla Costa <i>et al.</i> , 2016a; Ruet <i>et al.</i> , 2022; Minero <i>et al.</i> , 2018). Potential measure of positive affect.
QBA uneasy	Previously validated measures (AWIN, 2015; Dalla Costa <i>et al.</i> , 2016a; Ruet <i>et al.</i> , 2022). Potential measure of negative affect.
QBA alarmed	Previously validated measures (AWIN, 2015; Dalla Costa <i>et al.</i> , 2016a; Ruet <i>et al.</i> , 2022). Potential measure of negative affect.

LDA analysis was only run on phases ‘before’ and ‘during’ to avoid the impact of any carry over effect during phase ‘after’ which may have skewed the data (Cleophas, 1993). Nice separation between before and after phases was identified within the LDA plot (Figure 6.26) indicating distinct differences between the two phases. As such all variables were taken forward to Principal Component Analysis (PCA).

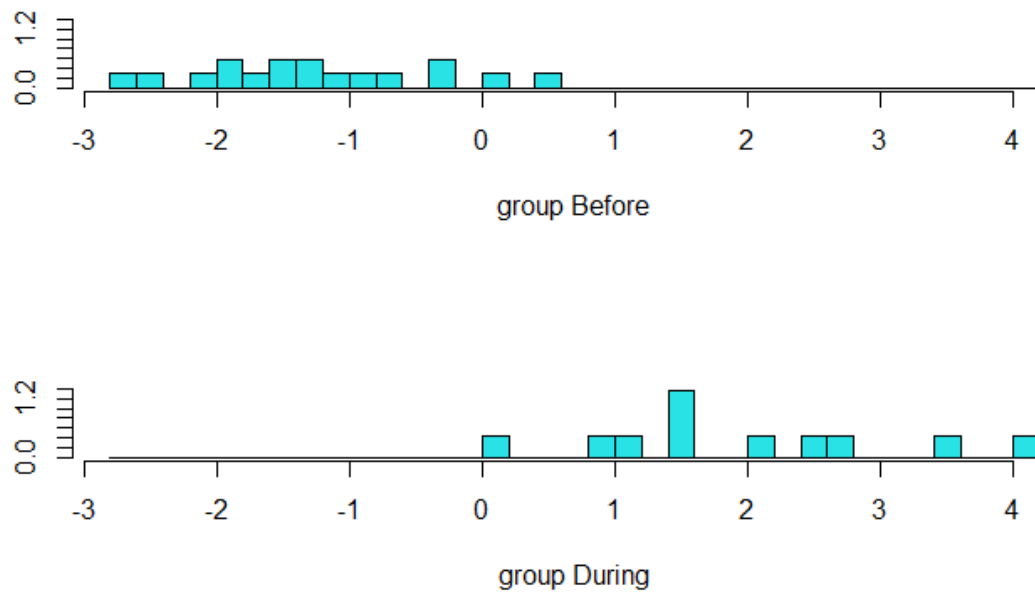


Figure 6.26 Separation of variables between the 'before' and 'during' phases of the scratch trial. With correlations of greater than 0.7 identified between variables 'SDRR' and 'RMSSD'.

This correlation was supported after data were normalised and missing data omitted (Figure 6.27).

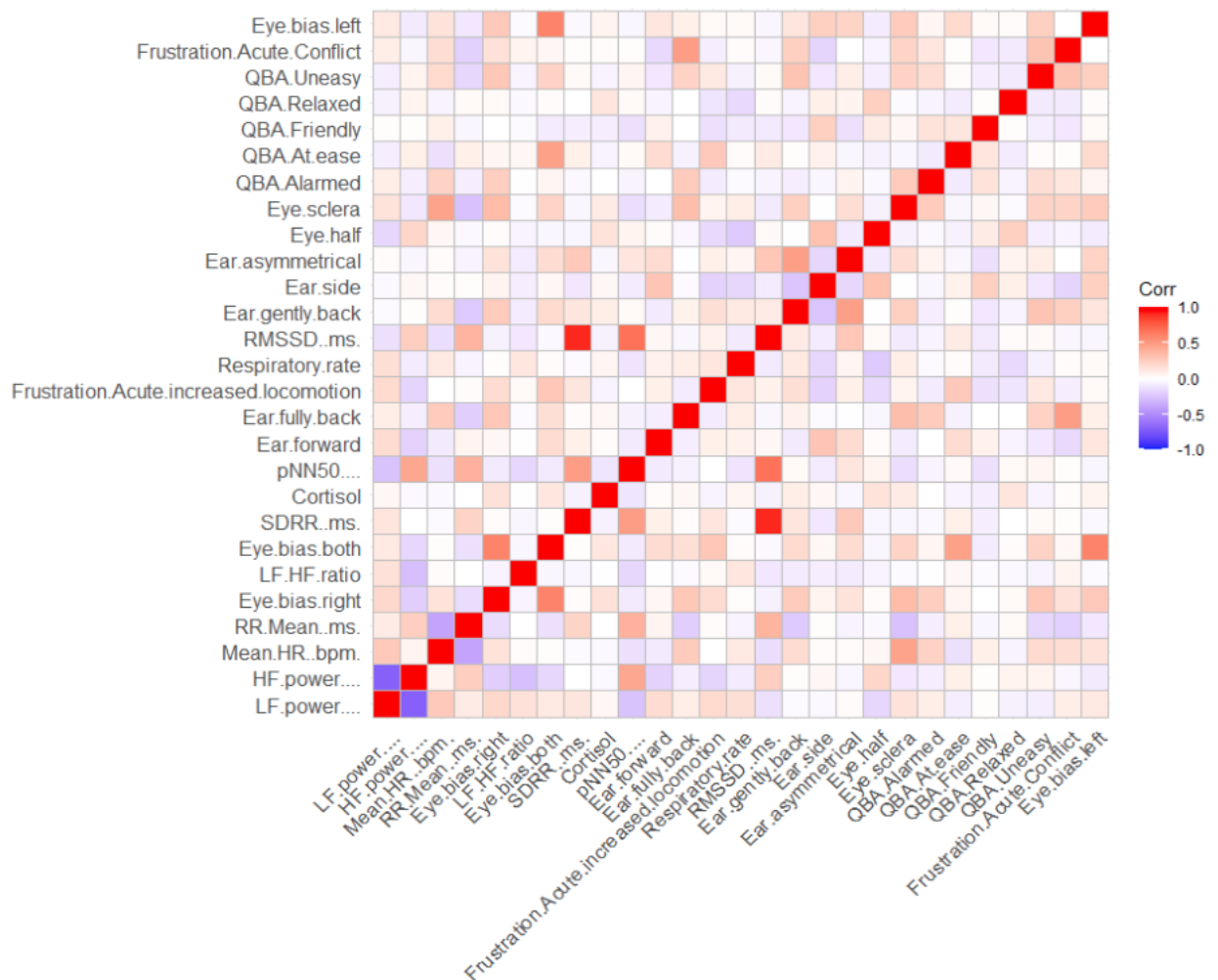


Figure 6.27 Correlation between all variables shown on a correlation plot. Positive correlation between SDRR and RMSSD identified by the bright red squares slightly above and below the central within variable correlation line.

PCA analysis identified that the first five components accounted for 71% of the variation within the data (Figure 6.28); therefore these five components were taken forward for further evaluation.

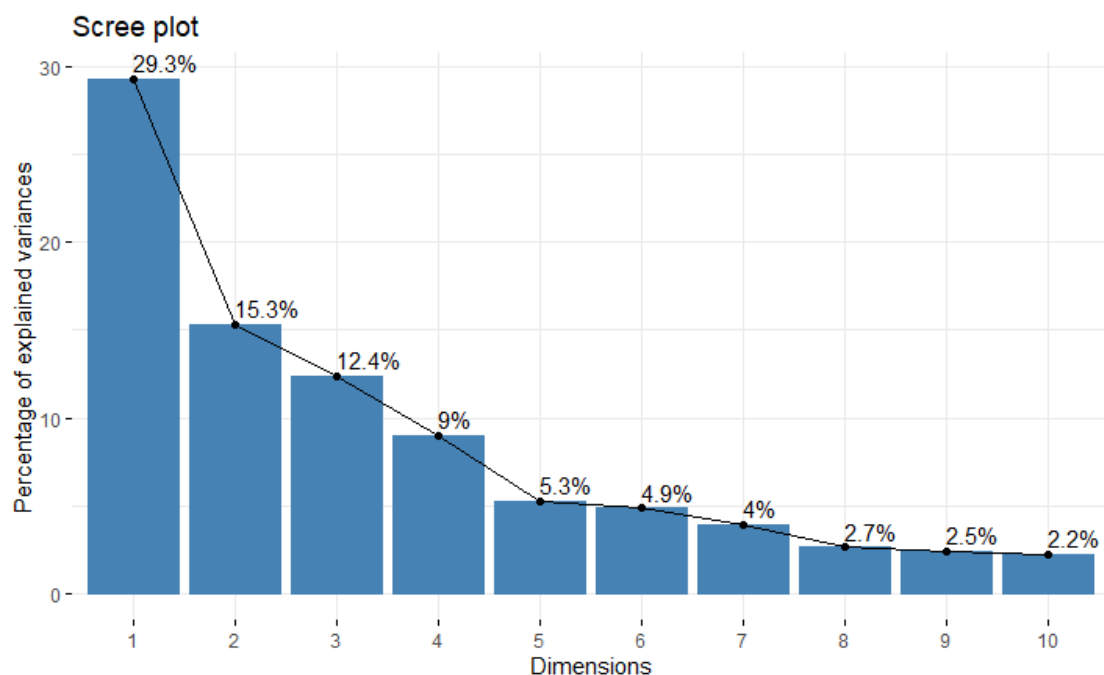


Figure 6.28 Scree plot of the first ten components from the PCA.

The loading plot produced via PCA was assessed with reference to the available evidence from the literature and via evaluation of the relationships identified between the selected measures and a proposed model was produced to explain the relevance of each variable to the measurement of affective state (Figure 6.29). Principal component (PC) 1 (Dim1) appears to be related to the arousal dimension with measures indicative of high arousal such as low frequency power, mean heart rate, cortisol, ears fully back, frustration and negative QBA measures loading positively on PC1 and measures indicative of low arousal including high frequency power, RR mean, ears forward and positive QBA measures, conversely loading negatively on PC1. PC 2 appears to be suggested as a valence dimension with indicators of potentially negative valence such as ears fully back, respiratory rate, ears asymmetrical, eye sclera, negative QBA measures and frustration responses loading negatively and indicators of potentially positive valence such as ears forward, ears side, positive QBA measures and eyes half loading positively (Figure 6.30).

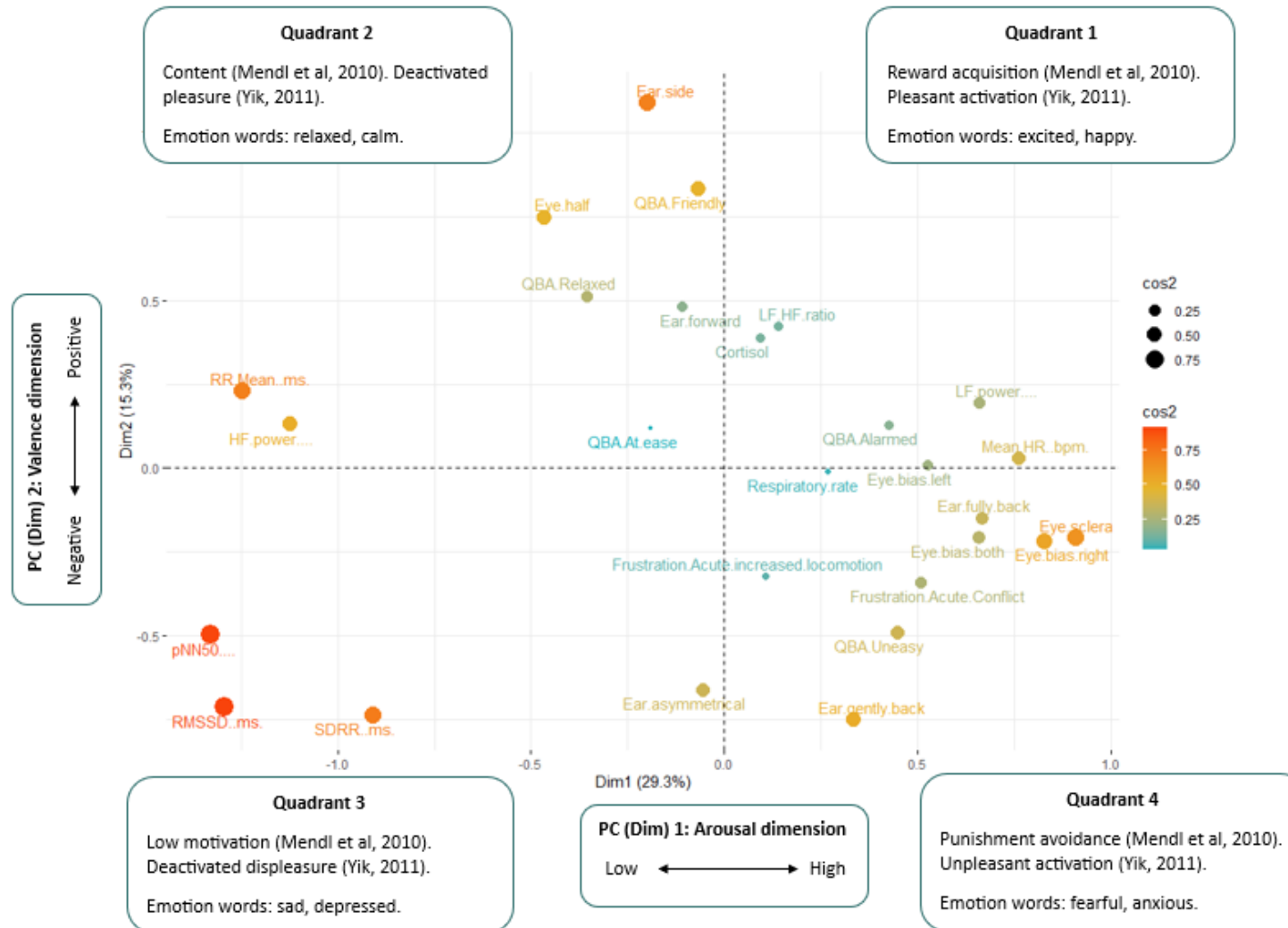


Figure 6.29 Proposed model for measures of valence and affect in equines based upon the framework of Mendl et al. (2010) following evaluation of measurement variable relationships. Dimension 1 is proposed as the arousal dimension and dimension 2 as the valence dimension.

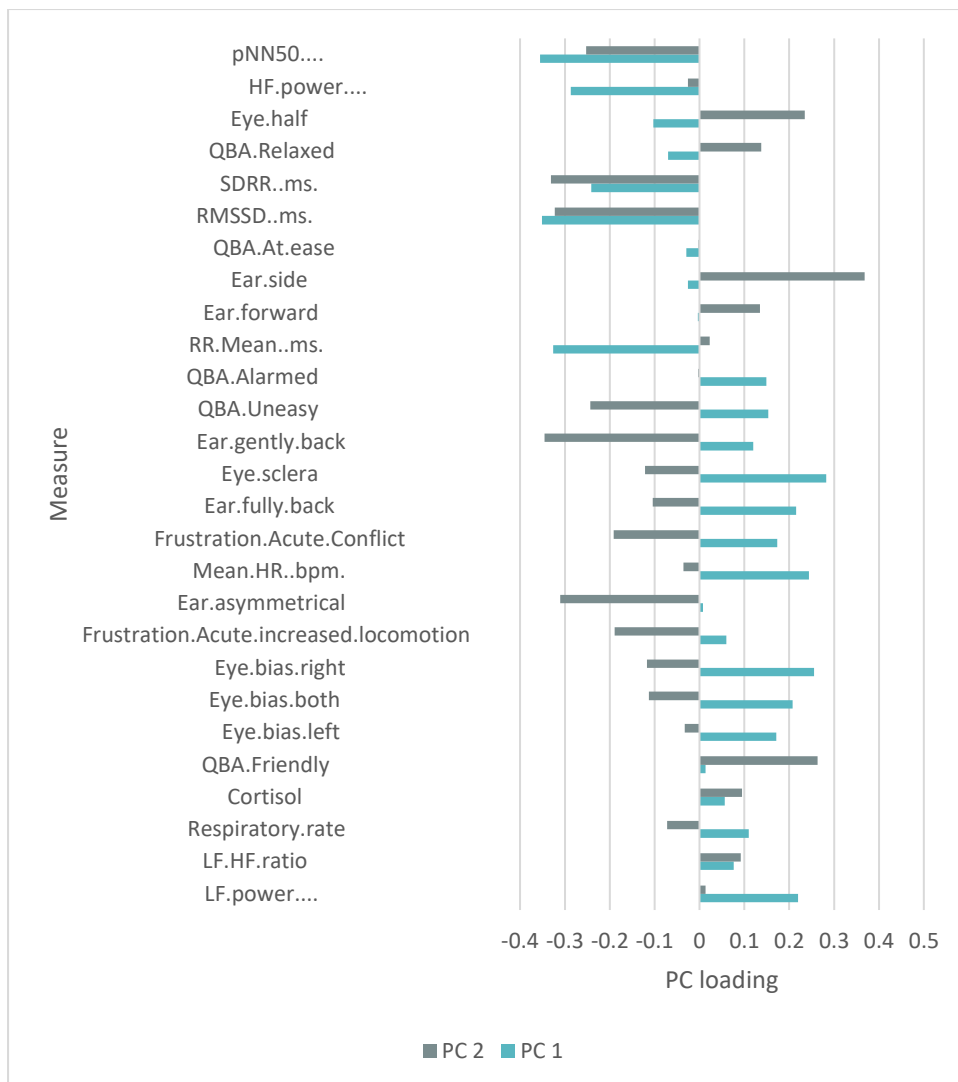


Figure 6.30 Loadings of each measure on principal components one and two.

Principal components three and four may represent elements of activation such as seeking and conflict. PC 3 appears more closely aligned with PC2 and may be indicative of appetitive seeking type states. PC 4 appears to display similar loadings to PC 1 across many variables and could be linked to more activated emotional states (Figure 6.31). PC 5 appears unimportant with no trends identified.

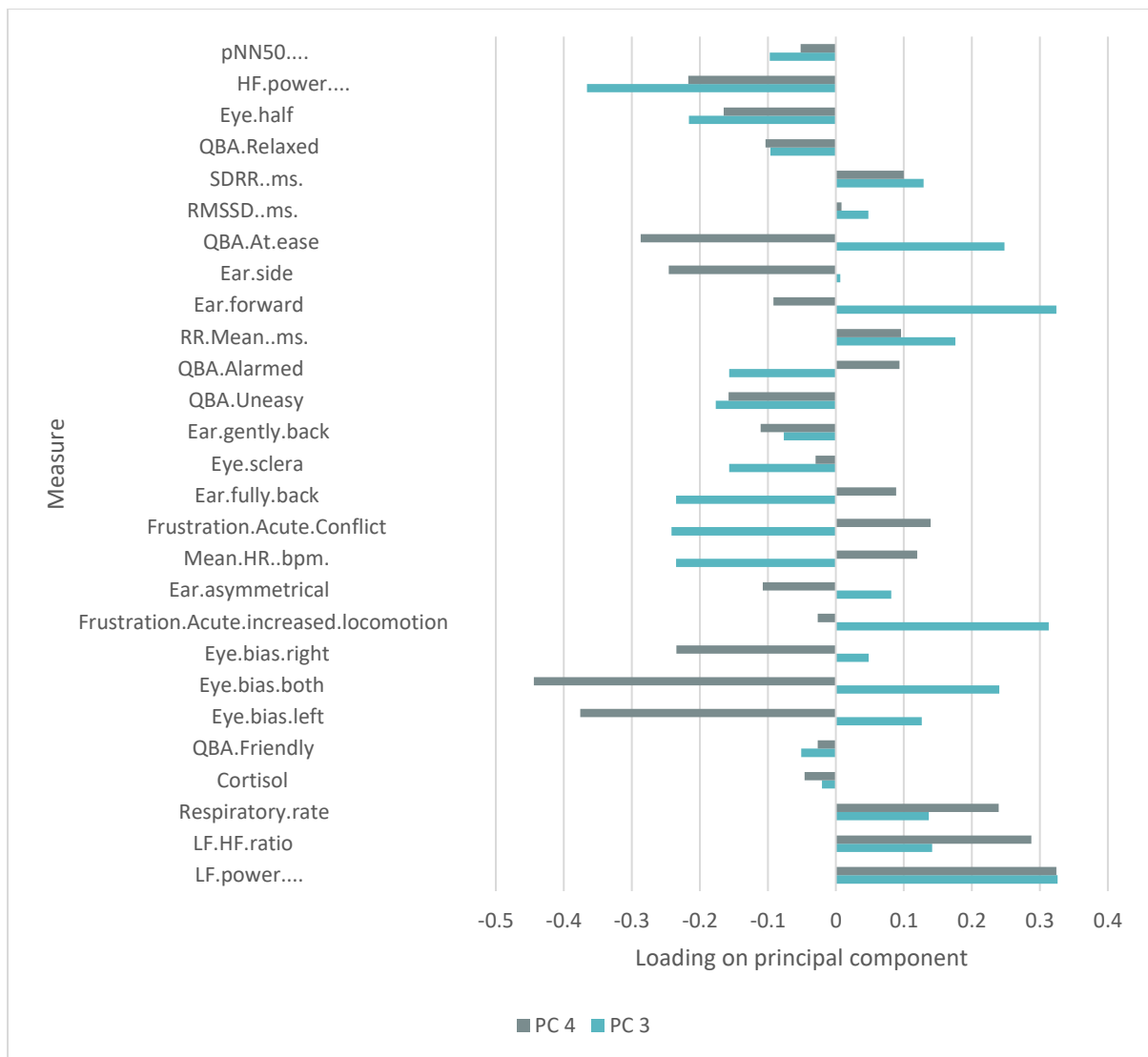


Figure 6.31 Loadings of measures on PCs 3 and 4. Physiological variables show a tendency to load onto both PC 3 and 4 either positively or negatively per measure apart from mean heart rate. Behavioural measures appear to load appositionally on each principal component apart from QBA friendly, which has very low contribution to each component; QBA uneasy, eye sclera and ears gently back also load negatively across both principal components.

When mapped onto the affective circumplex, markers for varying affective states were identified (Figure 6.32).

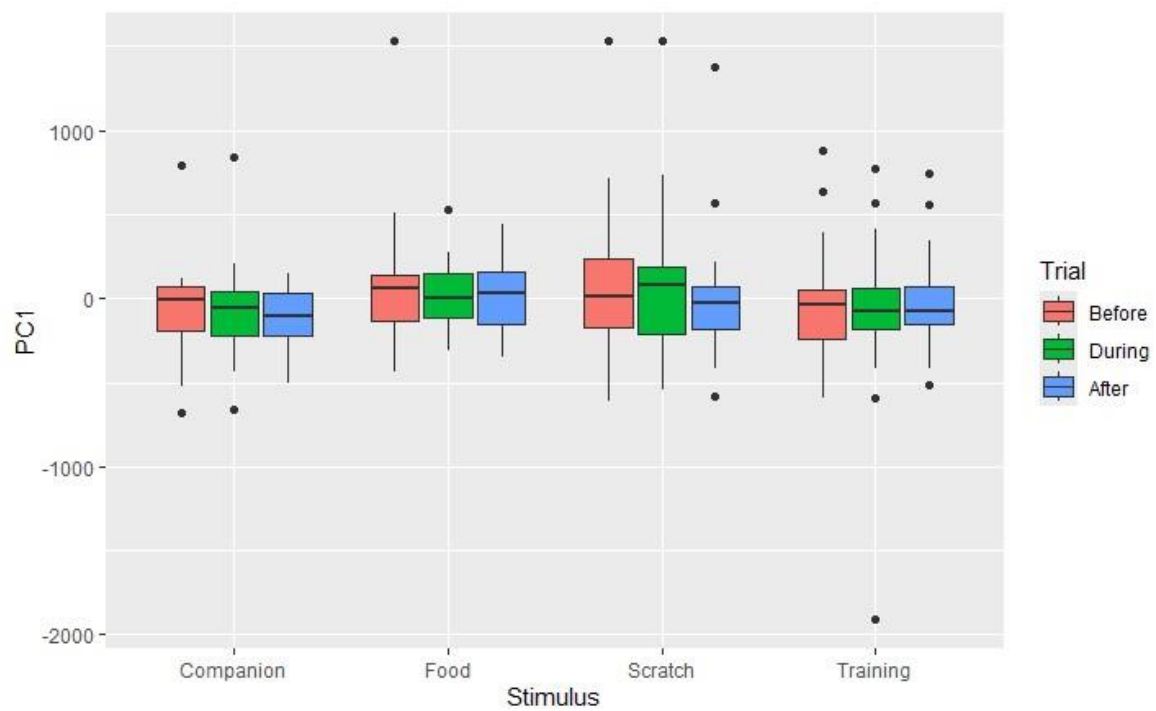


Figure 6.32 Mapped affect markers overlaid onto the affective circumplex with quadrants identified (Q1-4) and green arrows denoting increased pleasure and red aroused denoting decreased arousal. Positive affect markers are placed within green boxes and negative affect markers within peach boxes. Original 12-measure affect diagram from Yik *et al.* (2011) overlaid with the suggested affect markers from this study.

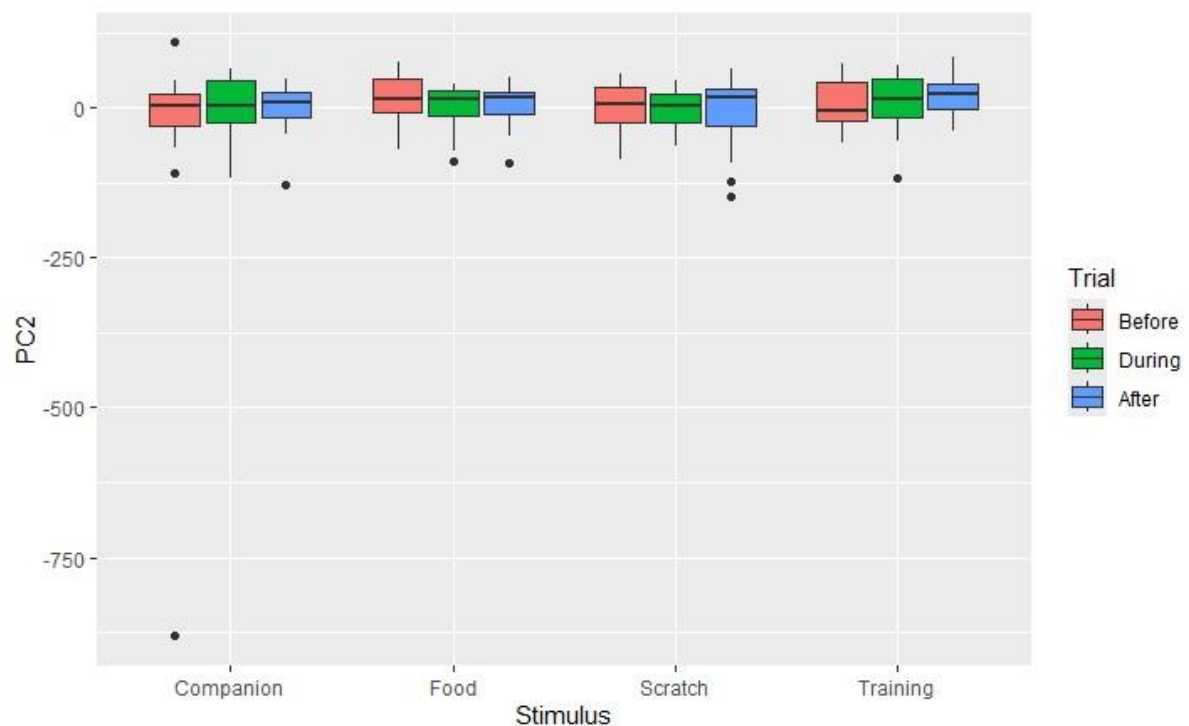
6.3.7 Linear modelling of PCA components by stimulus and trial

Linear models were run to identify any effect of trial phase and stimulus type on PCA component data. A small effect of stimulus type was identified within principal component 1

($F = 2.5609$, $df = 3$, $p = 0.1$; Figure 6.33a) but no effect of trial phase was found. No significant effects were identified for principal component 2 (Figure 6.33b).



a



b

Figure 6.33a Box and whisker plot displaying loadings on PC1 accounting for the effect of stimulus and trial and 6.33b Box and whisker plot displaying loadings on PC2 accounting for the effect of stimulus and trial.

A mixed model was run to consider the effect of individual horses (inputted as random effects) on principal component 1 with a significant difference between horses across all induction stimuli (Type iii ANOVA with Satterthwaite's method, Stimulus: $F = 4.3784$, $df\ 3$, $p = 0.005$) and a significant effect of horse found on the stimulus 'wither scratching' ($t = 2.702$, $df\ 193.554$, $p = 0.007$).

Within principal component 2 there was no significant difference between horses across all trial phases and induction stimuli overall (Type iii ANOVA with Satterthwaite's method, Trial: $F = 0.5811$, $df\ 193.44$, $p = 0.56$; Stimulus: $F = 1.3920$, $df\ 194.18$, $p = 0.25$). However there was a significant effect of horse on component 2 for 'trial during' ($t = 2.525$, $df\ 192.86$, $p = 0.01$), 'trial after' ($t = 2.130$, $df\ 193.42$, $p = 0.03$), 'stimulus food' ($t = 2.898$, $df\ 193.81$, $p = 0.004$), 'stimulus scratch' ($t = 2.406$, $df\ 193.54$, $p = 0.01$), 'stimulus training' ($t = 2.425$, $df\ 193.19$, $p = 0.01$), 'trial during stimulus food' ($t = -2.223$, $df\ 192.97$, $p = 0.02$), 'trial after stimulus food' ($t = -2.065$, $df\ 193.73$, $p = 0.04$) and 'trial during stimulus scratch' ($t = -1.995$, $df\ 192.86$, $p = 0.04$).

6.3.8 Linear modelling of PCA components by personality group

Linear models were run to consider the effect of personality on equine responses to trial phase and stimulus. Some significant relationships were identified as shown in Table 6.13, with increases in arousal displayed by horses within certain personality groups to differing induction stimuli. Additionally, increases in positive valence are suggested within some personality groups in response to differing induction stimuli. Unfriendly horses appear to display increased positive valence after induction trials, rather than during them.

Table 6.13 Results of linear models examining the effect of personality traits on responses to stimuli and trial phase. Results in bold indicate that the entire model was significant.

Principal component	Independent variable	Personality trait	Overall model result	Effect variable	Effect result	Suggested affective influence
PC1	Stimulus	Unfriendly	F = 5.8926, df = 3, $p < 0.001$	Scratch	t = 0.610, df = 199.44, $p < 0.001$	Increased arousal to scratch
PC1	Stimulus	Affiliative	F = 0.5222, df = 3, $p = 0.667$	Scratch	t = 2.056, df = 199.68, $p < 0.05$	Increased arousal to scratch
PC1	Stimulus	Bold	F = 0.6095, df = 3, $p = 0.6096$	Food	t = 1.993, df = 199.80, $p < 0.05$	Increased arousal to food
PC1	Stimulus	Unresponsive	F = 2.7937, df = 3, $p < 0.05$	Training	t = 2.889, df = 200.27, $p < 0.05$	Increased arousal to training
PC1	Stimulus	Nervous	F = 8.7709, df = 3, $p < 0.001$	Food	t = 2.765, df =	Increased arousal to food

					199.991, $p < 0.01$	
PC1	Stimulus	Nervous	F = 8.7709, df = 3, p <0.001	Scratch	t = 4.365, df = 200.533, $p < 0.001$	Increased arousal to scratch
PC2	Stimulus	Exuberant	F = 0.5272, df = 3, p = 0.664	Companion	t = - 2.630, df 74.142, $p = 0.01$	Decreased positive valence to companion
PC2	Stimulus	Bold	F = 2.3902, df = 3, p =0.06	Training	t = 2.633, df = 199.130, $p < 0.01$	Increased positive valence to training
PC2	Stimulus	Bold	F = 2.3902, df = 3, p =0.06	Companion	t = - 2.073, df = 84.244, $p < 0.05$	Decreased positive valence to companion
PC2	Stimulus	Unresponsive	F = 2.3713, df = 3, p = 0.07	Training	t = 2.542, df = 201.695, $p = 0.01$	Increased positive valence to training
PC2	Stimulus	Nervous	F = 1.7283, df = 3, p = 0.16	Companion	t = - 1.999, df =	Decreased positive valence to companion

					85.500, $p < 0.05$	
PC2	Stimulus	Nervous	F = 1.7283, df = 3, p = 0.16	Training	t = 2.180, df = 200.734, $p < 0.05$	Increased positive valence to training
PC2	Trial	Unfriendly	F = 2.4939, df = 2, p = 0.08	After	t = 2.184, df = 204.779, $p < 0.05$	Increased positive valence after induction stimuli

6.3.9 Further Linear Discriminant Analyses (LDA) on companion, food and training stimuli

It was considered that the second stage data reduction via the LDA on the wither scratching stimulus may have affected the results obtained later in PCA analysis and therefore LDA analyses were run separately on each of the three other induction stimuli to consider how results may differ if the LDA on the wither scratching stimulus had not been utilised as a secondary data sift.

For the model 'Training' the LDA was only run on phases 'before' and 'during' to avoid the impact of any carry over effect during phase 'after' which may have skewed the data (Cleophas, 1993). Generally, there was nice separation between before and after phases was identified within the LDA plot (Figure 6.34) indicating distinct differences between the two phases; however, there was some overlap identified.

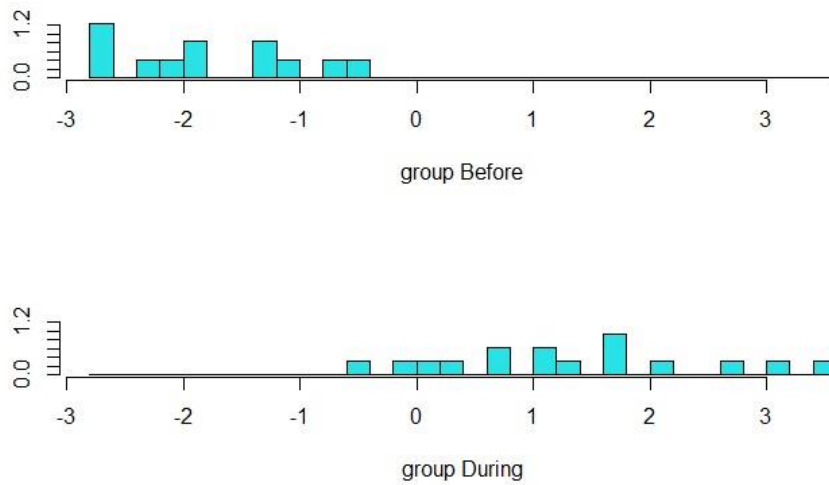


Figure 6.34 Separation of variables between the 'before' and 'during' phases of the training trial. With correlations of greater than 0.7 only identified between variables 'Mean heart rate', 'SDHR', 'Minimum heart rate' and 'Maximum heart rate'.

For the model 'Companion' the LDA was only run on phases 'before' and 'during' to avoid the impact of any carry over effect during phase 'after' which may have skewed the data (Cleophas, 1993). Some nice separation between before and after phases was identified within the LDA plot (Figure 6.35).

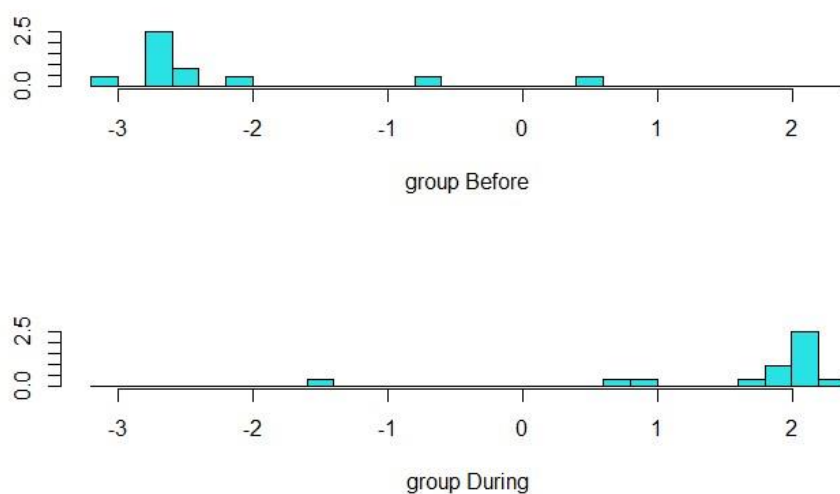


Figure 6.35 Separation of variables between the 'before' and 'during' phases of the companion trial. No correlations of greater than 0.7 were identified between variables.

For the model 'Food' the LDA was only run on phases 'before' and 'during' to avoid the impact of any carry over effect during phase 'after' which may have skewed the data (Cleophas, 1993). There was significant overlap between variables (Figure 6.36).

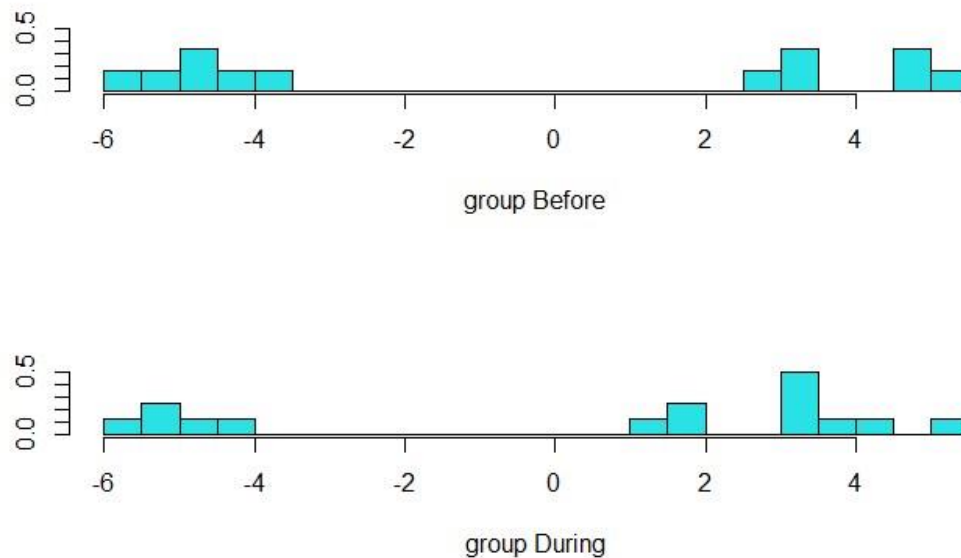


Figure 6.36 Separation of variables between the 'before' and 'during' phases of the food trial. With correlations of greater than 0.7 identified between variables 'Mean heart rate', 'Minimum heart rate' and 'Maximum heart rate'; 'Maximum heart rate' and 'SDHR'; 'Left eye temperature' and 'Left ear temperature'; 'Left ear' and 'Right ear' temperatures.

6.4 Discussion

The aim of this study was to induce positive affective state in a cohort of equines to enable identification of potential measures of positive affective state and to differentiate these from potential markers of negative affective state. It was recognised *a priori* that the induction of positive affect would likely differ between individuals, due in part to their personality traits, as well as the impact of prior life experience, current environment, and learning. The results of this study have identified that the measures utilised to attempt to measure affective state in equines appear to be reliable in terms of their inter and intra-rater reliability. Principal component analysis has identified clusters of variables that appear to be informative

regarding the affective state of the equines within this study and the variables appear to cluster along the axes of principal components that could be described as valence and arousal components. From the results it appears that some behavioural measures may be useful proxies for their physiological counterparts within the PCA clusters and there appears to be an effect of individual personality on equine responses to the four induction stimuli, and to a lesser degree to the trial phase.

6.4.1 Reliability of measures

Cortisol

Whilst the lateral flow and ELISA cortisol tests did not show strong correlation most measures were found to be within +/- 2 n/mol of each other, except for around 15 significant outliers across the dataset of 240 measures. It seems reasonable that were these outliers removed the agreement between these measures would be much higher, indeed 29 of the 80 groups assessed showed correlations of 0.7 or above (many of these 0.9 or above), a further six showed correlations of .6 or above. However, 27 groups showed negative correlations ranging from -.02 to -.92 which will have significantly reduced overall agreement. To ensure accuracy in identifying later correlations with other variables the choice to retain only the ELISA variables appeared valid.

Respiratory rate

Whilst intra-rater agreement was excellent across the respiratory rate measurements these values were not externally tested through inter-rater agreement and as such it is difficult to assess whether the values scored by other raters would match the values utilised within these analyses thereby reducing the external validity of this measure within this trial.

Behavioural data

It appears that the behavioural data measurement was extremely reliable, with excellent agreement across all five raters. The stringent methodology applied for the scoring of behavioural data, including training of the scorers, provision of a written method and the use of Cowlog software for scoring, is likely to have positively influenced the high agreement between scorers. Given the range of experience between the scorers (from behavioural specialists to scorers with no behavioural or equine specialist knowledge) these results appear

promising in terms of providing strong external validation for all the behavioural measures used including body and facial action scoring, QBA measures and frustration measures.

6.4.2 Data reduction

The decision to utilise LDA as a second data sift was undertaken to increase the robustness of subsequent analyses and the results obtained from them. The process was used to facilitate the later modelling of distinctions between groups by separating classes (in this case before and during stimulus application) to ensure as much as possible that variables included within later PCA modelling were distinct from each other, making it more likely that an effect would be found within the analyses. Prior assessment of covariance matrices helped to satisfy the assumptions of the LDA test making it likely that the LDA analysis did accurately classify the different classes, especially as LDA analysis can handle residual multicollinearity quite well (Tharwat *et al.*, 2017). Prior research has identified LDA analysis as an optimal method for data reduction when taking into account sensitivity, specificity and overall classification accuracy of both physiological and behavioural scientific measures (Maroco *et al.*, 2011).

It was considered, retrospectively, that LDA analysis undertaken solely on the wither scratching data may have inadvertently skewed later results by not considering the separation of variables identified across the other three induction methods. However, subsequent LDA analysis of the other induction methods appears in some ways to support the initial analysis methodology in that separation between variables across the other induction methods was not as good as that seen within the wither scratching LDA; separation of variables within the food trials was highlighted as being much poorer than for wither scratching, companion, and training trials. It could therefore be argued that utilising wither scratching as the gold standard induction method (due to its support within the literature) enabled sensitive and specific classification of variables making the later PCA analysis more reliable. Conversely, it could be considered that the variation in the separation of measures across the different induction methods may mean that the PCA results obtained are most reliable for measuring affective state induced by wither scratching compared with other methods. However later PCA analysis considered all methods across all the variables and because there was no overlap between variables within the wither scratching LDA this enabled all measurement variables to be considered within the final analyses. This likely avoided undetected effects that could have occurred if certain variables (seen as overlapping on companion, food and training LDA

results) had been removed prior to PCA analysis. As such it appears likely that the chosen LDA methodology facilitated more realistic later PCA analysis across all induction methods and variables.

6.4.3 Identifying potential markers of affective state – valence and arousal domains

Following the two-step data reduction process the first PCA analysis did not appear to contain enough variables to be meaningful in terms of measuring affective state using both physiological and behavioural measures. The inclusion of only five behavioural measures appeared to significantly impair interpretation of the results with limited clustering of behavioural and physiological measures, in addition there was a heavy weighting of measures on PC1 (43.5% of the data variation) compared to PC2 (14.4%). The addition of the further twelve behavioural variables enabled a more detailed PCA analysis and facilitated further interpretation of the data. Examination of the second loading plot output identified what appears to be clear separation between behavioural measures that are reportedly indicative of positive and negative affective state on PC2, likewise separation of physiological variables (predominantly heart rate analysis variables) that are reported in the literature to account for high and low levels of arousal was identified across PC1.

This study was designed around the premise of the core affect 2-dimensional model (Mendl *et al.*, 2010; Russell, 2003; Burgdorf and Panksepp, 2006) and the affective circumplex (Yik *et al.*, 2011; Russell, 1980) with the proposal that different physiological and behavioural measures may be identified which could be proposed as potential measures of varying affective states. The output derived from PCA plot two provides an initial starting point for further exploration of the identified measures of affective state across the four quadrants.

