

Research Article

Cite this article: Merkies K and Trudel K (2024). How well can you tell? Success of human categorisation of horse behavioural responses depicted in media. *Animal Welfare*, 33, e50, 1–11
<https://doi.org/10.1017/awf.2024.55>

Received: 29 July 2024
 Revised: 07 October 2024
 Accepted: 10 October 2024

Keywords:

affective state; animal welfare; human-horse interactions; interoception; overt behaviours; subtle behaviours

Corresponding author:

Katrina Merkies;
 Email: kmerkies@uoguelph.ca

Author contributions:

Conceptualisation: KM, KT; Data curation: KT, KM; Formal analysis: KM, KT; Investigation: KT, KM; Methodology: KM, KT; Project administration: KM; Supervision: KM; Writing – original draft: KT, KM; Writing – review & editing: KM, KT

© The Author(s), 2024. Published by Cambridge University Press on behalf of The Universities Federation for Animal Welfare. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



Twitter: @UFAW_1926
 webpage: <https://www.ufaw.org.uk/>

How well can you tell? Success of human categorisation of horse behavioural responses depicted in media

Katrina Merkies and Katelyn Trudel

Department of Animal Biosciences and Campbell Centre for the Study of Animal Welfare, University of Guelph Guelph, ON N1G 2W1, Canada

Abstract

Horses employ a range of subtle to overt behaviours to communicate their current affective state. Humans who are more cognisant of their own bodily sensations may be more attuned to recognising affective states in horses (*Equus caballus*) thereby promoting positive human-horse interactions. This study investigated human ability to categorise human-horse interactions depicted in media relative to equine behaviour experts and compared participant scores to their level of interoception. Using an online survey, participants (n = 534) categorised 31 photographs and videos as (overt) positive, likely (subtle) positive, neutral, likely (subtle) negative or (overt) negative human-horse interactions from the horse's point of view and completed the Multidimensional Assessment of Interoceptive Awareness questionnaire (MAIA-2) to assess their level of interoception. Demographic information was also collected (age, gender, education, level of experience with horses, location). Participants differed from expert categorisations of horse affective states across all categories, exactly matching experts only 52.5% of the time and approximately matching experts for positive and negative valence 78.5% of the time. The MAIA-2 did not predict participant ability to accurately categorise human-horse interactions. Women outperformed men in categorising overt positive, overt negative and subtle negative human-horse interactions. Increased levels of education and greater experience with horses were associated with improved categorisation of certain human-horse interactions. More training or awareness is needed to recognise behavioural indicators of horse affect to guide appropriate human-horse activities that impact horse welfare.

Introduction

The modern human-horse relationship spans a large spectrum of interactions which can be characterised by the context and duration of the contact (Goodwin 1999; Hausberger *et al.* 2008). Some human-horse interactions, such as veterinary care or farriery, are brief and infrequent — but still contribute to a horse's perception of human interactions (Hausberger *et al.* 2008). Conversely, more long-term human-horse interactions, like ownership, allow the development and maintenance of the human-horse bond (Hausberger *et al.* 2008; Merkies *et al.* 2024). Horses (*Equus caballus*) involved in riding lesson programmes or in equine-assisted services may experience a degree of consistency in their connection with students or clients across multiple sessions or programmes but also a regular turnover in clients creating a mixture of both long- and short-term human-horse interactions (Hausberger *et al.* 2008; Arrazola & Merkies 2020; Ekholm Fry 2021). Due to this repeated turnover of human participants, it is imperative to ensure that human-horse interactions are on the whole positive, as a horse that experiences frequent negative interactions with people may become fearful or resistant, making them more difficult to handle, risking both human and horse safety (Waiblinger *et al.* 2006; Luke *et al.* 2022).

From the horse's point of view, interactions with humans involve most of their senses: visual, tactile, olfactory and auditory (Rørvang *et al.* 2020; Merkies & Franzin 2021). Individual horses may perceive each interaction as positive, neutral or negative, however we cannot rely upon self-reporting to confirm these subjective emotions as we do in humans; instead we must rely upon the horse's outward expressions of behaviour as a proxy for their true emotional valence (Russell 2003). Horses are highly expressive animals, and their affective state can be inferred by analysing both observable and measurable qualities, such as their physical behaviours and vocalisations, or physiological markers such as heart rate (de Santis *et al.* 2017; Merkies & Franzin 2021). Physical messaging can range from obvious signals like biting or kicking to subtle signals like changing ear positioning or tightening of the lips (McDonnell & Haviland 1995; Pierard *et al.* 2019). Correctly deciphering these behaviours is central to inferring the valence of the horse's affective state (positive, neutral or negative) and thereby encouraging positive interactions for both the horse and the human while protecting human safety and improving the overall welfare of the horse (Ladewig 2019).

Affiliative behaviours are indicative of long-term positive affective states in farm animals (Hausberger & Muller 2002). They promote welfare through building the human-horse bond and reducing the incidence of aggressive behaviours (Boissy *et al.* 2007). Animals that exhibit affiliative behaviours may be labeled as 'relaxed', 'friendly', 'affectionate', 'playful' or 'social' by their caretakers (Boissy *et al.* 2007). Affiliative behaviours from horses directed toward humans can include allogrooming, approaching, olfactory investigation, following human direction, nudging or licking (Pierard *et al.* 2019; Arrazola & Merkies 2020). Within these physical indicators of positive affective states, there are subtleties that may not be immediately noticeable to someone unfamiliar with horse behaviour, like slight turning towards a human or loosening of the lower lip (Arrazola & Merkies 2020; de Oliveira *et al.* 2021).