Quadrant one: Reward acquisition, pleasant activation

Within quadrant one there are clusters of measures that are suggested within the literature to indicate a state of alertness such as 'LF power', indicative of sympathetic arousal (Scopa *et al.*, 2020) which appears statistically close to 'QBA alarmed'. Both measures are loaded positively on PC1 (the suggested arousal domain) (LF power 0.07, QBA Alarmed 0.15) however, whilst loaded positively onto PC2 (the suggested valence domain) they are only positively loaded at a low level (LF power 0.013; QBA Alarmed 0.002) suggesting that these

measures may be more indicative of arousal than valence. Likewise, 'mean heart rate' is loaded strongly positively on PC1 (0.24) but almost neutral on PC2 (-0.03) suggesting that it too is a measure more of arousal than valence, a premise supported by Hall *et al.* (2018). Within this study heart rate was measured over time in beats per minute and was not mathematically decomposed as per the method of Jansen *et al.* (2009) to identify physical and emotional components which limits the interpretation of the placement of this measure within the loading plot. Also within quadrant one, 'LF:HF ratio' is identified as having a positive loading on both PC's; 'LF:HF ratio' is identified within the literature as a measure to quantify the changing relationship between sympathetic and parasympathetic nervous activity indicating the level of parasympathetic dominance (Pagani *et al.*, 1984) with a low 'LF:HF ratio' suggested by many to be indicative of high parasympathetic dominance (Scopa *et al.*, 2020; Stucke *et al.*, 2015; Pham *et al.*, 2021). However there is disagreement regarding this by others such as Shaffer and Ginsberg (2017) and Billman (2013) who suggest that this oversimplifies the complex non-linear interactions between sympathetic and parasympathetic divisions of the autonomic nervous system; with 'LF power' being particularly difficult to disentangle from its indications within parasympathetic drive, sympathetic drive and other unidentified factors. However, 'HF power' has been strongly linked to parasympathetic activity and is much less disputed (Billman, 2013) and within this study is shown within the loading plot to be appositional to 'LF power' and to 'LF:HF ratio'. This appears to indicate that, within these results, 'LF power' indicates a higher sympathetic state and that it is clustered with increased 'LF:HF ratio' within quadrant 1. Supporting this is the clustering of 'cortisol' within quadrant one, close to 'LF:HF ratio'. This cluster supports the denotation of quadrant one as an alert, potentially reward acquisition, appetitive state with the possibility that rises in cortisol level are resultant of increased arousal emanating from a state of anticipation and excitation. However, it is important to evaluate the fact that cortisol does not only represent a measure of stress, rather it plays important roles in blood sugar regulation, inflammation reduction, metabolism regulation and memory formation meaning that it is possible that rises in cortisol may represent other elements aside from a stress response (Mormède *et al.*, 2011). This is further supported when considering that cortisol was seen to increase across all induction stimuli (Appendix I-2), which may have been the result of increased arousal or alternatively could be related to learning and memory processing (van Ast *et al.*, 2013). The final measure within this quadrant is 'eye bias left' which sits with a

positive loading on PC1 (0.17) and neutrally on PC 2 (-0.03) indicating that, similarly to mean heart rate, this measure may be indicative of arousal but with a valence more positively skewed than that of 'eye bias right'.

Quadrant two: Contentment, deactivated pleasure

As described above, 'HF power' sits marginally within quadrant two with negative loading on PC1 (-0.28) and almost neutral loading on PC2 (-0.02). 'RR mean', the time between successive heartbeats (also termed inter-beat intervals, IBI), should be appositional to heart rate, with increasing heart rate leading to shorter IBI's; within this analysis 'RR mean' was directly appositional to 'mean heart rate' loaded negatively on PC1 (-0.32) and almost neutral on PC2 (0.02) indicating that a high 'RR mean' may suggest low arousal (Shaffer and Ginsberg, 2017), potentially with positive valence depending on the other variables it is associated with. The QBA variable 'at ease' was loaded negatively on PC1 (-0.02) and neutral on PC2 (-0.002) indicating that it may represent low arousal in a relatively neutral valence (Barrett, 2017a), hence its placement within the 'contented' quadrant. Loading more highly on PC2 were the variables 'ears side' (0.36), 'eye half' (0.23), 'QBA relaxed' (0.13), 'ear forward' (0.13) and 'QBA friendly' (0.26). This cluster may indicate states of relaxation and as such could be proxy measures of RR mean and HF power, within positive valence. 'QBA friendly' appears to sit on the border between low and high arousal and therefore could also be a marker of appetitive state such as reward anticipation, supported by the data which shows increases in scores for 'QBA friendly' during the companion and training stimuli (Appendix I-3). Additionally, 'ears forwards' frequency rate, a factor previously associated within the literature as indicative of positive affect (Lansade *et al.*, 2018) was seen to increase during the training stimulus and on companion approach (Appendix I-4). Together these quadrant two variables may be indicative of an overall state of relaxation and contentment.

Quadrant three: Low motivation, deactivated displeasure

Within quadrant three 'pNN50' and 'RMSSD' were closely clustered with negative loadings on both PCs 1 and 2. 'pNN50' is the percentage of successive RR intervals that differ by more than 50ms, ergo, it indicates high heart rate variability. Similarly, 'RMSSD' is the root mean square of successive RR interval differences therefore it is logical that these measures correlate closely. Within quadrant three these measures may indicate an association with low heart rate variability (these variables are lower in Q3), indicative of a shutdown, low

motivation affective state. 'SDRR', the standard deviation of RR intervals, is also clustered close to both 'RMSSD' and 'pNN50' which would be expected given their similar properties (Shaffer and Ginsberg, 2017). All these variables sit with the time domain indices of HRV which quantify the amount of heart rate variability observed during the monitoring period (Shaffer and Ginsberg, 2017). 'pNN50' is closely correlated with parasympathetic activity alongside 'RMSSD' which is often the preferred metric by researchers (Otzenberger *et al.*, 1998), this is especially true of short-term HRV testing utilised within the design of this study (Ciccone *et al.*, 2017). Interestingly 'SDRR' and 'pNN50' are described within the literature as being more accurate measures of longer-term HRV over a 24-hour period therefore the finding that these measures are closely correlated within this study is interesting in terms of their prospective use over shorter measurement periods (Shaffer and Ginsberg, 2017). The only behavioural measure sitting within this quadrant is 'ear asymmetrical', loading neutrally onto PC1 (-0.007) and negatively on PC2 (-0.31) potentially indicating low levels of parasympathetic dominance (its placement is closer to the midline) in negative affect (strongly negatively loaded on the affect dimension). This measure may be indicative of a rising level of arousal which could move further towards quadrant 4 and potential frustration responses if subtle cues are not noticed by a handler and is reported within the literature to be an indicator of negative affect and/or divided attention towards stimuli (Maurício *et al.*, 2023).

Quadrant four: Punishment avoidance, unpleasant activation

Quadrant four is suggested to be representative of negative affective state in high arousal. Within this quadrant several previously identified behavioural markers of negative affect are closely clustered including 'ear fully back' (Lansade *et al.*, 2018), 'eye sclera' (Lansade *et al.*, 2018), and 'QBA uneasy' (Dalla Costa *et al.*, 2016a). Also clustered close to these measures are the eye bias measures 'right' and 'both' with each measure being highly positively loaded on PC1 (0.25 and 0.20 respectively) and negatively loaded on PC2 (-0.11 for both measures) which could be considered as potential measures of more negative valence in a highly aroused context. However, these variables are located relatively close to 'eye bias left' which also loads highly on the arousal dimension potentially signifying that eye bias may be indicative of arousal more generally, rather than specific to valence or that it may be specific within individuals leading to a variety of responses across this cohort as indicated previously by Leliveld *et al.* (2013). The 'respiratory rate' variable sits just within this quadrant, being

moderately positively loaded on PC1 (0.10) and just negatively loaded on PC2 (-0.07); although this variable is along the same plane within PC1 as 'eye bias left' and 'mean heart rate' with no visual relationship being identified when examining these variables in a paired manner (Appendix I-5). 'Respiratory rate' did not vary widely during the different stimuli or trials phases (Appendix I-6) and as such may be a less nuanced measure of the valence of affective state and of subtle increases in arousal.

Behavioural variables 'QBA uneasy', 'frustration acute locomotion', 'frustration acute conflict' and 'ears gently back' were relatively closely clustered within the middle of quadrant four (PC1 loadings 0.14, 0.05, 0.17 and 0.12; PC2 loadings -0.24, -0.18, -0.19 and -0.34 respectively). Logically, behaviours indicative of uncertainty such as conflict, increased movement and general uneasiness would be expected to be closely aligned (Draaisma, 2017) and of a lower level of arousal than more overt signs such as 'ears fully back' and 'eye sclera' that are loaded more significantly on PC1. 'Ears gently back' was not anticipated to be identified within this quadrant as prior research has indicated that this ear position, with the ears roughly in line with the nostrils, is indicative of positive emotions in equines (Maurício *et al.*, 2023). However, within this cohort this does not appear to be supported. Conversely it is possible that the coding of 'ear gently back' behaviour was inaccurate in the original video coding leading to erroneous attribution of this variable however due to the very high level of inter-rater agreement within the behavioural coding this seems unlikely. What is clear is that 'ear gently back' is indicated as a variable in lower arousal than 'ears fully back' and potentially of greater negative valence, perhaps indicative of general uneasiness given its relative statistical proximity to 'QBA uneasy'. Interestingly 'ear gently back' increased markedly during wither scratching as has been found previously (Lansade *et al.*, 2022; Lansade *et al.*, 2018) which may support its use as measure more attributable to quadrant 2; 'ear gently back' frequency rate varied hugely across the other three induction stimuli which could account for its ambiguous placement within the PCA analysis (Appendix I-7).

6.4.4 Consideration of seeking and conflicted emotions

It is postulated that components three and four may be linked to seeking-type motivation and conflicted emotion respectively. Variables loaded onto PC3 in a similar manner to PC 1 and were neutral or negatively loaded within quadrants 2 and 3, potentially indicating low seeking motivation; mixed within quadrant 4 and high within quadrant 1 potentially indicating higher

seeking motivation which could link to higher arousal. Experimental trials on rats have reported strong links between dopamine-mediated arousal and seeking behaviours; the administration of dopamine antagonists was seen to reduce both arousal and seeking behaviour with immediate reinstatement of the behaviour upon antagonist withdrawal (Kaźmierczak and Nicola, 2022). The effect of dopamine is likely widespread across brain systems, it is associated with activation of glutamatergic inputs on the ventral tegmental area (VTA) dopamine neurons (Enrico *et al.*, 1998) and is important in stimulus-reward learning whereby incentive salience is assigned to reward cues thereby powerfully motivating and controlling seeking behaviour and is implicated in instances of reduced impulse control and potentially within frustration (Flagel *et al.*, 2011; Fellous and Suri, 2002). It has been suggested that lower plasma dopamine concentration in horses is associated with higher levels of fearfulness and vice versa (Kim *et al.*, 2023), which suggests that further study of the importance of component three variables could include evaluation of plasma dopamine to identify whether high loadings on PC3 are associated with higher levels of plasma dopamine indicative of seeking behaviours or whether levels are lower, potentially indicating greater fearfulness and a more negative affective state.

Variables loading positively on to PC4 were predominantly located with quadrant 2 (contentment) and PC3 (low motivation), loadings were mixed in variables located within quadrants four and one. Deactivated states on the affective circumplex suggested positive loadings on the component whereas activated displeasure states (activation, unpleasant activation, and activated displeasure) were more mixed in their loading. Quadrant 2 states of pleasure, pleasant deactivation and deactivated pleasure were negatively loaded. These findings may indicate a state of emotional conflict within the mixed loading states (Q1 and 4) where the potential for reward acquisition and punishment avoidance could lead to uncertainty and conflict. Variables associated with the mixed loadings included 'eye bias both', 'eye bias right', 'frustration increased locomotion', 'ear fully back', 'eye sclera', 'ear gently back', 'QBA uneasy' and 'QBA alarmed', 'cortisol', 'LF:HF ratio', 'eye bias left' and 'mean heart rate' which potentially indicate a predominantly aroused state with suggestion of approach and avoidance behaviours indicative of conflicted emotions. This suggestion is supported by Christensen *et al.* (2021) who found increases in heart rate, cortisol, avoidance behaviours and startle responses in horses displaying conflict behaviour, and supported by

Kienapfel *et al.* (2014) and Jastrzębska *et al.* (2017) who identified similar behavioural and physiological markers in conflicted horses. Interestingly Phelipon *et al.* (2024) distinguished between horses displaying positive anticipation and frustration identifying frustrated horses as those with higher levels of 'ear fully back', 'eye sclera' and increases in locomotion and reduced visual focus on the handler in horses in positive anticipation. These findings could be built upon in further studies to aid differentiation between appetitive and aversive states of conflict, i.e. positive anticipation and frustration.

6.4.5 Limitations of the study

This exploratory study has several significant limitations which limit the reliability and generalisability of the results reported.

The sample

The sample size within this study is relatively small at 20 equines, considering the estimated required sample size calculated at 384 equines for the UK population of ~850000 equines, with a confidence level of 95%. However, many studies similar to this one have sample sizes within the same range (13-32 horses) including 27 horses in the study of facial expression and oxytocin as possible markers of positive emotions in horses (Lansade *et al.*, 2018), 16 horses in a judgement bias test to identify behavioural indicators of affective state (Hintze and Schanz, 2021), 13 horses within the PhD thesis examining positive emotion in horses (Stratton, 2022) and 32 horses within a PhD thesis examining emotional responses to reward and withheld reward in horses (Kappel, 2023). The sample within this study was broadly evenly distributed across sexes, age (young = 0-4 years, mature = 5-15 years and aged = 15+ years) and represented a range of breeds including pure and mixed-breed horses which aids generalisability of the results across equine demographics, but the low sample size prohibits further analysis of the effects of these demographic differences within the statistical modelling.

Using proxies of emotional state

The measurement of subjective, emotional experience in non-human animals is impeded by the absence of linguistic reporting and as such measurement must rely on proxies in the form of behavioural and physiological measures (Mendl *et al.*, 2010). However, emotional states are multifaceted and composed of physiological, behavioural, cognitive and subjective

components unique to the individual within a particular context and time continuum (Paul *et al.*, 2005b). Earlier research focussed heavily on the induction and measurement of discrete emotional states such as fear and anxiety and other negative emotional and welfare states such as pain leaving a gap in the literature regarding the measurement of positive affective states (Mendl *et al.*, 2010; Boissy *et al.*, 2007a). More recently there has been a suggested shift in focus to more dimensional approaches to the study of affect in animals within two, three or four dimensions (Mendl *et al.*, 2010; Russell and Barrett, 1999; Barrett, 2017b; Burgdorf and Panksepp, 2006; Cabanac, 2002) to assess the placement of affect in relation to valence (positivity/negativity) and arousal both in the short term and over longer periods of time. However, even within the dimensional approach there is disagreement between researchers as to how different affective dimensions are related to one another with some theorists proposing that positive and negative affective states are inversely proportional to each other (Russell, 1980) and other suggesting that they are independent of each other (Cacioppo and Tassinary, 1990; Watson *et al.*, 1988).

A limitation to any of these approaches is the lack of objective surety regarding the affective state we are attributing to an individual at a moment in time; even within the human literature, measuring a person's affective state is considered to be one of the most vexing problems in affective science (Mauss and Robinson, 2009) because it is impossible to ascribe with certainty an affective state to any living being without being the one directly experiencing that affective state. Even self-report of affective state in humans has been deemed unreliable in some contexts (Robinson and Clore, 2002). Furthermore, there is evidence within the literature regarding the differentiated display of emotional behaviour in humans (Hu *et al.*, 2022; Wagner, 1990) and animals (Snowdon, 2003) and the influence of individual personality on the experience and presentation of affective states (Ashton, 2022; Gartner *et al.*, 2016; Hausberger *et al.*, 2011). However, many of the measures now being utilised within measurements of affect in non-human animals are derived from measures that have previously been extensively verified in human trials. Indices such as heart rate variability have been extensively reviewed within the human affect literature (Cacioppo *et al.*, 2000) and HRV is deemed to be a useful measure of affective arousal when using it to gauge responding along the dimensional approach, rather than the discrete emotion approach (Mauss and Robinson, 2009). Within humans, facial actions have been shown to reliably indicate the

valence of affective state (Russell, 1994; Mauss *et al.*, 2005) however application to discrete emotions, compared with a dimensional approach has proved more challenging (Mauss and Robinson, 2009). Likewise, measures of affect via assessment of whole-body posture have identified differences in the types of emotions expressed via whole body posture changes versus facial action changes in humans. However, this may have a highly cognitive component with researchers reporting higher level emotions such as embarrassment, guilt, pride and shame being denoted by changes in whole body posture and 'simpler' emotions such as anger, fear and happiness displayed across variable facial actions (App *et al.*, 2007). Within this study it is possible that whole body postures such as frustration responses and QBA measures are describing broader affective states, particularly as these measures are not seen at the extremes of any of the principal components whereas facial actions may be indicative of more specific or extreme emotional states and were often seen to be loaded heavily on a principal component.

Omission of potential measures

This study was limited in the measures employed both intentionally and by availability of equipment, time and expertise. A core objective of the research was to consider methods of measuring affective state that could be used by the layperson without access to advanced or expensive equipment and expertise. As such the strategy to utilise several low-tech, low-skill methods of measurement alongside more scientifically advanced measures was intentional but limited the choice of measurement methods employed. Both time available and video recording parameters limited further behavioural measurements such as blink rate which has previously been utilised as a measure of startle response (increases in blink rate) and of approach motivation (lower blink rate) (Davis, 1989). This measure has been relatively extensively studied in horses in relation to learning (Evans *et al.*, 2024), stress (Lelláková *et al.*, 2023; Mott *et al.*, 2020; Merkies *et al.*, 2019) and attention (Cherry *et al.*, 2020; Phelipon *et al.*, 2024). Vocal characteristics were not selected for study, despite some prior evidence suggesting that changes in vocalisation may be suggestive of affective state (Briefer *et al.*, 2017b; Stomp *et al.*, 2018; Maigrot *et al.*, 2017), due to the hypothesised low frequency with which vocalisations were predicted to occur throughout the trials, a hypothesis which was supported within analysis of the data. The use of electroencephalography (EEG), a measure which is purported to be a useful measure of approach motivation (denoted by greater left-

frontal EEG activity) and avoidance motivation (greater right-sided activation), was outside the scope of this thesis due to the vast learning curve required to achieve proficiency in interpreting EEG data and the lack of availability of sensitive, non-invasive equipment. Likewise, facial electromyography (EMG) was outside the scope of this work with evidence limited in equines but more widely studied in humans (Mauss and Robinson, 2009).

Data pertaining to ear and eye thermography were omitted from later analyses due to the high correlation between all variables which indicated that there were not differentiated enough to be meaningful in later analyses. However, this omission may have resulted in missing information that could have been pertinent to measuring affective state, especially when considering their placement within the PCA loading plot when this was retrospectively analysed (Appendix I-8). Both Kappel (2023) and Stratton (2022) identified changes in eye and ear temperature between treatments in their studies of affective state in equines, however, similarly to this study, they did not identify differences between opposingly lateralised features. This potentially indicates that lateralised differences may be extremely subtle (potentially requiring more sensitive equipment to detect changes), or that lateralised differences are not valid measures of affective state in equines. Additionally, some of the QBA and frustration variables were omitted due to them not being observed frequently in the cohort of subjects. Whilst their inclusion was infeasible, due to the low number of data points, instances of these variables may be more evident in different cohorts of equines and should therefore be utilised in any subsequent data collection.

Some authors suggest a more cognitive approach to emotion assessment (Paul *et al.*, 2005b) through methods such as appraisal checks (Baragli *et al.*, 2009), attention bias tests (Kappel *et al.*, 2024; Rochais *et al.*, 2018), broader cognitive bias testing (Mendl *et al.*, 2009; Löckener *et al.*, 2016), memory bias tests (McLean, 2004; Sankey *et al.*, 2010a) and judgement bias tests (Freymond *et al.*, 2014). Biases in aspects such as judgement, memory and attention have been shown to vary according to emotional state and support from the human literature suggests that a cognitive approach inclusive of behavioural and physiological measurements may help to understand the cognition-emotion link and develop our assessment of affective states (Harding *et al.*, 2004; Spruijt *et al.*, 2001). Within this study formal cognitive assessment processes were not undertaken (aside from preference testing of stimuli) however further

evaluation of the video data may allow for assessment of aspects such as attention and judgement biases through evaluation of the equine responses to the four induction stimuli.

Omission of potential induction stimuli

Several potential induction stimuli were omitted from the methodological design due to the perceived difficulty in obtaining the selected behavioural and physiological measures during their implementation. The four selected induction stimuli had valid rationales for their inclusion however it is possible that some equines may have found some of the omitted induction stimuli more intrinsically reinforcing or less negatively valenced or arousing than those that were selected. In particular, some ethologically relevant induction stimuli such as access to graze and browse, turnout and horse-horse interaction at liberty may have provided further useful data in relation to the physiological and behavioural responses to these stimuli and further work could focus on the attempted induction of positive affect using only at liberty, ethologically relevant induction stimuli to reduce the potentially negative effects of confinement and restraint on the achievement of positive affect (Lansade *et al.*, 2012; Fureix *et al.*, 2009; Löckener *et al.*, 2016).

To address the potential confounds of variable responses to the induction stimuli, personality assessment of all horses was undertaken for use in the linear models. This process has allowed for consideration of individual personality within perception of the induction stimuli and results indicated that there were distinct differences between the responses of different personality types to the four stimuli. Unfriendly horses appeared to display increased arousal to the scratch stimulus, potentially suggesting that they disliked this induction stimulus; and were happier (increases in positive valence) after trials finished. Affiliative horses also displayed increased arousal to the scratch stimulus, potentially relating to positive arousal. Unresponsive horses showed greater arousal to the training stimulus; however, they may also have shown increased positive valence during the training task which is an interesting finding in terms of the potential to utilise this method to improve affect in these horses. Nervous horses appeared more aroused towards the food and scratch stimuli, more positively valenced towards training (again suggesting that this could be a useful method for improving affective state in nervous horses) and more negatively valenced towards a companion which may be related to erroneous selection of the companion horses or to other factors relating to the individuals or the environment. Bold horses appeared more positively valenced towards

the training compared with the companion stimuli, potentially indicating that these horses preferred the experience of a new positive training task to the addition of a conspecific. To further address confounds relating to stimuli selection, preference testing of horses prior to the food and scratch trials was undertaken to maximise the potential for positive emotional responses to these stimuli and selection of conspecifics was carefully considered prior to that trial. These tests were carried out in line with previously reported methodologies for preference testing in horses (Holzer *et al.*, 2022; Kieson *et al.*, 2020; Vinassa *et al.*, 2020) however these interventions do not account for the possibility that some horses may have found other unselected induction stimuli more appetitive than the four stimuli chosen for the study.

Selection of ‘gold standard’ induction and measurement methods

To inform the statistical analysis a perceived ‘gold standard’ induction method, wither scratching, was selected to allow for further data reduction. The selection was chosen due to the evidence weighted within the literature (Lansade *et al.*, 2022; Lansade *et al.*, 2018; Thorbergson *et al.*, 2016; Watson and McDonnell, 2018) and the high rank of the induction stimulus by the Delphi experts. Whilst the use of wither scratching within the data reduction strategy may have biased the choice of variables included within the later analyses, *post hoc* examination of the other three induction stimuli, via linear discriminant analysis, suggested that the selection of this induction stimulus as ‘gold standard’ was valid due to the absence of overlap of variables. However there remains a possibility that this methodological decision may have impacted the overall results reported.

Qualitative behaviour assessment (QBA) was selected as the ‘gold standard’ measurement method due to its validation within the literature (AWIN, 2015; Dalla Costa *et al.*, 2016a; Estevez *et al.*, 2017; Ruet *et al.*, 2022) however there is limited evidence in relation to the relationship of these behavioural measures to the physiological methods selected. The results of this study have however identified some interesting relationships between the QBA measures and other behavioural and physiological measures which provides interesting new information on these markers of affect. It is notable however that some QBA measures were not represented frequently within the sample and as such were omitted from later analysis which may have affected the results.

Overall, no variables were omitted based on their assessment in relation to the gold standard induction and measurement methods and as such it is unlikely that these have had significant bearing on the results obtained.

Scoring of data

Intra and inter-rater reliability was demonstrably excellent within this study suggesting that the definitions used to score video data were reliably applied by each rater. However, there is the potential that some definitions may not be entirely accurate leading to inadvertent misattribution of scores to specific measures. Of consideration are the measures 'QBA alarmed' and 'ear gently back' which showed surprising relationships within the PCA analysis. 'Ear gently back' was expected to relate to other measures of low arousal, positive affect however its placement within the moderate arousal, negative affect quadrant poses questions as to its indication as an affect marker. It is feasible that within this cohort it was an accurate measure of increasing arousal and negative affect but conversely it is plausible that the distinction between ears gently back and ears flat back was not adequately realised by the video scorers and as such its placement within the core affect model may not be a true reflection of the measure itself. 'QBA alarmed' was identified within high arousal positive affect (quadrant 1) which appears incommensurate with its stated definition as 'worried/tense, apprehensive, jumpy, nervous, watchful, on guard against possible threat or danger' with the description of a 'rigid stance, startled reaction, vigilance, moving ears' (AWIN, 2015). It is plausible that this measure was mistakenly attributed to horses in an alert anticipatory, appetitive state who may be vigilant to the environment. Tension in muscle tone and scanning for stimuli with the ears are responses reported within the equine literature associated with increasing arousal and startle responses (Forkman *et al.*, 2007; Noble *et al.*, 2013); however alertness was not closely related to other negative descriptors such as worried or stressed within a PCA analysis undertaken by (Fleming *et al.*, 2013).

Interpretation of the data

The results reported were interpreted considering their statistical relevance and the evidence within the existing literature. Interpretation of PCA data is somewhat subjective in terms of attributing meaning to the principal components produced however interpretations are based upon the evaluation of variables and their correlation on each relevant component. Relative loadings on components indicates that the variables vary together (Jolliffe and

Cadima, 2016). As discussed in section 6.4.3, the first two principal components are suggested to be indicative of arousal and valence respectively, this conclusion was reached through careful consideration of the strength of the relationship between variables and the evidence within the literature pertaining to the meaning of each variable in relation to affective state. However, it is feasible that the assumptions made may not be valid, either due to the over or under estimation of the significance of individual variables or due to inconsistencies that arise within the literature in relation to the potential meaning of variables. While some variables clustered relatively tightly others were more standalone which may be a reliable effect but could also be anomalous data. The interpretation of the principal components was then taken forward within their use as dependent variables for linear regression models and the assumptions utilised to consider the responses of individual horses and personality types on stimulus and trial phase, therefore any erroneous interpretation could also affect the interpretation of these models. However, due to the relative novelty of this study area the use of exploratory PCA analyses and subsequent linear models is warranted within the context and aims of this experiment (Jolliffe and Cadima, 2016).

An important consideration, which is a limitation within this study, is the circular argument that exists where a set of measures were selected to identify positive affective state produced using the selected induction methods, following which attempts have been made to validate the measures using the responses to those induction methods. Whilst there is no clear solution to this, the statistical methods utilised to interrogate the data attempted to tease apart these phenomena into discrete areas by evaluating the measures across all induction methods, followed by disentangling the differences between induction stimuli whilst considering the effects of both the individual subject differences and personality-based differences.

6.4.6 Suggestions for further study

The results of this exploratory study have provided some insight into the responses of a cohort of equines to four different induction stimuli and have afforded consideration of how behavioural and physiological measures of affective state cluster together to give insight into how simple behavioural measures may be utilised in the future as proxy measures of affective state. As with many exploratory studies, the results have also raised many questions that require further investigation to be able to answer them. Particularly

pertinent questions include how the results may differ in a larger cohort of equines, how equines may respond to different induction stimuli (notably stimuli that can be provided for equines at liberty) and whether the findings in terms of the clustering and interpreted meaning of the measurement variables can be replicated in further studies.

Initially it would be useful to collect further data using the same methodological design to add greater power to the study, analysis of a larger dataset may provide instructive evidence around the strength and stability of the observed measurement clusters and may report more meaningful data for variables that were excluded due to their lack of significance or low numbers such as eye and ear thermography and some QBA and frustration measures. It may also provide further insight into the effect of personality on affective state. Following this, further induction and measurement variables could be added to the study design to collect further data for comparison and to strengthen the evidence base for the purported affective domains. Lastly it would be useful to retest the horses utilised within the current study on at least two further occasions to consider the temporal stability of the results reported within this trial and the longer-term temporal stability of the personality trait assessment.

6.4.7 Conclusion

This study has generated a dataset which provides further advancement of our knowledge around the induction of positive affect and consideration of the meaning and usefulness of measures used to attempt to measure affective state in equines. The exploratory approach has facilitated an open-minded assessment of the data generated and allowed evaluation of novel approaches to the study of affective state in equines. The methodology of the study, while generally robust, has limitations in terms of the power of the study and the potential for recording and scoring errors which could have affected the results obtained. The statistical approach taken has minimised the likelihood of type one errors however this is somewhat counterbalanced by the low sample size. As an exploratory study, with the aim of better understanding how positive affect can be induced and characterised through behavioural and physiological measures, this dataset provides new information to researchers in this field that could be utilised as a basis for further explanatory work. The decision to include assessment of personality within the study design has further enhanced the robustness of the results by attempting to account for a significant factor that is likely to lead to variation in equine responses.

The findings of this work suggest that there is variation within the equine cohort in terms of their physiological and behavioural responses to the induction stimuli and that, to some extent this is influenced by personality, as well as by prior experience and their current environment. The results also suggest variation in the usefulness of the measurement methods for identifying different types of affective state and the dimensional approach taken has allowed for assessment of the broad arousal and valence components of affect without ascribing measures to discrete emotions such as fear and happiness. It is suggested that there is clustered loading of measurement variables across the dimensions of arousal and valence that are relevant across all four induction stimuli, however it is likely that responses to individual stimuli have influenced the clustering of these variables to greater and lesser degrees.

This study is relevant in the current climate where social license to operate has become a significant factor within equine welfare, especially within the competitive sphere. Furtherment of knowledge around ways in which can induce positive affect in equines is likely to enable enhancement of emotional welfare, in the longer term a factor that is likely to also enhance physical welfare through the avoidance of pre-pathological and pathological states emanating from chronic impaired emotional welfare. Furthermore, the ability to provide equine owners, carers and scientists in the field with simple measures to assess positive affective state is likely to facilitate greater consideration of this facet of the five domains of welfare, improving overall quality of life for these equines.

Chapter 7 General discussion

7.1 Summary

Throughout this thesis there has been a continual theme around attempting to consider equine specific methodologies for inducing a measuring positive affective state whilst also facilitating consideration of the individual animal, their personality, and their discrete affective experience. The thesis has taken an emergent approach to the understanding of equine positive affective state and in the development and evaluation of measures of equine personality, and the induction and measurement of affective state. The starting point for this work began with an extensive review of the available literature surrounding equine personality and affective state to develop a series of questions for the Delphi consultation of experts (Chapter 1), following this a three round Delphi (Chapter 3) provided a solid basis for the later experimental chapters (Chapters 4-6) via the examination of expert consensus. The rich qualitative and quantitative data generated from a Delphi approach allows for decision making and action whilst remaining circumspect about the nuances of the debate unfolding.

The input from Delphi experts provided guidance for the development of the Equine Personality Eight-factor Model (EPEM) (Chapter 4), which would allow for consideration of personality within affective state in later chapters. Subsequent chapters reported an extensive pilot study (Chapter 5), used to trial the methodology intended for use within the final experimental trial (Chapter 6). This final study was a focal experimental trial, building upon the work completed in previous chapters, and was undertaken utilising four perceivably positive induction stimuli and eight groups of methods designed to attempt to measure affective state. Within this final trial the assessment of the effect of personality on affective state was interwoven through the methodological design.

The results reported within this thesis have added useful and interesting data to the field of equine personality and affective state. The work has produced what appears to be a reliable personality index for equines, albeit this requires further validation with a greater sample size and over varied temporal scales; and has reported compelling relationships between physiological and behavioural variables that may be indicators of varying affective states in equines.

Discussion within the chapters of this thesis has critically evaluated the methodologies utilised and the findings emanating from the methodological design. As anticipated, this work has generated further questions around how we can measure affective state in equines in a sensitive and specific manner using measures that are repeatable, feasible and temporally stable.

7.2 The Delphi consultation

A Delphi consultation is a structured communication technique used to reach consensus agreement amongst a group of experts. A highly regarded technique, it enables structured forecasting and assumes that group judgements are more valid than individual judgements (Boulkedid *et al.*, 2011; Green *et al.*, 1999; Hasson *et al.*, 2000; Hohmann *et al.*, 2018; Hsu and Sandford, 2007a). Utilised for many years within business forecasting and health and medicine fields, Delphi consultations have recently gained popularity within the animal welfare sphere (DuBois *et al.*, 2018; Pannewitz and Loftus, 2023; Pearson *et al.*, 2021; Rioja-Lang *et al.*, 2020; Williams *et al.*, 2023; Yeoman *et al.*, 2017). This Delphi consultation was utilised to ensure that later experimental design was based on a strong foundation of consensus expert views, rather than through individual selection of measures by the experimenters. The consultation brought a wide and experienced range of equine and animal welfare scientists together in academic debate with a collegiate focus to disseminate and share knowledge to reach a consensus on the fundamental principles of affective state in equines and following this, measures to both induce and measure it. Many aspects quickly achieved consensus however the crucial aspects surrounding inducing and measuring affect drew much discussion and a wider range of views. The thematic qualitative data generated from free-text questions was valuable in facilitating deeper understanding of the ratings provided by the experts and helped to generate the questioning in the iterative Delphi rounds. The Delphi also highlighted the paucity of validated methods to induce and measure positive affective state in equines and identified the often-specialised fields that the specialists worked within; for example, clinical behavioural specialists often lacked depth of knowledge around many of the physiological methods available for use, whereas veterinary scientists often reported a lesser understanding of some of the behavioural nuances. This facet highlights the importance of multidisciplinary research to develop and validate methods surrounding

affective state that are reliable, valid, and repeatable due to the breadth and depth of knowledge of the team developing them.

7.3 Equine personality

From the preliminary planning stage of this work onwards, personality was identified as a factor that would be fundamental in our understanding of individual affective state (Ashton, 2022; Boissy and Erhard, 2014; Bolig *et al.*, 1992; Carter *et al.*, 2013; Duberstein and Gilkeson, 2010; Eysenck and Eysenck, 2013; Fortin *et al.*, 2018; Gosling, 2008). The predilections of the individual have been demonstrated to markedly affect their emotional responses to stimuli and the surgency of affective state experienced (Duberstein and Gilkeson, 2010; Fortin *et al.*, 2018; Hausberger *et al.*, 2004; Hausberger *et al.*, 2011; Ijichi *et al.*, 2014; Jolivald, 2021; Jolivald *et al.*, 2022; Lansade *et al.*, 2017; Lansade *et al.*, 2016; Sauer *et al.*, 2019; Suwała *et al.*, 2016). Within this, consideration of prior life experiences and current environment is important in terms of their influences on both short- and long-term affective state, albeit these aspects were only superficially examined within this work. This thesis has reported the iterative design of a new equine personality questionnaire, the Equine Eight-Factor Personality Model (EPEM), utilising the Delphi expert's consensus and real-world testing of variables. Careful consideration of personality traits and groupings over successive rounds allowed production of what appears to be a reliable and practically useful equine personality model that requires further validation in both its full and abridged forms. It is recognised that since the start of this thesis other researchers have worked on equine personality questionnaires with some useful and valid assessments produced (Jolivald, 2021; Jolivald *et al.*, 2022; Suwała *et al.*, 2016). Similarly to our EPEM, these researchers have set their work within the long-recognised and well validated Five Factor Model with equines demonstrated to express traits representative of all the five factors, a facet which supports the similarities in personality traits across mammalian species and possibly beyond (Jolivald *et al.*, 2022).

The EPEM is presented as a seemingly reliable model to assess personality traits in equines and to consider the grouping of traits across eight overarching groups with good construct validity, good levels of inter-rater reliability and excellent intra-rater reliability. The model encompasses 36 personality traits which were divided through statistical testing and expert *post hoc* evaluation, across the eight overarching groups. The groupings provide coverage of gross personality traits which can be clearly mapped on to the Big Five factor model. Further

testing of the EPEM over time and across a larger cohort would further reinforce its reliability and validity; additionally, it would be useful to test the association between scores using both the EPEM and Big Five model to further consider the attribution of groupings within the EPEM to the factors in the Big Five. To increase usability of the EPEM an abridged scale was designed and tested within a subset of the study population, initial findings suggest excellent agreement between scores in the full and abridged versions of the EPEM, however this facet requires wider validation.

7.4 Inducing and measuring positive affective state in equines

The preparatory work within the Delphi consultation and EPEM development culminated in the final experimental trial to test the expert derived methods in a real-life scenario. A detailed pilot study was undertaken to allow for testing of the planned methodological design within different environments to consider how closely trials could be replicated in different locations and to test specific measures prior to commencement of the full study. The pilot study provided useful data on some of the measures included to strengthen the case for their inclusion or exclusion in the later study, it also allowed for evaluation and refinement of the final study design. As with most experimental trials limitations to data collection were evident within this work. An important consideration is the aspect of affect induction whilst being gently restrained through the employment of a headcollar and lead rope. The design of the study (including the measures employed) required the equines to be visible throughout the trials and therefore they could not be free moving at liberty. This may have affected the results obtained if some horses were frustrated or anxious during this restriction however all equines were subjected to the same conditions meaning that individual variation in responses to stimuli could still be identified. However, this methodology, utilising light restraint, has been used previously in similar studies where clear video data of the subjects is required throughout the test period (Stratton, 2022; Lansade *et al.*, 2018). The sample size for the final trial was smaller than would be ideal and negatively impacts on the power of the study and the reliability of the results obtained; however, the equines utilised within the study were representative of the types of equines found across the equine population in terms of age, breed, sex and usage which strengthened the study. Further data collection utilising the same methodology would help to strengthen the data and provide evaluation of any changes in effect within a larger sample size. For some measures data were missing and/or omitted (ear

and eye thermography data, small sections of HRV data, some QBA and frustration variables) due to the technology employed and low response rates, and these missing data may have impacted the results obtained. Individual responses to induction stimuli varied widely as was anticipated and suggests that whilst every subject appeared to find at least one of the induction stimuli appetitive, further preference testing of each individual over a period of time is warranted within the methodological design of further experimental trials. Additionally greater preference testing could produce individual ranking data across each stimulus in terms of individual preference which could then be included as a random effect, or independent variable depending on how the data were to be considered, in further analyses.

Induction stimuli response appeared to be mediated to some degree by personality with horses high in 'unfriendliness' being more aroused during wither scratching, unresponsive horses engaging differently with training, nervous horses being more aroused to the food and scratch stimuli, bold horses responding differently to the training and companion stimuli. Further consideration of personality and induction methods could be investigated through a larger empirical trial. Responses to the test environment may have varied within individuals across test sessions likely related to time of day (am or pm), weather conditions and prior experiences over the preceding 12 hours. Some horses may have been averse to the cortisol collection method which may have impacted on the cortisol results obtained but also on other physiological and behavioural variables as well as overall mood. These facets were not empirically scored and therefore were not considered within the statistical modelling. Further examination of these facets in future trials seems warranted to account for the already considered effect of environment and experience on affective state however objective methods to measure these variables would require further consideration, potentially with further input from the Delphi collective.

The results suggests that some of the measures investigated may be valid potential methods for affective state measurement. Some methods appeared to be indicative of negative affective state whereas others seem to suggest positive affect. Interestingly some measures appear to be more ambiguous, potentially indicative of neutral affect however none appeared to truly sit neutrally across valence and arousal (Table 7.1). Stratton (2022) also suggested the potential for some markers to be indicative of a more neutral affective state, a proposition requiring further investigation especially due to its consideration by other authors across

species, including equines (Gasper and Danube, 2016; Gasper *et al.*, 2021; Gasper and Hackenbracht, 2015; Gasper *et al.*, 2019; Hintze and Schanz, 2021; Killgore and Yurgelun-Todd, 2007; Samson *et al.*, 2016; Scopa *et al.*, 2020).

Table 7.1 Proposed measures that may be indicative of positive, neutral and negative valence in high, neutral or low arousal states.

	Positive affect	Neutral affect	Negative affect
High arousal	LF/HF ratio Cortisol LF power	Eye bias left Respiratory rate Mean heart rate QBA alarmed	Increased locomotion Ears gently back QBA uneasy Acute conflict Ears fully back Eye sclera Eye bias right Eye bias both
Neutral arousal	QBA at ease		Ears asymmetrical
Low arousal	HF power RR mean Eyes half QBA relaxed QBA friendly Ears side Ears forward		pNN50 RMSSD SDRR

There are many measures which seem to give clearer indication of arousal levels, particularly physiological variables but give less information regarding valence, these include many of the heart rate variability measures, respiratory rate and potentially eye bias (lateralisation). Further testing of these measures across different arousal levels, and potentially across reinforcing and punishing events may help to disentangle the arousal component further.

Investigation of the correlation between behavioural and physiological measures has demonstrated some interesting possible behavioural proxies for the more invasive or specialised physiological measurements, a finding which could prove useful to lay users of affect assessment in their equines with further validation. For example, within this study measures suggested to be indicative of quadrant two states (low arousal, positive valence) included some QBA measures (relaxed, friendly and at ease) as well as eye half open and ears forward which appear correlated with HRV measures RR mean and HF power. This may indicate that the behavioural measures could, with further validation, be used as proxies for the HRV variables. However physiological measures can be ambiguous in relation to affective state especially in their meaning in relation to valence; even well validated methods such as heart rate variability analysis and cortisol sampling are complex in their interpretation and are affected by multifaceted dimensions that require careful extrapolation (Billman, 2013; Mormède *et al.*, 2011; van Ast *et al.*, 2013). The complex cross-emotion, multi-system responses to stimuli indicates that many measures may be sensitive to both positive and negative affect and as such may not be specific to either, this is particularly evident in the measures that appear to be indicative of arousal such as mean heart rate, respiratory rate and eye bias. Interestingly the behavioural measures (such as QBA measures, eye and ear position and frustration measures) appear to be more informative in terms of valence of affect, with these loading more extremely on the principal components. This could suggest that within affective state, arousal may be most accurately measured using physiological measures and valence most accurately defined using behavioural measures. Further investigation of this hypothesis is required to enable clarification.

A facet for further exploration is the consideration of quadrant two states (calm and content) which can be characterised by low arousal in positive valence. This state is considered optimal from a welfare perspective, for safety and for positive horse-human relationships (Hall *et al.*, 2018; Lesimple, 2020) and therefore the ability to quantify and measure the presence of this state may be practically beneficial to many engaging with equines in a variety of capacities (Mendl *et al.*, 2010; Watson and McDonnell, 2018). Several proxy behavioural measures (eye half, ear side, ear forward, QBA relaxed, friendly and at ease) were identified as potential markers of quadrant two state however very little physiological data paired with the behavioural data (aside from HRV variables) therefore further research to establish

relationships between physiological measures and behavioural measures for the quadrant two state would be useful.

Within quadrants one and three there appears to be a close interplay between arousal and valence with many of the measures clustered in proximity across the valence and arousal dimensions. Kuppens et al. (2017) considered how arousal and valence are related to each other with four main theories presented (Figure 7.1). This is an area that should receive greater attention to facilitate a better understanding of affective state from a theoretical (consideration of the structure of affective state), empirical (how affect is related to other psychological phenomena) and psychometric (can valence and arousal be measured or manipulated independently) perspectives.

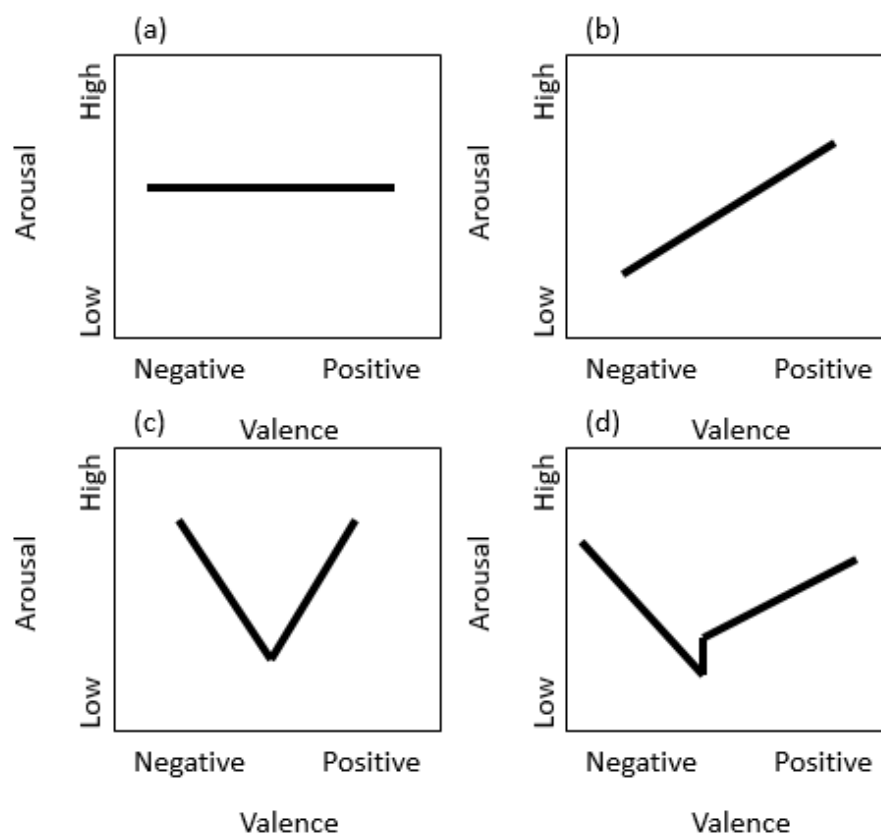


Figure 7.1 Overview of potential relationships between valence and arousal: (a) independence, (b) positive linear relation, (c) symmetric V-shaped relation, (d) asymmetric V-shaped relation including positivity offset and negativity bias. Figure adapted from Kuppens *et al.* (2017).

Relationship (a) suggests that pleasant or unpleasant valence conveys no information regarding arousal levels, relationship (b) suggests a strictly linear relationship between valence and arousal (pleasant valence increases or decreases with arousal levels), relationship (c) suggests increased states of positive or negative valence are associated with increased arousal and (d) suggests a positivity offset where feelings of arousal are higher at low levels of positive valence compared with negative valence and a negativity bias where negative feelings increase more strongly than positive feelings do with increasing arousal (Kuppens *et al.*, 2017). Within their PhD thesis investigating positive affect in horses Stratton (2022) suggested that relationship (c) may be applicable with the trough at the bottom of the V indicating neutral states. Within the current study there is some evidence for theories (c) and (d). Evidence from the clustering of variables suggests that increases in arousal are associated with increases in negative or positive affect as appears evident within quadrants one and four which are purported to be high arousal positive and high arousal negative states. However, what is also suggested within the strength of the relationships is that there may be a positivity offset where arousal may be higher (as depicted by higher loadings on PC2) at lower levels of positive affect (depicted by lower loadings on PC1) than at lower levels of negative affect (where variables indicating negative affect have lower arousal (PC1) loadings at lower negative affect levels (PC2)). Also worth considering is the greater number of measurement variables which appear to be available to measure negative affect. This could be due to the variables selected for PCA analysis within this study, or to the design of the study overall. Conversely, it is plausible that the greater availability of measures for negative affect may indicate a negativity bias where negative feelings increase more strongly than positive feelings during increased arousal, making these negative states easier to measure using behavioural and physiological measures. The possibility that positive affect is more subtle in its behavioural and physiological effects has been suggested by others (Sortheix and Weber, 2023; Kuppens *et al.*, 2017) and is a factor that is worth further consideration in the quest to define markers of its presence in animals, including equines. There is potential that previously validated metrics such as the Positive and Negative Affect Schedule (PANAS) developed by Watson *et al.* (1988), which has previously being further developed from the original human model for use in dogs (Savalli *et al.*, 2019), could be further validated for use in horses in order to aid further investigation of the differentiation between markers of positive and negative affective state in equines.

7.5 Limitations

Within each experimental chapter of this thesis several limitations to the methodological design or interpretation of the results have been evaluated and discussed. As with many studies of this nature, sample size is relatively low which has a significant impact upon the validity and generalisability of the results. In studies such as these there is clear support for study replication to gather a larger dataset to enable re-analysis of the data within a larger cohort which should enable evaluation as to whether the conclusions drawn within this study remain reliable when the study is replicated.

The basis of the experimental design of Chapters 4 and 6 relied heavily on expert Delphi consensus and whilst the sample size was high within the Delphi consultation it is plausible that the expert consensus does not reflect reality. However, the efforts undertaken to reach a high sample size across specialisms within the Delphi helps to offset some of these concerns. It may be useful to conduct further consultation with the general equestrian population to gain an understanding of their views on the facets discussed by the Delphi experts so that the results can be considered within the design of further studies. One facet for consideration is that where Delphi participants did not feel educated enough to answer questions on specific topics, they reported via the free-text comment boxes that they often opted to not give a response. This could have increased reliability of the results because only those with the higher levels of expertise in those areas answered, conversely, it could have negatively impacted the reliability in that only a subset of the cohort was responsible for the decisions made following analysis of their responses.

As discussed in Chapter 6, the interpretation of the PCA results and linear models, whilst based on objective statistical analysis, involved some subjective assessment of the meaning of the results reported. There are several ways in which the data could have been interpreted which may affect the conclusions reached within this thesis. However, the dimensional approach taken aimed to reduce the potential for incorrect attribution by avoiding applying the results to discrete emotions.

7.6 Recommendations

The questions raised by the conclusions drawn within this thesis suggest several future research pathways to further investigate the phenomena reported within this thesis. Further extrapolation of the relationship between valence and arousal appears needed within this

area of study. A greater understanding of the relationship, linear or otherwise, between these two aspects of affective state is fundamental to enabling the teasing apart of the meanings of variables that are used to attempt to measure positive affective states. Within this further consideration of personality, through further validation of the EPEM alongside the standardised assessment of the environment and prior experiences may help to provide contextual information regarding the perception of affect inducing stimuli by individual equines and could be factored into further analyses to enhance our understanding. It appears that it would be beneficial for researchers in this area to collaborate more widely on a large-scale project to collect a substantial dataset of wide-ranging measurement and induction methods across a much larger cohort of equines to allow for the generation of reliable evidence in this field.

7.7 Concluding remarks

At its conception this thesis was undertaken with the overarching aim of maximising positive welfare states for the equines that so many people spend their working or personal lives with. This was motivated by the realisation that many of our equines do not have their emotional welfare considered to the same degree as their physical welfare is monitored. Many equines live in environments that are totally at odds with their natural ethology and the effect of this on their mental wellbeing is often underappreciated. The objectives within this thesis of further understanding equine personality, investigating ways in which we may be able to make equines happy and measure this happiness; and to consider how personality may affect their experience and expression of happiness were set with the aim of providing information that is practically relevant to everyday equine carers. It is hoped that the findings within this thesis contribute in a small way to fulfilling this aim. It has been quoted by experienced horse people that “a good rider can hear their horse speak to them; a great rider can hear their horse whisper”, perhaps in the quest to understand and measure positive affect in equines we must pay attention to the whispers.

8. References

- Adolphs, R., Mlodinow, L. & Barrett, L. F. 2019. What is an emotion? *Current biology*, 29(20), pp R1060-R1064.
- Akiskal, H. S. 2013. An integrative perspective on recurrent mood disorders: The mediating role of personality. *Psychosocial aspects of depression*, pp 215-236.
- Alexander, S. L. & Irvine, C. H. G. 1998. The effect of social stress on adrenal axis activity in horses: the importance of monitoring corticosteroid-binding globulin capacity. *Journal of endocrinology*, 157(3), pp 425-432.
- Alexandrov, Y. I. & Sams, M. E. 2005. Emotion and consciousness: Ends of a continuum. *Cognitive Brain Research*, 25, pp 387-405.
- Alhaji, M., Zubair, M. & Farhana, A. 2023. Enzyme linked immunosorbent assay. *StatPearls*.
- Alharahsheh, H. H. & Pius, A. 2020. A review of key paradigms: Positivism VS interpretivism. *Global Academic Journal of Humanities and Social Sciences*, 2(3), pp 39-43.
- Amsel, A. 1992. *Frustration theory: An analysis of dispositional learning and memory*: Cambridge University Press.
- APA. 2024. Cultural bias. *APA Dictionary of Psychology*. Washington: APA.
- App, B., McIntosh, D. & Reed, C. A social-functional approach to emotion communication: How depends on why. Poster presented at the annual meeting of the Society for Personality and Social Psychology, 2007.
- Ashton, M. C. 2022. *Individual differences and personality*: Academic Press.
- Aureli, F., Preston, S. D. & de Waal, F. 1999. Heart rate responses to social interactions in free-moving rhesus macaques (*Macaca mulatta*): a pilot study. *Journal of comparative psychology*, 113(1), pp 59.
- Aurich, J., Wulf, M., Ille, N., Erber, R., Von Lewinski, M., Palme, R. & Aurich, C. 2015. Effects of season, age, sex, and housing on salivary cortisol concentrations in horses. *Domestic animal endocrinology*, 52(11-16).

- Austin, N. P. & Rogers, L. J. 2007. Asymmetry of flight and escape turning responses in horses. *Laterality*, 12(5), pp 464-474.
- Austin, N. P. & Rogers, L. J. 2012. Limb preferences and lateralization of aggression, reactivity and vigilance in feral horses, *Equus caballus*. *Animal Behaviour*, 83(1), pp 239-247.
- Austin, N. P. & Rogers, L. J. 2014. Lateralization of agonistic and vigilance responses in Przewalski horses (*Equus przewalskii*). *Applied Animal Behaviour Science*, 151(43-50).
- AWIN. 2015. AWIN Welfare assessment protocol for horses, http://dx.doi.org/10.13130/AWIN_HORSES_2015.
- Baldwin, J. D. & Baldwin, J. I. 2001. *Behaviour Principles in Everyday Life*, 4, Pennsylvania: Prentice Hall.
- Bär, K.-J., de la Cruz, F., Schumann, A., Koehler, S., Sauer, H., Critchley, H. & Wagner, G. 2016. Functional connectivity and network analysis of midbrain and brainstem nuclei. *Neuroimage*, 134, pp 53-63.
- Baragli, P., Gazzano, A., Martelli, F. & Sighieri, C. 2009. How do horses appraise humans' actions? A brief note over a practical way to assess stimulus perception. *Journal of Equine Veterinary Science*, 29(10), pp 739-742.
- Baragli, P., Padalino, B. & Telatin, A. 2015. The role of associative and non-associative learning in the training of horses and implications for the welfare (a review). *Annali dell'Istituto superiore di sanita*, 51, pp 40-51.
- Baragli, P., Scopa, C., Felici, M. & Reddon, A. R. 2021. Horses show individual level lateralisation when inspecting an unfamiliar and unexpected stimulus. *Plos one*, 16(8), pp e0255688. <https://doi.org/10.1371/journal.pone.0255688>.
- Bär, K.-J., de la Cruz, F., Schumann, A., Koehler, S., Sauer, H., Critchley, H. & Wagner, G. 2016. Functional connectivity and network analysis of midbrain and brainstem nuclei. *Neuroimage*, 134, pp 53-63.
- Barbas, H. 2015. General cortical and special prefrontal connections: principles from structure to function. *Annual review of neuroscience*, 38, pp 269-289.
- Bard, P. 1934. *Emotion: I. The Neuro-humoral Basis of Emotional Reactions*, Worcester, MA, US: Clark University Press.

- Barnes, S. A., Young, J. W. & Neill, J. C. 2012. Rats tested after a washout period from sub-chronic PCP administration exhibited impaired performance in the 5-Choice Continuous Performance Test (5C-CPT) when the attentional load was increased. *Neuropharmacology*, 62(3), pp 1432-1441.
- Barrett, L. 2018. *How Emotions Are Made.* , London: Pan Macmillan.
- Barrett, L. F. 2017a. *How emotions are made: The secret life of the brain.* London: Pan Macmillan.
- Barrett, L. F. 2017b. The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, 12(1), pp 1-23.
- Barrett, L. F., Gross, J., Christensen, T. C. & Benvenuto, M. 2001. Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. . *Cognition & Emotion*, 15(6), pp 713-724.
- Bartolomé, E. & Cockram, M. S. 2016. Potential effects of stress on the performance of sport horses. *Journal of Equine Veterinary Science*, 40, pp 84-93.
- Baruch, Y. & Holtom, B. C. 2008. Survey response rate levels and trends in organizational research. . *Human Relations*, 61(8), pp 1139-1160.
- Baumgartner, H. & Weijters, B. 2015. Response biases in cross-cultural measurement. *Handbook of culture and consumer behavior.* Oxford University Press New York.
- Beerda, B., Schilder, M. B., van Hooff, J. A., de Vries, H. W. & Mol, J. A. 1999. Chronic stress in dogs subjected to social and spatial restriction. I. Behavioral responses. *Physiol Behav*, 66(2), pp 233-42.
- Bell, A. M. 2007. Future directions in behavioural syndromes research. *Proc Biol Sci*, 274(1611), pp 755-61.
- Bell, A. M., Hankison, S. J. & Laskowski, K. L. 2009. The repeatability of behaviour: a meta-analysis. *Animal behaviour*, 77(4), pp 771-783.
- Bell, C. 1847. *The Anatomy and Philosophy of Expression as Connected with the Fine Arts*, 3, London: John Murray.

- Bell, C., Rogers, S., Taylor, J. & Busby, D. 2019. Improving the Recognition of Equine Affective States. *Animals*, 9(12), pp 1124.
- Bell, E. T. 1965. *Men of Mathematics*, New York: Simon and Schuster.
- Benedetti, F., Radaelli, D., Poletti, S., Locatelli, C., Dallspezia, S., Lorenzi, C., Pirovano, A., Colombo, C. & Smeraldi, E. 2011. Association of the C(-1019)G 5-HT1A promoter polymorphism with exposure to stressors preceding hospitalization for bipolar depression. *Journal of affective disorders*, 132(1-2), pp 297-300.
- Ben-Hur, A. & Guyon, I. 2003. Detecting stable clusters using principal component analysis. *Functional Genomics: Methods and Protocols*, pp 159-182.
- Bennett-Wimbush, K., Amstutz, M. & Duthie, M. 2020. Effects of overcheck use on stress parameters and welfare implications in driving horses. *Journal of applied animal welfare science*, 23(1), pp 83-94.
- Berchtold, A. 2016. Test–retest: agreement or reliability? *Methodological Innovations*, 9, <https://doi.org/10.1177/2059799116672875>.
- Bergner, R. M. 2020. What is personality? Two myths and a definition. *New Ideas in Psychology*, 57, pp e100759.
- Berridge, K. C. & O'Doherty, J. P. 2014. From experienced utility to decision utility. *Neuroeconomics*. Elsevier.
- BETA. 2019. *National Equestrian Survey 2019*. [Online]. <https://www.beta-uk.org/pages/industry-information/market-information.php>. [Accessed 02.10.22 2022].
- Billman, G. E. 2013. The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Front Physiol*, 4: 26, <https://doi.org/10.3389%2Ffphys.2013.00026>.
- Birke, L., Hockenhull, J., Creighton, E., Pinno, L., Mee, J. & Mills, D. 2011. Horses' responses to variation in human approach. *Applied Animal Behaviour Science*, 134(1-2), pp 56-63.
- Blasius, J. 2012. Comparing Ranking Techniques in Web Surveys. *Field Methods*, 24(4), pp 382-398.

- Bleidorn, W., Schwaba, T., Zheng, A., Hopwood, C. J., Sosa, S. S., Roberts, B. W. & Briley, D. A. 2022. Personality stability and change: A meta-analysis of longitudinal studies. *Psychological bulletin*, 148(7-8), pp 588.
- Blumenthal, A. L. 2002. *A reappraisal of Wilhelm Wundt*, Washington, DC, US: American Psychological Association.
- Boccia, M. L., Reite, M. & Laudenslager, M. 1989. On the physiology of grooming in a pigtail macaque. *Physiology & Behavior*, 45(3), pp 667-670.
- Bogen, J. E. 1985. *The callosal syndromes*, New York: Oxford University Press.
- Boissy, A., Arnould, C., Chaillou, E., Désiré, L., Duvaux-Ponter, C., Greiveldinger, L., Leterrier, C., Richard, S., Roussel, S. & Saint-Dizier, H. 2007a. Emotions and cognition: a new approach to animal welfare. *Animal Welfare*, 16(S1), pp 37-43.
- Boissy, A. & Erhard, H. W. 2014. How Studying Interactions Between Animal Emotions, Cognition, and Personality Can Contribute to Improve Farm Animal Welfare. In: Grandin, T. & Deesing, M. J. (eds.) *Genetics and the Behavior of Domestic Animals*. 2 ed. London: Academic Press.
- Boissy, A., Manteuffel, G., Jensen, M. B., Moe, R. O., Spruijt, B. M., Keeling, L. J., Winckler, C., Forkman, B., Dimitrov, I. & Langbein, J. 2007b. Assessment of positive emotions in animals to improve their welfare. *Physiology & behavior*, 92(3), pp 375-397.
- Boiten, F. A. 1998. The effects of emotional behaviour on components of the respiratory cycle. *Biological psychology*, 49(1-2), pp 29-51.
- Boiten, F. A., Frijda, N. H. & Wientjes, C. J. E. 1994. Emotions and respiratory patterns: review and critical analysis. *International journal of psychophysiology*, 17(2), pp 103-128.
- Bolig, R., Price, C. S., O'Neill, P. L. & Suomi, S. J. 1992. Subjective assessment of reactivity level and personality traits of rhesus monkeys. *International Journal of Primatology*, 13(3), pp 287-306.
- Borman, W. C. 1977. Consistency of rating accuracy and rating errors in the judgment of human performance. *Organizational Behavior and Human Performance*, 20(2), pp 238-252.

- Bornmann, T., Randle, H. & Williams, J. 2021. Investigating equestrians' perceptions of horse happiness: An exploratory study. *Journal of Equine Veterinary Science*, 104, 103697, <https://doi.org/10.1016/j.jevs.2021.103697>.
- Borstel, U. K. V., Visser, E. K. & Hall, C. 2017. Indicators of stress in equitation. . *Applied Animal Behaviour Science*, 190, pp 43-56, <https://doi.org/10.1016/j.applanim.2017.02.018>.
- Boukdedid, R., Abdoul, H., Loustau, M., Sibony, O. & Alberti, C. 2011. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PloS one*, 6(6), pp p.e20476.
- Bourne, V. J. 2010. How are emotions lateralised in the brain? Contrasting existing hypotheses using the Chimeric Faces Test. *Cognition and Emotion*, 24(5), pp 903-911.
- Bouwknicht, J. 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), pp 41-59.
- Bowen, I. M. 2010. Ambulatory electrocardiography and heart rate variability. *Cardiology of the Horse*, 2, pp 127-137.
- Bradley, A., Mennie, N., Bibby, P. A. & Cassaday, H. J. 2020. Some animals are more equal than others: Validation of a new scale to measure how attitudes to animals depend on species and human purpose of use. *PloS one*, 15(1), pp e0227948.
- Brambell, F. W. R. 1965. Report of the technical committee to enquire into the welfare of animals kept under intensive livestock husbandry systems, H.M.S.O (London).
- Breland, K. & Breland, M. 1961. The misbehavior of organisms. *American psychologist*, 16(11), pp 681.
- Briefer, E. F., Mandel, R., Maigrot, A.-L., Briefer Freymond, S., Bachmann, I. & Hillmann, E. 2017a. Perception of emotional valence in horse whinnies. *Frontiers in Zoology*, 14(1), <https://doi.org/10.1186/s12983-017-0193-1>.
- Briefer, E. F., Mandel, R., Maigrot, A., Freymond, S. B., Bachmann, I. & Hillmann, E. 2017b. Perception of emotional valence in horse whinnies. *Frontiers in Zoology*, 14(8), <https://doi.org/10.1186/s12983-017-0193-1> .

- Briefer, E. F., Tettamanti, F. & McElligott, A. G. 2015. Emotions in goats: mapping physiological, behavioural and vocal profiles. *Animal Behaviour*, 99, pp 131-143.
- Brinks, V., van der Mark, M. H., de Kloet, E. R. & Oitzl, M. S. 2007. Differential MR/GR activation in mice results in emotional states beneficial or impairing for cognition. *Neural plasticity*, 2007, pp e90163, <https://doi.org/10.1155/2007/90163>.
- British Equestrian. 2022. *Equality and Diversity* [Online]. Available: <https://www.britishequestrian.org.uk/getInvolved/equality-diversity> [Accessed October 2022].
- Broom, D. M. & Kirkden, R. D. 2004. Welfare, stress, behaviour and pathophysiology. *Veterinary pathophysiology*, pp 337-369.
- Broom, D. M. & Kennedy, M. J. 1993. Stereotypies in horses: their relevance to welfare and causation. *Equine Veterinary Education*, 5, pp 151-151.
- Broom, D. 1987. Applications of Neurobiological Studies to Farm Animal Welfare. In: Wiepkema, P.R., Van Adrichem, P.W.M. (eds) *Biology of Stress in Farm Animals: An Integrative Approach*. Current Topics in Veterinary Medicine and Animal Science, vol 42. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-3339-2_8.
- Browne, K. 2005. Snowball sampling: using social networks to research non-heterosexual women. *International journal of social research methodology*, 8(1), pp 47-60.
- Browning, J. R., Whiteman, A. C., Leung, L. Y., Lu, X. M. & Shear, D. A. 2017. Air-puff induced vocalizations: A novel approach to detecting negative affective state following concussion in rats. *J Neurosci Methods*, 275, pp 45-49.
- Brucks, D., Härterich, A. & König von Borstel, U. 2022. Horses wait for more and better rewards in a delay of gratification paradigm. *Frontiers in psychology*, 13, pp e954472.
- Buchanan, T. W., al'Absi, M. & Lovallo, W. R. 1999. Cortisol fluctuates with increases and decreases in negative affect. *Psychoneuroendocrinology*, 24(2), pp 227-241.
- Bucik, V., Brenk, K. & Vodopivec, B. 1995. Emotions profile index: stability and dimensionality of the structure. *Physics*. <https://api.semanticscholar.org/CorpusID:193621044>.

- Budzyńska, M. 2014. Stress reactivity and coping in horse adaptation to environment. *Journal of Equine Veterinary Science*, 34(8), pp 935-941.
- Bulens, A., Sterken, H., Van Beirendonck, S., Van Thielen, J. & Driessen, B. 2015. The use of different objects during a novel object test in stabled horses. *Journal of Veterinary Behavior*, 10(1), pp 54-58.
- Bullmore, E. & Sporns, O. 2012. The economy of brain network organization. *Nature Reviews Neuroscience*, 13(5), pp 336-349.
- Burgdorf, J., Knutson, B. & Panksepp, J. 2000. Anticipation of rewarding electrical brain stimulation evokes ultrasonic vocalization in rats. *Behavioral neuroscience*, 114(2), pp 320.
- Burgdorf, J. & Panksepp, J. 2006. The neurobiology of positive emotions. *Neurosci Biobehav Rev*, 30(2), pp 173-87.
- Burghardt, G. 2005. *The Genesis of Animal Play: Testing the Limits*, Cambridge, MA: MIT Press, 10.7551/mitpress/3229.001.0001.
- Burman, O. H. P., McGowan, R., Mendl, M., Norling, Y., Paul, E. S., Rehn, T. & Keeling, L. 2011. Using judgement bias to measure positive affective state in dogs. *Applied Animal Behaviour Science*, 132(3-4), pp 160-168.
- Burn, C., Dennison, T. & Whay, H. R. 2010. Relationships between behaviour and health in working horses, donkeys, and mules in developing countries. *Applied Animal Behaviour Science*, 126, pp 109-118.
- Buss, D. M. 1997. Evolutionary foundations of personality. *Handbook of personality psychology*. London: Elsevier.
- Buwalda, B., Stubbendorff, C., Zickert, N. & Koolhaas, J. 2013. Adolescent social stress does not necessarily lead to a compromised adaptive capacity during adulthood: a study on the consequences of social stress in rats. *Neuroscience*, 249, pp 258-270.
- Cabanac, M. 2002. What is emotion? *Behavioural Processes*, 60, pp 60-84.
- Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., Ito, T. A., Lewis, M. & Haviland-Jones, J. M. 2000. Handbook of emotions. *New York: Guilford*, pp 173-91.

- Cacioppo, J. T. & Tassinary, L. G. 1990. Inferring psychological significance from physiological signals. *American psychologist*, 45(1), pp 16.
- Cajori, F. 1926. Origins of Fourth Dimension Concepts. *The American Mathematical Monthly* 33(8), pp 397-406.
- Cameron, E. Z., Linklater, W. L., Stafford, K. J. & Minot, E. O. 2003. Social grouping and maternal behaviour in feral horses (*Equus caballus*): the influence of males on maternal protectiveness. *Behavioral Ecology and Sociobiology*, 53(2), pp 92-101.
- Canessa, C., Vavvos, A., Triliva, S., Kafkalas, I., Vracholi, M. & Sauer, J. 2022. Implementing a combined Delphi and Focus Group qualitative methodology in Nexus research designs—The case of the WEFE Nexus in Apokoronas, Crete. *Plos one*, 17(7), pp e0271443.
- Cannon, W. B. 1927. The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *The American Journal of Psychology*, 39(1/4), pp 106-124.
- Cardinal, R. N., Parkinson, J. A., Hall, J. & Everitt, B. J. 2002. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, 26(3), pp 321-352.
- Careau, V., Thomas, D., Humphries, M. & Réale, D. 2008. Energy metabolism and animal personality. *Oikos*, 117(5), pp 641-653.
- Carenzi, C. & Verga, M. 2009. Animal welfare: review of the scientific concept and definition. *Italian Journal of Animal Science*, 8(sup1), pp 21-30.
- Carney, R. M., Freedland, K. E. & Veith, R. C. 2005. Depression, the autonomic nervous system, and coronary heart disease. *Psychosomatic medicine*, 67, pp S29-S33.
- Carter, A. J., Feeney, W. E., Marshall, H. H., Cowlshaw, G. & Heinsohn, R. 2013. Animal personality: what are behavioural ecologists measuring? *Biological reviews*, 88(2), pp 465-475.
- Cary, M. A., Plamondon, K., Banner-Lukaris, D., Oelke, N., Sibley, K. M., Baxter, K., Vis-Dunbar, M., Hoens, A. M., Wick, U. & Bigsby, S. 2023. Building consensus in research partnerships: a scoping review of consensus methods. *Evidence & Policy*, 19(3), pp 485-511.

- Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A. & Poulton, R. 2003. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*, 301(5631), pp 386-9.
- Cavallone, E., Di Giancamillo, M., Secchiero, B., Belloli, A., Pravettoni, D. & Rimoldi, E. M. 2002. Variations of serum cortisol in Argentine horses subjected to ship transport and adaptation stress. *Journal of equine veterinary science*, 22(12), pp 541-545.
- Cayado, P., Muñoz-Escassi, B., Dominguez, C., Manley, W., Olabarri, B., De La Muela, M. S., Castejon, F., Maranon, G. & Vara, E. 2006. Hormone response to training and competition in athletic horses. *Equine Veterinary Journal*, 38(S36), pp 274-278.
- Cerasoli, F., Podaliri Vulpiani, M., Saluti, G., Conte, A., Ricci, M., Savini, G. & D'Alterio, N. 2022. Assessment of welfare in groups of horses with different management, environments and activities by measuring cortisol in horsehair, using liquid chromatography coupled to hybrid orbitrap high-resolution mass spectrometry. *Animals*, 12(14), pp 1739.
- Chamove, A. S., Crawley-Hartrick, O. J. E. & Stafford, K. J. 2002. Horse reactions to human attitudes and behavior. *Anthrozoös*, 15(4), pp 323-331.
- Chamove, A. S., Eysenck, H. J. & Harlow, H. F. 1972. Personality in monkeys: Factor analyses of rhesus social behaviour. *The Quarterly Journal of Experimental Psychology*, 24(4), pp 496-504.
- Cherry, R. L., Adair, H. S., Chen, T., Hendrix, D. V. H. & Ward, D. A. 2020. Effect of attentional focus levels on spontaneous eyeblink rate in horses. *Veterinary ophthalmology*, 23(4), pp 690-695.
- Chiavacci, D. How to analyse data: An introduction to methods of data analysis in qualitative Social Science research. Studying Japan, 2020. Nomos Verlagsgesellschaft mbH & Co. KG, 300-309.
- 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), pp 44-52.
- Christensen, J. W., Malmkvist, J., Nielsen, B. L. & Keeling, L. J. 2008a. Effects of a calm companion on fear reactions in naive test horses. *Equine Veterinary Journal*, 40(1), pp 46-50.

- Christensen, J. W., Ahrendt, L. P., Malmkvist, J. & Nicol, C. J. 2021a. Exploratory behaviour towards novel objects is associated with enhanced learning in young horses. *Scientific reports*, 11(1), pp 1428.
- Christensen, J. W., Munk, R., Hawson, L., Palme, R., Larsen, T., Egenvall, A., König von Borstel, U. U. & Rørvang, M. V. 2021b. Rider effects on horses' conflict behaviour, rein tension, physiological measures and rideability scores. *Applied Animal Behaviour Science*, 234, pp e105184.
- Christensen, J. W., Rundgren, M. & Olsson, K. 2006. Training methods for horses: habituation to a frightening stimulus. *Equine Veterinary Journal*, 38(5), pp 439-443.
- Christensen, J. W., Zharkikh, T. & Ladewig, J. 2008b. Do horses generalise between objects during habituation? *Applied Animal Behaviour Science*, 114(3-4), pp 509-520.
- Chun Tie, Y., Birks, M. & Francis, K. 2019. Grounded theory research: A design framework for novice researchers. *SAGE Open Med*, 7, pp e2050312118822927.
- Ciccone, A. B., Siedlik, J. A., Wecht, J. M., Deckert, J. A., Nguyen, N. D. & Weir, J. P. 2017. Reminder: RMSSD and SD1 are identical heart rate variability metrics. *Muscle & nerve*, 56(4), pp 674-678.
- Clark, D. K., Friend, T. H. & Dellmeier, G. 1993. The effect of orientation during trailer transport on heart rate, cortisol and balance in horses. *Applied Animal Behaviour Science*, 38(3-4), pp 179-189.
- Clayton, N. S., Bussey, T. J. & Dickinson, A. 2003. Can animals recall the past and plan for the future? *Nature Reviews Neuroscience*, 4(8), pp 685-691.
- Clayton, N. S. & Dickinson, A. 1998. Episodic-like memory during cache recovery by scrub jays. *Nature*, 395(6699), pp 272-274.
- Clayton, N. S. & Dickinson, A. 1999. Scrub jays (*Aphelocoma coerulescens*) remember the relative time of caching as well as the location and content of their caches. *Journal of Comparative Psychology*, 113(4), pp 403.
- Cleophas, T. J. 1993. Carryover bias in clinical investigations. *J Clin Pharmacol*, 33(9), pp 799-804.
- Cloninger, C. R. 1986. A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatric developments*, 3(2), pp 167-226.

- Cockburn, A., Smith, M., Rusbridge, C., Fowler, C., Paul, E. S., Murrell, J. C., Blackwell, E. J., Casey, R. A., Whay, H. R. & Mendl, M. 2018. Evidence of negative affective state in Cavalier King Charles Spaniels with syringomyelia. *Applied Animal Behaviour Science*, 201, pp 77-84.
- Collins, J., Hanlon, A., More, S. J., Wall, P. G. & Duggan, V. 2009. Policy Delphi with vignette methodology as a tool to evaluate the perception of equine welfare. *The Veterinary Journal*, 181(1), pp 63-69.
- Contalbrigo, L., Borgi, M., De Santis, M., Collacchi, B., Tuozi, A., Toson, M., Redaelli, V., Odore, R., Vercelli, C. & Stefani, A. 2021. Equine-assisted interventions (EAls) for children with autism spectrum disorders (ASD): behavioural and physiological indices of stress in domestic horses (*Equus caballus*) during riding sessions. *Animals*, 11(6), pp 1562.
- Contreras-Aguilar, M. D., Lamy, E., Escribano, D., Cerón, J. J., Tecles, F., Quiles, A. J. & Hevia, M. L. 2020. Changes in salivary analytes of horses due to circadian rhythm and season: A pilot study. *Animals*, 10(9), pp 1486.
- Cook, N., Schaefer, A. L. C., Warren, L. K., Burwash, L., Anderson, M. & Baron, V. S. 2001. Adrenocortical and metabolic responses to ACTH injection in horses An assessment by salivary cortisol and infrared thermography of the eye. *Canadian Journal of Animal Science*, 81, pp 621.
- Cooper, J. & McGreevy, P. 2007. Stereotypic behaviour in the stabled horse: causes, effects and prevention without compromising horse welfare. *The welfare of horses*. New York: Springer.
- Cooper, J. O., Heron, T. E. & Heward, W. L. 2020. *Applied behavior analysis*, London: Pearson UK.
- Cooper, J. R. 1998. Comparative learning theory and its application in the training of horses. *Equine veterinary journal. Supplement*, 30, pp 39-43.
- Corr, P. J. 2008. *Reinforcement Sensitivity Theory (RST): Introduction*, New York, NY, US: Cambridge University Press.
- Costa, P. T. & McCrae, R. R. 1989. The NEO-PI/NEO-FFI Manual Supplement., Odessa (Florida).

- Costa, P. T. & McCrae, R. R. 1995. Domains and Facets: Hierarchical Personality Assessment Using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, 64(1), pp 21-50.
- Cousillas, H., Oger, M., Rochais, C., Pettoello, C., Ménoret, M., Henry, S. & Hausberger, M. 2017. An ambulatory electroencephalography system for freely moving horses: an innovating approach. *Frontiers in veterinary science*, 4, pp 57.
- Covalesky, M. E., Russoniello, C. R. & Malinowski, K. 1992. Effects of show-jumping performance stress on plasma cortisol and lactate concentrations and heart rate and behavior in horses. *Journal of equine veterinary science*, 12(4), pp 244-251.
- Cozzi, A., Sighieri, C., Gazzano, A., Nicol, C. J. & Baragli, P. 2010. Post-conflict friendly reunion in a permanent group of horses (*Equus caballus*). *Behavioural processes*, 85(2), pp 185-190.
- Csoltova, E. & Mehinagic, E. 2020. Where do we stand in the domestic dog (*Canis familiaris*) positive-emotion assessment: a state-of-the-art review and future directions. *Frontiers in psychology*, 11, pp e525721.
- d'Ingeo, S., Quaranta, A., Siniscalchi, M., Stomp, M., Coste, C., Bagnard, C., Hausberger, M. & Cousillas, H. 2019. Horses associate individual human voices with the valence of past interactions: a behavioural and electrophysiological study. *Scientific reports*, 9(1), pp e11568.
- da Luz, M. P. F., Narciso, M. H. P. M., Filho, J. N. P. P., Arruda, L. A. d. S., Delagrancia, M. F., Maia, C. M. & Baragli, P. 2022. Motivation and frustration of horses and mules: behavioral and physiological differences. CABI, pp 317-333.
- Dalla Costa, E., Bracci, D., Dai, F., Lebelt, D. & Minero, M. 2017. Do different emotional states affect the horse grimace scale score? A pilot study. *Journal of Equine Veterinary Science*, 54(114-117).
- Dalla Costa, E., Dai, F., Lebelt, D., Scholz, P., Barbieri, S., Canali, E., Zanella, A. J. & Minero, M. 2016a. Welfare assessment of horses: the AWIN approach. *Animal Welfare*, 25(4), pp 481-488.
- 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), pp e92281.

- Dalla Costa, E., Minero, M., Lebelt, D., Stucke, D., Canali, E. & Leach, M. C. 2014b. Development of the Horse Grimace Scale (HGS) as a Pain Assessment Tool in Horses Undergoing Routine Castration. *PLoS ONE*, 9(3), pp e92281.
- Dalla Costa, E., Murray, L., Dai, F., Canali, E. & Minero, M. 2014c. Equine on-farm welfare assessment: A review of animal-based indicators. *Animal Welfare*, 23 (3), pp 323-341, doi:10.7120/09627286.23.3.323.
- Dalla Costa, E., Stucke, D., Dai, F., Minero, M., Leach, M. C. & Lebelt, D. 2016b. Using the horse grimace scale (HGS) to assess pain associated with acute laminitis in horses (*Equus caballus*). *Animals*, 6(8), pp 47.
- Dallmann, R., Steinlechner, S., Von Hörsten, S. & Karl, T. 2006. Stress-induced hyperthermia in the rat: comparison of classical and novel recording methods. *Laboratory animals*, 40(2), pp 186-193.
- Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L. B., Parvizi, J. & Hichwa, R. D. 2000. Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature neuroscience*, 3(10), pp 1049-1056.
- Daniel, B. K. & Harland, T. 2017. *Higher education research methodology: A step-by-step guide to the research process*, London: Routledge.
- Darwin, C. 1872. The expression of emotions in man and animals. Cambridge: Cambridge University Press, <https://doi.org/10.1017/CBO9781139833813>.
- Davidson, R. J. 1992. Emotion and affective style: Hemispheric substrates. *Psychological Science*, 3(1), pp 39-43.
- Davis, M. 1989. Neural systems involved in fear-potentiated startle. *Annals of the New York Academy of Sciences*, 563, pp 165-183.
- Davis, A. K., Maney, D. L. & Maerz, J. C. 2008. The use of leukocyte profiles to measure stress in vertebrates: a review for ecologists. *Functional ecology*, 22(5), pp 760-772.
- Davis, K., Panksepp, J. & Normansell, L. 2014. The Affective Neuroscience Personality Scales: Normative Data and Implications. *Neuropsychanalysis: An Interdisciplinary Journal for Psychoanalysis and the Neurosciences*, 5, pp 57-69.

- De Mira, M. C., Lamy, E., Santos, R., Williams, J., Pinto, M. V., Martins, P. S., Rodrigues, P. & Marlin, D. 2021. Salivary cortisol and eye temperature changes during endurance competitions. *BMC Veterinary Research*, 17(1), pp 1-12.
- De Santis, M, Contalbrigo, L., Borgi, M., Cirulli, F., Luzi, F., Redaelli, V., Stefani, A., Toson, M., Odore, R., Vercelli, C., Valle, E., Farina, L. 2017. Equine Assisted Interventions (EAI): Methodological Considerations for Stress Assessment in Horses. *Veterinary Sciences*, 4 (3), pp 44, <https://doi.org/10.3390/vetsci4030044>.
- DEFRA. 2004. *A report on research on the horse industry in Great Britain*. [Online]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/69259/pb9255-bhic-report-040318.pdf. [Accessed 02.10.2022 2022].
- DeGrazia, D. & Beauchamp, T. L. 2019. Beyond the 3 Rs to a more comprehensive framework of principles for animal research ethics. *ILAR journal*, 60(3), pp 308-317.
- Des Roches, A. D. B., Richard-Yris, M. A., Henry, S., Ezzaouia, M. & Hausberger, M. 2008. Laterality and emotions: Visual laterality in the domestic horse (*Equus caballus*) differs with objects' emotional value. *Physiology & Behavior*, 94(3), pp 487-490.
- Désiré, L., Boissy, A. & Veissier, I. 2002. Emotions in farm animals:: a new approach to animal welfare in applied ethology. *Behavioural Processes*, 60(2), pp 165-180.
- Destrez, A., Deiss, V., Leterrier, C., Calandreau, L. & Boissy, A. 2014. Repeated exposure to positive events induces optimistic-like judgment and enhances fearfulness in chronically stressed sheep. *Applied Animal Behaviour Science*, 154, pp 30-38.
- Devaney, L. & Henchion, M. 2018. Who is a Delphi 'expert'? Reflections on a bioeconomy expert selection procedure from Ireland. *Futures*, 99, pp 45-55.
- Dhabhar, F. S. 2009. Enhancing versus suppressive effects of stress on immune function: implications for immunoprotection and immunopathology. *Neuroimmunomodulation*, 16(5), pp 300-317.
- Diaz-Uriarte, R. 2001. The analysis of cross-over trials in animal behavior experiments: review and guide to the statistical literature. *Ramon Diaz-Uriarte, University of Wisconsin, Madison, WI*.
- Dittrich, H. 1898. *Handbuch der Anatomie der Tiere für Künstler*. Leipzig: Dieterichsche Verlagsbuchhandlung.

- Dixon, J. 1970. The horse: a dumb animal?... neigh. *Thoroughbred Rec*, 192(19), pp 1654-1657.
- Dockès, A. C. & Kling-Eveillard, F. 2006. Farmers' and advisers' representations of animals and animal welfare. *Livestock science*, 103(3), pp 243-249.
- Dolcos, F., Iordan, A. D. & Dolcos, S. 2011. Neural correlates of emotion–cognition interactions: A review of evidence from brain imaging investigations. *Journal of Cognitive Psychology*, 23(6), pp 669-694.
- Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H., Sears, R. R., Ford, C. S., Hovland, C. I. & Sollenberger, R. T. 2013. *Frustration and aggression*: Abingdon: Routledge.
- Dolnicar, S. & Leisch, F. 2012. One legacy of Mazanec: binary questions are a simple, stable and valid measure of evaluative beliefs. *International Journal of Culture, Tourism and Hospitality Research*, 6(4), pp 316-325.
- Dorey, N. R. & Cox, D. J. 2018. Function matters: a review of terminological differences in applied and basic clicker training research. *PeerJ*, 6, pp e5621.
- Douglas, C., Bateson, M., Walsh, C., Bédoué, A. & Edwards, S. A. 2012. Environmental enrichment induces optimistic cognitive biases in pigs. *Applied Animal Behaviour Science*, 139(1-2), pp 65-73.
- Douglas, J., Owers, R. & Campbell, M. H. 2022. Social licence to operate: what can equestrian sports learn from other industries? *Animals*, 12(15), pp 1987.
- 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), pp 349-352.
- Draaisma, R. 2017. *Language signs and calming signals of horses: recognition and application*, New York: CRC Press.
- Drevets, W. C., Gautier, C., Price, J. C., Kupfer, D. J., Kinahan, P. E., Grace, A. A., Price, J. L. & Mathis, C. A. 2001. Amphetamine-induced dopamine release in human ventral striatum correlates with euphoria. *Biological psychiatry*, 49(2), pp 81-96.
- Dror, O. E. 2014. The Cannon–Bard thalamic theory of emotions: A brief genealogy and reappraisal. *Emotion Review*, 6(1), pp 13-20.

- Duberstein, K. J. & Gilkeson, J. A. 2010. Determination of sex differences in personality and trainability of yearling horses utilizing a handler questionnaire. *Applied animal behaviour science*, 128(1-4), pp 57-63.
- DuBois, C., Hambly Odame, H., Haley, D. & Merckies, K. 2018. An exploration of industry expert perception of Canadian equine welfare using a modified Delphi technique. *PLoS One*, 13(7), pp e0201363.
- Duman, R. S. & Monteggia, L. 2006. A Neurotrophic Model for Stress-Related Mood Disorders. *Biological psychiatry*, 59(11)16-27.
- Dunbar, R. I. 2010. The social role of touch in humans and primates: behavioural function and neurobiological mechanisms. . *Neuroscience & Biobehavioral Reviews*, 34(2), pp 260-268.
- Duncan, E., O'Cathain, A., Rousseau, N., Croot, L., Sworn, K., Turner, K. M., Yardley, L. & Hoddinott, P. 2020. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. *BMJ open*, 10(4), pp e033516.
- Duran, M. C., Janz, D. M., Waldner, C. L., Campbell, J. R. & Marques, F. J. 2017. Hair cortisol concentration as a stress biomarker in horses: Associations with body location and surgical castration. *Journal of equine veterinary science*, 55, pp 27-33.
- Durston, T. 2022. Emotions, stress, and stressor stacking. *Emotional wellbeing for animal welfare professionals*. Walingford: CABI.
- Dutton, D. M. 2008. Subjective assessment of chimpanzee (*Pan troglodytes*) personality: reliability and stability of trait ratings. *Primates*, 49(4), pp 253-259.
- Dyson, S. 2022. The ridden horse pain ethogram. *Equine Veterinary Education*, 34(7), pp 372-380.
- Dyson, S., Thomson, K., Quiney, L., Bondi, A. & Ellis, A. D. 2020. Can veterinarians reliably apply a whole horse ridden ethogram to differentiate nonlame and lame horses based on live horse assessment of behaviour? *Equine Veterinary Education*, 32, pp 112-120.
- Dyson, S. & Van Dijk, J. 2020. Application of a ridden horse ethogram to video recordings of 21 horses before and after diagnostic analgesia: Reduction in behaviour scores. *Equine Veterinary Education*, 32, pp 104-111.

- Dyson, S., Berger, J. M., Ellis, A. D. & Mullard, J. 2017. Can the presence of musculoskeletal pain be determined from the facial expressions of ridden horses (FEReq)? *Journal of veterinary behavior*, 19, pp 78-89.
- Edelman, G. M. 1989. *The remembered present: a biological theory of consciousness*: Basic Books.
- Edgar, J. L., Nicol, C. J., Clark, C. A. A. & Paul, E. S. 2012. Measuring empathic responses in animals. *Applied Animal Behaviour Science*, 138(3-4), pp 182-193.
- Edmond, S. N., Turk, D. C., Williams, D. A. & Kerns, R. D. 2019. Considerations of trial design and conduct in behavioral interventions for the management of chronic pain in adults. *Pain Rep*, 4(3), pp e655.
- Eldridge, S. M., Lancaster, G. A., Campbell, M. J., Thabane, L., Hopewell, S., Coleman, C. L. & Bond, C. M. 2016. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. *PloS one*, 11(3), pp e0150205.
- Elia, J. B., Erb, H. N. & Houpt, K. A. 2010. Motivation for hay: effects of a pelleted diet on behavior and physiology of horses. *Physiol Behav*, 101(5), pp 623-7.
- Ellenberg, U., Mattern, T. & Seddon, P. J. 2013. Heart rate responses provide an objective evaluation of human disturbance stimuli in breeding birds. *Conservation Physiology*, 1(1), pp cot013.
- Ellis, S. & Greening, L. 2016. Positively reinforcing an operant task using tactile stimulation and food—A comparison in horses using clicker training. *J. Vet. Behav. Clin. Appl. Res*, 15(S 78), pp S78.
- Enrico, P., Bouma, M., De Vries, J. B. & Westerink, B. H. C. 1998. The role of afferents to the ventral tegmental area in the handling stress-induced increase in the release of dopamine in the medial prefrontal cortex: a dual-probe microdialysis study in the rat brain. *Brain research*, 779(1-2), pp 205-213.
- Equine Business Association. 2022. The equine industry: A global perspective. Available from: <https://equinebusinessassociation.com/equine-industry-statistics/> [Accessed 02.10.2022].
- Erber, R., Wulf, M., Aurich, J., Rose-Meierhöfer, S., Hoffmann, G., von Lewinski, M., Möstl, E. & Aurich, C. 2013. Stress response of three-year-old horse mares to changes in

- husbandry system during initial equestrian training. *Journal of equine veterinary science*, 33(12), pp 1088-1094.
- Estevez, I., Battini, M., Canali, E., Ruiz, R., Stilwell, G., Ferrante, V., Minero, M., Marchewka, J., Barbieri, S., Mattiello, S., Beltrán de Heredia, I., Dwyer, C. M. & Zanella, A. J. 2017. AWIN mobile apps; animal welfare assessment at your fingertips. *Journal of Animal Science*, 95, pp 6a.
- Evans, D. L. & Rose, R. J. 1986. Method of investigation of the accuracy of four digitally-displaying heart rate meters suitable for use in the exercising horse. *Equine veterinary journal*, 18(2), pp 129-132.
- Evans, L., Cameron-Whytock, H. & Ijichi, C. 2024. Eye understand: Physiological measures as novel predictors of adaptive learning in horses. *Applied Animal Behaviour Science*, 271, pp e106152.
- Eysenck, H. J. 1963. Biological basis of personality. *Nature*, 199, pp 1031-4.
- Eysenck, H. J. & Eysenck, S. B. G. 1984. Eysenck personality questionnaire-revised. Available at <https://scales.arabpsychology.com/s/eysenck-personality-questionnaire-revised-epq-r/>.
- Eysenck, H. J. & Eysenck, S. B. G. 2013. The biological basis of personality. *Personality Structure and Measurement (Psychology Revivals)*. New York: Routledge.
- Farcomeni, A. & Greco, L. 2016. *Robust methods for data reduction*: Florida: CRC press.
- Farmer, K., Krueger, K. & Byrne, R. W. 2010. Visual laterality in the domestic horse (*Equus caballus*) interacting with humans. *Animal Cognition*, 13, pp 229-238.
- Fazio, E., Medica, P., Aronica, V., Grasso, L. & Ferlazzo, A. 2008. Circulating β -endorphin, adrenocorticotrophic hormone and cortisol levels of stallions before and after short road transport: stress effect of different distances. *Acta Veterinaria Scandinavica*, 50, pp 1-7.
- Fazio, E., Medica, P., Cravana, C. & Ferlazzo, A. 2013. Cortisol response to road transport stress in calm and nervous stallions. *Journal of Veterinary Behavior*, 8(4), pp 231-237.
- Feh, C. & de Mazières, J. 1993. Grooming at a preferred site reduces heart rate in horses. *Animal Behaviour*, 46(6), pp 1191-1194.

- Fellous, J.-M. & Suri, R. E. 2002. *A Roles of Dopamine*: Eidgenössische Technische Hochschule Zürich.
- Fenner, K., Yoon, S., White, P., Starling, M. & McGreevy, P. 2016. The effect of noseband tightening on horses' behavior, eye temperature, and cardiac responses. *PLoS One*, 11(5), pp 1-20.
- Figueredo, A. J., Cox, R. L. & Rhine, R. J. 1995. A Generalizability Analysis of Subjective Personality Assessments in the Stumptail Macaque and the Zebra Finch. *Multivariate Behavioral Research*, 30(2), pp 167-197.
- Finlayson, K., Lampe, J. F., Hintze, S., Würbel, H. & Melotti, L. 2016. Facial indicators of positive emotions in rats. *PloS one*, 11(11), pp e0166446.
- Fiske, J. C. & Potter, G. D. 1979. Discrimination Reversal Learning in Yearling Horses. *Journal of Animal Science*, 49(2), pp 583-588.
- Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E. M. & Akil, H. 2011. A selective role for dopamine in stimulus–reward learning. *Nature*, 469(7328), pp 53-57.
- Flannery, B. 1997. Relational discrimination learning in horses. *Applied Animal Behaviour Science*, 54(4), pp 267-280.
- Fleming, P. A., Clarke, T., Wickham, S. L., Stockman, C. A., Barnes, A. L., Collins, T. & Miller, D. W. 2016. The contribution of qualitative behavioural assessment to appraisal of livestock welfare. *Animal Production Science*, 56(10), pp 1569-1578.
- Fleming, P. A., Paisley, C. L., Barnes, A. L. & Wemelsfelder, F. 2013. Application of Qualitative Behavioural Assessment to horses during an endurance ride. *Applied Animal Behaviour Science*, 144(1-2), pp 80-88.
- Fordham, D. P., Al-Gahtani, S., Durotoye, L. A. & Rodway, R. G. 1991. Changes in plasma cortisol and β -endorphin concentrations and behaviour in sheep subjected to a change of environment. *Animal Science*, 52(2), pp 287-296.
- 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. *Physiol Behav*, 92(3), pp 340-74.

- Fortin, M., Valençon, 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, pp 130-140.
- Fowler, V. L., Kennedy, M. & Marlin, D. J. 2012. A comparison of the Monty Roberts technique with a conventional UK technique for initial training of riding horses. *Anthrozoös*, 25(3), pp 301-321.
- Fraser, D. & Duncan, I. J. H. 1998. 'Pleasures', 'pains' and animal welfare: toward a natural history of affect. *Animal welfare*, 7(4), pp 383-396.
- Fraser, A. 2010. *The behaviour and welfare of the horse*, Oxford: CABI.
- Fraser, D. 2009. Animal behaviour, animal welfare and the scientific study of affect. *Applied Animal Behaviour Science*, 118(3-4), pp 108-117.
- Freymond, S. B., 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, pp 34-45.
- Frippiat, T., van Beckhoven, C., Moyse, E. & Art, T. 2021. Accuracy of a heart rate monitor for calculating heart rate variability parameters in exercising horses. *Journal of equine veterinary science*, 104, pp e103716.
- Funder, D. C. 1995. On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102(4), pp 652-670.
- Funder, D. C. 1999. *Personality judgment: A realistic approach to person perception*: London: Elsevier.
- Funder, D. C. & Dobroth, K. M. 1987. Differences between traits: properties associated with interjudge agreement. *Journal of personality and social psychology*, 52(2), pp 409.
- Fureix, C. & Meagher, R. K. 2015. What can inactivity (in its various forms) reveal about affective states in non-human animals? A review. *Applied Animal Behaviour Science*, 171(8-24).
- Fureix, C., Jegou, P., Henry, S., Lansade, L. & Hausberger, M. 2012. Towards an ethological animal model of depression? A study on horses. *PLoS One*, 7(6), pp e39280.

- Fureix, C., Pagès, M., Bon, R., Lassalle, J.-M., Kuntz, P. & Gonzalez, G. 2009. A preliminary study of the effects of handling type on horses' emotional reactivity and the human–horse relationship. *Behavioural processes*, 82(2), pp 202-210.
- Gabor, V. & Gerken, M. 2018. Study into long-term memory of a complex learning task in Shetland ponies (*Equus caballus*). *Applied animal behaviour science*, 198, pp 60-66.
- Gallegos, J. M. & Gasper, K. 2018. Differential effects of rejection and acceptance on feeling shocked, numb, and neutral. . *Emotion*, 18, pp 536-550.
- Garamszegi, L. Z., Markó, G. & Herczeg, G. 2012. A meta-analysis of correlated behaviours with implications for behavioural syndromes: mean effect size, publication bias, phylogenetic effects and the role of mediator variables. *Evolutionary Ecology*, 26, pp 1213-1235.
- Gartner, M. C., Powell, D. M. & Weiss, A. 2016. Comparison of Subjective Well-Being and Personality Assessments in the Clouded Leopard (*Neofelis nebulosa*), Snow Leopard (*Panthera uncia*), and African Lion (*Panthera leo*). *Journal of Applied Animal Welfare Science*, 19(3), pp 294-302.
- Gasper, K. & Danube, C. L. 2016. The scope of our affective influences: when and how naturally occurring positive, negative, and neutral affects alter judgment. *Personality and Social Psychology Bulletin* 42, pp 385-399.
- Gasper, K., Danube, C. L. & Hu, D. 2021. Making room for neutral affect: Evidence indicating that neutral affect is independent of and co-occurs with eight affective states. *Motivation and Emotion*, 45, pp 103-121.
- Gasper, K. & Hackenbracht, J. 2015. Too busy to feel neutral: reducing cognitive resources attenuates neutral affective states. *Motivation and Emotion*, 39, pp 458-466.
- Gasper, K., Spencer, L. A. & Hu, D. 2019. Does Neutral Affect Exist? How Challenging Three Beliefs About Neutral Affect Can Advance Affective Research. *Front Psychol*, 10, pp 2476.
- Genon, S., Reid, A., Langner, R., Amunts, K. & Eickhoff, S. B. 2018. How to Characterize the Function of a Brain Region. *Trends in Cognitive Sciences*, 22(4), pp 350-364.
- Giebel, H.-D. 1958. Visuelles Lernvermögen bei Einhufern. *Zool. Jb., Physiol.* 67, pp 487–520.

- Gleerup, K. B., Forkman, B., Lindegaard, C. & Andersen, P. H. 2015. An equine pain face. *Vet Anaesth Analg*, 42(1), pp 103-14.
- Gleerup, K. B. & Lindegaard, C. 2016. Recognition and quantification of pain in horses: A tutorial review. *Equine Veterinary Education*, 28(1), pp 47-57.
- Gmel, A. I., Gmel, G., von Niederhäusern, R., Weishaupt, M. A. & Neuditschko, M. 2020. Should We Agree to Disagree? An Evaluation of the Inter-Rater Reliability of Gait Quality Traits in Franches-Montagnes Stallions. *Journal of Equine Veterinary Science*, 88, pp e102932.
- Goodwin, D. 2007. Horse behaviour: evolution, domestication and feralisation. *The welfare of horses*. Netherlands: Springer.
- Goodwin, D. 1999. The importance of ethology in understanding the behaviour of the horse. *Equine Veterinary Journal*, 31(S28), pp 15-19.
- Górecka, A., Bakuniak, M., Chruszczewski, M. H. & Jezierski, T. A. 2007. A note on the habituation to novelty in horses: handler effect. *Animal Science Papers and Reports*, 25(3), pp 143-152.
- Goschke, T. & Bolte, A. 2014. Emotional modulation of control dilemmas: The role of positive affect, reward, and dopamine in cognitive stability and flexibility. *Neuropsychologia*, 62, pp 403-423.
- Gosling, S. D. 2008. Personality in non-human animals. *Social and Personality Psychology Compass*, 2(2), pp 985-1001.
- Gosling, S. D. & Bonnenburg, A. V. 1998. An integrative approach to personality research in anthrozoology: Ratings of six species of pets and their owners. *Anthrozoös*, 11(3), pp 148-156.
- Gosling, S. D. & John, O. P. 1999. Personality dimensions in nonhuman animals: A cross-species review. *Current directions in psychological science*, 8(3), pp 69-75.
- Gosling, S. D., Kwan, V. S. Y. & John, O. P. 2003a. A dog's got personality: a cross-species comparative approach to personality judgments in dogs and humans. *Journal of personality and social psychology*, 85(6), pp 1161.
- Gosling, S. D., Rentfrow, P. J. & Swann Jr, W. B. 2003b. A very brief measure of the Big-Five personality domains. *Journal of Research in personality*, 37(6), pp 504-528.

- Gray, J. A. 1981. A critique of Eysenck's theory of personality. *A model for personality*. New York: Springer.
- Green, B., Jones, M., Hughes, D. & Williams, A. 1999. Applying the Delphi technique in a study of GP's information requirements. *Health Society Care Community*, 7, pp 198-205.
- Greenacre, M., Groenen, P. J. F., Hastie, T., d'Enza, A. I., Markos, A. & Tuzhilina, E. 2022. Principal component analysis. *Nature Reviews Methods Primers*, 2(1), pp 100.
- Greening, L., Downing, J., Amiouny, D., Lekang, L. & McBride, S. D. 2021. The effect of altering routine husbandry factors on sleep duration and memory consolidation in the horse. *Applied Animal Behaviour Science*, 236, pp e105229.
- Grint, N. J., Johnson, C. B., Lorena, S. D., Luna, S., Hussni, C. A., Whay, H. R. & Murrell, J. C. 2014. Electroencephalographic responses to a noxious surgical stimulus in mules, horses, and ponies. *Journal of Equine Veterinary Science*, 34(8), pp 955-962.
- Gross, J. J. 2015. Emotion regulation: Current status and future prospects. *Psychological Inquiry*, 26(1), pp 1-26.
- Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q. K., Yu, A. B., Kahn, A. E., Medaglia, J. D., Vettel, J. M., Miller, M. B. & Grafton, S. T. 2015. Controllability of structural brain networks. *Nature communications*, 6(1), pp 8414.
- Guis, E. 2023. CATMA. Zenodo. Available <https://catma.de/>.
- Gunnar Götestam, K., Svebak, S. & Naper Jensen, E. 2008. The role of personality, mood, subjective health, and stress in depressive symptoms among high school students. *The European journal of psychiatry*, 22(3), pp 121-129.
- Guthrie Yarwood, M. 2022. Classical Theories of Emotion: Cannon-Bard Theory. *Psychology of human emotion: An open access textbook*. Pressbooks: Penn State University.
- Hajjar, S. 2018. Statistical analysis: Internal-consistency reliability and construct validity. *International Journal of Quantitative and Qualitative Research Methods*, 6(1), pp 27-38.
- Hall, C. & Heleski, C. 2017. The role of the ethogram in equitation science. *Applied Animal Behaviour Science*, 190, pp 102-110.

- Hall, C., Kay, R. & Yarnell, K. 2014. Assessing ridden horse behavior: Professional judgment and physiological measures. *Journal of Veterinary Behavior*, 9(1), pp 22-29.
- Hall, C., Randle, H., Pearson, G., Preshaw, L. & Waran, N. 2018. Assessing equine emotional state. *Applied Animal Behaviour Science*, 205, pp 183-193.
- Hall, C., Goodwin, D., Heleski, C., Randle, H. & Waran, N. 2008. Is there evidence of learned helplessness in horses? *J Appl Anim Welf Sci*, 11(3), pp 249-66.
- Hanggi, E. B. 1999. Interocular transfer of learning in horses (*Equus caballus*). *Journal of Equine Veterinary Science*, 19(8), pp 518-524.
- Hanggi, E. B. 2010. Short-term Memory Testing in Domestic Horses: Experimental Design Plays a Role. *Journal of Equine Veterinary Science*, 30(11), pp 617-623.
- Hanggi, E. B. & Ingersoll, J. F. 2009. Long-term memory for categories and concepts in horses (*Equus caballus*). *Animal Cognition*, 12(3), pp 451-462.
- Harding, E. J., Paul, E. S. & Mendl, M. 2004. Cognitive bias and affective state. *Nature*, 427(6972), pp 312-312.
- Hare, B. 2007. From Nonhuman to Human Mind:What Changed and Why? *Current Directions in Psychological Science*, 16(2), pp 60-64.
- Hartmann, E., Christensen, J. W. & Keeling, L. J. 2011. Training young horses to social separation: Effect of a companion horse on training efficiency. *Equine veterinary journal*, 43(5), pp 580-584.
- Haselton, M. G., Nettle, D. & Andrews, P. W. 2015. The evolution of cognitive bias. *The handbook of evolutionary psychology*, pp 724-746.
- Hasson, F., Keeney, S. & McKenna, H. 2000. Research guidelines for the Delphi survey technique. *Journal of advanced nursing*, 32(4), pp 1008-1015.
- Hausberger, M., Bruderer, C., Le Scolan, N. & Pierre, J.-S. 2004. Interplay between environmental and genetic factors in temperament/personality traits in horses (*Equus caballus*). *Journal of Comparative Psychology*, 118(4), pp 434.
- Hausberger, M., Fureix, C. & Lesimple, C. 2016. Detecting horses' sickness: In search of visible signs. *Applied Animal Behaviour Science*, 175, pp 41-49.

- Hausberger, M., Gautier, E., Biquand, V., Lunel, C. & Jégo, P. 2009. Could work be a source of behavioural disorders? A study in horses. *PloS one*, 4(10), pp e7625.
- Hausberger, M., Muller, C. & Lunel, C. 2011. Does work affect personality? A study in horses. *PloS one*, 6(2), pp e14659.
- Hausberger, M., Roche, H., Henry, S. & Visser, E. K. 2008. A review of the human–horse relationship. *Applied Animal Behaviour Science*, 109(1), pp 1-24.
- 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, pp 540-559.
- Heath, R. G. 1996. *Exploring the Mind-Brain Relationship*. Michigan: Moran Printing.
- Hebb, D. O. & Williams, K. 1946. A method of rating animal intelligence. *J Gen Psychol*, 34, pp 59-65.
- Heitor, F., do Mar Oom, M. & Vicente, L. 2006. Social relationships in a herd of Sorraia horses Part II. Factors affecting affiliative relationships and sexual behaviours. *Behav Processes*, 73(3), pp 231-9.
- Heleski, C. R. 2023. Social license to operate—why public perception matters for horse sport—some personal reflections. *Journal of Equine Veterinary Science*, 124, pp e104266.
- Heleski, C. R. & Bello, N. M. 2010. Evaluating memory of a learning theory experiment one year later in horses. *Journal of Veterinary Behavior: Clinical Applications and Research*, 4(5), pp 213.
- Heleski, C., Bauson, L. & Bello, N. 2008. Evaluating the Addition of Positive Reinforcement for Learning a Frightening Task: A Pilot Study With Horses. *Journal of Applied Animal Welfare Science*, 11(3), pp 213-222.
- Hemsworth, P. H., Mellor, D. J., Cronin, G. M. & Tilbrook, A. J. 2015. Scientific assessment of animal welfare. *New Zealand veterinary journal*, 63(1), pp 24-30.
- Hendrickson, A. R., Massey, P. D. & Cronan, T. P. 1993. On the test-retest reliability of perceived usefulness and perceived ease of use scales. *MIS quarterly*, pp 227-230.

- Henley, E. 2019. Stress triggers and how to deal with them. *The Veterinary Nurse*, 10(2), pp 120-120.
- 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, pp 1-15.
- Hintze, S., Murphy, E., Bachmann, I., Wemelsfelder, F. & Würbel, H. 2017. Qualitative Behaviour Assessment of horses exposed to short-term emotional treatments. *Applied Animal Behaviour Science*, 196, pp 44-51.
- Hintze, S. & Schanz, L. 2021. Using the Judgment Bias Task to Identify Behavioral Indicators of Affective State: Do Eye Wrinkles in Horses Reflect Mood? *Front Vet Sci*, 8, pp e676888.
- 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), pp e0164017.
- Hockenhull, J. & Creighton, E. 2014. Management practices associated with owner-reported stable-related and handling behaviour problems in UK leisure horses. *Applied Animal Behaviour Science*, 155, pp 49-55.
- Hockenhull, J. & Creighton, E. 2013. Training horses: Positive reinforcement, positive punishment, and ridden behavior problems. *Journal of Veterinary Behavior*, 8(4), pp 245-252.
- Hofstee, W. K., de Raad, B. & Goldberg, L. R. 1992. Integration of the big five and circumplex approaches to trait structure. *J Pers Soc Psychol*, 63(1), pp 146-63.
- Hogan, R. R. 2002. The Hogan personality inventory. *Big five assessment*, pp 329-351.
- Hohmann, E., Cote, M. P. & Brand, J. C. 2018. Research pearls: expert consensus based evidence using the Delphi method. *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, 34(12), pp 3278-3282.
- Holzer, S., Herholz, C., Tanadini, L. G., Ineichen, S. & Julliand, S. 2022. Hay preferences in horses versus selection by their owners. *Livestock Science*, 258, pp e104896.
- Hotchkiss, J., W., , Reid, S. W. & Christley, R. 2006. Construction and validation of a risk-screening questionnaire for the investigation of recurrent airway obstruction in

epidemiological studies of horse populations in Great Britain. *Preventive Veterinary Medicine*, 75(1-2), pp 8-21.

Howard, D. L., Lancaster, B. & de Grauw, J. 2024. Development and Preliminary Validation of an Equine Brief Pain Inventory for Owner Assessment of Chronic Pain Due to Osteoarthritis in Horses. *Animals*, 14(2), pp 181.

Hsu, C.-C. & Sandford, B. A. 2007a. The Delphi technique: making sense of consensus. *Practical assessment, research, and evaluation*, 12(1), <https://doi.org/10.7275/pdz9-th90>.

Hsu, C. C. & Sandford, B. A. 2007b. Minimising non-response in the Delphi process: how to respond to non-response. *Practical Assessment, Research and Evaluation*, 14(17), pp 17.

Hu, X., Wang, F. & Zhang, D. 2022. Similar brains blend emotion in similar ways: Neural representations of individual difference in emotion profiles. *NeuroImage*, 247, pp e118819.

Hughes, D. J., Kratsiotis, I. K., Niven, K. & Holman, D. 2020. Personality traits and emotion regulation: A targeted review and recommendations. *Emotion*, 20(1), pp 63-67.

IBM. 2021. IBM SPSS Statistics for Windows. 28.0 ed. Armonk, New York: IBM Corp.

Ijichi, C., Collins, L. M. & Elwood, R. W. 2014. Pain expression is linked to personality in horses. *Applied Animal Behaviour Science*, 152, pp 38-43.

Ijichi, C., Evans, L., Woods, H. & Yarnell, K. 2020. The Right Angle: Validating a standardised protocol for the use of infra-red thermography of eye temperature as a welfare indicator. *Animal Welfare*, 29(2), pp 123-131.

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), pp 341-346.

Innes, L. & McBride, S. D. 2008. Negative versus positive reinforcement: An evaluation of training strategies for rehabilitated horses. *Applied Animal Behaviour Science*, 112(3-4), pp 357-368.

- Iwanicki, S. & Lehmann, J. 2015. Behavioral and trait rating assessments of personality in common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, 129(3), pp 205.
- James, W. 1884. What is an Emotion? *Mind*, 9(34), pp 188-205.
- Janczarek, I. & Kędzierski, W. 2011. Emotional Response to Novelty and to Expectation of Novelty in Young Race Horses. *Journal of Equine Veterinary Science*, 31(9), pp 549-554.
- Janczarek, I., Kędzierski, W., Stachurska, A., Wilk, I., Kolstrung, R. & Strzelec, K. 2017. Autonomic nervous system activity in purebred Arabian horses evaluated according to the low frequency and high frequency spectrum versus racing performance. *Acta Veterinaria Brno*, 85(4), pp 355-362.
- Janczarek, I., Stachurska, A., Wilk, I., Krakowski, L., Przetacznik, M., Zastrzeżyńska, M. & Kuna-Broniowska, I. 2018. Emotional excitability and behaviour of horses in response to stroking various regions of the body. *Animal Science Journal*, 89(11), pp 1599-1608.
- Jansen, F., Van der Krogt, J., Van Loon, K., Avezzi, V., Guarino, M., Quanten, S. & Berckmans, D. 2009. Online detection of an emotional response of a horse during physical activity. *The veterinary journal*, 181(1), pp 38-42.
- Jansson, A., Lindgren, G., Velie, & B.D., S., M. 2021. An investigation into factors influencing basal eye temperature in the domestic horse (*Equus caballus*) when measured using infrared thermography in field conditions. *Physiology & Behavior*, 228, pp e113218.
- Jastrzębska, E., Wolska, A., Minero, M., Ogłuszka, M., Earley, B., Wejer, J. & Górecka-Bruzda, A. 2017. Conflict Behavior in Show Jumping Horses: A Field Study. *Journal of Equine Veterinary Science*, 57, pp 116-121.
- John, O. P. & Robins, R. W. 1993. Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of personality*, 61(4), pp 521-551.
- Johnson, J. A. & Ostendorf, F. 1993. Clarification of the Five-Factor Model With the Abridged Big Five Dimensional Circumplex. *Journal of Personality and Social Psychology*, 65, pp 563-576.
- Johnson, P. J., Janvier, V., Luh, W.-M., FitzMaurice, M., Southard, T. & Barry, E. F. 2019. Equine Stereotaxic Population Average Brain Atlas With Neuroanatomic Correlation. *Frontiers in Neuroanatomy*, 13, pp 89.

- Joiner, J., Piva, M., Turrin, C. & Chang, S. W. C. 2017. Social learning through prediction error in the brain. *NPJ science of learning*, 2(1), pp 8.
- Jolivald, A. 2021. *An Investigation into Physiological Correlates of Equine Personality*. Ph.D., Nottingham Trent University (United Kingdom), <https://irep.ntu.ac.uk/id/eprint/46515/>.
- Jolivald, A., Ijichi, C., Hall, C. & Yarnell, K. 2023. The mane factor: Compliance is associated with increased hair cortisol in the horse. *Applied Animal Behaviour Science*, 258, pp e105819.
- Jolivald, A., Yarnell, K., Hall, C. & Ijichi, C. 2022. Do you see what I see? Investigating the validity of an equine personality questionnaire. *Applied Animal Behaviour Science*, 248, pp e105567.
- Jolliffe, I. T. & Cadima, J. 2016. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*, 374(2065), pp e20150202.
- Jöreskog, K. G. 1988. Analysis of covariance structures. *Handbook of multivariate experimental psychology*. New York: Springer.
- Junger, S., Payne, S. A., Brine, J., Radbruch, L. & Brearley, S. G. 2017. Guidance on Conducting and Reporting Delphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. . *Palliative Medicine*, 31(8), pp 684-706.
- Jungilligens, J., Paredes-Echeverri, S., Popkirov, S., Barrett, L. F. & Perez, D. L. 2022. A new science of emotion: implications for functional neurological disorder. *Brain*, 145(8), pp 2648-2663.
- Kang, O. D. & Lee, W. S. 2016. Changes in Salivary Cortisol Concentration in Horses during Different Types of Exercise. *Asian-Australas J Anim Sci*, 29(5), pp 747-52.
- Kappel, S. 2023. *Towards the Validation of Novel Cognitive and Physiological Indicators of Affective State in Horses (Equus caballus)*. PhD, University of Plymouth, <https://pearl.plymouth.ac.uk/bitstream/handle/10026.1/20644/2023kappel10627596phd.pdf?sequence=2&isAllowed=y>.

- Kappel, S., Ramirez Montes De Oca, M. A., Collins, S., Herborn, K., Mendl, M. T. & Fureix, C. 2024. Measuring Affect-Related Attention Bias to Emotionally Valenced Visual Stimuli in Horses. *Applied Animal Behaviour Science*, 275, pp e106303.
- Karenina, K. & Giljov, A. 2022. Lateralization in feeding is food type specific and impacts feeding success in wild birds. *Ecology and Evolution*, 12(2), pp e8598.
- Karenina, K., Giljov, A., de Silva, S. & Malashichev, Y. 2018. Social lateralization in wild Asian elephants: visual preferences of mothers and offspring. *Behavioral ecology and sociobiology*, 72, pp 1-11.
- Kay, R. 2012. *The use of heart rate variability measurements as a non-invasive method of assessing affective state in horses*: Nottingham Trent University (United Kingdom), <https://irep.ntu.ac.uk/id/eprint/25/>.
- Kaźmierczak, M. & Nicola, S. M. 2022. The Arousal-motor Hypothesis of Dopamine Function: Evidence that Dopamine Facilitates Reward Seeking in Part by Maintaining Arousal. *Neuroscience*, 499, pp 64-103.
- Keckeis, K., Lepschy, M., Schöpper, H., Moser, L., Troxler, J. & Palme, R. 2012. Hair cortisol: a parameter of chronic stress? Insights from a radiometabolism study in guinea pigs. *Journal of Comparative Physiology B*, 182, pp 985-996.
- Keeney, S., Hasson, F. & McKenna, H. 2006. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, 53, pp 205-212.
- Kellerman, H. & Plutchik, R. 1968. Emotion-trait interrelations and the measurement of personality. *Psychol Rep*, 23(3), pp 1107-14.
- Kennedy, J. S. 1992. *The new anthropomorphism*, Cambridge: Cambridge University Press.
- Kenrick, D. T. & Funder, D. C. 1988. Profiting from controversy: Lessons from the person-situation debate. *American Psychologist*, 43(1), pp 23-34.
- Keverne, E. B., Martensz, N. D. & Tuite, B. 1989. Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. *Psychoneuroendocrinology*, 14(1-2), pp 155-161.

- Kienapfel, K., Link, Y. & König v. Borstel, U. 2014. Prevalence of different head-neck positions in horses shown at dressage competitions and their relation to conflict behaviour and performance marks. *PloS one*, 9(8), pp e103140.
- Kieson, E., Felix, C., Webb, S. & Abramson, C. I. 2020. The effects of a choice test between food rewards and human interaction in a herd of domestic horses of varying breeds and experiences. *Applied Animal Behaviour Science*, 231, pp e105075.
- Killgore, W. D. S. & Yurgelun-Todd, D. A. 2007. The right-hemisphere and valence hypotheses: could they both be right (and sometimes left)? *Social Cognitive and Affective Neuroscience*, 2(3), pp 240-250.
- Kim, J., Jung, H. & Yoon, M. 2023. Relationship between plasma dopamine concentration and temperament in horses. *Domestic Animal Endocrinology*, 83, pp e106788.
- Kim, S.-M. & Cho, G.-J. 2021. Validation of Eye Temperature Assessed Using Infrared Thermography as an Indicator of Welfare in Horses. *Applied Sciences*, 11(16), pp 7186.
- Kimura, R. 1998. Mutual grooming and preferred associate relationships in a band of free-ranging horses. *Applied Animal Behaviour Science*, 59(4), pp 265-276.
- King, J. E. & Figueredo, A. J. 1997. The five-factor model plus dominance in chimpanzee personality. *Journal of research in personality*, 31(2), pp 257-271.
- King, J. E., Weiss, A. & Farmer, K. H. 2005. A chimpanzee (*Pan troglodytes*) analogue of cross-national generalization of personality structure: zoological parks and an African sanctuary. *J Pers*, 73(2), pp 389-410.
- Kinnunen, S., Laukkanen, R., Haldi, J., Hanninen, O. & Atalay, M. 2006. Heart rate variability in trotters during different training periods. *Equine veterinary journal*, 38(S36), pp 214-217.
- Kirk, R. G. W. 2018. Recovering The Principles of Humane Experimental Technique: The 3Rs and the Human Essence of Animal Research. *Science, Technology, & Human Values*, 43(4), pp 622-648.
- Koch, G., Biemann, C., Eiser, I., Fischer, T., Schneider, F., Stumpf, T. & Tijerina García, A. D-WISE Tool Suite for the Sociology of Knowledge Approach to Discourse. International Conference on Human-Computer Interaction, 2022. Springer, pp 68-83.

- Komárková, M. & Bartošová, J. 2013. Lateralized suckling in domestic horses (*Equus caballus*). *Animal cognition*, 16, pp 343-349.
- König v. Borstel, U., Visser, E. K. & Hall, C. 2017. Indicators of stress in equitation. *Applied Animal Behaviour Science*, 190, pp 43-56.
- Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H., De Jong, I. C., Ruis, M. A. W. & Blokhuis, H. J. 1999. Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience & Biobehavioral Reviews*, 23(7), pp 925-935.
- Koski, S. E. 2011. How to Measure Animal Personality and Why Does It Matter? Integrating the Psychological and Biological Approaches to Animal Personality. In: Inoue-Murayama, M., Kawamura, S. & Weiss, A. (eds.) *From Genes to Animal Behavior: Social Structures, Personalities, Communication by Color*. Tokyo: Springer Japan.
- Kovac, M., Ippolitova, T. V., Pozyabin, S., Aliev, R., Lobanova, V., Drakul, N. & Rutland, C. S. 2022. Equine Stress: Neuroendocrine Physiology and Pathophysiology. *Updates on Veterinary Anatomy and Physiology*. IntechOpen.
- Kreibig, S. D. 2010. Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84(3), pp 394-421.
- 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, pp 273-286.
- Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J. & Russell, J. A. 2017. The relation between valence and arousal in subjective experience varies with personality and culture. *Journal of personality*, 85(4), pp 530-542.
- Lacombe, V. A., Podell, M., Furr, M., Reed, S. M., Oglesbee, M. J., Hinchcliff, K. W. & Kohn, C. W. 2001. Diagnostic validity of electroencephalography in equine intracranial disorders. *Journal of veterinary internal medicine*, 15(4), pp 385-393.
- Laister, S., Stockinger, B., Regner, A.-M., Zenger, K., Knierim, U. & Winckler, C. 2011. Social licking in dairy cattle—Effects on heart rate in performers and receivers. *Applied Animal Behaviour Science*, 130(3-4), pp 81-90.
- Lambert, H. & Carder, G. 2019. Positive and negative emotions in dairy cows: Can ear postures be used as a measure? *Behavioural processes*, 158, pp 172-180.

- Lambert, H. S. & Carder, G. 2017. Looking into the eyes of a cow: Can eye whites be used as a measure of emotional state? *Applied animal behaviour science*, 186, pp 1-6.
- Lange, C. G. 1922. *The Emotions*: Cambridge Scholars Publishing.
- Lange, C. G. & James, W. 1922. *The emotions, Vol. 1*, Baltimore, MD, US: Williams & Wilkins Co.
- Lansade, L., Lemarchand, J., Reigner, F., Arnould, C. & Bertin, A. 2022. Automatic brushes induce positive emotions and foster positive social interactions in group-housed horses. *Applied animal behaviour science*, v. 246, pp. 105538--2022.
- Lansade, L. & Calandreau, L. 2018. A conditioned reinforcer did not help to maintain an operant conditioning in the absence of a primary reinforcer in horses. *Behavioural processes*, 146(61-63).
- 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), pp e0171010.
- Lansade, L., Neveux, C. & Levy, F. 2012. A few days of social separation affects yearling horses' response to emotional reactivity tests and enhances learning performance. *Behavioural processes*, 91(1), pp 94-102.
- Lansade, L., Nowak, R., Lainé, A. L., Leterrier, C., Bonneau, C., Parias, C. & Bertin, A. 2018. Facial expression and oxytocin as possible markers of positive emotions in horses. . *Scientific reports*, 8(11), pp 1-11.
- Lansade, L., Philippon, P., Hervé, L. & Vidament, M. 2016. Development of personality tests to use in the field, stable over time and across situations, and linked to horses' show jumping performance. *Applied Animal Behaviour Science*, 176, pp 43-51.
- Lansade, L. & Simon, F. 2010. Horses' learning performances are under the influence of several temperamental dimensions. *Applied Animal Behaviour Science*, 125(1-2), pp 30-37.
- Larose, C., Richard-Yris, M.-A., H., M., & Rogers, L. 2006. Laterality of horses associated with emotionality in novel situations. *Laterality: Asymmetries of Body, Brain, and Cognition*, 11(4), pp 355–367.

- Lawrence, A. B., Vigors, B. & Sandoe, P. 2019. What is so positive about positive animal welfare? – A critical review of the literature. *Animals*, 9(10), pp 703, <https://doi.org/10.3390/ani9100783>.
- Laxman, K. 2015. An integrated 'within' mixed research methodological approach: applying grounded theory to survey design. *International Journal of Innovation and Learning*, 18(1), pp 123-132.
- Lea, S. E. G. 2010. Concept learning in nonprimate mammals: In search of evidence. *The making of human concepts*. New York, NY, US: Oxford University Press.
- Leach, M. C., Klaus, K., Miller, A. L., Scotto di Perrotolo, M., Sotocinal, S. G. & Flecknell, P. A. 2012. The assessment of post-vasectomy pain in mice using behaviour and the Mouse Grimace Scale. *PloS one*, 7(4), pp e35656.
- Lebelt, D., Zanella, A. J. & Unshelm, J. 1998. Physiological correlates associated with cribbing behaviour in horses: changes in thermal threshold, heart rate, plasma β -endorphin and serotonin. *Equine veterinary journal*, 30(S27), pp 21-27.
- LeDoux, J. E. 2017. Semantics, surplus meaning, and the science of fear. *Trends in cognitive sciences*, 21(5), pp 303-306.
- Lee, K. E., Kim, J. G., Lee, H. & Kim, B. S. 2021. Behavioral and cardiac responses in mature horses exposed to a novel object. *Journal of Animal Science and Technology*, 63(3), pp 651-661.
- Leiner, L. & Fendt, M. 2011. Behavioural fear and heart rate responses of horses after exposure to novel objects: Effects of habituation. *Applied Animal Behaviour Science*, 131(3-4), pp 104-109.
- Leliveld, L. M., 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(1-2), pp 1-14.
- Lelláková, M., Pavlák, A., Lešková, L., Florián, M., Skurková, L., Mesarčová, L., Kottferová, L., Takáčová, D. & Kottferová, J. 2023. Monitoring blinks and eyelid twitches in horses to assess stress during the samples collection process. *Journal of Applied Animal Welfare Science*, 26(4), pp 530-539.
- Lenoir, A., Trachsel, D. S., Younes, M., Barrey, E. & Robert, C. 2017. Agreement between electrocardiogram and heart rate meter is low for the measurement of heart rate

- variability during exercise in young endurance horses. *Frontiers in Veterinary Science*, 4, pp 170.
- Lesimple, C. 2020. Indicators of Horse Welfare: State-of-the-Art. *Animals*, 10(2), pp 294.
- Lesimple, C. & Hausberger, M. 2014. How accurate are we at assessing others' well-being? The example of welfare assessment in horses. *Frontiers in psychology*, 5, pp e69278.
- Leslie, A. 2001. Learning: Association or Computation? Introduction to a Special Section. *Current Directions in Psychological Science*, 10, pp 124-127.
- Letzring, T. D. 2008. The good judge of personality: Characteristics, behaviors, and observer accuracy. *Journal of Research in Personality*, 42(4), pp 914-932.
- Letzring, T. D. & Adamcik, L. A. 2015. Personality traits and affective states: Relationships with and without affect induction. *Personality and Individual Differences*, 75, pp e114-120.
- Letzring, T. D., Colman, D. E., Krzyzaniak, S. L. & Roberts, B. W. 2020. Realistic Accuracy Model. *The Wiley Encyclopedia of Personality and Individual Differences*, Chichester: Wiley.
- Levenson, R. W. 2014. The autonomic nervous system and emotion. *Emotion review*, 6(2), pp 100-112.
- Levy, J., Heller, W., Banich, M. T. & Burton, L. A. 1983. Asymmetry of perception in free viewing of chimeric faces. *Brain and cognition*, 2(4), pp 404-419.
- Liotti, M. 2004. On the neural nature of human emotions and implications for biological psychiatry. In Panksepp J. *Textbook of Biological Psychiatry*. Hoboken, NJ: Wiley, 33-74.
- Liu, M. & Wronski, L. 2018. Examining Completion Rates in Web Surveys via Over 25,000 Real-World Surveys. *Social Science Computer Review*, 36(1), pp 116-124.
- Lloyd, A. S., Martin, J. E., Bornett-Gauci, H. L. I. & Wilkinson, R. G. 2007. Evaluation of a novel method of horse personality assessment: Rater-agreement and links to behaviour. *Applied Animal Behaviour Science*, 105(1), pp 205-222.
- Lochmiller, C. R. 2021. Conducting thematic analysis with qualitative data. *The Qualitative Report*, 26(6), pp 2029-2044.

- 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, pp 50-55.
- Loftus, L. 2023. Post Covid-19 behaviour problems: it's not all about pandemic puppies! *In Practice*, 45(6), pp 352-358.
- Loftus, L., Marks, K., Jones-McVey, R., Gonzales, J. L. & Fowler, V. L. 2016. Monty Roberts' public demonstrations: Preliminary report on the heart rate and heart rate variability of horses undergoing training during live audience events. *Animals*, 6(9), pp 55.
- Lynn, C. W. & Bassett, D. S. 2019. The physics of brain network structure, function and control. *Nature Reviews Physics*, 1(5), pp 318-332.
- MacKay, J. R. & Haskell, M. J. 2015. Consistent Individual Behavioral Variation: The Difference between Temperament, Personality and Behavioral Syndromes. *Animals (Basel)*, 5(3), pp 455-78.
- MacLean, P. D. 1990. *The triune brain in evolution: Role in paleocerebral functions*: Springer Science & Business Media.
- Maigrot, A.-L., Hillmann, E., Anne, C. & Briefer, E. F. 2017. Vocal expression of emotional valence in Przewalski's horses (*Equus przewalskii*). *Scientific reports*, 7(1), pp 8779.
- Malik, M., John Camm, A., Thomas Bigger, J., Breithardt, G., Cerutti, S. & Cohen, R. J. 1996. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*, 93(5), pp 1043-65.
- Malmkvist, J., Poulsen, J. M., Luthersson, N., Palme, R., Christensen, J. W. & Søndergaard, E. 2012. Behaviour and stress responses in horses with gastric ulceration. *Applied animal behaviour science*, 142(3-4), pp 160-167.
- Marinier, S. & Alexander, A. 1994. The use of a maze in testing learning and memory in horses. *Applied Animal Behaviour Science*, 39(2), pp 177-182.
- Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I. & de Mendonça, A. 2011. Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Research Notes*, 4(1), pp 299.

- Marsbøll, A. F., Henriksen, B. I. F. & Møller, S. H. 2019. The representativeness of a semi-random sampling method for animal welfare assessments on mink farms. *Animal Welfare*, 28(3), pp 307-315.
- Marsh, L. E., McGreevy, P., Hazel, S., Santos, L., Hebart, M. & Franklin, S. 2019. The effect of tongue-tie application on stress responses in resting horses. *BioRxiv*, p e634717, <https://doi.org/10.1101/634717>.
- Martau, P. A., Caine, N. G. & Candland, D. K. 1985. Reliability of the emotions profile index, primate form, with *Papio hamadryas*, *Macaca fuscata*, and two *Saimiri* species. *Primates*, 26(4), pp 501-505.
- Martin, J., E. 2005. The influence of rearing on personality ratings of captive chimpanzees (*Pan troglodytes*). *Applied Animal Behaviour Science*, 90(2), pp 167-181.
- Martin, P. & Bateson, P. 2007. *Measuring Behaviour: An introductory guide*, Cambridge: Cambridge University Press.
- Martin, S. & Friedman, S. G. Blazing clickers. Animal Behavior Management Alliance Conference, 2011 Denver. Co, available <https://www.behaviorworks.org/files/journals/blazing%20clickers.pdf>.
- Martínez Caro, L. & Martínez García, J. A. 2007. Cognitive–affective model of consumer satisfaction. An exploratory study within the framework of a sporting event. *Journal of Business Research*, 60(2), pp 108-114.
- Masko, M., Domino, M., Lewczuk, D., Jasinski, T. & Gajewski, Z. 2020. Horse behavior, physiology and emotions during habituation to a treadmill. *Animals*, 10(6), pp 921.
- Matheson, S. M., Asher, L. & Bateson, M. 2008. Larger, enriched cages are associated with ‘optimistic’ response biases in captive European starlings (*Sturnus vulgaris*). *Applied Animal Behaviour Science*, 109(2-4), pp 374-383.
- Matsuzawa, T. 2010. Cognitive development in chimpanzees: A trade-off between memory and abstraction? In: Mareschal, D., Quinn, P. C. & Lea, S. E. G. (eds.) *The Making of Human Concepts*. Oxford: Oxford University Press.
- Matsuzawa, T. 2017. Horse cognition and behavior from the perspective of primatology. *Primates*, 58(4), pp 473-477.

- Maurício, L. S., Leme, D. P. & Hötzel, M. J. 2023. How to Understand Them? A Review of Emotional Indicators in Horses. *Journal of Equine Veterinary Science*, 126, pp e104249.
- Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H. & Gross, J. J. 2005. The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, 5(2), pp 175.
- Mauss, I. B. & Robinson, M. D. 2009. Measures of emotion: A review. *Cogn Emot*, 23(2), pp 209-237.
- Mazzola, S. M., Colombani, C., Pizzamiglio, G., Cannas, S., Palestini, C., Dalla Costa, E., Gazzonis, A. L., Bionda, A. & Crepaldi, P. 2021. Do you think I am living well? A four-season hair cortisol analysis on leisure horses in different housing and management conditions. *Animals*, 11(7), pp 2141.
- McAfee, L. M., Mills, D. & Cooper, J. J. 2002. The use of mirrors for the control of stereotypic weaving behaviour in the stabled horse. *Applied Animal Behaviour Science*, 78(2-4), pp 159-173.
- McArthur, R. & Borsini, F. 2006. Animal models of depression in drug discovery: A historical perspective. *Pharmacology, biochemistry, and behavior*, 84, pp 436-52.
- McBride, S. D. & Cuddeford, D. 2001. The putative welfare-reducing effects of preventing equine stereotypic behaviour. *Animal Welfare*, 10(2), pp 173-189.
- 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), pp 76-81.
- McCall, C. A., Salters, M. A. & Simpson, S. M. 1993. Relationship between number of conditioning trials per training session and avoidance learning in horses. *Applied Animal Behaviour Science*, 36(4), pp 291-299.
- McCall, C. A., Potter, G. D., Friend, T. H. & Ingram, R. S. 1981. Learning Abilities in Yearling Horses Using the Hebb-Williams Closed Field Maze. *Journal of Animal Science*, 53(4), pp 928-933.
- McCrae, R. R. & Costa, P. T. 1987. Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), pp 81-90.

- McCrae, R. R. & Costa, P. T. 1989. Rotation to maximize the construct validity of factors in the NEO Personality Inventory. *Multivariate Behavioral Research*, 24(1), pp 107-124.
- McCrae, R. R. & Costa, P. T. 1992. Discriminant validity of NEO-PIR facet scales. *Educational and psychological measurement*, 52(1), pp 229-237.
- McCrae, R. R. & John, O. P. 1992. An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, 60(2), pp 175-215.
- McDonnell, S. M. 2003. *The equid ethogram: a practical field guide to horse behavior*. Lexington: Eclipse Press. ISBN 1-58150-090-4.
- McDonnell, S. M. & Poulin, A. 2002. Equid play ethogram. *Applied Animal Behaviour Science*, 78(2-4), pp 263-290.
- McDuffee, L., Carr, L. & Montepare, W. 2022. An observational evaluation of stress in horses during therapeutic riding sessions. *Journal of Veterinary Behavior*, 49, pp 53-64.
- McDuffee, L., Mills, M., McNiven, M. & Montepare, W. 2019. Establishing statistical stability for heart rate variability in horses. *Journal of veterinary behavior*, 32, pp 30-35.
- McGreevy, P. 2012. *Equine behavior: A guide for veterinarians and equine scientist*, 2, London: Elsevier.
- McGreevy, P. & Nicol, C. J. 1998. Physiological and behavioral consequences associated with short-term prevention of crib-biting in horses. *Physiology & behavior*, 65(1), pp 15-23.
- McGreevy, P. D., Oddie, C., Burton, F. L. & McLean, A. N. 2009. The horse-human dyad: can we align horse training and handling activities with the equid social ethogram? *Vet J*, 181(1), pp 12-8.
- McGrogan, C., Hutchison, M. D. & King, J. E. 2008. Dimensions of horse personality based on owner and trainer supplied personality traits. *Applied Animal Behaviour Science*, 113(1-3), pp 206-214.
- McHugh, M. L. 2012. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*, 22(3), pp 276-82.

- McKenna, H. P. 1994. The Delphi technique: a worthwhile research approach for nursing? . *Journal of Advanced Nursing*, 19(6), pp 1221-1225.
- McLean, A. N. 2004. Short-term spatial memory in the domestic horse. *Applied Animal Behaviour Science*, 85(1), pp 93-105.
- McLean, A. N. & McGreevy, P. D. 2010. Horse-training techniques that may defy the principles of learning theory and compromise welfare. *Journal of Veterinary Behavior*, 5(4), pp 187-195.
- McLean, A. N. & Christensen, J. W. 2017. The application of learning theory in horse training. *Applied Animal Behaviour Science*, 190, pp 18-27.
- 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, pp 19-26.
- McNiff, J. 2013. *Action research: Principles and practice*. London: Routledge.
- McVey, A., Wilkinson, A. & Mills, D. S. 2018. Social learning in horses: The effect of using a group leader demonstrator on the performance of familiar conspecifics in a detour task. *Applied Animal Behaviour Science*, 209, pp 47-54.
- Meagher, R. K. 2009. Observer ratings: Validity and value as a tool for animal welfare research. *Applied Animal Behaviour Science*, 119(1), pp 1-14.
- Mellor, D. J., Patterson-Kane, E. & Stafford, K. J. 2009. *The sciences of animal welfare*: John Wiley & Sons.
- Mellor, D. J. 2012. Animal emotions, behaviour and the promotion of positive welfare states. . *New Zealand Veterinary Journal*, 60, pp 1-8.
- Mellor, D. J. 2015a. Enhancing animal welfare by creating opportunities for positive affective engagement. *New Zealand veterinary journal*, 63(1), pp 3-8.
- Mellor, D. J. 2015b. Positive animal welfare states and reference standards for welfare assessment. *New Zealand veterinary journal*, 63(1), pp 17-23.
- Mellor, D. J. 2016. Updating animal welfare thinking: Moving beyond the “Five Freedoms” towards “a Life Worth Living”. *Animals*, 6(3), pp 21.

- Mellor, D. J., Beausoleil, N. J., Littlewood, K. E., McLean, A. N., McGreevy, P. D., Jones, B. & Wilkins, C. 2020. The 2020 Five Domains Model: Including human-animal interactions in assessments of animal welfare. *Animals*, 10, pp 1870.
- Menchetti, L., Dalla Costa, E., Minero, M. & Padalino, B. 2021. Development and validation of a test for the classification of horses as broken or unbroken. *Animals*, 11(8), pp 2303.
- Mendl, M. & Paul, E. S. 2004. Consciousness, Emotion and Animal Welfare: Insights from Cognitive Science. *Animal Welfare*, 13, pp (S1):S17-S25. doi:10.1017/S0962728600014330
- 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), pp 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*, 277, pp 2895-2904.
- Mendl, M. & Paul, E. S. 2020a. Animal affect and decision-making. *Neuroscience & Biobehavioral Reviews*, 112, pp 144-163.
- Mendl, M. & Paul, E. S. 2020b. Assessing affective states in animals. *Mental health and well-being in animals*. Wallingford UK: CABI.
- Meng, X.-L. 2020. Reproducibility, replicability, and reliability. *Harvard Data Science Review*, 2(4), pp 10.
- Mengoli, M., Pageat, P., Lafont-Lecuelle, C., Monneret, P., Giacalone, A., Sighieri, C. & Cozzi, A. 2014. Influence of emotional balance during a learning and recall test in horses (*Equus caballus*). *Behavioural processes*, 106, pp 141-150.
- Merkies, K., Ready, C., Farkas, L. & Hodder, A. 2019. Eye blink rates and eyelid twitches as a non-invasive measure of stress in the domestic horse. *Animals*, 9(8), pp 562.
- Messina, I., Grecucci, A. & Viviani, R. 2021. Neurobiological models of emotion regulation: A meta-analysis of neuroimaging studies of acceptance as an emotion regulation strategy. *Social Cognitive and Affective Neuroscience*, 16(3), pp 257-267.

- Miller, A. L., Kitson, G. L., Skalkoyannis, B., Flecknell, P. A. & Leach, M. C. 2016. Using the mouse grimace scale and behaviour to assess pain in CBA mice following vasectomy. *Applied Animal Behaviour Science*, 181, pp 160-165.
- Miller, J. D., McCain, J., Lynam, D. R., Few, L. R., Gentile, B., MacKillop, J. & Campbell, W. K. 2014. A comparison of the criterion validity of popular measures of narcissism and narcissistic personality disorder via the use of expert ratings. *Psychological Assessment*, 26(3), pp 958-969.
- Mills, D. S. 2022. Clinical Animal Behaviour: Paradigms, Problems and Practice. *Animals (Basel)*, 12(22), pp. 3103. doi: 10.3390/ani12223103.
- Mills, D. S. 1998a. Applying learning theory to the management of the horse: the difference between getting it right and getting it wrong. *Equine veterinary journal. Supplement*, 27, pp 44-48.
- Mills, D. S. 1998b. Personality and individual differences in the horse, their significance, use and measurement. *Equine Veterinary Journal*, 30(S27), pp 10-13.
- Minero, M., Dalla Costa, E., Dai, F., Canali, E., Barbieri, S., Zanella, A., Pascuzzo, R. & Wemelsfelder, F. 2018. Using qualitative behaviour assessment (QBA) to explore the emotional state of horses and its association with human-animal relationship. *Applied Animal Behaviour Science*, 204, pp 53-59.
- Minitab, L. 2021. Minitab. 18 ed. <https://www.minitab.com>: Minitab.
- Mirkó, E., Dóka, A. & Miklósi, Á. 2013. Association between subjective rating and behaviour coding and the role of experience in making video assessments on the personality of the domestic dog (*Canis familiaris*). *Applied Animal Behaviour Science*, 149(1-4), pp 45-54.
- Moberg, G. P. 2000. Biological response to stress: implications for animal welfare. *The biology of animal stress: Basic principles and implications for animal welfare*.: CABI publishing Wallingford UK.
- Mohammadi, G., Lin, K. & Vuilleumier, P. Towards Understanding Emotional Experience in a Componential Framework. 2019. 8th International Conference on Affective Computing and Intelligent Interaction (ACII), 3-6 Sept, pp 123-129.
- Mormède, P., Andanson, S., Aupérin, B., Beerda, B., Guémené, D., Malmkvist, J., Manteca, X., Manteuffel, G., Prunet, P. & van Reenen, C. G. 2007. Exploration of the

- hypothalamic–pituitary–adrenal function as a tool to evaluate animal welfare. *Physiology & behavior*, 92(3), pp 317-339.
- Mormède, P., Foury, A., Terenina, E. & Knap, P. W. 2011. Breeding for robustness: the role of cortisol. *animal*, 5(5), pp 651-657.
- Morris, P. H., Gale, A. & Duffy, K. 2002. Can judges agree on the personality of horses? *Personality and Individual differences*, 33(1), pp 67-81.
- Mortillaro, M. & Mehu, M. 2015. Emotions: Methods of Assessment. *International Encyclopedia of the Social & Behavioral Sciences*, pp 519-525.
- Mota-Rojas, D., Miranda-Cortés, A., Casas-Alvarado, A., Mora-Medina, P., Boscatto-Funes, L. & Hernández-Ávalos, I. 2021. Neurobiology and modulation of stress-induced hyperthermia and fever in animal. *Abanico veterinario*, 11, <https://doi.org/10.21929/abavet2021.11>.
- Mott, R. O., Hawthorne, S. J. & McBride, S. D. 2020. Blink rate as a measure of stress and attention in the domestic horse (*Equus caballus*). *Scientific Reports*, 10(1), pp e21409. <https://doi.org/10.1038/s41598-020-78386-z>.
- Mullan, S., Szmaragd, C., Hotchkiss, J. & Whay, H. R. 2014. The welfare of long-line tethered and free-ranging horses kept on public grazing land in South Wales. *Animal Welfare*, 23(1), pp 25-37.
- Mullard, J., Berger, J. M., Ellis, A. D. & Dyson, S. 2017. Development of an ethogram to describe facial expressions in ridden horses (FEReq). *Journal of veterinary behavior*, 18, pp 7-12.
- Mullen, P. M. 2003. Delphi: myths and reality. *J Health Organ Manag*, 17(1), pp 37-52.
- Munsters, C. C. B. M., Visser, E. J., van den Broek, J. & van Oldruitenborgh-Oosterbaan, M. M. S. 2013. Physiological and behavioral responses of horses during police training. *animal*, 7(5), pp 822-827.
- Munsters, C. C. B. M., Visser, E. K., van den Broek, J. & van Oldruitenborgh-Oosterbaan, M. M. S. 2012. The influence of challenging objects and horse-rider matching on heart rate, heart rate variability and behavioural score in riding horses. *The Veterinary Journal*, 192(1), pp 75-80.
- Murphy, J. 2009. Assessing equine prospective memory in a Y-maze apparatus. *The Veterinary Journal*, 181(1), pp 24-28.

- Murphy, J. & Arkins, S. 2007. Equine learning behaviour. *Behav Processes*, 76(1), pp 1-13.
- Muthen, B. & Asparouhov, T. 2012. Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, 17(3), pp 313-335.
- Mweshi, G. K. & Sakyi, K. 2020. Application of sampling methods for the research design. *Archives of Business Review—Vol*, 8(11), pp 180-193, <https://doi.org/10.14738/abr.811.9042>.
- Nath, L. C., Forbes, G., Elliott, A. D., Tomren, V., Ryan, A. & Franklin, S. H. 2022. Application of an electrocardiography device (iECG) for heart rhythm analysis after exercise in Thoroughbred horses. *Australian Veterinary Journal*, 100(3), pp 114-120.
- Neethirajan, S. 2022. Affective state recognition in livestock—artificial intelligence approaches. *Animals*, 12(6), pp 759.
- Nelis, D., Quoidbach, J., Hansenne, M. & Mikolajczak, M. 2011. Measuring Individual Differences in Emotion Regulation: The Emotion Regulation Profile-Revised (ERP-R). *Psychologica Belgica*, 51(1), pp 49–91. <https://doi.org/10.5334/pb-51-1-49>.
- Nesse, R. M. & Ellsworth, P. C. 2009. Evolution, emotions, and emotional disorders. *American Psychologist*, 64(2), pp 129.
- Nestler, E. J., Gould, E., Manji, H., Buncan, M., Duman, R. S., Greshenfeld, H. K., Hen, R., Koester, S., Lederhendler, I., Meaney, M., Robbins, T., Winsky, L. & Zalcman, S. 2002. Preclinical models: status of basic research in depression. *Biol Psychiatry*, 52(6), pp 503-28.
- Netzer, L., Gutentag, T., Kim, M. Y., Solak, N. & Tamir, M. 2018. Evaluations of emotions: Distinguishing between affective, behavioral and cognitive components. *Personality and Individual Differences*, 135, pp 13-24.
- Nicol, C. J. 1996. Farm animal cognition. *Animal Science*, 62(3), pp 375-391.
- Nicol, C. J. 2002. Equine learning: Progress and suggestions for future research. *Applied Animal Behaviour Science*, 78(2-4), pp 193-208.
- Ninomiya, S., Anjiki, A., Nishide, Y., Mori, M., Deguchi, Y. & Satoh, T. 2013. Polymorphisms of the dopamine D4 receptor gene in stabled horses are related to differences in behavioral response to frustration. *Animals*, 3(3), pp 663-669.

- 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), pp 465-469, <https://doi.org/10.1111/j.1740-0929.2004.00214.x>.
- Ninomiya, S., Mitsumasu, T., Aoyama, M. & Kusunose, R. 2007. A note on the effect of a palatable food reward on operant conditioning in horses. *Applied Animal Behaviour Science*, 108(3), pp 342-347.
- Nishimuta, M., Rosales-Ruiz, J., Will, S. M. & Hunter, M. 2022. Equines do not live for grass alone: Training equines using social interaction as a reinforcer. *Journal of the Experimental Analysis of Behavior*, 118(2), pp 261-277.
- Noble, G. K., Blackshaw, K. L., Cowling, A., Harris, P. A. & Sillence, M. N. 2013. An objective measure of reactive behaviour in horses. *Applied Animal Behaviour Science*, 144(3), pp 121-129.
- Normando, S., Haverbeke, A., Meers, L., Ödberg, F. O., Ibáñez Talegón, M. & Bono, G. 2003. Effect of manual imitation of grooming on riding horses' heart rate in different environmental situations. *Veterinary Research Communications*, 27, pp 615-617.
- Nyumba, O. T., Wilson, K., Derrick, C. J. & Mukherjee, N. 2018. The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and evolution*, 9(1), pp 20-32.
- O'Connor, C. M., Norris, D. R., Crossin, G. T. & Cooke, S. J. 2014. Biological carryover effects: linking common concepts and mechanisms in ecology and evolution. *Ecosphere*, 5(3), pp 1-11.
- Ohl, F., Arndt, S. S. & van der Staay, F. J. 2008. Pathological anxiety in animals. *The veterinary journal*, 175(1), pp 18-26.
- Okada, K., Takemura, K. & Sato, S. 2013. Investigation of Various Essential Factors for Optimum Infrared Thermography. *Journal of Veterinary Medical Science*, 75(10), pp 1349-1353.
- Ossorio, P. G. 2006. *The Behaviour of Persons: The Collected Works of Peter G. Ossorio*, Michigan: Descriptive Psychology Press.
- Otzenberger, H., Gronfier, C., Simon, C., Charloux, A., Ehrhart, J., Piquard, F. & Brandenberger, G. 1998. Dynamic heart rate variability: a tool for exploring sympathovagal balance continuously during sleep in men. *American Journal of Physiology-Heart and Circulatory Physiology*, 275(3), pp H946-H950.

- Pagani, M., Lombardi, F., Guzzetti, S., Sandrone, G., Rimoldi, O., Malfatto, G., Cerutti, S. & Malliani, A. 1984. Power spectral density of heart rate variability as an index of sympatho-vagal interaction in normal and hypertensive subjects. *Journal of hypertension. Supplement: official journal of the International Society of Hypertension*, 2(3), pp S383-5.
- Panksepp, J. 1998. The periconscious substrates of consciousness: Affective states and the evolutionary origins of the self. *Journal of consciousness studies*, 5(5-6), pp 566-582.
- Panksepp, J. 2003. At the interface of the affective, behavioral, and cognitive neurosciences: Decoding the emotional feelings of the brain. *Brain and cognition*, 52(1), pp 4-14.
- Panksepp, J. 2005. Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and cognition*, 14(1), pp 30-80.
- Panksepp, J. 2006. Emotional endophenotypes in evolutionary psychiatry. *Prog Neuropsychopharmacol Biol Psychiatry*, 30(5), pp 774-84.
- Panksepp, J. 2007. Neurologizing the Psychology of Affects: How Appraisal-Based Constructivism and Basic Emotion Theory Can Coexist. *Perspect Psychol Sci*, 2(3), pp 281-96.
- Panksepp, J. 2011a. The basic emotional circuits of mammalian brains: Do animals have affective lives? *Neuroscience and Biobehavioural Reviews*, 35, pp 1791-1804.
- Panksepp, J. 2011b. Cross-species affective neuroscience decoding of the primal affective experiences of humans and related animals. *PLoS one*, 6(9), pp e21236.
- Panksepp, J. 2011c. The “Dynamic Unconscious” May Be Experienced: Can We Discuss Unconscious Emotions When There Are no Adequate Measures of Affective Change? *Neuropsychoanalysis*, 13(1), pp 51-59.
- Panksepp, J., Lane, R. D., Solms, M. & Smith, R. 2017. Reconciling cognitive and affective neuroscience perspectives on the brain basis of emotional experience. *Neurosci Biobehav Rev*, 76(Pt B), pp 187-215.
- Pannewitz, L. & Loftus, L. 2023. Frustration in horses: Investigating expert opinion on behavioural indicators and causes using a delphi consultation. *Applied Animal Behaviour Science*, 258, pp e105818, <https://doi.org/10.1016/j.applanim.2022.105818>.

- Papez, J. W. 1937. A proposed mechanism of emotion. *Archives of Neurology & Psychiatry*, 38, pp 725-743.
- Park, Y. S., Konge, L. & Artino Jr, A. R. 2020. The positivism paradigm of research. *Academic medicine*, 95(5), pp 690-694.
- Parker, M., Goodwin, D., Eager, R. A., Redhead, E. S. & Marlin, D. J. 2009. Comparison of Polar® heart rate interval data with simultaneously recorded ECG signals in horses. *Comparative Exercise Physiology*, 6(4), pp 137-142.
- Pastell, M. 2016. CowLog – Cross-Platform Application for Coding Behaviours from Video. *Journal of Open Research Software.*, 4(1), pp 15.
- Paul, E. S., Harding, E. J. & Mendl, M. 2005a. Measuring emotional processes in animals: the utility of a cognitive approach. *Neuroscience & Biobehavioral Reviews*, 29(3), pp 469-491.
- Paul, E. S., Harding, E. J. & Mendl, M. 2005b. Measuring emotional processes in animals: the utility of a cognitive approach. *Neurosci Biobehav Rev*, 29(3), pp 469-91.
- 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, pp 749-770.
- Pawluski, J., Jegou, P., Henry, S., Bruchet, A., Palme, R., Coste, C. & Hausberger, M. 2017. Low plasma cortisol and fecal cortisol metabolite measures as indicators of compromised welfare in domestic horses (*Equus caballus*). *PLoS One*, 12(9), pp e0182257.
- Peabody, D. & Goldberg, L. R. 1989. Some determinants of factor structures from personality-trait descriptors. *J Pers Soc Psychol*, 57(3), pp 552-67.
- Pearson, G., Waran, N., Reardon, R. J. M., Keen, J. & Dwyer, C. 2021. A Delphi study to determine expert consensus on the behavioural indicators of stress in horses undergoing veterinary care. *Applied Animal Behaviour Science*, 237, pp e105291, <https://psycnet.apa.org/doi/10.1016/j.applanim.2021.105291>.
- Peeters, M., Sulon, J., Beckers, J. F., Ledoux, D. & Vandenheede, M. 2011. Comparison between blood serum and salivary cortisol concentrations in horses using an adrenocorticotrophic hormone challenge. *Equine veterinary journal*, 43(4), pp 487-493.

- Pelle, A. J., Kupper, N., Mols, F. & de Jonge, P. 2013. What is the use? Application of the short form (SF) questionnaires for the evaluation of treatment effects. *Quality of Life Research*, 22, pp 1225-1230.
- 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), pp 60-69.
- Perea, C. S., Paternina, A. C., Gomez, Y. & Lattig, M. C. 2012. Negative affectivity moderated by BDNF and stress response. *J Affect Disord*, 136(3), pp 767-74.
- Pessoa, L. 2017. A network model of the emotional brain. *Trends in cognitive sciences*, 21(5), pp 357-371.
- Pessoa, L. 2022. *The entangled brain*, Michigan: The MIT Press.
- Pessoa, L. & Hof, P. R. 2015. From Paul Broca's great limbic lobe to the limbic system. *The Journal of comparative neurology*, 523(17), pp 2495.
- Peterson, C. B., Crosby, R. D., Wonderlich, S. A., Joiner, T., Crow, S. J., Mitchell, J. E., Bardone-Cone, A. M., Klein, M. & Le Grange, D. 2007. Psychometric properties of the eating disorder examination-questionnaire: Factor structure and internal consistency. *International Journal of Eating Disorders*, 40(4), pp 386-389.
- Pfaller-Sadovsky, N., Hurtado-Parrado, C., Cardillo, D., Medina, L. G. & Friedman, S. G. 2020. What's in a click? the efficacy of conditioned reinforcement in applied animal training: a systematic review and meta-analysis. *Animals*, 10(10), pp 1757.
- Pham, T., Lau, Z. J., Chen, S. H. A. & Makowski, D. 2021. Heart Rate Variability in Psychology: A Review of HRV Indices and an Analysis Tutorial. *Sensors*, 21(12), pp 3998.
- Phelipon, R., Ascione, A., Ruet, A., Bertin, A., Biau, S., Arnould, C., Boichot, L. & Lansade, L. 2024. Differences in behaviour, facial expressions and locomotion between positive anticipation and frustration in horses. *Applied Animal Behaviour Science*, 274, pp e106262.
- Phillips, C. 2007. Animal welfare: A construct of positive and negative affect? *Veterinary Journal (London, England: 1997)*, 175(3), pp 291-292.

- Physick-Sheard, P. W., Marlin, D. J., Thornhill, R. & Schroter, R. C. 2000. Frequency domain analysis of heart rate variability in horses at rest and during exercise. *Equine Veterinary Journal*, 32(3), pp 253-262.
- Pierard, M., McGreevy, P. & Geers, R. 2017. Developing behavioral tests to support selection of police horses. *Journal of veterinary behavior*, 19, pp 7-13.
- Pierard, M., McGreevy, P. & Geers, R. 2019. Reliability of a descriptive reference ethogram for equitation science. *Journal of veterinary behavior*, 29, pp 118-127.
- Pinel, J. P. J. 1990. *Biopsychology*, Needham Heights, MA, US: Allyn & Bacon.
- Platzer, J. & Feuerbacher, E. N. 2022. Reinforcer efficacy of grain for horses. *Journal of the Experimental Analysis of Behavior*, 118(2), pp 302-315.
- Plutchik, R. 1980. A general psychoevolutionary theory of emotion. *Theories of emotion*. Elsevier.
- Plutchik, R. & Kellerman, H. 1974. *Emotions Profile Index*. Los Angeles, CA: Western Psychological Services.
- PolarElectro. 2024. Polar Flow. Available <https://flow.polar.com/>.
- Porges, S. W. 1995. Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A polyvagal theory. *Psychophysiology*, 32(4), pp 301-318.
- Powell, C. 2003. The Delphi technique: myths and realities. *J Adv Nurs*, 41(4), pp 376-82.
- Proops, L., Grounds, K., Smith, A. V. & McComb, K. 2018. Animals Remember Previous Facial Expressions that Specific Humans Have Exhibited. *Current Biology*, 28(9), pp 1428-1432.e4.
- Proops, L. & McComb, K. 2010. Attributing attention: the use of human-given cues by domestic horses (*Equus caballus*). *Animal cognition*, 13, pp 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), pp 947-951.

- Proudfoot, K. & Habing, G. 2015. Social stress as a cause of diseases in farm animals: current knowledge and future directions. *The Veterinary Journal*, 206(1), pp 15-21.
- Puppe, B., Ernst, K., Schön, P. C. & Manteuffel, G. 2007. Cognitive enrichment affects behavioural reactivity in domestic pigs. *Applied Animal Behaviour Science*, 105(1-3), pp 75-86.
- Quadt, L., Critchley, H. & Nagai, Y. 2022. Cognition, emotion, and the central autonomic network. *Autonomic Neuroscience*, 238, pp e102948.
- Qualtrics. 2020. Qualtrics. Provo, UT, USA: Qualtrics software. Available <https://www.qualtrics.com/uk/>.
- Raby, C. R., Alexis, D. M., Dickinson, A. & Clayton, N. S. 2007. Planning for the future by western scrub-jays. *Nature*, 445(7130), pp 919-921.
- Radiya-Dixit, A. 2017. *Visualizing the fourth dimension*. [Online]. Available: <https://researchblog.duke.edu/2017/04/26/visualizing-the-fourth-dimension/>, accessed October 2022.
- Rajmohan, V. & Mohandas, E. 2007. The limbic system. *Indian J Psychiatry*, 49(2), pp 132-9.
- Ralph, C. R. & Tilbrook, A. J. 2016. INVITED REVIEW: The usefulness of measuring glucocorticoids for assessing animal welfare. *Journal of Animal Science*, 94(2), pp 457-470.
- Ramirez Montes De Oca, M. A. 2021. Temperature asymmetries in facial areas as indicators of affective state in dairy calves and horses. *Laterality: Asymmetries of Body, Brain and Cognition*, 12, pp 464-474.
- Ramos, B. P., Stark, D., Verduzco, L., van Dyck, C. H. & Arnsten, A. F. T. 2006. α 2A-adrenoceptor stimulation improves prefrontal cortical regulation of behavior through inhibition of cAMP signaling in aging animals. *Learning & Memory*, 13(6), pp 770-776.
- Randle, H. D. 2015. Personality and performance: the influence of behaviour. In: Williams, J.M., Evans, D. (Eds.), *Training for equestrian performance*. Wageningen Academic Publishers, The Netherlands.
- Range, F., Bugnyar, T. & Kotrschal, K. 2008. The performance of ravens on simple discrimination tasks: a preliminary study. *acta ethologica*, 11(1), pp 34-41.

- Rankins, E. M. & Wickens, C. L. 2020. A systematic review of equine personality. *Applied Animal Behaviour Science*, 231, pp e105076.
- Ransom, J. I. & Cade, B. S. 2009. Quantifying equid behavior—a research ethogram for free-roaming feral horses. Reston, Virginia: *Publications of the US Geological Survey*, 26, pp 1-33.
- Rawlings, C. A., Garner, H. E., White, N. A. & Moore, J. N. 1982. Ear temperatures in the horse during recovery from anesthesia. *Journal of equine veterinary science*, 2(1), pp 12-15.
- RCoreTeam. 2023. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, available <https://www.r-project.org/>.
- Ready, R. E., Clark, L. A., Watson, D. & Westerhouse, K. 2000. Self-and peer-reported personality: Agreement, trait ratability, and the “self-based heuristic”. *Journal of research in personality*, 34(2), pp 208-224.
- Redaelli, V., Luzi, F., Mazzola, S., Bariffi, G. D., Zappaterra, M., Nanni Costa, L. & Padalino, B. 2019. The Use of Infrared Thermography (IRT) as Stress Indicator in Horses Trained for Endurance: A Pilot Study. *Animals*, 9(3), pp 84.
- RedGorilla. 2024. *Gorilla Tub Mini Shallow 5L* [Online]. Available: <https://www.redgorilla.red/products/gorilla-tub-mini-5l?variant=47370075570486>.
- Redish, A. D. 2013. *The mind within the brain: How we make decisions and how those decisions go wrong*, New York: Oxford University Press, USA.
- Reefmann, N., Kaszas, F. B., Wechsler, B. & Gyga, L. 2009. Physiological expression of emotional reactions in sheep. *Physiology & behavior*, 98(1-2), pp 235-241.
- Reid, J., Wiseman-Orr, M. L., Scott, E. M. & Nolan, A. M. 2013. Development, validation and reliability of a web-based questionnaire to measure health-related quality of life in dogs. *Journal of Small Animal Practice*, 54(5), pp 227-233.
- Reimert, I., Webb, L. E., van Marwijk, M. A. & Bolhuis, J. E. 2023. Review: Towards an integrated concept of animal welfare. *animal*, 17, pp e100838.
- Reisenzein, R. & Stephan, A. 2014. More on James and the physical basis of emotion. *Emotion Review*, 6(1), pp 35-46.

- Ricci-Bonot, C. & Mills, D. S. 2023. Recognising the facial expression of frustration in the horse during feeding period. *Applied Animal Behaviour Science*, 265, pp e105966.
- Rich, J. L., Chojenta, C. & Loxton, D. 2013. Quality, Rigour and Usefulness of Free-Text Comments Collected by a Large Population Based Longitudinal Study - ALSWH. *PLOS ONE*, 8(7), pp e68832.
- Richards, J. M. & Gross, J. J. 1999. Composure at any cost? The cognitive consequences of emotion suppression. *Personality and Social Psychology Bulletin*, 25(8), pp 1033-1044.
- 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, pp 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), pp 121-136.
- Riffenburgh, R. H. 2012. *Statistics in medicine*, London: Academic press.
- Rindfleisch, A., Malter, A. J., Ganesan, S. & Moorman, C. 2008. Cross-sectional versus longitudinal survey research: Concepts, findings, and guidelines. *Journal of marketing research*, 45(3), pp 261-279.
- Ringhofer, M., Inoue, S., Mendonça, R. S., Pereira, C., Matsuzawa, T., Hirata, S. & Yamamoto, S. 2017. Comparison of the social systems of primates and feral horses: data from a newly established horse research site on Serra D'Arga, northern Portugal. *Primates*, 58(4), pp 479-484.
- Rioja-Lang, F. C., Connor, M., Bacon, H. & Dwyer, C. M. 2020. Determining a welfare prioritization for horses using a Delphi method. *Animals*, 10(4), pp 647, <https://doi.org/10.3390/ani10040647>.
- Roberts, B. W., Walton, K. E. & Viechtbauer, W. 2006. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychol Bull*, 132(1), pp 1-25.
- Robinson, E., Bevelander, K. E., Field, M. & Jones, A. 2018a. Methodological and reporting quality in laboratory studies of human eating behavior. *Appetite*, 125, pp 486-491.

- Robinson, L. M., Coleman, K., Capitanio, J. P., Gottlieb, D. H., Handel, I. G., Adams, M. J., Leach, M. C., Waran, N. K. & Weiss, A. 2018b. Rhesus macaque personality, dominance, behavior, and health. *American Journal of Primatology*, 80(2), pp e22739.
- Robinson, L. M., Waran, N. K., Handel, I. & Leach, M. C. 2021. Happiness, welfare, and personality in rhesus macaques (*Macaca mulatta*). *Applied Animal Behaviour Science*, 236, pp e105268.
- Robinson, L. M., Waran, N. K., Leach, M. C., Morton, F. B., Paukner, A., Lonsdorf, E., Handel, I., Wilson, V. A. D., Brosnan, S. F. & Weiss, A. 2016. Happiness is positive welfare in brown capuchins (*Sapajus apella*). *Applied Animal Behaviour Science*, 181, pp 145-151.
- Robinson, M. D. & Clore, G. L. 2002. Episodic and semantic knowledge in emotional self-report: evidence for two judgment processes. *Journal of personality and social psychology*, 83(1), pp 198.
- Rochais, C., Henry, S., Sankey, C., Nassur, F., Gorecka-Bruzda, A. & Hausberger, M. 2014. Visual attention, an indicator of human-animal relationships? A study of domestic horses (*Equus caballus*). *Frontiers in psychology*, 5, pp e47331.
- Rochais, C., Sébilleau, M., Ménoret, M., Oger, M., Henry, S., Hausberger, M. & Cousillas, H. 2018. Attentional state and brain processes: state-dependent lateralization of EEG profiles in horses. *Scientific reports*, 8(1), pp e10153.
- Roche, D. G., Careau, V. & Binning, S. A. 2016. Demystifying animal 'personality'(or not): why individual variation matters to experimental biologists. *Journal of Experimental Biology*, 219(24), pp 3832-3843.
- Rogers, L. J. & Andrew, R. 2002. *Comparative vertebrate lateralization.*: Cambridge University Press.
- Rogers, L. J., Vallortigara, G. & Andrew, R. J. 2013. *Divided brains: the biology and behaviour of brain asymmetries*: Cambridge University Press.
- Rokade, P. B. 2011. Release of endorphin hormone and its effects on our body and moods: A review. International Conference on Chemical, Biological and Environment Sciences, pp 436-438.

- Rolls, E. T. 2005. *Emotion explained: Affective Science*. Oxford: Oxford University Press, ISBN: 9780198570042.
- Rørvang, M. V., Christensen, J. W., Ladewig, J. & McLean, A. 2018. Social Learning in Horses- Fact or Fiction? *Front Vet Sci*, 5, pp 212.
- Rørvang, M. V., Ničová, K., Sassner, H. & Nawroth, C. 2021. Horses' (*Equus caballus*) ability to solve visible but not invisible displacement tasks is associated with frustration behavior and heart rate. *Frontiers in Behavioral Neuroscience*, 15, pp e792035.
- Ross, E. D. 2021. Differential Hemispheric Lateralization of Emotions and Related Display Behaviors: Emotion-Type Hypothesis. *Brain Sciences*, 11(8), pp 1034.
- Ruet, A., Arnould, C., Lemarchand, J., Parias, C., Mach, N., Moisan, M. P., Foury, A., Briant, C. & Lansade, L. 2022. Horse welfare: A joint assessment of four categories of behavioural indicators using the AWIN protocol, scan sampling and surveys. *Animal Welfare*, 31(4), pp 455-466.
- Russell, J. A. 1980. A circumplex model of affect. *Journal of personality and social psychology*, 39(6), pp 1161.
- Russell, J. A. 1994. Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological bulletin*, 115(1), pp 102.
- Russell, J. A. 2003. Core affect and the psychological construction of emotion. *Psychological review*, 110(1), pp 145.
- Russell, J. A. & Barrett, L. F. 1999. Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76, pp 805-819.
- Samson, A. C., Kreibig, S. D., Soderstrom, B., Wade, A. A. & Gross, J. J. 2016. Eliciting positive, negative and mixed emotional states: A film library for affective scientists. *Cogn Emot*, 30(5), pp 827-56.
- Samuel, D. B., Lynam, D. R., Widiger, T. A. & Ball, S. A. 2012. An expert consensus approach to relating the proposed DSM-5 types and traits. 3, pp 1-16.
- Sankey, C., Henry, S., Clouard, C., Richard-Yris, M.-A. & Hausberger, M. 2011. Asymmetry of behavioral responses to a human approach in young naive vs. trained horses. *Physiology & behavior*, 104(3), pp 464-468.

- Sankey, C., Richard-Yris, M.-A., Leroy, H., Henry, S. & Hausberger, M. 2010a. Positive interactions lead to lasting positive memories in horses, *Equus caballus*. *Animal Behaviour*, 79(4), pp 869-875.
- Sankey, C., Richard-Yris, M. A., Henry, S., Fureix, C., Nassur, F. & Hausberger, M. 2010b. Reinforcement as a mediator of the perception of humans by horses (*Equus caballus*). *Animal Cognition*, 13, pp 753-764.
- Saslow, C. A. 2002. Understanding the perceptual world of horses. . *Applied Animal Behaviour Science*, 78(2-4), pp 209-224.
- Sauer, F. J., Gerber, V., Frei, S., Bruckmaier, R. M. & Groessl, M. 2020. Salivary cortisol measurement in horses: immunoassay or LC-MS/MS? *Domestic animal endocrinology*, 72, pp e106445.
- Sauer, F. J., Hermann, M., Ramseyer, A., Burger, D., Riemer, S. & Gerber, V. 2019. Effects of breed, management and personality on cortisol reactivity in sport horses. *PLoS One*, 14(12), pp e0221794.
- Savalli, C., Albuquerque, N., Vasconcellos, A. S., Ramos, D., de Mello, F. T. & Mills, D. S. 2019. Assessment of emotional predisposition in dogs using PANAS (Positive and Negative Activation Scale) and associated relationships in a sample of dogs from Brazil. *Scientific Reports*, 9(1), pp e18386.
- Scheid, C. & Noë, R. 2010. The performance of rooks in a cooperative task depends on their temperament. *Animal Cognition*, 13, pp 545-553.
- Scherer, K. R. 2009. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), pp 3459-3474.
- Schmidt, A., Aurich, J., Möstl, E., Müller, J. & Aurich, C. 2010a. Changes in cortisol release and heart rate and heart rate variability during the initial training of 3-year-old sport horses. *Hormones and behavior*, 58(4), pp 628-636.
- Schmidt, A., Hödl, S., Möstl, E., Aurich, J., Müller, J. & Aurich, C. 2010b. Cortisol release, heart rate, and heart rate variability in transport-naïve horses during repeated road transport. *Domestic animal endocrinology*, 39(3), pp 205-213.

- Schmidt, A., Möstl, E., Wehnert, C., Aurich, J., Müller, J. & Aurich, C. 2010c. Cortisol release and heart rate variability in horses during road transport. *Hormones and behavior*, 57(2), pp 209-215.
- Schumacher, J., Jamra, R. A., Becker, T., Ohlraun, S., Klopp, N., Binder, E. B., Schulze, T. G., Deschner, M., Schmä, C., Höfels, S., Zobel, A., Illig, T., Propping, P., Holsboer, F., Rietschel, M., Nöthen, M. M. & Cichon, S. 2005. Evidence for a relationship between genetic variants at the brain-derived neurotrophic factor (BDNF) locus and major depression. *Biol Psychiatry*, 58(4), pp 307-14.
- Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T. & Oitzl, M. S. 2012. Stress effects on memory: An update and integration. *Neuroscience & Biobehavioral Reviews*, 36(7), pp 1740-1749.
- 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, pp e582759, <https://doi.org/10.3389/fvets.2020.582759>.
- Scopa, C., Contalbrigo, L., Greco, A., Lanatà, A., Scilingo, E. P. & Baragli, P. 2019. Emotional transfer in human–horse interaction: new perspectives on equine assisted interventions. *Animals*, 9(12), pp 1030.
- Scopa, C., Palagi, E., Sighieri, C. & Baragli, P. 2018. Physiological outcomes of calming behaviors support the resilience hypothesis in horses. *Scientific Reports*, 8(1), pp 17501.
- Selye, H. 1973. The Evolution of the Stress Concept: The originator of the concept traces its development from the discovery in 1936 of the alarm reaction to modern therapeutic applications of syntoxic and catatoxic hormones. *American scientist*, 61(6), pp 692-699.
- Shaffer, F. & Ginsberg, J. P. 2017. An Overview of Heart Rate Variability Metrics and Norms. *Front Public Health*, 5, pp 258.
- Shamay-Tsoory, S. G., Lavidor, M. & Aharon-Peretz, J. 2008. Social learning modulates the lateralization of emotional valence. *Brain and Cognition*, 67(3), pp 280-291.
- Shariff, N. 2015. Utilizing the Delphi survey approach: A review. *J Nurs Care*, 4(3), pp 246.
- Shettleworth, S. J. 2009. *Cognition, evolution, and behavior*: Oxford university press.

- Shirtcliff, E. A., Buck, R. L., Laughlin, M. J., Hart, T., Cole, C. R. & Slowey, P. D. 2015. Salivary cortisol results obtainable within minutes of sample collection correspond with traditional immunoassays. *Clin Ther*, 37(3), pp 505-14.
- Sigurjónsdóttir, H., Van Dierendonck, M. C., Snorrason, S. & Thórhallsdóttir, A. G. 2003. Social relationships in a group of horses without a mature stallion. *Behaviour*, 140, pp 783-804.
- Sih, A., Mathot, K. J., Moiron, M., Montiglio, P.-O., Wolf, M. & Dingemanse, N. J. 2015. Animal personality and state-behaviour feedbacks: a review and guide for empiricists. *Trends in ecology & evolution*, 30(1), pp 50-60.
- Sih, A. 2011. Behavioral syndromes: A behavioral ecologist's view on the evolutionary and ecological implications of animal personalities. *Personality and temperament in nonhuman primates*, pp 313-336.
- Sikorska, U., Maśko, M., Ciesielska, A., Zdrojkowski, Ł. & Domino, M. 2023. Role of Cortisol in Horse's Welfare and Health. *Agriculture*, 13(12), pp 2219.
- Singh, A. S. & Masuku, M. B. 2014. Sampling techniques & determination of sample size in applied statistics research: An overview. *International Journal of economics, commerce and management*, 2(11), pp 1-22.
- Siniscalchi, M., d'Ingeo, S. & Quaranta, A. 2021. Lateralized emotional functioning in domestic animals. *Applied Animal Behaviour Science*, 237, pp e105282, <https://doi.org/10.1016/j.applanim.2021.105282>.
- Smith, A. V., Proops, L., Grounds, K., Wathan, J., Scott, S. K. & McComb, K. 2018. Domestic horses (*Equus caballus*) discriminate between negative and positive human nonverbal vocalisations. *Scientific Reports*, 8(1), pp e13052, <https://doi.org/10.1038/s41598-018-30777-z>.
- Smith, G. T., McCarthy, D. M. & Anderson, K. G. 2000. On the Sins of Short-Form Development. *Psychological Assessment*, 12(1), pp 102-111.
- Smithson, J. 2008. *Focus groups*. In Alasuutari, P., Bickman, L., Brannen, J. (Eds), *The Sage handbook of social research methods*, pp 357-370, London: Sage Publications Ltd, <https://doi.org/10.4135/9781446212165>.
- Smyth, J. D., Dillman, D. A., Christian, L. M. & Stern, M. J. 2006. Comparing check-all and forced-choice question formats in web surveys. *Public opinion quarterly*, 70(1), pp 66-77.

- Snowdon, C. T. 2003. Expression of emotion in nonhuman animals. *In: Davidson, R. J., Scherer, K. R. & Goldsmith, H. H. (eds.) Handbook of affective sciences*. Oxford: Oxford University Press.
- Søndergaard, E. & Halekoh, U. 2003. Young horses' reactions to humans in relation to handling and social environment. *Applied Animal Behaviour Science*, 84(4), pp 265-280.
- Sorg, B. A. & Whitney, P. 1992. The effect of trait anxiety and situational stress on working memory capacity. *Journal of research in personality*, 26(3), pp 235-241.
- Sortheix, F. M. & Weber, W. 2023. Comparability and reliability of the positive and negative affect scales in the European Social Survey. *Front Psychol*, 14, pp e1034423.
- Spinka, M. 2012. Social dimension of emotions and its implication for animal welfare. *Applied Animal Behaviour Science*, 138, pp 170-181.
- Spruijt, B. M., van den Bos, R. & Pijlman, F. T. A. 2001. A concept of welfare based on reward evaluating mechanisms in the brain: anticipatory behaviour as an indicator for the state of reward systems. *Applied Animal Behaviour Science*, 72(2), pp 145-171.
- Stamps, J. & Groothuis, T. G. G. 2010. The development of animal personality: relevance, concepts and perspectives. *Biological Reviews*, 85(2), pp 301-325.
- Starling, M. J., Branson, N., Cody, D. & McGreevy, P. D. 2013. Conceptualising the Impact of Arousal and Affective State on Training Outcomes of Operant Conditioning. . *Animals*, 11(3), pp 300-317.
- Stevenson-Hinde, J. & Hinde, C. A. 2011. Individual characteristics: Weaving psychological and ethological approaches. *Personality and temperament in nonhuman primates*. New York, NY, US: Springer Science + Business Media.
- Stevenson-Hinde, J. & Zunz, M. 1978. Subjective assessment of individual rhesus monkeys. . *Primates*, 19, pp 473-482.
- Stewart, M., Stafford, K. J., Dowling, S. K., Schaefer, A. L. & Webster, J. R. 2008. Eye temperature and heart rate variability of calves disbudded with or without local anaesthetic. *Physiol Behav*, 93(4-5), pp 789-97.

- Stomp, M., D'ingeo, S., Henry, S., Cousillas, H. & Hausberger, M. 2021. Brain activity reflects (chronic) welfare state: Evidence from individual electroencephalography profiles in an animal model. *Applied Animal Behaviour Science*, 236, pp e105271.
- Stomp, M., Leroux, M., Cellier, M., Henry, S., Lemasson, A. & Hausberger, M. 2018. An unexpected acoustic indicator of positive emotions in horses. *PloS one*, 13(7), pp e0197898.
- Storbeck, J. & Clore, G. L. 2007. On the interdependence of cognition and emotion. *Cognition and Emotion*, 21(6), pp 1212-1237.
- Storm, C. & Storm, T. 1987. A taxonomic study of the vocabulary of emotions. *Journal of Personality and Social Psychology*, 53, pp 805-816.
- Strand, S. C., Tiefenbacher, S., Haskell, M. J., Hosmer, T., McDonnell, S. M. & Freeman, D. A. 2002. Behavior and physiologic responses of mares to short-term isolation. *Applied Animal Behaviour Science*, 78(2-4), pp 145-157.
- Stratton, R. 2022. *Assessment of positive emotion in horses*. PhD, Massey university, University of New Zealand, available <https://mro.massey.ac.nz/items/333c4d5b-718d-47a8-bd8e-2f48ddc4fcb2>.
- Stubbs. 2024a. *Jumbo hook on feeder* [Online]. Stubbs England. Available: https://www.stubbsengland.com/index.php?route=product/product&path=0_63&product_id=670.
- Stubbs. 2024b. *Plastic feed scoop* [Online]. Stubbs England. Available: https://www.stubbsengland.com/index.php?route=product/product&product_id=431#:~:text=The%20recognised%20measure%20of%20the,supplied%20by%20Connolly's%20RED%20MILLS.
- Stucke, D., Ruse, M. G. & Lebelt, D. 2015. Measuring heart rate variability in horses to investigate the autonomic nervous system activity—Pros and cons of different methods. *Applied animal behaviour science*, 166, pp 1-10.
- Stull, C. L. & Rodiek, A. V. 2002. Effects of cross-tying horses during 24 h of road transport. *Equine Veterinary Journal*, 34(6), pp 550-555.
- Suwała, M., Górecka-Bruzda, A., Walczak, M., Ensminger, J. & Jezierski, T. 2016. A desired profile of horse personality—A survey study of Polish equestrians based on a new approach to equine temperament and character. *Applied animal behaviour science*, 180, pp 65-77.

- Swanson, L. W. 2012. *Brain architecture: understanding the basic plan*: Oxford University Press, USA.
- Swedberg, R. 2020. Exploratory research. *The production of knowledge: Enhancing progress in social science*, 2(1), pp 17-41.
- Taghipoorreyneh, M. 2023. Mixed methods and the Delphi method. In: Tierney, R. J., Rizvi, F. & Ercikan, K. (eds.) *International Encyclopedia of Education (Fourth Edition)*. Oxford: Elsevier.
- Tariq, M. U. 2015. Hypothetico-deductive method: A comparative analysis. *Journal of Basic and Applied Research International*, 7(4), pp 228-231.
- Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-aho, P. O. & Karjalainen, P. A. 2014. Kubios HRV – Heart rate variability analysis software. *Computer Methods and Programs in Biomedicine*, 113(1), pp 210-220.
- Tay, S. C. 2011. *The psychometric principles of affect: are they ideal?*. Doctor of Philosophy, University of Illinois at Urbana-Champaign, available https://www.researchgate.net/publication/277121116_The_psychometric_principles_of_affect_are_they_ideal.
- Teddle, C. & Tashakkori, A. 2011. Mixed methods research. *The Sage handbook of qualitative research*, 4(285-300).
- TeledyneFlir. 2024. FLIR Ignite. Teledyne Flir, available <https://www.flir.co.uk/>.
- Temple, D., Manteca, X., Velarde, A. & Dalmau, A. 2011. Assessment of animal welfare through behavioural parameters in Iberian pigs in intensive and extensive conditions. *Applied Animal Behaviour Science*, 131(1-2), pp 29-39.
- Ter Woort, F., Dubois, G., Didier, M. & Van Erck-Westergren, E. 2021. Validation of an equine fitness tracker: heart rate and heart rate variability. *Comparative Exercise Physiology*, 17(2), pp 189-198.
- Thangaratnam, S. & Redman, C. W. E. 2005. The delphi technique. *The obstetrician & gynaecologist*, 7(2), pp 120-125.
- Tharwat, A., Gaber, T., Ibrahim, A. & Hassanien, A. E. 2017. Linear discriminant analysis: A detailed tutorial. *AI communications*, 30(2), pp 169-190.

- The Brooke. 2022. *Gaps in population data put working horses and donkeys at threat*. [Online]. <https://www.thebrooke.org/news/gaps-population-data-put-working-horses-and-donkeys-threat#:~:text=There%20are%20an%20estimated%20116,world's%20donkeys%20residing%20in%20LMICs>. [Accessed 02.10.2022 2022].
- Thorbergson, Z. W., Nielsen, S. G., Beaulieu, R. J. & Doyle, R. E. 2016. Physiological and behavioral responses of horses to wither scratching and patting the neck when under saddle. *Journal of Applied Animal Welfare Science*, 19(3), pp 245-259.
- Toepoel, V. & Funke, F. 2018. Sliders, visual analogue scales, or buttons: Influence of formats and scales in mobile and desktop surveys. *Mathematical Population Studies*, 25(2), pp 112-122.
- Tomberg, C., Petagna, M. & de Selliers de Moranville, L.-A. 2023. Horses (*Equus caballus*) facial micro-expressions: insight into discreet social information. *Scientific Reports*, 13(1), pp 8625.
- Torcivia, C. & McDonnell, S. M. 2021. Equine Discomfort Ethogram. *Animals*, 11(2), pp 580, <https://doi.org/10.3390/ani11020580>.
- 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, pp 80-87.
- Travain, T. & Valsecchi, P. 2021. Infrared Thermography in the Study of Animals' Emotional Responses: A Critical Review. *Animals*, 11(9), pp 2510.
- Trevethan, R. 2017. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. *Frontiers in public health*, 5, pp e308890.
- Trondrud, L. M., Pigeon, G., Albon, S., Arnold, W., Evans, A. L., Irvine, R. J., Król, E., Ropstad, E., Stien, A. & Veiberg, V. 2021. Determinants of heart rate in Svalbard reindeer reveal mechanisms of seasonal energy management. *Philosophical Transactions of the Royal Society B*, 376(1831), pp e20200215.
- 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), pp 643-653.

- Tulppo, M. P., Makikallio, T. H., Seppänen, T., Laukkanen, R. T. & Huikuri, H. V. 1998. Vagal modulation of heart rate during exercise: effects of age and physical fitness. *American Journal of Physiology-Heart and Circulatory Physiology*, 274(2), pp H424-H429.
- Tupes, P.D. & Christal, R.E. 1961. Recurrent Personality Factors Based on Trait Ratings. Technical Report ASD-TR-61-97, Texas: United States Air Force Personnel Department, available <https://apps.dtic.mil/sti/tr/pdf/AD0267778.pdf>.
- Tupes, P. D. & Christal, R. E. 1992. Recurrent personality factors based on trait ratings. *Journal of Personality*, 60(2), pp 225-51, <https://doi.org/10.1111/j.1467-6494.1992.tb00973.x>.
- Turkheimer, E., Pettersson, E. & Horn, E. E. 2014. A phenotypic null hypothesis for the genetics of personality. *Annual review of psychology*, 65, pp 515-540.
- Uddin, J., Phillips, C. J., Goma, A. A. & McNeill, D. M. 2019. Relationships between infrared temperature and laterality. *Applied Animal Behaviour Science*, 220, pp e104855, <https://doi.org/10.1016/j.applanim.2019.104855>.
- Uher, J., Asendorpf, J. B. & Call, J. 2008. Personality in the behaviour of great apes: temporal stability, cross-situational consistency and coherence in response. *Animal Behaviour*, 75(1), pp 99-112.
- Urbaniak, G. C. & Plous, S. 2013. Research randomizer. 4 ed, available <https://randomizer.org/>.
- Vaidya, J. G., Gray, E. K., Haig, J. & Watson, D. 2002. On the temporal stability of personality: evidence for differential stability and the role of life experiences. *Journal of personality and social psychology*, 83(6), pp 1469.
- Valenchon, M., Lévy, F., Moussu, C. & Lansade, L. 2017. Stress affects instrumental learning based on positive or negative reinforcement in interaction with personality in domestic horses. *PloS one*, 12(5), pp e0170783.
- Valenchon, M., Lévy, F., Fortin, M., Leterrier, C. & Lansade, L. 2013. Stress and temperament affect working memory performance for disappearing food in horses, *Equus caballus*. *Animal Behaviour*, 86(6), pp 1233-1240.
- van Ast, V. A., Cornelisse, S., Marin, M.-F., Ackermann, S., Garfinkel, S. N. & Abercrombie, H. C. 2013. Modulatory mechanisms of cortisol effects on emotional learning and memory: Novel perspectives. *Psychoneuroendocrinology*, 38(9), pp 1874-1882.

- Van Den Berg, M., Giagos, V., Lee, C., Brown, W. Y. & Hinch, G. N. 2016. Acceptance of novel food by horses: The influence of food cues and nutrient composition. *Applied Animal Behaviour Science*, 183, pp 59-67.
- van Oldruitenborgh-Oosterbaan, M. M. S., Van den Hoven, R. & Breukink, H. J. 1988. The accuracy of three different heart rate meters used for studies in the exercising horse. *Zentralblatt fur Veterinarmedizin. Reihe A*, 35(9), pp 665-672.
- VanDierendonck, M. C. & Spruijt, B. M. 2012. Coping in groups of domestic horses – Review from a social and neurobiological perspective. *Applied Animal Behaviour Science*, 138(3), pp 194-202.
- Vasconcellos, A. d. S., Viranyi, Z., Range, F., Ades, C., Scheidegger, J. K., Moestl, E. & Kotrschal, K. 2016. Training reduces stress in human-socialised wolves to the same degree as in dogs. *PLoS One*, 11(9), pp e0162389.
- Vazire, S., Gosling, S. D., Dickey, A. S. & Schapiro, S. J. 2007. Measuring personality in nonhuman animals. *Handbook of research methods in personality psychology*, pp 190-206.
- Verma, J. P. 2015. *Repeated measures design for empirical researchers*: Oxford: John Wiley & Sons.
- Villas-Boas, J. D., Dias, D. P. M., Trigo, P. I., A.N.A., d. S., de Almeida, F. Q. & de Medeiros, M. A. 2016. Behavioural, endocrine and cardiac autonomic responses to a model of startle in horses. *Applied Animal Behaviour Science*, 174, pp 76-82.
- Vinassa, M., Cavallini, D., Galaverna, D., Baragli, P., Raspa, F., Nery, J. & Valle, E. 2020. Palatability assessment in horses in relation to lateralization and temperament. *Applied Animal Behaviour Science*, 232, pp e105110.
- Visser, E., Van Reenen, C., Engel, B., Schilder, M., Barneveld, A. & Blokhuis, H. 2003. The association between performance in show-jumping and personality traits earlier in life. *Applied Animal Behaviour Science*, 82(4), pp 279-295.
- 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), pp 521-533.
- Visser, E. K., Van Reenen, C. G., Van der Werf, J. T. N., Schilder, M. B. H., Knaap, J. H., Barneveld, A. & Blokhuis, H. J. 2002. Heart rate and heart rate variability during a novel

- object test and a handling test in young horses. *Physiology & behavior*, 76(2), pp 289-296.
- Visser, E. K., Van Reenen, C. G., Hopster, H., Schilder, M. B. H., Knaap, J. H., Barneveld, A. & Blokhuis, H. J. 2001. Quantifying aspects of young horses' temperament: consistency of behavioural variables. *Applied animal behaviour science*, 74(4), pp 241-258.
- Volkow, N. D., Fowler, J. S., Wang, G.-J. & Goldstein, R. Z. 2002. Role of dopamine, the frontal cortex and memory circuits in drug addiction: insight from imaging studies. *Neurobiology of learning and memory*, 78(3), pp 610-624.
- Von Borell, E. 1995. Neuroendocrine integration of stress and significance of stress for the performance of farm animals. *Applied Animal Behaviour Science*, 44(2-4), pp 219-227.
- Von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant, J., Marchant-Forde, R., Minero, M., Mohr, E. & Prunier, A. 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), pp 293-316.
- von Borstel, U. K., Visser, E. K. & Hall, C. 2017. Indicators of stress in equitation. *Applied Animal Behaviour Science*, 190, pp 43-56.
- von Lewinski, M., Biau, S., Erber, R., Ille, N., Aurich, J., Faure, J.-M., Möstl, E. & Aurich, C. 2013. Cortisol release, heart rate and heart rate variability in the horse and its rider: Different responses to training and performance. *The Veterinary Journal*, 197(2), pp 229-232.
- Voudouris, K., Crosby, M., Beyret, B., Hernández-Orallo, J., Shanahan, M., Halina, M. & Cheke, L. G. 2022. Direct human-AI comparison in the animal-AI environment. *Frontiers in Psychology*, 13, pp e711821.
- Wagner, H. L. 1990. The spontaneous facial expression of differential positive and negative emotions. *Motivation and Emotion*, 14(1), pp 27-43.
- Waiyavutti, C., Johnson, W. & Deary, I. J. 2012. Do personality scale items function differently in people with high and low IQ? *Psychological assessment*, 24(3), pp 545.
- Walker, D. R., Walker, J. C., Cavnar, P. J., Taylor, J. L., Pickel, D. H., Hall, S. B. & Suarez, J. C. 2006. Naturalistic quantification of canine olfactory sensitivity. *Applied Animal Behaviour Science*, 97(2), pp 241-254.

- Waran, N. & Randle, H. 2017. What we can measure, we can manage: The importance of using robust welfare indicators in Equitation Science. *Applied Animal Behaviour Science*, 190, pp 74-81.
- Waring, G. H. 1983. *Horse behavior: The behavioral traits and adaptations of domestic and wild horses, including ponies*. Park Ridge, NJ: Noyes Publications.
- Wascher, C. A., Kotrschal, K. & Arnold, W. 2018. Free-living greylag geese adjust their heart rates and body core temperatures to season and reproductive context. *Scientific reports*, 8(1), pp 2142.
- Wascher, C. A. F. 2021. Heart rate as a measure of emotional arousal in evolutionary biology. *Philos Trans R Soc Lond B Biol Sci*, 376(1831), pp e20200479.
- Wathan, J., Burrows, A., Waller, B. M. & McComb, K. 2015. EquiFACS: The equine facial action coding system. *PLoS one*, 10(8), pp e0131738.
- Wathan, J., Proops, L., Grounds, K. & McComb, K. 2016. Horses discriminate between facial expressions of conspecifics. *Scientific Reports*, 6(1), pp e38322.
- Watrin, L., Geiger, M., Spengler, M. & Wilhelm, O. 2019. Forced-choice versus Likert responses on an occupational big five questionnaire. *Journal of Individual Differences*, 40(3), pp 134-148, <https://psycnet.apa.org/doi/10.1027/1614-0001/a000285>.
- Watson, D., Clark, L. A. & Tellegen, A. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. 54, pp 1063-1070.
- Watson, J. C. & McDonnell, S. M. 2018. Effects of three non-confrontational handling techniques on the behavior of horses during a simulated mildly aversive veterinary procedure. *Applied Animal Behaviour Science*, 203, pp 19-23.
- Watson, D., Wiese, D., Vaidya, J. & Tellegen, A. 1999. The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of personality and social psychology*, 76(5), pp 820.
- Webster, J. 2005. *Animal Welfare: limping towards eden: A practical approach to redressing the problem of our dominion over the animals*, Oxford: Blackwell Publishing.
- Weiss, A. 2018. Personality traits: A view from the animal kingdom. *Journal of Personality*, 86(1), pp 12-22.

- Weiss, A., Adams, M. J., Widdig, A. & Gerald, M. S. 2011. Rhesus macaques (*Macaca mulatta*) as living fossils of hominoid personality and subjective well-being. *Journal of Comparative Psychology*, 125, pp 72-83.
- Weiss, A., Inoue-Murayama, M., Hong, K. W., Inoue, E., Usono, T., Ochiai, T., Matsuzawa, T., Hirata, S. & King, J. E. 2009. Assessing chimpanzee personality and subjective well-being in Japan. *Am J Primatol*, 71(4), pp 283-92.
- Weiss, A., King, J. E. & Hopkins, W. D. 2007. A cross-setting study of chimpanzee (*Pan troglodytes*) personality structure and development: zoological parks and Yerkes National Primate Research Center. *American Journal of Primatology*, 69(11), pp 1264-1277.
- Weiss, J. M. Influence of psychological variables on stress-induced pathology. 1972. Ciba Foundation Symposium 8-Physiology, Emotion and Psychosomatic Illness. Wiley Online Library, pp 253-279.
- Wemelsfelder, F., Hunter, E. A., Mendl, M. T. & Lawrence, A. B. 2000. The spontaneous qualitative assessment of behavioural expressions in pigs: first explorations of a novel methodology for integrative animal welfare measurement. *Applied animal behaviour science*, 67(3), pp 193-215.
- Wemelsfelder, F., Hunter, T. E. A., Mendl, M. T. & Lawrence, A. B. 2001. Assessing the 'whole animal': a free choice profiling approach. *Animal Behaviour*, 62(2), pp 209-220.
- Whishaw, I. Q., Sacrey, L.-A. R. & Gorny, B. 2009. Hind limb stepping over obstacles in the horse guided by place-object memory. *Behavioural brain research*, 198(2), pp 372-379.
- Whitacre, J. & Bender, A. 2010. Degeneracy: a design principle for achieving robustness and evolvability. *Journal of theoretical biology*, 263(1), pp 143-153.
- Whittaker, A. L. & Marsh, L. E. 2019. The role of behavioural assessment in determining 'positive' affective states in animals. *CABI Reviews*, 14(10), pp 1-13.
- Wiepkema, P. R. & Koolhaas, J. M. 1993. Stress and animal welfare. *Animal welfare*, 2(3), pp 195-218.
- 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), pp 630-638.

- Williams, I., Harlock, J., Robert, G., Mannion, R., Brearley, S. G. & Hall, K. 2017a. Findings from the focus groups and Delphi study involving citizens, patient/service user representatives, carers and representatives from third-sector organisations and community groups. *Decommissioning health care: identifying best practice through primary and secondary research—a prospective mixed-methods study*. NIHR Journals Library, available <https://europepmc.org/article/NBK/nbk441776>.
- Williams, I., Hoppitt, W. & Grant, R. 2017b. The effect of auditory enrichment, rearing method and social environment on the behavior of zoo-housed psittacines (Aves: Psittaciformes); implications for welfare. *Applied Animal Behaviour Science*, 186, pp 85-92.
- Williams, J. M., Berg, L. C., Clayton, H. M., Kirsch, K., Marlin, D. J., Randle, H., Roepstroff, L., Oldruitenborgh-Oosterbaan, M. S., Weishaupt, M. A. & Munsters, C. 2023. A Delphi Study to Determine International and National Equestrian Expert Opinions on Domains and Sub-Domains Essential to Managing Sporthorse Health and Welfare in the Olympic Disciplines. *Animals*, 13(21), pp 3404.
- 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), pp 625-640.
- Wolff, A. & Hausberger, M. 1996. Learning and memorisation of two different tasks in horses: the effects of age, sex and sire. *Applied Animal Behaviour Science*, 46(3), pp 137-143.
- Yarnell, K., Hall, C., Royle, C. & Walker, S. L. 2015. Domesticated horses differ in their behavioural and physiological responses to isolated and group housing. . *Physiology & Behavior*, 143, pp 51-57.
- Yeoman, S. B. A., Burford, J. H., England, G. C. W. & Freeman, S. L. 2017. Horse Owners' Opinion On The Recognition And Diagnosis Of Colic: Outcomes From A Delphi Consensus Process. . *Equine Veterinary Education*, 29, pp 33.
- Yik, M., Russell, J. A. & Steiger, J. H. 2011. A 12-point circumplex structure of core affect. *Emotion*, 11(4), pp 705.
- Yoon, J.-H., Kang, M.-Y., Jeung, D. & Chang, S.-J. 2017. Suppressing emotion and engaging with complaining customers at work related to experience of depression and anxiety symptoms: a nationwide cross-sectional study. *Industrial health*, 55(3), pp 265-274.

- Young, T. J. 2015. *Questionnaires and Surveys*. In Hua, Z (Ed), *Research Methods in Intercultural Communication: A Practical guide*. Oxford: John Wiley & Sons Inc.
- Zapf, D. 2002. Emotion work and psychological well-being: A review of the literature and some conceptual considerations. *Human resource management review*, 12(2), pp 237-268.
- Zelenski, J. M. & Larsen, R. J. 2000. The distribution of basic emotions in everyday life: a state and trait perspective from experience sampling data. *Journal of Research in Personality*, 34, pp 178-197.
- Zevon, M. A. & Tellegen, A. 1982. The structure of mood change: An idiographic/nomothetic analysis. 43, pp 111-122.
- Zhao, D., Wang, Y. & Li, B. 2020. Eye Preference for Emotional Stimuli in Sichuan Snub-Nosed Monkeys. . *Folia Primatologica*, 91(6), pp 630-642.
- Zippia. 2022. *Equestrian demographics and statistics in the US*. [Online]. Available: <https://www.zippia.com/equestrian-jobs/demographics/> [Accessed October 2022].
- Zuckerman, M. 1991. *Psychobiology of personality*, Cambridge: Cambridge University Press.
- Zuckerman, M. 1998. *Psychobiological theories of personality*. In: Barone, D.F., Hersen, M., Van Hasselt, V.B. (eds) *Advanced Personality. The Plenum Series in Social/Clinical Psychology*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-8580-4_6.

Appendix A – Chapter 3 - Delphi participant information and consent form

Delphi Study Online Framework

Aims

This Delphi study 'Equine personality and positive affective state' aims to achieve the following:

1. Capture the qualified opinions and experience of a range of equine experts internationally
2. Identify views around equine personality and specific personality traits
3. Identify views around positive affective state in equines
4. Identify views and experience around inducing positive affective state in equines
5. Identify views around measuring positive affective state in equines
6. Reach consensus agreement on the role of equine personality in affect and methods to induce and measure affect.

Objectives

The aims of this study will be satisfied via the following objectives:

1. Preparation of a 1st round study including multiple choice and ranking questions to identify the viewpoints of a range of experts
2. Preparation of a 1st round study including open text questions for participants to identify methods for discussion
3. Take identified methods forward to a 2nd round study and facilitate a ranking based discussion to identify consensus
4. Finalise consensus and obtain additional information from a small group of selected experts with equal weighting across all fields represented.

Methodology

This study will be conducted online using Qualtrics software and each participant will be provided with a link to join, password and joining instructions following giving informed consent to participate in the study.

The study is likely to be comprised of three rounds (however it will continue until such a time as the 70% consensus level is reached):

Round one - Participants will complete a Qualtrics survey which will gather individual data (no personal data will be recorded), individual views on equine personality and positive affect and views on methods to induce and measure positive affect. There will be a 4 week deadline placed on this round.

Round two - The data gathered during round one will be analysed and the results fed back to participants anonymously via a second Qualtrics survey. Participants will then be asked to

rank data in order of agreement to reach a baseline consensus. Again there will be a 4 week deadline placed on this round.

Round three - An expert panel will be convened with a selection of participants from a range of backgrounds in order to reach final consensus and inform further study. This will be facilitated via online video conferencing.

Further rounds - Further panel consensus meetings as required to reach consensus level.

Inclusion criteria

In order to be considered for conclusion in this study participants must satisfy at least one of the following criteria:

- Be an equine veterinary surgeon
- Be an equine behaviourist
- Be an equine trainer
- Be an academic with significant expertise in the field of equine behaviour, science or welfare
- Be an academic with significant experience in the field of personality or affective state in animals
- Be an accredited member of an equine or professional body such as BEVA, BHA, ABTC, RCVS, HEA, BHS, ABRS, UKCC, BVNA, APBC, ASAB, ISES, ABTC, ARCVS, BE, BD, BSJA, AAEP The Jockey Club.

And

- Have at least 5 years' experience in their field of expertise.

Exclusion criteria

The following exclusion criteria apply to selection of participants:

- Less than 5 years' experience in the field
- No formal equine or field related scientific or professional qualification
- Those who advocate for aversive training methods
- Those who have been retired from or not working in the field for more than 5 years
- Those who do not fully complete surveys
- Those below the age of 18

Consent form

Dear Participant,

Delphi Consultation on Equine Personality and Positive Affect

Please read through the information below and confirm your name, the date and sign your acceptance where indicated.

Following receipt of this form you will be provided with joining details for the Delphi consultation process.

Confidentiality and anonymity

If you consent to taking part in this study your name will not be disclosed during the Delphi rounds. Following the Delphi rounds there may be opportunity for expert panel discussion/workshop which will break this anonymity however by this point a high level of consensus will have been reached within the group, a decision on whether to be part of this panel is entirely up to you and you can withdraw at any point.

All information will be handled and stored in compliance with General Data Protection Regulations 2018.

Further information

Once you have agreed to participate in this study further information will be sent to you via email. If you have any questions or queries you can address these at any time to i.l.loftus@newcastle.ac.uk.

Consent statement

I have read this form, and the research study has been explained to me. I have been given the opportunity to ask questions and my questions have been answered. If I have additional questions, I have been informed whom to contact. I am over the age of 18 and have capacity to give informed consent. I agree to participate in the research study described above and will receive a countersigned copy of this consent form after I sign it.

Participant name and signature

Date

Name and signature of person obtaining consent

Date

Delphi on equine personality and positive affect

Q1 Dear Participant,

You have been invited to participate in this Delphi consultation as an expert in one or more of the areas within its remit.

This is round one of the consultation, please read each question carefully and answer in relation to your own knowledge and experience. Once all participants have completed round one, the data gathered will be analysed and fed into round two where upon you will receive an invitation to complete the next round. A smaller subset of participants will then be selected to complete round three and any subsequent rounds until a consensus of 70% is reached. The data generated from this consultation will be utilised to instruct the design of experimental studies to measure positive affective state in equines which will form the remainder of my PhD studies.

Your participation is very much valued as this consultation forms part of my PhD studies under the supervision of Dr Matthew Leach and Dr Lucy Asher at Newcastle University.

A glossary of terms can be found at [this link](#) which gives definitions of the technical terms used within this consultation. Please consult this document prior to answering the questions which follow.

Section one

In this section you will complete six demographic questions which are designed to be used for analysis purposes.

Q1 In which equine area(s) do you work? (Please select all that apply)

- ☐ Veterinary (1)
 - ☐ Behaviour (2)
 - ☐ Academia (3)
 - ☐ Training (4)
 - ☐ Charity (5)
 - ☐ Competition (6)
 - ☐ Rehabilitation (7)
 - ☐ Other (8) _____
-

Q2 How many years' experience do you have in your field of expertise?

5 12 18 25 31 38 44 51 57 64 70

Years experience ()	

Q3 Please select the age category you are in:

20 30 40 50 60 70 80 90 100

Age ()	

Q4 What is your gender?

☐ Male (1)

☐ Female (2)

☐ Other (4) _____

Q5 Which country do you currently work in?

Q6 Which qualifications do you possess? (Please select all that apply)

- ☐ Undergraduate degree at level 6 e.g. Bachelor of Science (1)
 - ☐ Postgraduate degree at level 7 e.g. Masters (2)
 - ☐ Postgraduate degree at level 8 e.g. PhD (3)
 - ☐ Veterinary license to practice (4)
 - ☐ Clinical animal behaviour accreditation (5)
 - ☐ No formal higher education qualification (6)
 - ☐ Professional qualification e.g. British Horse Society qualifications (7)
-

Section 2

This section contains questions around equine personality.

Q7 Do you believe that equines possess individual personality traits?

- ☐ Definitely yes (1)
 - ☐ Probably yes (2)
 - ☐ Do not know/Unsure (3)
 - ☐ Probably not (4)
 - ☐ Definitely not (5)
-

Q8 Are you aware of current research into equine personality?

- ☐ Yes and I read new research findings as they are published (1)
 - ☐ Yes, however I do not follow new research closely (2)
 - ☐ No, however I am aware it is an area that is being studied (3)
 - ☐ No, I was not aware it is an area under study (4)
-

Q9 Please rate your percentage agreement with the following definitions where 0% indicates no agreement and 100% indicates complete agreement.

0 10 20 30 40 50 60 70 80 90 100

Personality is the individual's general behaviour compared with those found in the general population. ()	
Temperament is the animal's behavioural response in a single context measured on some form of scale, for example flight distance. ()	
Behavioural syndromes describe variation within populations, not individuals. They describe the distribution of personalities within a population, for example horses which crib bite. ()	

Q10 Please indicate (by checking the boxes for each trait) if you feel the following personality trait terms are clear and unambiguous, present in equines and reliably identifiable in equines. Please refer to the trait definitions defined within [this](#) document when considering your answers.

	Clarity (1)	Present (2)	Reliably identifiable (3)
Fearful (1)	☐	☐	☐
Dominant (2)	☐	☐	☐
Persistent (3)	☐	☐	☐
Cautious (4)	☐	☐	☐
Stable (5)	☐	☐	☐
Autistic (6)	☐	☐	☐
Curious (7)	☐	☐	☐
Thoughtless (8)	☐	☐	☐
Greedy (9)	☐	☐	☐
Jealous (10)	☐	☐	☐
Individualistic (11)	☐	☐	☐
Reckless (12)	☐	☐	☐
Sociable (13)	☐	☐	☐
Distractable (14)	☐	☐	☐
Timid (15)	☐	☐	☐

Sympathetic (16)	☐	☐	☐
Playful (17)	☐	☐	☐
Solitary (18)	☐	☐	☐
Vulnerable (19)	☐	☐	☐
Innovative (20)	☐	☐	☐
Active (21)	☐	☐	☐
Helpful (22)	☐	☐	☐
Bullying (23)	☐	☐	☐
Aggressive (24)	☐	☐	☐
Manipulative (25)	☐	☐	☐
Gentle (26)	☐	☐	☐
Affectionate (27)	☐	☐	☐
Excitable (28)	☐	☐	☐
Impulsive (29)	☐	☐	☐
Inquisitive (30)	☐	☐	☐

Submissive (31)	☐	☐	☐
Cool (32)	☐	☐	☐
Dependent (33)	☐	☐	☐
Irritable (34)	☐	☐	☐
Unperceptive (35)	☐	☐	☐
Predictable (36)	☐	☐	☐
Decisive (37)	☐	☐	☐
Depressed (38)	☐	☐	☐
Conventional (39)	☐	☐	☐
Sensitive (40)	☐	☐	☐
Defiant (41)	☐	☐	☐
Intelligent (42)	☐	☐	☐
Protective (43)	☐	☐	☐
Quitting (44)	☐	☐	☐
Inventive (45)	☐	☐	☐

Clumsy (46)	☐	☐	☐
Erratic (47)	☐	☐	☐
Friendly (48)	☐	☐	☐
Anxious (49)	☐	☐	☐
Lazy (50)	☐	☐	☐
Disorganised (51)	☐	☐	☐
Unemotional (52)	☐	☐	☐
Imitative (53)	☐	☐	☐
Independent (54)	☐	☐	☐

Section 3

This section asks you questions relating to animal/equine emotions. It includes questions to gather your views on emotions in animals as well as specific questions regarding inducing and measuring affective state in equines.

Q11 Please rate your percentage agreement with the following statement where 0% indicates no agreement and 100% indicates complete agreement.

0 10 20 30 40 50 60 70 80 90 100

Following the recognition of animals as sentient beings their emotional welfare has become an important issue due to their capacity to experience affective experiences and thus experience stress and fear or conversely enjoy positive feelings. ()



Q12 When assessing emotional feelings (affect) please select the factor(s) you identify as important measures from the list below.

	Always important (1)	Sometimes important (2)	Not important (3)
Valence (feelings of goodness and badness i.e. positive and negative affects) (1)	☐	☐	☐
Arousal (intensity of feelings) (2)	☐	☐	☐
Surgency or power (how much does the feeling fill the mental life of the subject) (3)	☐	☐	☐

Q13 Please identify your level of agreement with the following statements.

	Strongly disagree (1)	Somewhat disagree (3)	Don't know/unsure (5)	Somewhat agree (6)	Strongly agree (7)
Emotional processes include motor-expressive, sensory-perceptual, autonomic-hormonal, cognitive-attentional and affective-feeling aspects (1)	☐	☐	☐	☐	☐
It is difficult to distinguish between the cognitive and affective aspects unambiguously when undertaking field-based research (2)	☐	☐	☐	☐	☐
Equines can experience negative affect (negative emotional state). (3)	☐	☐	☐	☐	☐
Equines can experience neutral affect (neutral emotional state). (4)	☐	☐	☐	☐	☐

Equines can experience positive affect (positive emotional state). (5)

☐☐☐☐☐

Q14 Do you think that an individual equine's personality influences the intensities of different emotional states within the individual? For example, a pessimistic personality is more likely to experience a greater intensity of negative emotional state and a lower intensity of positive emotional state.

- ☐ Definitely yes (1)
- ☐ Unsure (2)
- ☐ Definitely no (3)

Q15 In order to attempt to measure positive affective state it must be reliably induced. Please identify your level of agreement regarding the reliability of the following methods to induce a positive affective state in equines. Note: Please use the free text box in question 21 to provide additional suggestions of methods to induce positive affective state in equines.

	Strongly disagree (1)	Somewhat disagree (2)	Don't know/unsure (3)	Somewhat agree (4)	Strongly agree (5)
Provision of high value (as perceived by the individual animal) feed e.g. carrots, apples, concentrate feed such as coarse mix (1)	☐	☐	☐	☐	☐
Provision of medium to low value (as perceived by the individual animal) concentrate feed e.g. pony nuts, high fibre cubes (2)	☐	☐	☐	☐	☐
Provision of roughage feed e.g. hay, chaff, haulage (3)	☐	☐	☐	☐	☐
Wither scratching (also termed scritchng) (4)	☐	☐	☐	☐	☐
Base of tail/sides of tail scratching (5)	☐	☐	☐	☐	☐
Forehead rubbing (6)	☐	☐	☐	☐	☐
Removal of restraint (7)	☐	☐	☐	☐	☐

Addition of conspecific companionship (8)	☐	☐	☐	☐	☐
Neck scratching (9)	☐	☐	☐	☐	☐
Verbal praise (10)	☐	☐	☐	☐	☐
Turnout in field (11)	☐	☐	☐	☐	☐
Access to graze and browse (12)	☐	☐	☐	☐	☐

Q16 Please identify in the free text box below any methods that you feel can induce positive affect in equines that have not been described in question 20 above.

Q17 Do you consider a multi-measure approach (which assesses behavioural and physiological indices) is optimal for measuring affective state in equines?

- ☐ Yes (1)
- ☐ Don't know/unsure (2)
- ☐ No (3)

Q18 Please drag and drop to rank the following behavioural measures for positive affective state in equines in order of preference.

- _____ Stereotypy performance (1)
 - _____ Free choice profiling (2)
 - _____ Cognitive bias testing (3)
 - _____ Attention bias testing (4)
 - _____ Novel object tests (5)
 - _____ Facial action scoring (6)
 - _____ Human-animal relationship assessment via avoidance distance, voluntary approach, forced approach tests (7)
 - _____ Time budget analysis (8)
-

Q19 Please identify other behavioural measures you feel could be valuable in assessing positive affective state in equines.

Q20 Please drag and drop to rank the following physiological measures for positive affective state in equines in order of preference.

- _____ Infrared thermography (1)
 - _____ Heart rate (2)
 - _____ Heart rate variability (3)
 - _____ Blood cortisol (4)
 - _____ Hair cortisol (5)
 - _____ Salivary cortisol (6)
 - _____ Urinary cortisol (7)
 - _____ Corticotropin releasing factor challenge (8)
 - _____ Electroencephalogram (9)
 - _____ Immune system function testing (immune status, IgA secretion and inflammatory marker identification) (10)
 - _____ Catecholamine levels (11)
 - _____ Serotonin levels (12)
 - _____ Noradrenalin levels (13)
 - _____ Body temperature (14)
-

Q21 Please identify other physiological measures you feel could be valuable in assessing positive affective state in equines.

Delphi Consultation Equine Positive Affective State Round Two

Q1 Dear Participant,

You have been invited to participate in this Delphi consultation as an expert in one or more of the areas within its remit.

This is round two of the consultation, it is shorter than round one as a number of aspects achieved consensus in round one (these have been brought for ratification in section 1 of this round). Those results not achieving consensus in round one, and suggestions made by participants in round one, have been brought forward into round two.

Please read each question carefully and answer in relation to your own knowledge and experience. Once all participants have completed round two, the data gathered will be analysed, a smaller subset of participants may then be selected to complete round three if required and any subsequent rounds until a consensus of 70% is reached. The data generated from this consultation will be utilised to instruct the design of experimental studies to measure positive affective state in equines which will form the remainder of the study.

Your participation is very much valued as this consultation forms part of my PhD studies under the supervision of Dr Matthew Leach and Professor Lucy Asher at Newcastle University.

A glossary of terms can be found at [this link](#) which gives definitions of the technical terms used within this consultation. Please consult this document prior to answering the questions which follow.

Q1 Please state your full name

Section 1

This section contains all questions and statements that achieved consensus in round 1. Please indicate that you agree with this consensus by selecting yes after each question. If you do not agree please select no and provide a written response explaining your reasons for disagreement. The glossary of terms can be found at [this link](#) which gives definitions of the technical terms used within this consultation.

1. It was agreed that equines possess individual personality traits

☐ Yes (1)

☐ No, please state why? (2)

2. It was agreed that personality is the individual's general behaviour compared with those found in the general population.

☐ Yes (1)

☐ No, please state why? (2)

3. It was agreed that 'Following the recognition of animals as sentient beings their emotional welfare has become an important issue due to their capacity to experience affective experiences and thus experience stress and fear or conversely enjoy positive feelings.'

☐ Yes (1)

☐ No, please state why? (2)

4. It was agreed that valence is always important when assessing affective state

☐ Yes (1)

☐ No, please state why? (2)

5. It was agreed that level of arousal is always important when assessing affective state

☐ Yes (1)

☐ No, please state why? (2)

6. It was agreed that 'Emotional processes include motor-expressive, sensory-perceptual, autonomic-hormonal, cognitive-attentional and affective-feeling aspects'

☐ Yes (1)

☐ No, please state why? (2)

7. It was agreed that 'It is difficult to distinguish between the cognitive and affective aspects unambiguously when undertaking field-based research'

☐ Yes (1)

☐ No, please state why? (2)

8. It was agreed that 'Equines can experience negative affect (negative emotional state)'.

☐ Yes (1)

☐ No, please state why? (2)

9. It was agreed that 'Equines can experience neutral affect (neutral emotional state)'.

☐ Yes (1)

☐ No, please state why? (2)

10. It was agreed that 'Equines can experience positive affect (positive emotional state)'.

☐ Yes (1)

☐ No, please state why? (2)

11. It was agreed that 'an individual equine's personality influences the intensities of different emotional states within the individual? For example, a pessimistic personality is more likely to experience a greater intensity of negative emotional state and a lower intensity of positive emotional state'.

☐ Yes (1)

☐ No, please state why? (2)

12. The reliability of methods to induce positive affective state were ranked and agreed as follows. Please indicate that you agree with this order by selecting 'Yes' below.

1. High value food provision
2. Wither scratching
3. Addition of a liked conspecific companion
4. Access to graze and browse
5. Medium value food provision
6. Provision of turnout
7. Provision of roughage (hay or similar)
8. Neck scratching
9. Scratching of the base or side of tail
10. Verbal praise

☐ Yes (1)

☐ No, please state why? (2)

13. It was agreed that 'a multi-measure approach which assesses both behavioural and physiological indices is optimal for measuring positive affective state'.

☐ Yes (1)

☐ No, please state why? (2)

14. The reliability of methods to measure positive affective state from a behavioural perspective were ranked and agreed as follows. Please indicate that you agree with this order.

1. Free choice profiling/Qualitative Behaviour Assessment
2. Facial action scoring
3. Cognitive bias testing
4. Human animal relationship tests
5. Attention bias testing
6. Novel object tests
7. Time budget analysis
8. Absence of stereotypy performance

☐ Yes (1)

☐ No, please state why? (2)

15. The reliability of methods to measure positive affective state from a physiological perspective were ranked and agreed as follows. Please indicate that you agree with this order.

1. Heart rate variability
2. Heart rate
3. Infra-red thermography
4. Salivary cortisol
5. Blood cortisol
6. Electroencephalogram
7. Hair cortisol
8. Serotonin levels

- 9. Urinary cortisol
- 10. Corticotrophin releasing factor challenge
- 11. Catecholamine levels
- 12. Immune function tests
- 13. Body temperature

☐ Yes (1)

☐ No , please state why? (2)

Section 2

This section contains questions around equine personality. Please rate your level of agreement with the statements and groupings presented and provide any feedback where indicated. The glossary of terms can be found at [this link](#) which gives definitions of the technical terms used within this consultation.

16. Please rate your percentage agreement with the following definition where 0% indicates no agreement and 100% indicates complete agreement.

0 10 20 30 40 50 60 70 80 90 100

Temperament is the animal's behavioural response in a single context measured on some form of scale, for example flight distance. ()	
Behavioural syndromes describe variation within populations, not individuals. They describe the distribution of personalities within a population, for example horses which crib bite. ()	

17. Following analysis of the results relating to the 54 personality traits gathered in the previous round and within a concurrent field experiment, the following personality groupings have been extracted via Principal Component Analysis. For each grouping please rank your level of agreement where 0% indicates no agreement and 100% indicates complete agreement.

0 10 20 30 40 50 60 70 80 90 100

DEMOTIVATED - includes traits: Lazy, depressed, unemotional and quitting ()	
STEADY - includes traits: Gentle & predictable ()	
CONTROLLING - includes traits: Dominant, aggressive, independent, thoughtless and greedy ()	
AFFILIATIVE - includes traits: Friendly, helpful, stable, sociable, sympathetic ()	
ANTAGONISTIC - includes traits: Irritable, defiant, unperceptive and bullying ()	
LONER - includes traits: Timid and solitary ()	
NERVOUS - includes traits: Cautious, sensitive, submissive, fearful, imitative and erratic ()	
PLAYFUL - includes traits: Excitable, playful, inquisitive, active and erratic ()	
COERCIVE - includes traits: Individualistic, manipulative, cool, jealous, decisive, reckless and impulsive, disorganised ()	
SUPPORTIVE - includes traits: Conventional, persistent and affectionate ()	
HESITANT - includes traits: Vulnerable, protective, curious ()	
GOOFY - includes traits: Clumsy and distractible ()	
INTELLECTUAL - includes traits: Intelligent, innovative, autistic and inventive ()	

18. Please suggest any revisions to the personality group names presented in the question above, explaining why you feel these names may be more suitable.

19. If you disagree with the groupings identified in the personality groupings, please indicate your reasons below.

Section 3 This sections asks you questions relating to equine emotions. It includes questions to gather your views on emotions in equines as well as specific questions regarding inducing and measuring their affective state. The glossary of terms can be found at [this link](#) which gives definitions of the technical terms used within this consultation.

20. When assessing emotional feelings (affect) the surgency or power of the feeling (how much the feeling fills the mental life of the subject) is

- ☐ Important (1)
- ☐ Neither important nor unimportant (4)
- ☐ Unimportant (2)

21. In order to attempt to measure positive affective state it must be reliably induced. Please drag and drop to rank the suitability of the following methods, suggested by respondents in the last round, to induce a positive affective state in equines. 1 is high and 9 is low.

- _____ Human-horse interaction through training (e.g. clicker training) (1)
- _____ Access to a substrate to roll in (e.g. sand) (2)
- _____ Grooming (3)
- _____ Equine massage (4)
- _____ Problem solving activities (5)
- _____ Food enrichment toys (6)
- _____ Aromatherapy (7)
- _____ Play toys (e.g. Horseball) (8)
- _____ Introduction of a non-frightening novel object (9)

22. Please drag and drop to rank the following behavioural measures, identified by respondents in the last round, for positive affective state in equines in order of preference. 1 is high and 12 is low.

- _____ Absence of frustration (1)
 - _____ Scoring of social behaviours/interactions (2)
 - _____ Scoring of play behaviours (3)
 - _____ Eye gaze analysis (4)
 - _____ Spontaneous blink rate analysis (5)
 - _____ Body language scoring (including muscle tone) (6)
 - _____ EBARQ (7)
 - _____ Time spent in lateral recumbency (8)
 - _____ Time spent in REM sleep (9)
 - _____ Assessment of vocalisations (10)
 - _____ Low facial grimace score (12)
 - _____ Biomechanical analysis (e.g. of step rate & frequency (use of accelerometers) and/or head and neck position) (13)
-

23. Please drag and drop to rank the following physiological measures, identified by respondents in the last round, for positive affective state in equines in order of preference. 1 is high and 11 is low.

- _____ Faecal cortisol (1)
 - _____ MRI scans (2)
 - _____ Blood oxytocin (3)
 - _____ Opioids and metabolites (4)
 - _____ Tryptophan (5)
 - _____ Respiratory rate (6)
 - _____ Skin conductivity/Sweat rate (7)
 - _____ Eye temperature (8)
 - _____ Computed tomography (9)
 - _____ Gut microflora assessment (10)
 - _____ Tympanic membrane temperature (11)
-

Delphi consultation: Equine personality and positive affect round 3

Dear Participant,

You have been invited to participate in this Delphi consultation as an expert in one or more of the areas within its remit.

This is the third and final round of the consultation, it is shorter than rounds one and two as a number of aspects achieved consensus in the previous rounds. Those aspects not achieving consensus in previous rounds, and suggestions made by participants in round two, have been brought forward into round three.

Please read each question carefully and answer in relation to your own knowledge and experience. Once all participants have completed round three, the data gathered will be analysed to identify which aspects have achieved the required 70% consensus agreement. The data generated from this consultation will be utilised to instruct the design of experimental studies to measure positive affective state in equines which will form the remainder of the study.

Your participation is very much valued as this consultation forms part of my PhD studies under the supervision of Dr Matthew Leach and Professor Lucy Asher at Newcastle University.

A glossary of terms can be found at [this link](#) which gives definitions of the technical terms used within this consultation. Please consult this document prior to answering the questions which follow.

Q1 Please state your full name

Section 1

This section contains all of the amendments made to personality groupings and traits following suggestions by participants in the previous round. Please state whether you agree with each grouping and the trait names within it.

Q1 Please select 'agree' or 'disagree' for each group name (given in capital letters) and the personality traits (given in lower case) placed within it.

	Agree (1)	Disagree (2)
EXUBERANT: Excitable, Playful, Active (1)	☐	☐
NERVOUS: Cautious, Erratic, Fearful, Acquiescent, Vulnerable, Sensitive (2)	☐	☐
UNRESPONSIVE: Lazy, Depressed, Unemotional, Quitting (3)	☐	☐
STEADY: Predictable, Stable, Conventional (4)	☐	☐
BOLD: Assertive, Confrontational, Independent (5)	☐	☐
AFFILIATIVE: Friendly, Helpful, Sociable, Sympathetic, Affectionate, Gentle, Protective (6)	☐	☐
HOSTILE: Irritable, Defiant, Unperceptive, Bullying (7)	☐	☐
SHY: Timid, Solitary (8)	☐	☐
DETERMINED: Decisive, Impulsive, Persistent (9)	☐	☐
ALOOF: Individualistic, Cool (10)	☐	☐
INATTENTIVE: Clumsy, Distractible (11)	☐	☐
INVESTIGATORY: Intelligent, Inventive, Innovative, Curious (12)	☐	☐

Q2 If you disagree with any of the groupings please state which one(s) and your rationale/suggestions here.

Section 2

This section contains questions around measures to induce positive affective state in horses.

Q3 Please state whether you agree that each of the measures listed could induce positive affective state in horses. <i>Please note that it is already established and accepted via previous rounds of this consultation that the intensity and valence of affective state induced will vary between individual horses dependent on their prior experiences, current environment, physiology, personality and individual predilections.</i>	Agree (1)	Disagree (2)	Don't know (4)
High value food provision fed by human (e.g. sweet mix, novel food items) (1)	☐	☐	☐
Medium value food provision fed by human (e.g. succulents, horse nuts) (2)	☐	☐	☐
Low value food provision fed by human (e.g. long stalk forage, high fibre cubes, chaff) (3)	☐	☐	☐
Wither scratching (human scratching of the wither area with fingers) (4)	☐	☐	☐

Addition of a known, affiliative conspecific (e.g. field mate, stable mate) (5)	☐	☐	☐
Access to graze and browse (in-hand) (6)	☐	☐	☐
Provision of turnout (loose in field) (7)	☐	☐	☐
Provision of loose roughage not hand fed by human (hay or similar) (8)	☐	☐	☐
Neck scratching (scratching of the neck area, not withers, by human) (9)	☐	☐	☐
Scratching of the base or sides of tail (delivered by human) (10)	☐	☐	☐
Verbal praise (e.g. Good, Yes) (11)	☐	☐	☐
Human-Horse interaction through positive training methods correctly undertaken (e.g. clicker training) (12)	☐	☐	☐
Access to a substrate to roll in at liberty (e.g. sand) (13)	☐	☐	☐
Grooming (standard grooming with appropriate brushes) (14)	☐	☐	☐
Equine massage (undertaken by a trained professional) (15)	☐	☐	☐

Problem solving activities (e.g. free work, scent work) (16)	☐	☐	☐
Food enrichment toys (e.g. handing fruit and vegetables, puzzle feeder) (17)	☐	☐	☐
Aromatherapy (such as lavender, provided in the environment) (18)	☐	☐	☐
Play toys (e.g. Horseball) (19)	☐	☐	☐
Introduction of a non-frightening novel object (placed in the environment for the horse to investigate at liberty) (20)	☐	☐	☐

Q4 Please make any comments that you would like to in relation to your answers to question 3.

Section 3 This section contains questions around methods to measure positive affective state in horses.

Q5 Please state whether, in your experience, you feel that each of the following measures **could** reliably be used to measure the behaviour of horses' indicative of positive emotional state. To do this, please state whether you think the measures should be included or excluded for the purposes of further field trials.

	Should be included (1)	Should be excluded (2)	Don't know (4)
Free choice profiling/Qualitative Behaviour Assessment (via one of two methods: observers are free to choose qualitative descriptors for an animal or a fixed terms approach using predefined descriptors for observers to select between) (1)	☐	☐	☐
Facial action scoring (scoring of facial action units such as orbital tightening, ear position, cheek muscle tension etc on an ordinal scale) (4)	☐	☐	☐
Cognitive bias testing [presenting the animal with cues that predict positive events and those that predict negative events as well as presenting ambiguous cues; then testing the animal's response to the ambiguous cues as to whether it is positive (optimistic) or negative (pessimistic)] (5)	☐	☐	☐
Human-animal relationship tests (such as avoidance distance and voluntary approach tests) (6)	☐	☐	☐

Attention bias testing (measuring the animal's attention between positive and negative stimuli) (7)	☐	☐	☐
Novel object tests (measuring the animal's response to the presentation of a novel object) (8)	☐	☐	☐
Time budget analysis (quantifying the behaviours exhibited by an animal during a specified period of time) (9)	☐	☐	☐
Absence of repetitive behaviour (stereotypy) performance (including crib biting, wind sucking, weaving, box walking, fence walking, headshaking) (10)	☐	☐	☐
Absence of frustration (identifying whether there are behaviours exhibited that could indicate frustration such as door kicking, pawing at the ground, negative attention to barriers) (11)	☐	☐	☐
Scoring of social behaviour/interactions (scoring both affiliative and agonistic interactions with conspecifics) (12)	☐	☐	☐
Scoring of play behaviours (13)	☐	☐	☐

Eye gaze analysis (also termed eye lateralisation, scoring which eye is used to focus on stimuli) (14)	☐	☐	☐
Spontaneous blink rate analysis (scoring of blink rate patterns and palpebral aperture features) (15)	☐	☐	☐
Body language scoring (ethograms) including muscle tone (16)	☐	☐	☐
EBARQ (Equine Behaviour Assessment and Research Questionnaire) (17)	☐	☐	☐
Time spent in lateral recumbency (18)	☐	☐	☐
Time spent in REM sleep (identified through rapid eye moved under closed eyes, loss of reflexes and loose muscle tone) (19)	☐	☐	☐
Assessment of vocalisations (identifying the frequency(Hz), patterns and time spend vocalising) (20)	☐	☐	☐
Low facial grimace score (using published Equine Facial Grimace Scoring technique to identify absence of pain) (21)	☐	☐	☐

Biomechanical analysis
(e.g. of step rate &
frequency, use of
accelerometers and/or
head and neck
position) (2)



Q6 Please make any comments that you would like to in relation to your answers to question 5.

Q7 Please state whether, in your experience, you feel that each of the following measures **could** reliably be used to measure the physiology of horses' indicative of positive emotional state. To do this, please state whether you think the measures should be included or excluded for the purposes of further field trials.

	Should be included (1)	Should be excluded (2)	Don't know (4)
Heart rate variability (assessing the balance between the sympathetic and parasympathetic nervous system via a heart rate monitor) (1)	☐	☐	☐
Heart rate (via a heart rate monitor) (2)	☐	☐	☐
Infra-red thermography (e.g. of the eye or ear using a thermographic camera) (4)	☐	☐	☐
Salivary cortisol (testing cortisol levels using saliva collected and run through a calibrated test system such as ELISA) (5)	☐	☐	☐
Blood cortisol (testing cortisol levels using blood samples collected and run through a calibrated test system such as ELISA) (6)	☐	☐	☐
Electroencephalogram (utilising resting or ambulatory EEG technology to assess brain activity) (7)	☐	☐	☐
Hair cortisol (testing long term cortisol levels using hair collected and run through a calibrated test system such as ELISA) (8)	☐	☐	☐

Serotonin levels
(testing serotonin
levels using blood
collected and run
through a calibrated
test system such as
ELISA) (9)



Urinary cortisol
(testing cortisol levels
using urine samples
collected and run
through a calibrated
test system such as
ELISA) (10)



Corticotrophin
releasing factor
challenge (assessing
physiological
response, via ACTH
and cortisol levels, to
intravenous infusion
with corticotrophin-
releasing hormone)
(11)



Catecholamine levels
(via blood tests for
adrenalin and
noradrenalin levels)
(12)



Immune function tests
(testing for
immunoglobulin
deficiencies/disorders)
(13)



Body temperature
(rectal temperature)
(14)



Faecal cortisol (testing cortisol levels using faecal samples collected and run through a calibrated test system such as ELISA) (15)	☐	☐	☐
Blood oxytocin (via blood samples) (16)	☐	☐	☐
Opioids and their metabolites (via testing of plasma samples by ELISA) (17)	☐	☐	☐
Tryptophan levels (measured in plasma via mass spectrometry) (18)	☐	☐	☐
MRI scans (magnetic resonance imaging) (19)	☐	☐	☐
Skin conductivity/sweat rate (measured using electrodes and sweat rate sensors on the skin) (20)	☐	☐	☐
Eye temperature (via thermography) (21)	☐	☐	☐
Computed tomography (a series of x-ray images taken at different angles and processed to create cross sectional images) (22)	☐	☐	☐
Gut microflora assessment (via faecal samples) (23)	☐	☐	☐

Tympanic membrane temperature
(measurement of the temperature of the tympanic membrane within the ear using a specialised device)
(24)

Respiratory rate (via counting the rise and fall of the flanks) (25)



Q8 Please make any comments that you would like to in relation to your answers to question 6.

Glossary of Terms

Affect - any experience of feeling or emotion, ranging from suffering to elation, from the simplest to the most complex sensations of feeling, and from the most normal to the most pathological emotional reactions.

Affective experiences – Affective (aka emotional) experiences are not always unequivocally positive or negative in nature. Occasionally, affective episodes may entail a blend of various affective states of opposing valence.

Affective-feeling – mental phenomenon of experiencing an emotion.

Arousal - Arousal is the physiological and psychological state of being awoken or of sense organs. It can also be defined as Extrinsic and Intrinsic processes responsible for monitoring, evaluating, and modifying emotional reactions.

Autonomic-hormonal – hormonal responses associated with the autonomic nervous system (control system that acts largely unconsciously and regulates bodily functions such as the heart rate, digestion, respiratory rate, pupillary response, urination, and sexual arousal), can be associated with dysfunction of this system.

Behavioural indices - Set of indices used to assess behaviour. "Indices" is originally a Latin plural, while "Indexes" has taken the English way of making plurals, using –s or –es. Though both are still widely used, they take on different usage in their senses. "Indices" is used when referring to mathematical, scientific and statistical contexts.

Behavioural syndrome - A behavioural syndrome is a suite of correlated behaviours expressed either within a given behavioural context (e.g., correlations between foraging behaviours in different habitats) or across different contexts (e.g., correlations among feeding, antipredator, mating, aggressive, and dispersal behaviours).

Cognitive-attentional – thoughts and emotion pattern which is persistent and often negative.

Emotional welfare - feeling happy and secure, especially because you are being properly looked after and cared for.

Fear - an unpleasant emotion caused by the threat of danger, pain, or harm.

Floor and ceiling effect - A floor effect is when measurements of the dependent variable (intensity of affect felt in this case) result in very low scores on the measurement scale. A ceiling effect is the opposite, where measurements of the dependent variable result in very high scores.

Induced – bring about or give rise to.

Intensity - the quality or state of being intense especially extreme degree of strength, force, energy, or feeling.

Motor-expressive – expressed via behavioural reactions.

Multi-measure approach - the use of multiple indicators and sources of evidence of varying kinds, often gathered at multiple points in time.

Negative affect - the experience of negative emotions and poor self-concept.

Neutral affect - feeling indifferent, nothing in particular, and a lack of preference one way or the other.

Personality - the combination of characteristics or qualities that form an individual's distinctive character.

Pessimistic – adjective: tending to see the worst aspect of things or believe that the worst will happen.

Physiological indices – Set of indices used to assess physiology.

Positive affect - characteristic that describes how much pain animal experiences positive sensations, emotions, sentiments; and as a consequence, how they interact with others and with their surroundings.

Positive feelings - any feeling where there is a lack of negativity, such that no pain or discomfort is felt.

Sensory-perceptual – perceptual capability utilising the senses.

Sentient - able to perceive or feel things.

Stress - when psychological it is a state of mental or emotional strain or tension resulting from adverse or demanding circumstances and/or Physiological stress, is any external or internal condition that challenges the homeostasis of a cell or an organism.

Surgency – level of power, rising, surging or swelling.

Temperament - a person's or animal's nature, especially as it permanently affects their behaviour.

Trait - a distinguishing quality or characteristic, typically one belonging to an individual.

Valence - Valence, as used in psychology, especially in discussing emotions, means the intrinsic attractiveness/"good"-ness (positive valence) or averseness/"bad"-ness (negative valence) of an event, object, or situation.

Appendix C – Chapter 4 - Hominoid Personality Questionnaire from Weiss *et al.*, 2011

MONKEY PERSONALITY TRAIT ASSESSMENT

Monkey personality assessments can be made with this questionnaire by assigning a numerical score for all of the personality traits listed on the following pages. Make your judgments on the basis of your own understanding of the trait guided by the short clarifying definition following each trait. The monkey's own behaviors and interactions with other monkeys should be the basis for your numerical ratings. Use your own subjective judgment of typical monkey behavior to decide if the monkey you are scoring is above, below, or average for a trait. The following seven point scale should be used to make your ratings.

1. Displays either total absence or negligible amounts of the trait.
2. Displays small amounts of the trait on infrequent occasions.
3. Displays somewhat less than average amounts of the trait.
4. Displays about average amounts of the trait.
5. Displays somewhat greater than average amounts of the trait.
6. Displays considerable amounts of the trait on frequent occasions.
7. Displays extremely large amounts of the trait.

Please give a rating for each trait even if your judgment seems to be based on a purely subjective impression of the monkey and you are somewhat unsure about it. Indicate your rating by placing a cross in the box underneath the chosen number. ☐

Finally, do not discuss your rating of any particular monkey with anyone else. As explained in the handout accompanying this questionnaire, this restriction is necessary in order to obtain valid reliability coefficients for the traits.

MONKEY PERSONALITY TRAIT ASSESSMENT

Monkey's full name: _____

Rater's full name: _____

Date (Mon/Day/Yr): _____

FEARFUL: Subject reacts excessively to real or imagined threats by displaying behaviors such as screaming, grimacing, running away or other signs of anxiety or distress.

least most

1 2 3 4 5 6 7

DOMINANT: Subject is able to displace, threaten, or take food from other monkeys. Or subject may express high status by decisively intervening in social interactions.

least most

1 2 3 4 5 6 7

PERSISTENT: Subject tends to continue in a course of action, task, or strategy for a long time or continues despite opposition from other monkeys.

least most

1 2 3 4 5 6 7

CAUTIOUS: Subject often seems attentive to possible harm or danger from its actions. Subject avoids risky behaviors.

least most

1 2 3 4 5 6 7

STABLE: Subject reacts to its environment including the behavior of other monkeys in a calm, equable, way. Subject is not easily upset by the behaviors of other monkeys.

least												most
	1	2	3	4	5	6	7					
	☐	☐	☐	☐	☐	☐	☐					

AUTISTIC: Subject often displays repeated, continuous, and stereotyped behaviors such as rocking or self claspings.

least												most
	1	2	3	4	5	6	7					
	☐	☐	☐	☐	☐	☐	☐					

CURIOUS: Subject has a desire to see or know about objects, devices, or other monkeys. This includes a desire to know about the affairs of other monkeys that do not directly concern the subject.

least												most
	1	2	3	4	5	6	7					
	☐	☐	☐	☐	☐	☐	☐					

THOUGHTLESS: Subject often behaves in a way that seems imprudent or forgetful.

least												most
	1	2	3	4	5	6	7					
	☐	☐	☐	☐	☐	☐	☐					

STINGY/GREEDY: Subject is excessively desirous or covetous of food, favored locations, or other resources in the enclosure. Subject is unwilling to share these resources with others.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

JEALOUS: Subject is often troubled by others who are in a desirable or advantageous situation such as having food, a choice location, or access to social groups. Subject may attempt to disrupt activities of advantaged monkeys.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

INDIVIDUALISTIC: Subject's behavior stands out compared to that of the other individuals in the group. This does not mean that it does not fit or is incompatible with the group.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

RECKLESS: Subject is rash or unconcerned about the consequences of its behaviors.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

SOCIABLE: Subject seeks and enjoys the company of other monkeys and engages in amicable, affable, interactions with them.

least 1 2 3 4 5 6 7 most

☐	☐	☐	☐	☐	☐	☐

DISTRACTIBLE: Subject is easily distracted and has a short attention span.

least 1 2 3 4 5 6 7 most

☐	☐	☐	☐	☐	☐	☐

TIMID: Subject lacks self confidence, is easily alarmed and is hesitant to venture into new social or non-social situations.

least 1 2 3 4 5 6 7 most

☐	☐	☐	☐	☐	☐	☐

SYMPATHETIC: Subject seems to be considerate and kind towards others as if sharing their feelings or trying to provide reassurance.

least 1 2 3 4 5 6 7 most

☐	☐	☐	☐	☐	☐	☐

PLAYFUL: Subject is eager to engage in lively, vigorous, sportive, or acrobatic behaviors with or without other monkeys.

least 1 2 3 4 5 6 7 most

☐	☐	☐	☐	☐	☐	☐

SOLITARY: Subject prefers to spend considerable time alone not seeking or avoiding contact with other monkeys.

least most

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

VULNERABLE: Subject is prone to be physically or emotionally hurt as a result of dominance displays, highly assertive behavior, aggression, or attack by another monkey.

least most

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

INNOVATIVE: Subject engages in new or different behaviors that may involve the use of objects or materials or ways of interacting with others.

least most

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

ACTIVE: Subject spends little time idle and seems motivated to spend considerable time either moving around or engaging in some overt, energetic behavior.

least most

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

HELPFUL: Subject is willing to assist, accommodate, or cooperate with other monkeys.

least most

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

BULLYING: Subject is overbearing and intimidating towards younger or lower ranking monkeys.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

AGGRESSIVE: Subject often initiates fights or other menacing and agonistic encounters with other monkeys.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

MANIPULATIVE: Subject is adept at forming social relationships for its own advantage, especially using alliances and friendships to increase its social standing. Monkey seems able and willing to use others.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

GENTLE: Subject responds to others in an easy-going, kind, and considerate manner. Subject is not rough or threatening.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

AFFECTIONATE: Subject seems to have a warm attachment or closeness with other monkeys. This may entail frequently grooming, touching, embracing, or lying next to others.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

EXCITABLE: Subject is easily aroused to an emotional state. Subject becomes highly aroused by situations that would cause less arousal in most monkeys.

least												most
	1	2	3	4	5	6	7					
	☐	☐	☐	☐	☐	☐	☐					

IMPULSIVE: Subject often displays some spontaneous or sudden behavior that could not have been anticipated. There often seems to be some emotional reason behind the sudden behavior.

least												most
	1	2	3	4	5	6	7					
	☐	☐	☐	☐	☐	☐	☐					

INQUISITIVE: Subject seems drawn to new situations, objects, or animals. Subject behaves as if it wishes to learn more about other monkeys, objects, or persons within its view.

least												most
	1	2	3	4	5	6	7					
	☐	☐	☐	☐	☐	☐	☐					

SUBMISSIVE: Subject often gives in or yields to another monkey. Subject acts as if it is subordinate or of lower rank than other monkeys.

least												most
	1	2	3	4	5	6	7					
	☐	☐	☐	☐	☐	☐	☐					

COOL: Subject seems unaffected by emotions and is usually undisturbed, assured, and calm.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

DEPENDENT/FOLLOWER: Subject often relies on other monkeys for leadership, reassurance, touching, embracing and other forms of social support.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

IRRITABLE: Subject often seems in a bad mood or is impatient and easily provoked to anger exasperation and consequent agonistic behavior.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

UNPERCEPTIVE: Subject is slow to respond or understand moods, dispositions, or behaviors of others.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

PREDICTABLE: Subject's behavior is consistent and steady over extended periods of time. Subject does little that is unexpected or deviates from its usual behavioral routine.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

DECISIVE: Subject is deliberate, determined, and purposeful in its activities.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

DEPRESSED: Subject does not seek out social interactions with others and often fails to respond to social interactions of other monkeys. Subject often appears isolated, withdrawn, sullen, brooding, and has reduced activity.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

CONVENTIONAL: Subject seems to lack spontaneity or originality. Subject behaves in a consistent manner from day to day and stays well within the social rules of the group.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

SENSITIVE: Subject is able to understand or read the mood, disposition, feelings, or intentions of other monkeys often on the basis of subtle, minimal cues.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

DEFIANT: Subject is assertive or contentious in a way inconsistent with the usual dominance order. Subject maintains these actions despite unfavorable consequences or threats from others.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

INTELLIGENT: Subject is quick and accurate in judging and comprehending both social and non-social situations. Subject is perceptive and discerning about social relationships.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

PROTECTIVE: Subject shows concern for other monkeys and often intervenes to prevent harm or annoyance from coming to them.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

QUITTING: Subject readily stops or gives up activities that have recently been started.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

INVENTIVE: Subject is more likely than others to do new things including novel social or non-social behaviors. Novel behavior may also include new ways of using devices or materials.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

CLUMSY: Subject is relatively awkward or uncoordinated during movements including but not limited to walking, acrobatics, and play.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

ERRATIC: Subject is inconsistent, indefinite, and widely varying in its behavior and moods.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

FRIENDLY: Subject often seeks out contact with other monkeys for amiable, genial activities. Subject infrequently initiates hostile behaviors towards other monkeys.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

ANXIOUS: Subject often seems distressed, troubled, or is in a state of uncertainty.

least								most
	1	2	3	4	5	6	7	
	☐	☐	☐	☐	☐	☐	☐	

LAZY: Subject is relatively inactive, indolent, or slow moving and avoids energetic activities.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

DISORGANIZED: Subject is scatterbrained, sloppy, or haphazard in its behavior as if not following a consistent goal.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

UNEMOTIONAL: Subject is relatively placid and unlikely to become aroused, upset, happy, or sad.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

IMITATIVE: Subject often mimics, or copies behaviors that it has observed in other monkeys.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

INDEPENDENT: Subject is individualistic and determines its own course of action without control or interference from other monkeys.

least							most
	1	2	3	4	5	6	7
	☐	☐	☐	☐	☐	☐	☐

Appendix D – Chapter 4 - Equine Personality Eight-Factor Model and Abridged Questionnaire

Equine Personality Eight-Factor Model (EPEM)

Equine personality assessments can be made with this questionnaire by assigning a numerical score for all the personality traits listed on the following pages. Make your judgments based on your own understanding of the trait guided by the short clarifying definition following each trait. The equine's own behaviours and interactions with other equines should be the basis for your numerical ratings. Use your own subjective judgment of typical equine behaviour to decide if the equine you are scoring is above, below, or average for a trait. The following seven-point scale should be used to make your ratings.

Scoring system

1	2	3	4	5	6	7
Least or negligible amount of the trait	Low amount of the trait	Less than average amount of the trait	Average amount of the trait	Slightly higher than average amount of the trait	High amount of the trait	Very high amount of the trait

Please give a rating for each trait even if your judgment seems to be based on a purely subjective impression of the equine and you are somewhat unsure about it. Indicate your rating by placing a cross in the box underneath the chosen number as shown,

X

DISTRACTIBLE: Equine is easily distracted and has a short attention span.

Least	1	2	3		4	5	6	7	Most

PLAYFUL: Equine is eager to engage in lively, vigorous, sportive, or acrobatic behaviours with or without other horses.

Least	1	2	3	4	5	6	7	Most

ACTIVE: Equine spends little time idle and seems motivated to spend considerable time either moving around or engaging in some overt, energetic behaviour.

Least	1	2	3	4	5	6	7	Most

EXCITABLE: Equine is easily aroused to an emotional state. Equine becomes highly aroused by situations that would cause less arousal in most horses.

Least	1	2	3	4	5	6	7	Most

IMPULSIVE: Equine often displays some spontaneous or sudden behaviour that could not have been anticipated. There often seems to be some emotional reason behind the sudden behaviour.

Least	1	2	3	4	5	6	7	Most

FEARFUL: Equine reacts excessively to real or imagined threats by displaying behaviours such as snorting, shying, moving away or other signs of anxiety or distress.

Least	1	2	3	4	5	6	7	Most

CAUTIOUS: Equine often seems attentive to possible harm or danger from its actions. Equine avoids risky behaviours.

Least	1	2	3	4	5	6	7	Most

TIMID: Equine lacks self-confidence, is easily alarmed and is hesitant to venture into new social or non-social situations.

Least	1	2	3	4	5	6	7	Most

VULNERABLE: Equine is prone to be physically or emotionally hurt as a result of highly assertive behaviour, aggression, or attack by another horse.

Least	1	2	3	4	5	6	7	Most

ACQUIESCENT: Equine often gives in or yields to another horse. Equine acts as if it is subordinate or defers to other horses.

Least	1	2	3	4	5	6	7	Most

SENSITIVE: Equine can understand or read the mood, disposition, feelings, or intentions of other horses often on the basis of subtle, minimal cues.

Least	1	2	3	4	5	6	7	Most

SOLITARY: Equine prefers to spend considerable time alone not seeking or avoiding contact with other horses.

Least	1	2	3	4	5	6	7	Most

UNPERCEPTIVE: Equine is slow to respond or understand moods, dispositions, or behaviours of others.

Least	1	2	3	4	5	6	7	Most

DEPRESSED: Equine does not seek out social interactions with others and often fails to respond to social interactions of other horses. Equine often appears isolated, withdrawn and has reduced activity.

Least	1	2	3	4	5	6	7	Most

UNEMOTIONAL: Equine is relatively placid and unlikely to become aroused, upset, happy, or sad.

Least	1	2	3	4	5	6	7	Most

COOL: Equine seems unaffected by emotions and is usually undisturbed, assured, and calm.

Least	1	2	3	4	5	6	7	Most

PREDICTABLE: Equine's behaviour is consistent and steady over extended periods of time. Equine does little that is unexpected or deviates from its usual behavioural routine.

Least	1	2	3	4	5	6	7	Most

CONVENTIONAL: Equine seems to lack spontaneity or originality. Equine behaves in a consistent manner from day to day and stays well within the social rules of the group.

Least	1	2	3	4	5	6	7	Most

ASSERTIVE: Equine is able to displace, threaten, or take resources from other horses. Or equine may express high status by decisively intervening in social interactions.

Least	1	2	3	4	5	6	7	Most

INDIVIDUALISTIC: Equine's behaviour stands out compared to that of the other individuals in the group. This does not mean that it does not fit or is incompatible with the group.

Least	1	2	3	4	5	6	7	Most

PERSISTENT: Equine tends to continue in a course of action, task, or strategy for a long time or continues despite opposition from other horses or humans.

Least	1	2	3	4	5	6	7	Most

DECISIVE: Equine is deliberate, determined, and purposeful in its activities.

Least	1	2	3	4	5	6	7	Most

INDEPENDENT: Equine is individualistic and determines its own course of action without control or interference from other horses.

Least	1	2	3	4	5	6	7	Most

SOCIABLE: Equine seeks and enjoys the company of other horses and engages in amicable, affable, interactions with them.

Least	1	2	3	4	5	6	7	Most

SYMPATHETIC: Equine seems to be considerate and kind towards others as if sharing their feelings or trying to provide reassurance.

Least	1	2	3	4	5	6	7	Most

HELPFUL: Equine is willing to assist, accommodate, or cooperate with other horses.

Least	1	2	3	4	5	6	7	Most

GENTLE: Equine responds to others in an easy-going, kind, and considerate manner. Equine is not rough or threatening.

Least	1	2	3	4	5	6	7	Most

AFFECTIONATE: Equine seems to have a warm attachment or closeness with other horses or humans. This may entail frequently grooming, touching, or spending time next to others.

Least	1	2	3	4	5	6	7	Most

FRIENDLY: Equine often seeks out contact with other horses for amiable, genial activities. Equine infrequently initiates hostile behaviours towards other horses.

Least	1	2	3	4	5	6	7	Most

BULLYING: Equine is overbearing and intimidating towards younger or more acquiescent horses.

Least	1	2	3	4	5	6	7	Most

CONFRONTATIONAL: Equine often initiates fights or other agonistic encounters with other horses.

Least	1	2	3	4	5	6	7	Most

IRRITABLE: Equine often seems in a bad mood or is impatient and easily provoked to anger, exasperation, and consequent agonistic behaviour.

Least	1	2	3	4	5	6	7	Most

DEFIANT: Equine is assertive or contentious in a way inconsistent with the usual dominance order. Equine maintains these actions despite unfavourable consequences or threats from others.

Least	1	2	3	4	5	6	7	Most

CURIOUS: Equine has a desire to see or know about objects, people, or other horses. This includes a desire to know about the affairs of other horses or activities that do not directly concern the equine.

Least	1	2	3	4	5	6	7	Most

INNOVATIVE: Equine engages in new or different behaviours that may involve the use of objects or materials or ways of interacting with others.

Least	1	2	3	4	5	6	7	Most

INVENTIVE: Equine is more likely than others to do new things including novel social or non-social behaviours. Novel behaviour may also include innovative ways of using objects or materials.

Least	1	2	3	4	5	6	7	Most

INTELLIGENT: Equine is quick and accurate in judging and comprehending both social and non-social situations. Equine is perceptive and discerning about social relationships.

Least	1	2	3	4	5	6	7	Most

Short form of the Equine Personality Eight-factor Model (EPEM-R)

Equine personality group assessments can be made with this questionnaire by assigning a numerical score for all the personality groups listed on the following pages. Make your judgments based on your own understanding of the group guided by the list of included traits. The equine's own behaviours and interactions with other equines should be the basis for your numerical ratings. Use your own subjective judgment of typical equine behaviour to decide if the equine you are scoring is above, below, or average for a group. The following seven-point scale should be used to make your ratings.

Scoring system

1	2	3	4	5	6	7
Least or negligible amount of the trait	Low amount of the trait	Less than average amount of the trait	Average amount of the trait	Slightly higher than average amount of the trait	High amount of the trait	Very high amount of the trait

Please give a rating for each group even if your judgment seems to be based on a purely subjective impression of the equine and you are somewhat unsure about it. Indicate your rating by placing a cross in the box underneath the chosen number as shown,

X

EXUBERANT: Includes Excitable, Playful, Active, Distractible, Impulsive.

Least	1	2	3	4	5	6	7	Most

NERVOUS: Includes Cautious, Fearful, Acquiescent, Vulnerable, Sensitive, Timid.

Least	1	2	3	4	5	6	7	Most

UNRESPONSIVE: Includes Depressed, Unemotional, Solitary, Unperceptive.

Least	1	2	3	4	5	6	7	Most

STEADY: Includes Predictable, Conventional, Cool.

Least	1	2	3	4	5	6	7	Most

BOLD: Includes Assertive, Independent, Decisive, Persistent, Individualistic.

Least	1	2	3	4	5	6	7	Most

AFFILIATIVE: Includes Friendly, Helpful, Sociable, Sympathetic, Gentle, Affectionate.

Least	1	2	3	4	5	6	7	Most

UNFRIENDLY: Includes Irritable, Defiant, Bullying, Confrontational.

Least	1	2	3	4	5	6	7	Most

INVESTIGATORY: Includes Inventive, Innovative, Curious.

Least	1	2	3	4	5	6	7	Most

Appendix E – Chapter 6 - Study participant information sheet

PARTICIPANT INFORMATION

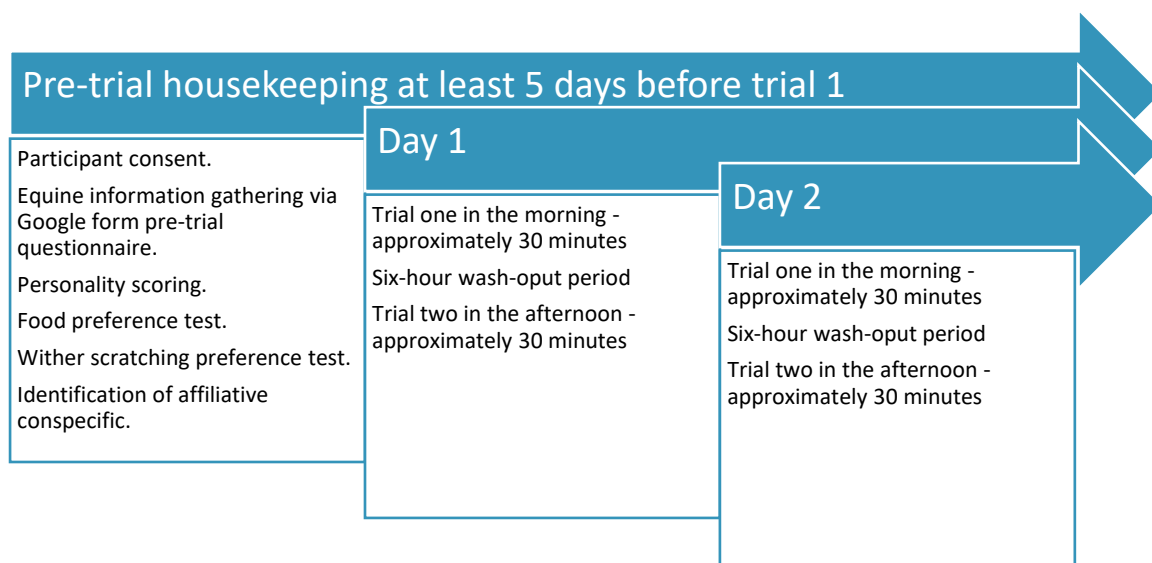
Dear Participant,

Thank you for agreeing to participate in our study which we hope will further the welfare of equines. This study has full ethical approval from Newcastle University's Animal Welfare Ethical Review Board (AWERB project ID: 1025) and does not involve any invasive measures. If you have any questions about the study, please email the lead researcher Loni Loftus at i.l.loftus@newcastle.ac.uk.

In order to participate in the study, you will need to:

- Have at least one equine that you own or care for
- Have another friendly equine available for use during the trial
- Confirm that your equine is physically healthy
- Confirm that your equine will be comfortable wearing a heart rate monitor surcingle band
- Confirm that your equine will be comfortable having saliva swabs taken
- Confirm that you have access to a stable or tying up area
- Be over the 18 years of age and have the capacity to consent to your involvement and the use of the equine for study purposes
- Complete the pre-trial questionnaire at least 24 hours prior to the allocated study period
- Be available on two consecutive days for a period of time in the morning and afternoon of both days in order to facilitate data collection (each trial lasts around 20-30 minutes per equine and each equine will undertake 4 trials as specified below)
- Be comfortable handling your equine

The planned process for the trial is shown below.



The trials will all follow the same format. Your equine will be habituated to the test environment, and we will ensure that they are comfortable in the place being used. A friendly equine will be placed nearby for company. The video recording equipment will be in place and after ten minutes of habituation a heart rate monitor will be fitted to your equine. The equine will then be allowed to explore the area for a ten-minute period. Following this the equine will be loosely cross tied using quick release safety ties, or held if more appropriate, and a saliva swab will be taken; following this the first trial will begin. **Each trial will last no longer than two minutes**, following which the equipment will be removed and salivary swabs taken immediately after, and the equine allowed free movement in the area for 15 minutes after the end of the trial. Following this time period a third salivary swab will be taken and the equine will then be returned to their normal environment. **A washout period (gap) of at least 6 hours is required before progressing on to the next trial.**

There are four trials in total which will be allocated to each equine in random order, the trials are:

- **Provision of a high value food source** (as ascertained via the earlier preference testing).
- **Withers scratching** (also known as scritch) in their pre-tested preferred format.
- A short **target training** session.
- **Proximity to a known friendly equine**, identified by you as one that your equine chooses to spend time with as per the criteria that they have been observed allogrooming or

interacting positively with the other subject equine on at least five occasions within the last month.

During each trial the following data will be collected from your equine:

- Heart rate data (including heart rate variability)
- Respiratory rate
- Salivary cortisol samples
- Thermography video of the head
- Video footage of the whole equine (from the front and rear)

Following the data collection, the data will be anonymised prior to analysis. At this point it will no longer be possible to withdraw from the trial as the data will be unidentifiable.

After reading this participant information sheet, please confirm your consent and answer the questions within the pre-trial questionnaire. Please complete a separate questionnaire for each equine that you own or care for that is participating in the trial. You will also be sent a brief short-form personality questionnaire to complete for each equine prior to the trial.

Thank you for giving your valuable time to this research.

Appendix F – Chapter 6 - Pre-trial questionnaire

Measuring Positive Affective State in Equines: Pre-trial equine questionnaire

This form is used to collect relevant information regarding your equine prior to the trial sessions. The data provided will allow for careful analysis of the results while considering the individual situation of each equine and their personality facets. Please do complete the questionnaire as fully and accurately as possible. We are very grateful for your participation in this project which will help horse owners and professionals to understand how happy their equine is.

The project has full ethical approval from Newcastle University's Animal Welfare Ethical Review Board (AWERB project ID: 1025). If you have any questions about the study please email the lead researcher Loni Loftus at i.j.loftus@newcastle.ac.uk

Please read the [study information document](#) which gives you an overview of the purpose of the study and the methods and processes we will be using over the two data collection days spent with you and your equine. Following this, please check the 'I consent' box below to consent to your equine's data being collected and used in the study.

* Indicates required question

1. Please state your full name *

2. Please provide a contact phone number

3. Please provide a contact email address *

4. I consent to my equine's data being collected and used within the study. *
- I understand that once data are collected they will be anonymised and therefore I will be unable to withdraw the data from the study.
- I understand and consent to images of my equine potentially being used within publications and presentations relevant to the study.

Mark only one oval.

☐ I consent

Equine information

This section collects demographic information on your equine and their environment.

5. Equine name *

6. Equine age in years (in number digits) *

7. Equine breed *

8. Sex *

Mark only one oval.

☐ Mare

☐ Gelding

☐ Stallion

9. Equine activities (select all that apply) *

Check all that apply.

- ☐ Non-ridden
- ☐ Elite competition (such as SJ, BD, BE, Racing, Driving)
- ☐ Affiliated competition
- ☐ Unaffiliated competition
- ☐ Hacking /pleasure riding
- ☐ Breeding
- ☐ Education (such as riding school, equine assisted therapy)
- ☐ Other: _____

10. How long has this equine been in your ownership or care? *

Mark only one oval.

- ☐ Less than 1 year
- ☐ 1-2 years
- ☐ 3-4 years
- ☐ 5-6 years
- ☐ 6 years +

11. Has anything changed in relation to your equines environment (location, other horses, routine etc) in the last 6 weeks? *

Mark only one oval.

- ☐ Yes
- ☐ No

12. If you answered yes to the question above regarding your equine's environment, please state how this has changed below.

13. Is your equine healthy to the best of your knowledge? *

Mark only one oval.

- ☐ Yes
☐ No
☐ Unsure

14. If you stated No or Unsure regarding your equines general health please give further details of medical conditions here.

15. Does your equine have any dietary intolerances or are they on a specialised diet * for health purposes?

Mark only one oval.

- ☐ Yes
☐ No

16. If you answered Yes to dietary issues/specialised diet please give further details here

17. Does your equine have another equine friend at their location? *Please note that access to another friendly horse is required for this research. *

Mark only one oval.

☐ Yes

☐ No

18. Has your horse ever experienced target training? *

Mark only one oval.

☐ Yes

☐ No

☐ Unsure

19. Do you think that your horse enjoys scratches (such as wither scratches)? *

Mark only one oval.

☐ Yes

☐ No

20. How do you reward your horse for correct behaviours? Please select all that apply. *

Check all that apply.

- ☐ Food treats
- ☐ Verbal praise (e.g. good boy/girl)
- ☐ Scratches
- ☐ Pressure release
- ☐ Patting
- ☐ Stroking
- ☐ Head rubbing
- ☐ No reward

21. Is your equine used to wearing a girth, surcingle or roller? *Please note that this trial requires the horse to wear a Polar equine heart rate monitor band. *

Mark only one oval.

- ☐ Yes
- ☐ No

22. Will your horse be comfortable with having salivary cortisol swabs taken? This is required for the trial. Three swabs per test will be taken (12 in total across the four tests) with a small cotton collecting stick inserted into the mouth at the corner where the bit usually sits. *

Mark only one oval.

- ☐ Yes
- ☐ No

23. Does your horse have any behavioural issues? Please select all that apply. *

Check all that apply.

- ☐ Repetitive behaviours (such as crib biting, wind sucking, box walking, weaving, fence pacing)
- ☐ Self mutilation
- ☐ Aggression towards humans (bite, kick behaviours)
- ☐ Aggression towards other equines (bite, kick behaviours)
- ☐ Aggression towards other species (dogs, sheep etc)
- ☐ Fears or phobias (such as aversion to noises, headshyness etc)
- ☐ Handling issues (such as barging, pulling away from the handler etc)
- ☐ Ridden issues (bucking, rearing, bolting)
- ☐ Resource guarding (for example food)
- ☐ None

24. Is your horse comfortable being tied up? *

Mark only one oval.

- ☐ Yes, totally comfortable
- ☐ For short periods of time
- ☐ No

Your equine's personality

Please answer the question below to rate your equines personality.

(Irritable, Hostile, defiant, bullying, confrontational) (Irritable, defiant, bullying, confrontational)	☐	☐	☐	☐	☐	☐	☐
(Intelligent, Investigative, inventive, innovative, curious)	☐	☐	☐	☐	☐	☐	☐

Appendix G – Chapter 6 - Trial order randomisation results

Equine number	Stimulus presentation order			
	Stimulus Food	Stimulus Training	Stimulus Scratch	Stimulus Companion
1	4	1	2	3
2	1	4	3	2
3	4	3	1	2
4	1	4	3	2
5	1	4	2	3
6	2	4	1	3
7	1	4	3	2
8	2	1	3	4
9	1	4	2	3
10	1	3	4	2
11	2	3	1	4
12	4	2	1	3
13	3	4	1	2
14	4	3	2	1
15	2	1	4	3
16	3	4	1	2
17	1	2	4	3
18	2	1	3	4
19	2	4	3	1
20	2	3	4	1

Appendix H – Chapter 6 - Cowlog instructions for raters

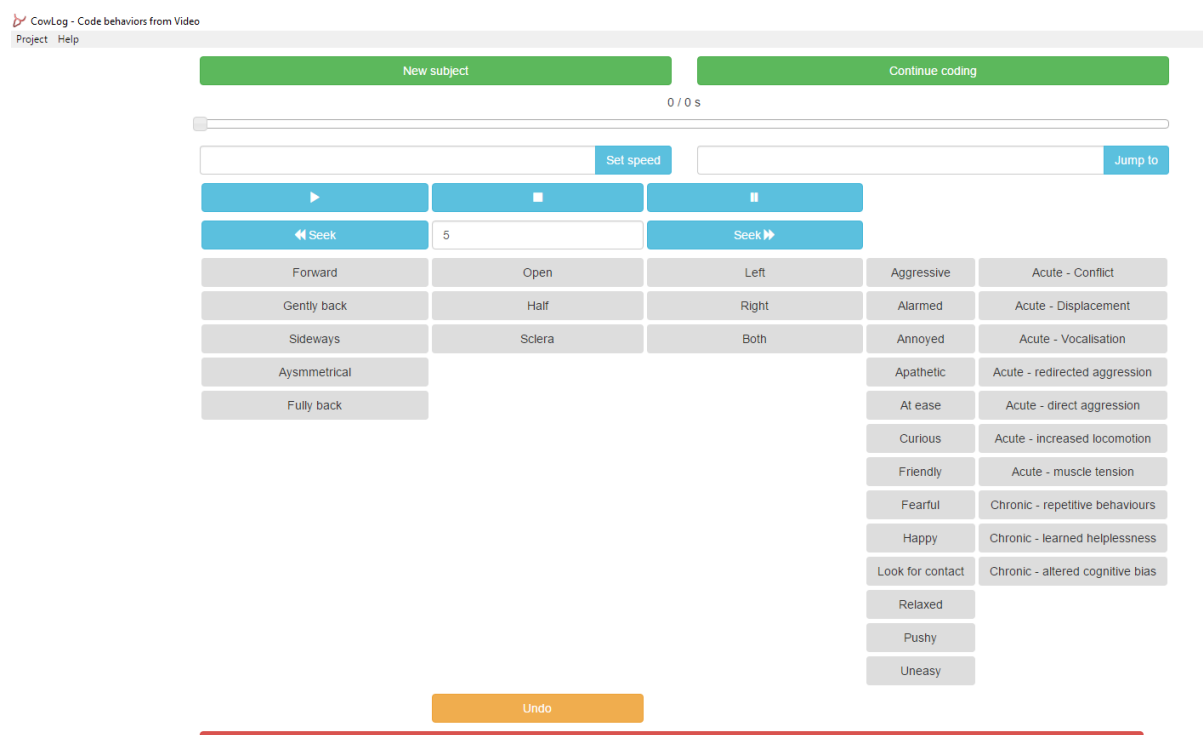
Video scoring instructions

Setting up Cowlog

1. Please download Cowlog from <https://cowlog.org/> and install on your PC/laptop
2. Save the attached file 'PA trial final_cowlog_project.json' on your PC/Laptop
3. Open cowlog on your PC/Laptop, it will look like this:

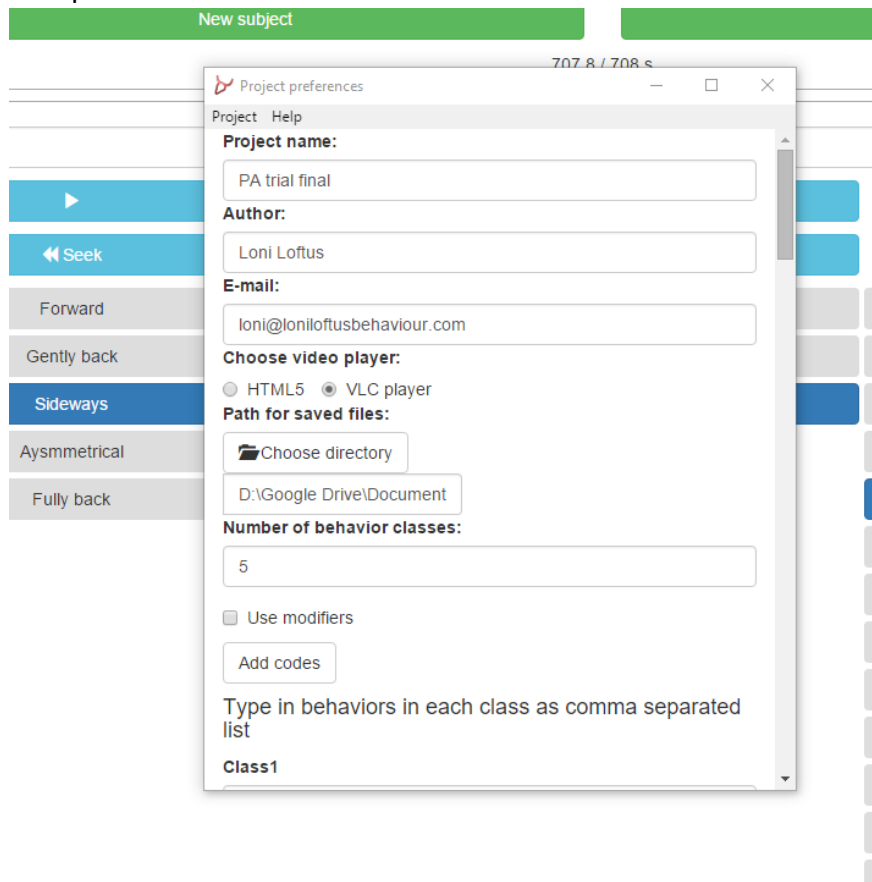


4. In the menu bar go to Project and then Load project
5. When the file explorer window opens select 'PA trial final_cowlog_project.json'. This will add all of the behaviour coding buttons to your Cowlog window and will now look like this:

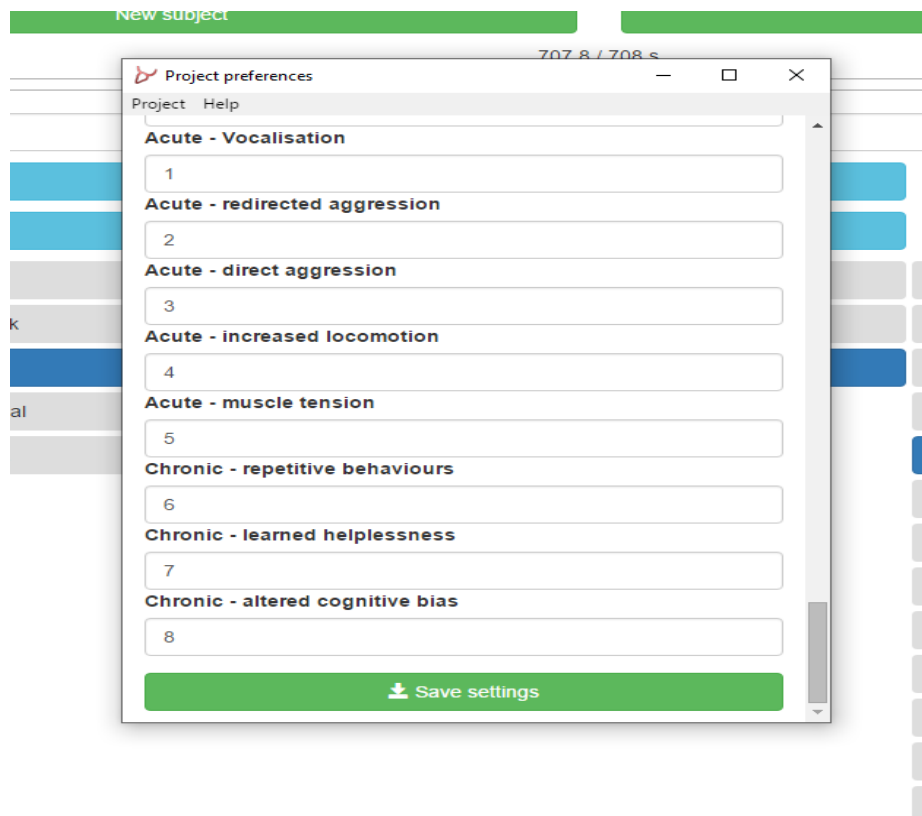


6. Now go back to 'Project' top left hand corner and click edit project, this will open a pop-up window which is in file explorer and you need to navigate to the 'PA trial

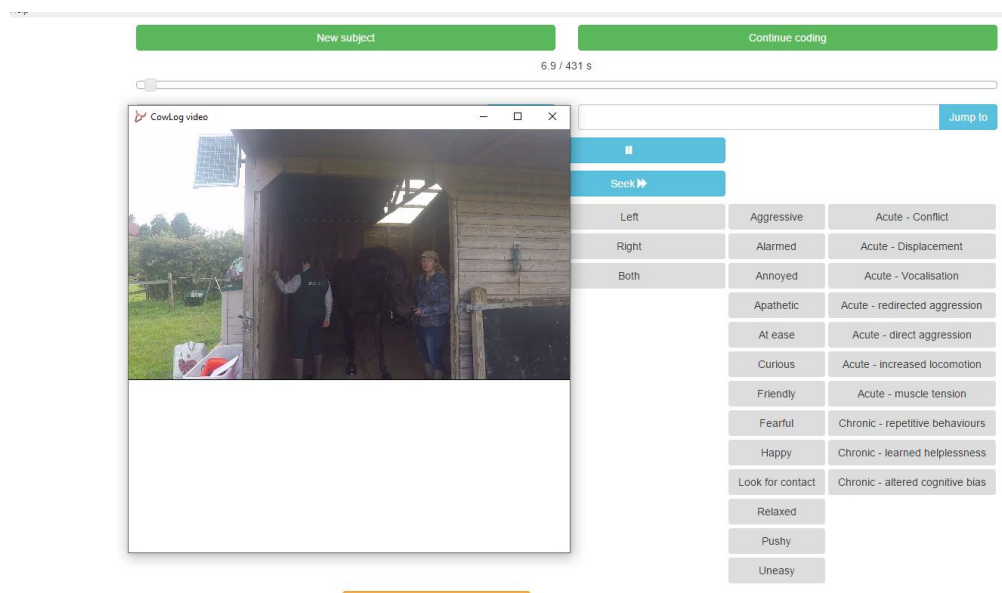
final_cowlog_project.json' file you saved earlier. Double click on it. Another pop-up window will open as shown below:



7. Click on 'Choose directory' this will bring up another file explorer window which should be to the place where you have stored the cowlog file. Select where you would like Cowlog to save the files it generates. This may be your desktop or another folder that you make. Choose the folder and click select folder in the bottom right of the pop-up window. This will then close the file explorer window and take you back to the previous screen (as in the above picture) in Cowlog.
8. Now scroll down to the bottom of that project preferences window and click save settings (see pic below). It will ask you if you want to replace the existing Cowlog file you have saved. Select Yes. You are now ready to load a new subject and start coding! Cowlog will automatically save the output excel sheets to the folder you have stipulated in the preferences you have just set up.



9. Click on New subject (green button at the top) and in the window that pops up, name the new video coding the same name as the video you are scoring e.g. H1S1 or H5T3, set the video recording start time as 0, then click choose video. In the window that opens select the video that you want to score and then click Start coding. You will get a new pop-up window with the video you are scoring, like this:



10. Arrange the video and the scoring window on your screen so that you can see both at the same time (or use two screens).

Scoring the videos

Within each video you will see that there are three parts to it:

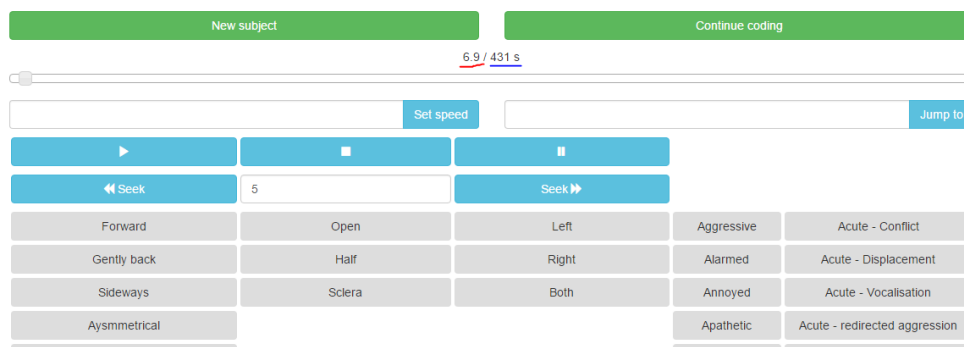
- I. The before phase which begins when the video starts and where the horse is just standing (you may see salivary swabs been taken or hay being eaten) with none of the treatments being applied,
- II. The during phase when the horse is experiencing the treatment (there are 4 treatments: Wither scratching, food, target training and a companion horse),
- III. and the after phase when the treatment stops which lasts until the end of the video.

Please make a note of the timings for:

The total video length in seconds (underlined in blue in the example below)

The end of phase 1/beginning of phase 2 (example underlined in red below) – when the stimulus is introduced to the horse (i.e. the food is given, the companion arrives, the training starts or the wither scratching starts)

The end of phase 2/ beginning of phase 3 – when the stimulus is removed (i.e. the food is removed, the companion is taken away, the training stops or the wither scratching stops).



You will need to watch each video 5 times to score:

- a. Ear position (forward, gently back, sideways etc)
- b. Eye position (Open, half and sclera)
- c. Eye bias (left, right, both)
- d. QBA assessment (Aggressive, alarmed etc)
- e. Frustration markers (Acute – conflict, Acute – displacement etc)

Please score these factors in this order so starting with ear position and ending with frustration markers.

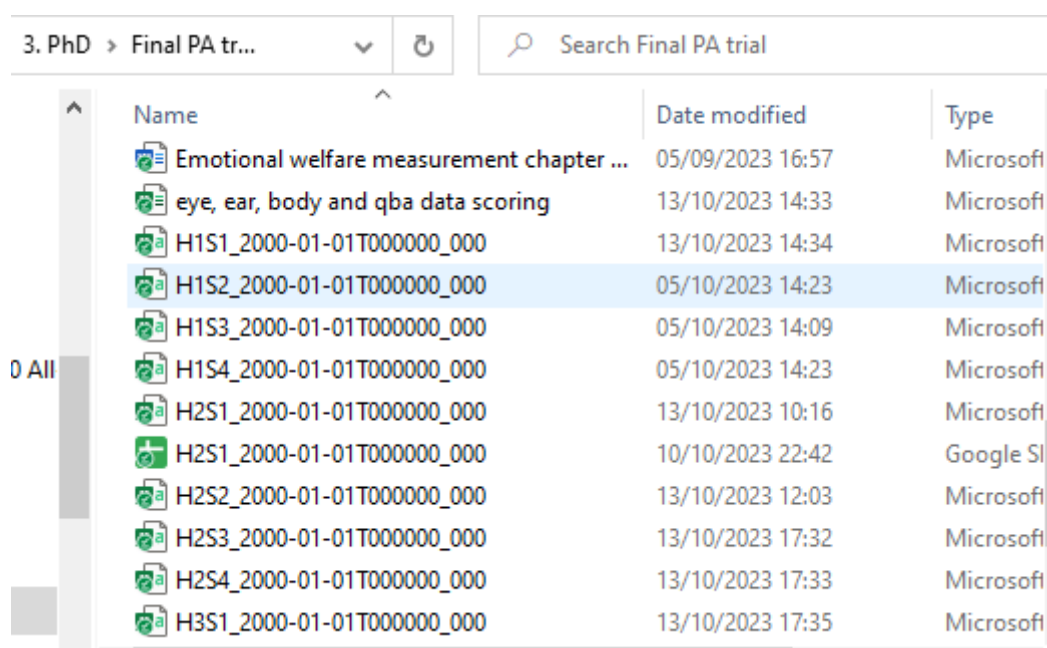
To start scoring press the play button (blue button) and the video will start playing. Start scoring immediately clicking the appropriate behavioural code for the category you are scoring so for example if scoring ear position and the ears are forwards click 'forward'. If the horse changes ear position click the relevant button, e.g. 'gently

back' and continue in this fashion, at the end of each part of the video (before, during and after) click the red 'End coding' button at the bottom of the scoring screen and then continue scoring the next phase.

When the video ends you can simply run the marker on the white progress bar at the top (under the green buttons) back to the beginning and restart the video to score the next set of behaviours.

When you have scored all of the behaviours for that video ensure that you have clicked 'end coding'.

You will see that an Microsoft excel file has automatically been generated and saved in the same location that you saved the original project file (PA trial final-_cowlog_project.json) and the programme will have named it the name you gave when you uploaded the video (e.g. H1S1 or H5T2) plus some more alphanumerical digits (see example below)



Name	Date modified	Type
Emotional welfare measurement chapter ...	05/09/2023 16:57	Microsoft
eye, ear, body and qba data scoring	13/10/2023 14:33	Microsoft
H1S1_2000-01-01T000000_000	13/10/2023 14:34	Microsoft
H1S2_2000-01-01T000000_000	05/10/2023 14:23	Microsoft
H1S3_2000-01-01T000000_000	05/10/2023 14:09	Microsoft
H1S4_2000-01-01T000000_000	05/10/2023 14:23	Microsoft
H2S1_2000-01-01T000000_000	13/10/2023 10:16	Microsoft
H2S1_2000-01-01T000000_000	10/10/2023 22:42	Google Sl
H2S2_2000-01-01T000000_000	13/10/2023 12:03	Microsoft
H2S3_2000-01-01T000000_000	13/10/2023 17:32	Microsoft
H2S4_2000-01-01T000000_000	13/10/2023 17:33	Microsoft
H3S1_2000-01-01T000000_000	13/10/2023 17:35	Microsoft

When you open the file it should look like this:

	A	B	C	D	E	F	G	H	I	J
1	time	code	class							
2	3.525	Gently ba	1							
3	5.84	Gently ba	1							
4	8.739	Forward	1							
5	10.523	Forward	1							
6	11.803	Forward	1							
7	14.879	Forward	1							
8	16.923	Asymmet	1							
9	18.203	Asymmet	1							
10	22.043	Sideways	1							
11	28.699	Forward	1							
12	30.5	Forward	1							
13	37.159	Gently ba	1							
14	39.199	Gently ba	1							
15	39.963	Asymmet	1							
16	47.643	Gently ba	1							
17	57.883	Forward	1							
18	65.059	Asymmet	1							
19	68.639	Forward	1							
20	76.839	Asymmet	1							
21	85.539	Forward	1							
22	94.239	Fully back	1							
23	95.003	Gently ba	1							
24	102.439	Forward	1							
25	105.499	Asymmet	1							

With a long list of all of your scores and associated times running down the page interspersed with END where you have ended the coding between each phase of the video (before, during and after) and between each set of behaviours scores (ears, eyes, QBA etc).

This is the file you need to return to me for each video with your list of the timings for each section and the total video.

Other scoring information

For the behavioural scores for ear position, eye position, eye bias and QBA please click the relevant buttons as you see the behaviours occur moving from one behaviour/position to another as you go through the video as the horse will always be showing one of the behaviours on the list.

For the frustration scoring the horse may not be showing any of the behaviours at all or for large segments of the video. Therefore for this set of coding please click the relevant button (e.g. Acute – Conflict) when you start to see the behaviour and click the same button again when the behaviour stops (this allows me to work out the duration of the behaviour performance later on). For example – if the horse starts to show some conflict behaviour click the Acute-Conflict button when you see it begin and again when it stops and the horse returns to normal behaviour.

Definitions for scoring:

Ear position

Forward	Both ears facing forwards, inner ear can be seen
Gently back	Both ears rotated backwards but relaxed, inner ear not seen
Full back	Both ears rotated fully backwards and flattened
Aysmetrical	Each ear facing different direction, may be forward, side or back
Sideways	Both ears facing to either side, relaxed

Eye position

Open	The eye is fully open with no eye white showing
Half	The eye is half closed with the upper eyelid partially covering the eye
Sclera	Some white of the eye is showing

Eye bias

Both	The horse is focussing on the stimulus with both eyes equally or is looking straight ahead
Left	The horse angles or turns the head to focus (or try to focus) more with the left eye than the right eye
Right	The horse angles or turns the head to focus (or try to focus) more with the right eye than the left eye

QBA

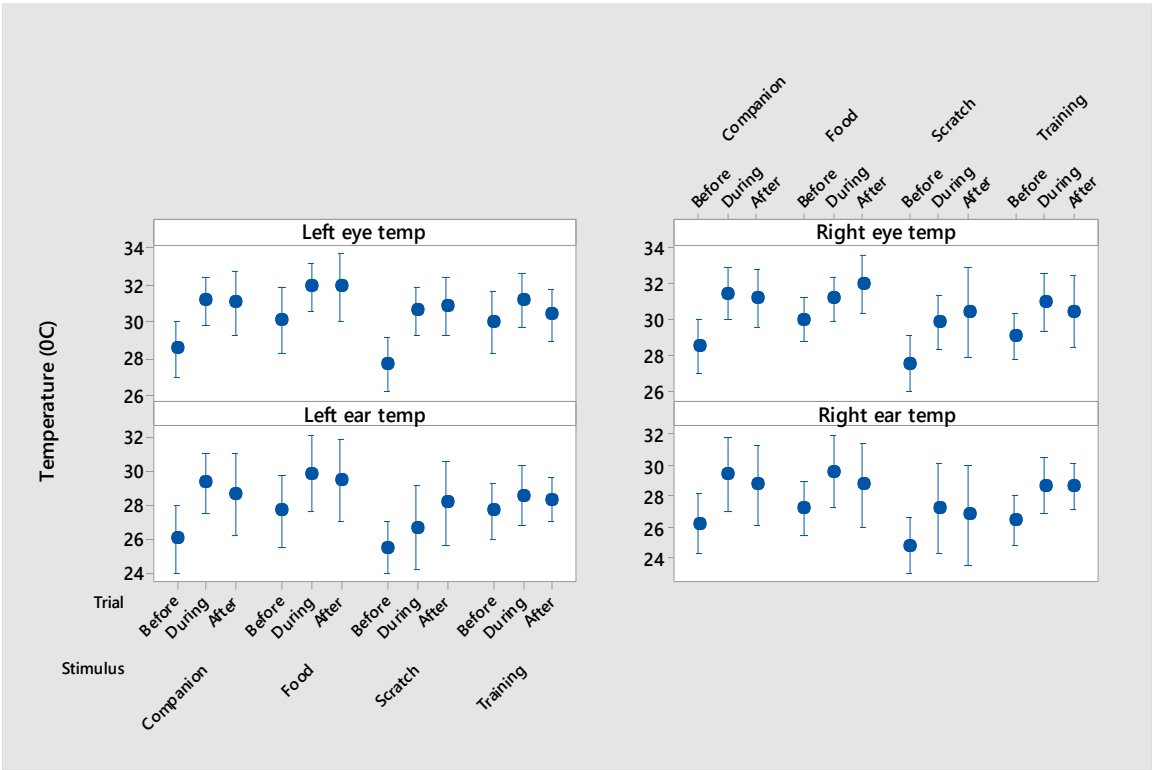
Aggressive	Hostile, attacking, wants to fight/attack, dominance, defensive aggression, (i.e. may display the following: bite/kick, position of ears flat-back against head, dilated nostrils, turns the hindquarters towards object of aggression, intention to harm, tail-swishing)
Alarmed	Worried/tense, apprehensive, jumpy, nervous, watchful, on guard against a possible threat/danger (i.e. rigid stance, startled reaction to loud noise, looking around/vigilant, moving ears)
Annoyed	Irritated, displeased, bothered by something, disturbed, upset, troubled, exasperated (i.e. may display rapid tail-swishing, stomping)
Apathetic	Having or showing little or no emotion; disinterested, indifferent, isolated, depressed, unresponsive, motionless
At ease	Calm, carefree, peaceful
Curious	Inquisitive, desire to investigate (i.e. approach person/object of curiosity, engaged in exploratory behaviour; possibly displaying head and neck extended toward object of curiosity, with ears pricked forward)
Friendly	Affectionate, kind, not hostile, receptive, positive feelings toward people, confident (i.e. the horse approaches the person, may sniff or interact in some way)
Fearful	Afraid, hesitant, timid, not confident, not necessarily linked with something going on in the environment (i.e. you may see the body tremble, flared nostrils, tail clamped)
Happy	Feeling, showing or expressing joy, pleased, lively, playful, satisfied
Look for contact	Actively looking for interaction, interested, close proximity, eager to approach
Relaxed	Not tense or rigid, easy-going, tranquil
Pushy	Assertive or forceful (i.e. not leaving space, head butting out of the way, exhibits dominant behaviour, may be mouthy or nippy)
Uneasy	Afflicted, uncomfortable, unsettled, restless

Frustration

Acute - Conflict	The horse shows some resistance to the stimulus or a person including avoiding interaction, trying to move away, being hesitant
Acute - Displacement	The horse engages in another behaviour such as bed eating, increased eating or drinking, active looking away, mouthing items or people, undertaking crib-biting, wind sucking, weaving or repetitive pacing) rather than engaging with the stimulus
Acute - Vocalisation	The horse produces a sound
Acute - redirected aggression	The horse is aggressive to something that is not the stimulus being presented
Acute - direct aggression	The horse is aggressive to the stimulus being presented
Acute - increased locomotion	The horse moved around (this may be anything from one step to an entire movement sequence)
Acute - muscle tension	The horse becomes tense in its posture
Chronic - repetitive behaviours	The horse engages in new repetitive behaviours for a prolonged period of time
Chronic - learned helplessness	The horse is unresponsive and inappetent for a prolonged period of time
Chronic - altered cognitive bias	The horse shows a prolonged change in judgement about the stimulus

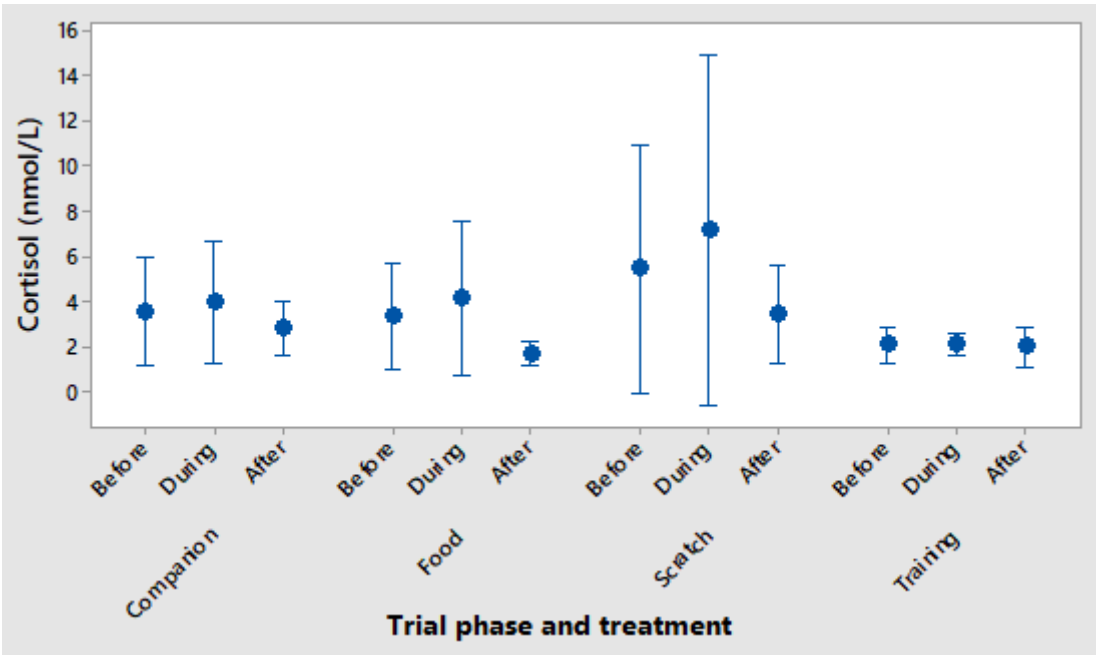
Appendix I – Chapter 6 - Additional graphs

I-1



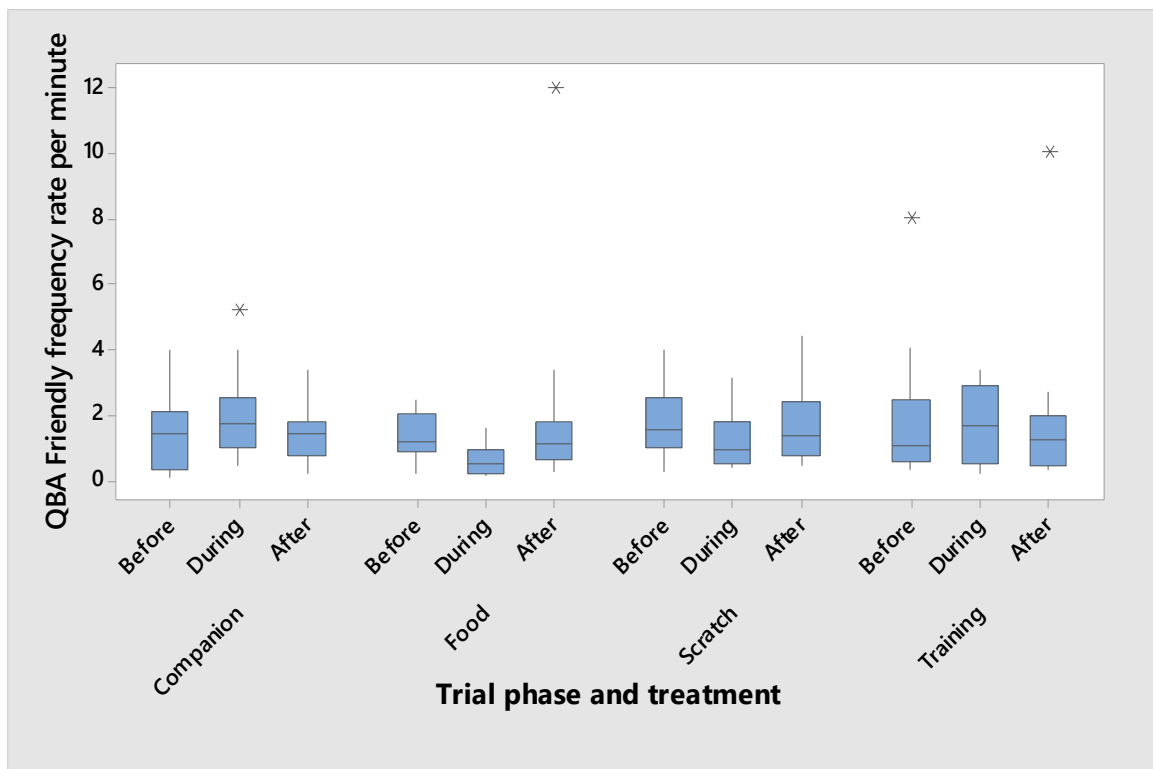
Interval plots with 95% confidence interval of left and right eye and ear temperatures over trial phase and treatment.

I-2



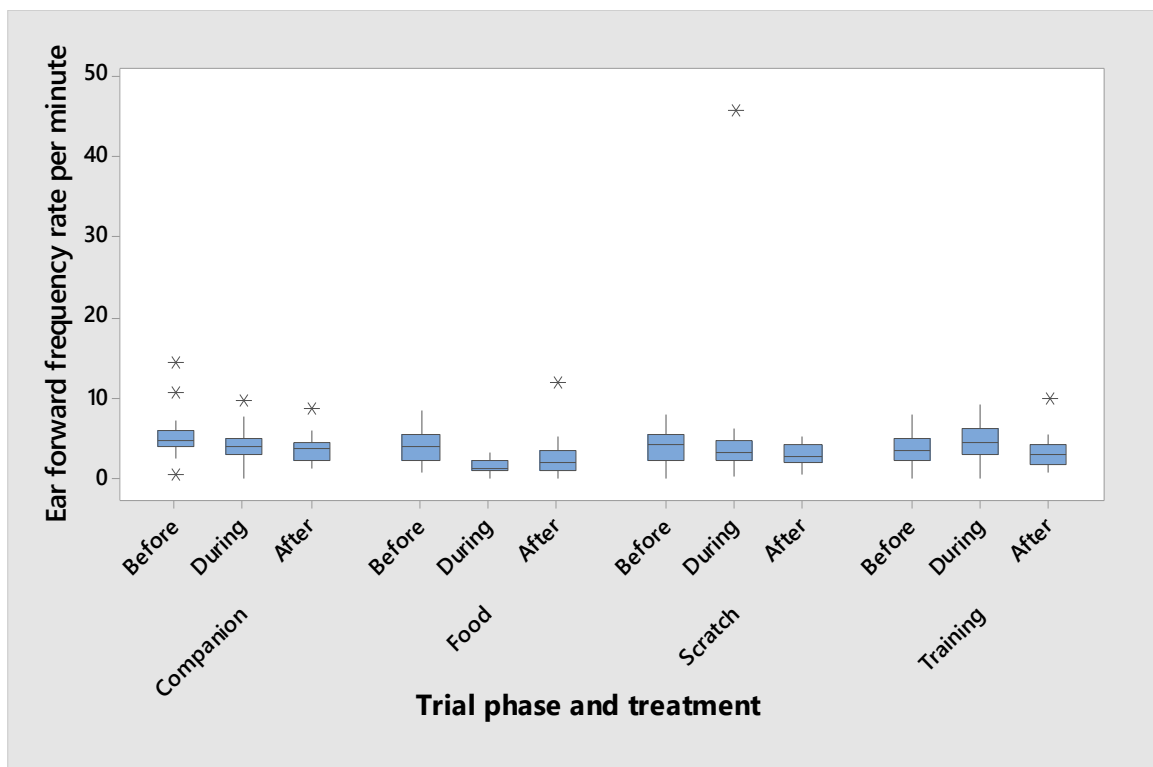
Interval plot with 95% confidence intervals of cortisol data over trial phase and treatment

I-3



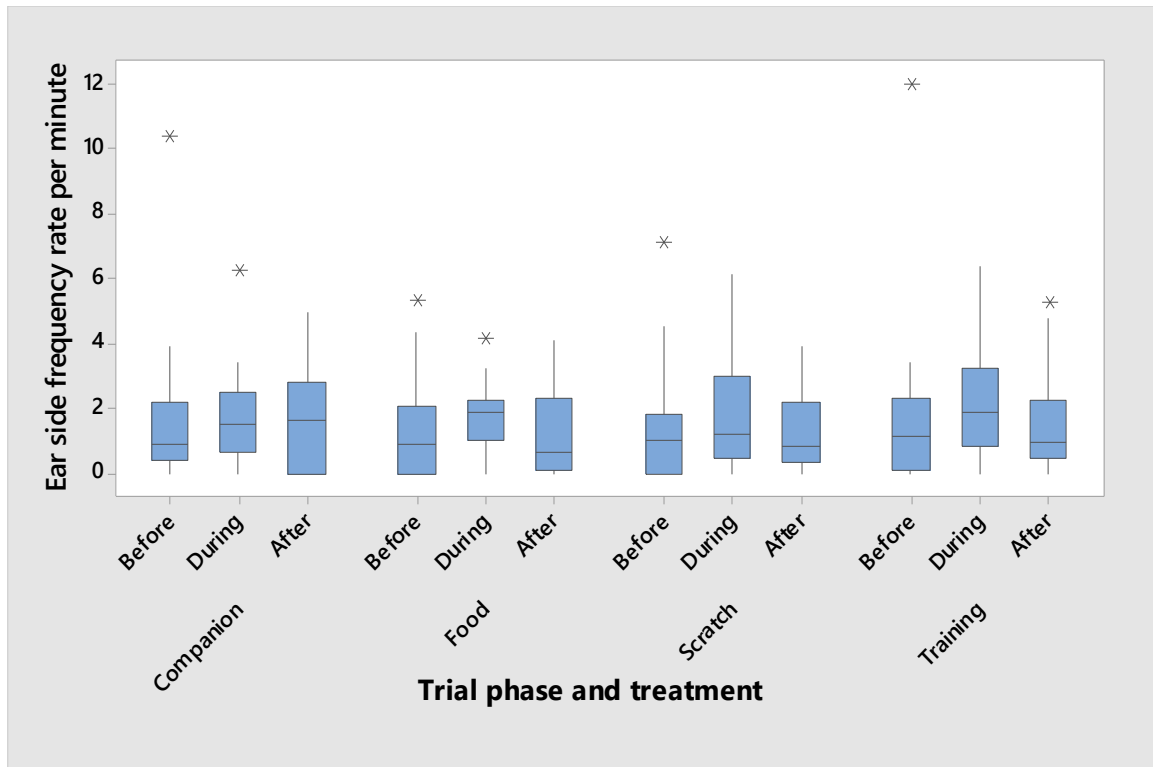
Boxplot of QBA friendly over trial phase and treatment

I-4



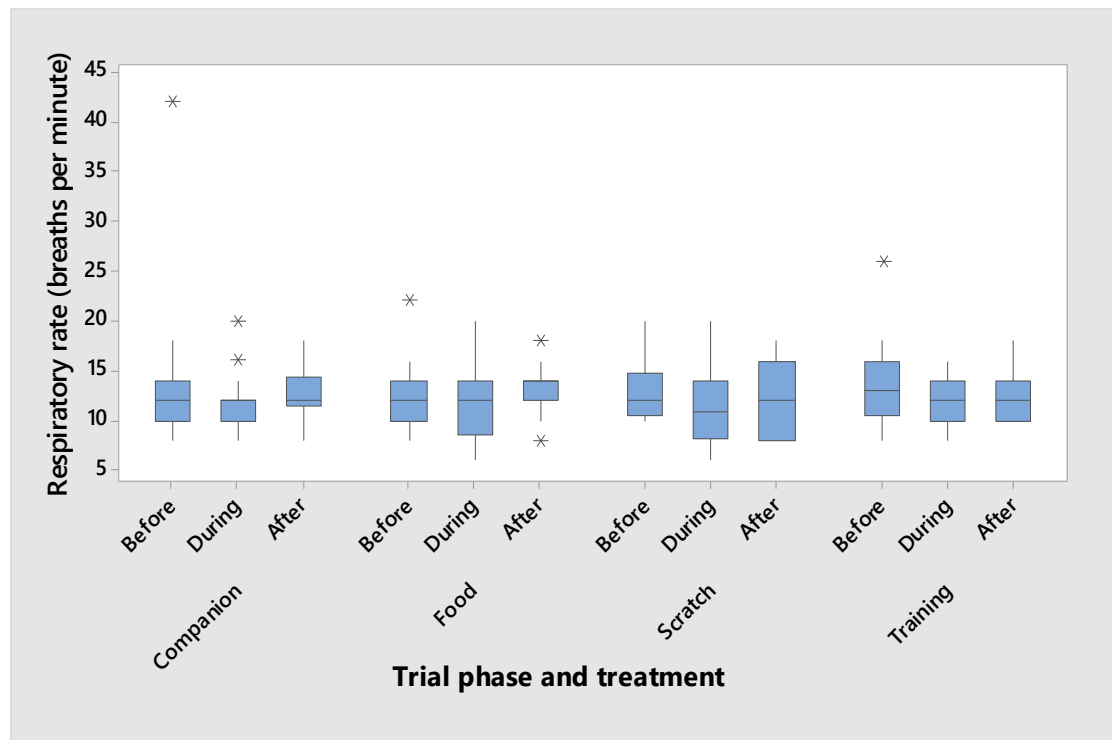
Boxplot of ears forward frequency over trial phase and treatment

I-5



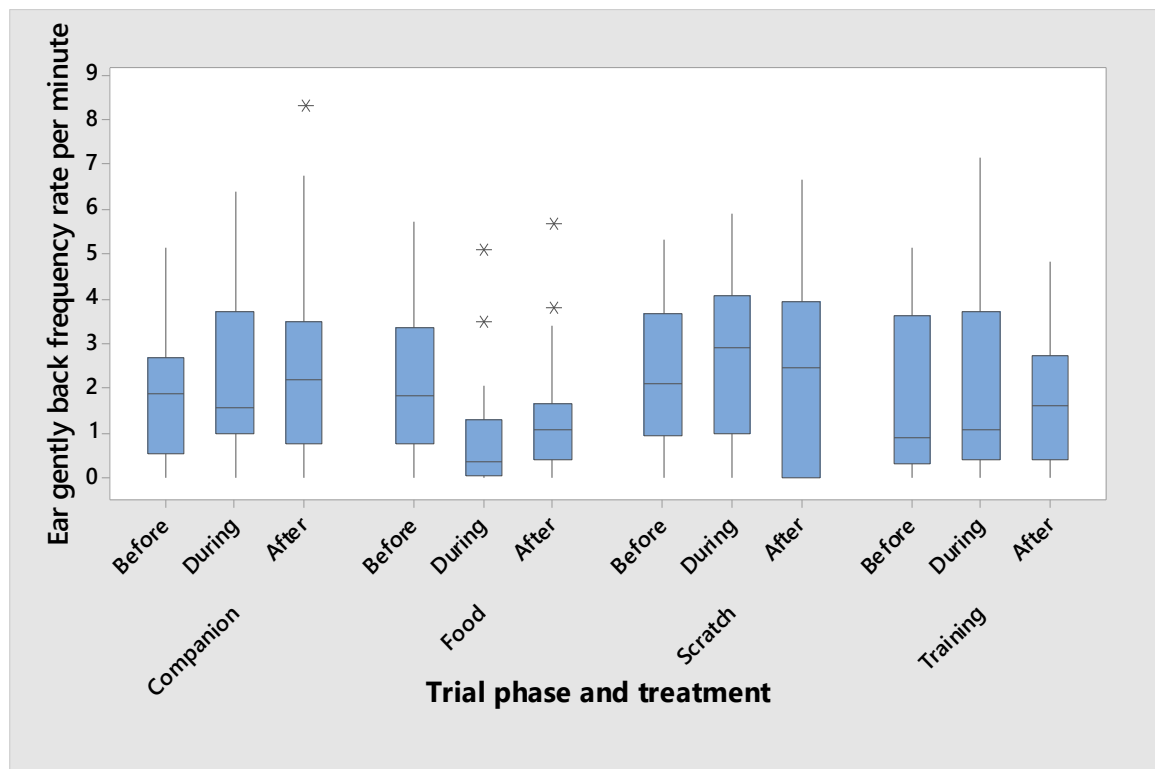
Boxplot of ears side frequency over trial phase and treatment

I-6



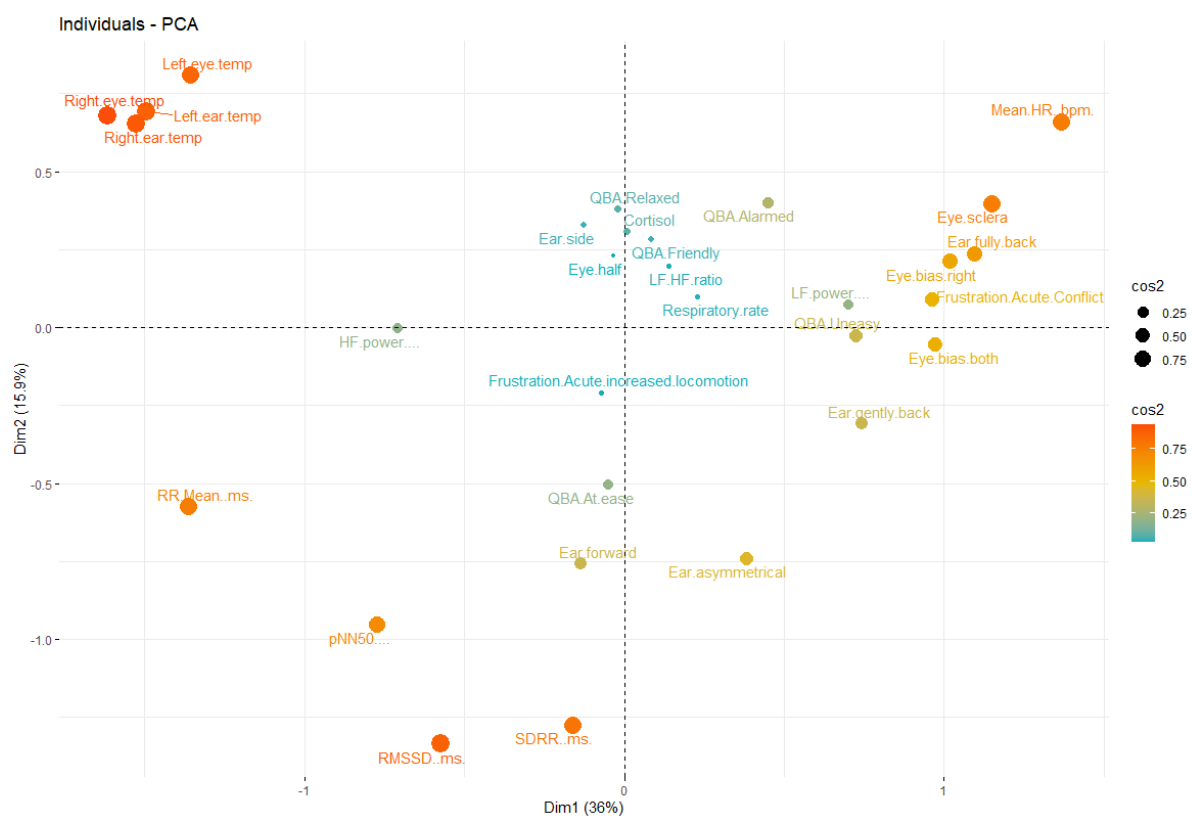
Respiratory rate across trial phase and treatment

I-7



Boxplot of ears gently back frequency across trial phase and treatment

I-8



PCA loading plot with ear and eye temperature added in retrospectively.