Agonistic behaviours are characterised by the use of force or aggression to terminate an interaction and when observed it can then be assumed that the current interaction was negative or unpleasant for the instigating horse (McDonnell & Haviland 1995; Briefer Freymond *et al.* 2013; Ekholm Fry 2021). Typically, horses only exhibit agonistic behaviours to the minimum degree required to terminate an unpleasant interaction prior to returning to a normal state, but behaviours will escalate as required until the negative interaction ceases (Briefer Freymond *et al.* 2013; Burla *et al.* 2016). Agonistic behaviours can include both the action itself or the threat of the same action (McDonnell & Haviland 1995; McGreevy 2012). For example, a bite is a common agonistic behaviour across a variety of species (e.g. dogs, cats, horses), but the threat of a bite also represents an agonistic behaviour even if no contact is made (McDonnell & Haviland 1995; McGreevy 2012). Other agonistic behaviours include: (threat of) kicking, (threat of) striking, alert posture and approach, pinned ears, head threat (head lowered with neck extended and ears pinned) or tightened facial musculature (McDonnell & Haviland 1995; McGreevy 2012; Pierard *et al.* 2019). More passive avoidance strategies to cease an interaction may be indicated by subtle behaviours like turning away to increase distance between the horse and the stressor, even though the horse may not be fully removed from the situation (McDonnell & Haviland 1995; McGreevy 2012; Pierard *et al.* 2019; Ekholm Fry 2021). Both agonistic and avoidant behavioural responses exhibited by horses are assumed to be indicative of a negative affective state and embody the physical display of the fight or flight response (Ekholm Fry 2021).

The physical expression of positive or negative affective states are subjective to each individual horse's temperament but dampened by training which acts to suppress unwanted behaviours (Seaman *et al.* 2002; McGreevy *et al.* 2009). Within these individual differences horses may be categorised as either 'active' or 'passive' copers (Wechsler 1995; Seaman *et al.* 2002). The major difference between coping styles is that passive copers do not present major outward signs of aversion to stimuli, with agonistic or avoidant behaviours being muted (Wechsler 1995; Seaman *et al.* 2002). In this passive and often immobile state, animals are attempting to go undetected while waiting for the aversive stimuli to pass (Wechsler 1995). This passive reaction may be mistaken for a calm state, as no obvious negative behaviours are detected, but may not be a true representation of the horse's affective state (Wechsler 1995; Squibb *et al.* 2018). On the other hand, active copers may overtly display their response to a stressor (Budzyńska 2014) which can alert the human to change their approach to a situation.

It is advantageous for humans to be aware of various behaviours their horse displays, nonetheless they are often oblivious to them, especially the more subtle behaviours (Bell *et al.* 2019). When assessing an animal's emotional state, individuals may bias their

evaluation through their personal, subjective lens and imbue their observations with anthropomorphism (Bradshaw & Casey 2007; Thompson & Clarkson 2019). The complexity and subjectivity when assessing an animal's affective state has been flagged as an issue, and even when using accepted behavioural markers, many studies and welfare assessments have struggled with inter-observer reliability (Green & Mellor 2011; Yeates 2011).

Empathy is the ability to relate to another being's emotions which implies that one is aware of their own bodily sensations to be able to resonate with another being (Marson *et al.* 2024). Humans more readily empathise with other beings viewed as similar or that they can relate to, and anthropomorphism may act as a catalyst to endow other beings with human qualities thereby making them more like us (Vanutelli & Balconi 2015). As Thompson and Clarkson (2019) argue, the application of anthropomorphism may actually strengthen interspecies relationships as we can see that horses are like us in some ways, although unlike us in others. Empathy involves processing and recognising one's own internal states, making interoception an important component of empathy (Ernst *et al.* 2013). Interoception is defined as the ability in which one can sense and interpret what is going on within their own body, which can include either conscious or subconscious processes (Mehling *et al.* 2018). Neural pathways involved in affective states are linked to neural pathways involved in physiological responses to the point that interoceptive awareness is believed to drive emotional behaviours meant to fulfill social needs (Craig 2016). Those possessing a higher level of bodily awareness are likely to be more empathetic (Li *et al.* 2024). Specifically, those with a higher interoceptive sensibility were better able to recognise facial expressions of emotions in other people (Hübner *et al.* 2021).

This study investigated whether humans with a higher self-awareness of their own bodily state are better able to recognise a horse's affective state. To the authors' knowledge, the influence of interoception on the evaluation of affective states of horses has not been researched. Using a survey methodology whereby participants viewed a number of media depicting horses interacting with humans in various scenarios, their ability to recognise behavioural cues indicating positive, neutral and negative affective states in horses was compared to expert evaluations of each scenario. The percent agreement of participant responses was then related to their scores on the Multidimensional Assessment of Interoceptive Awareness (Version 2) (MAIA-2). The MAIA-2 is one of the most popular validated psychometric scales that serves as a self-reported measure of the participant's interoception (Vig *et al.* 2022). The 37 MAIA-2 items represent eight subcategories, termed factors, of interoception: noticing, not-distracting, not-worrying, attention regulation, emotional awareness, self-regulation, body listening and trust (Mehling *et al.* 2018). It was hypothesised that those scoring higher in the MAIA-2 would better categorise interactions between humans and horses because they are more aware of their own emotional responses and affective states.

Materials and methods

Ethical approval

This research protocol was approved by the institutional Research Ethics Board for the use of humans in research (REB21-12-026).

Survey development: Media selection and categorisation

All media (photographs and video) were collected from researchers' personal files. Media depicted positive, negative and neutral horse

Table 1. Brief descriptions of selected media as agreed upon by equine behaviour experts ($n = 7$) and included in the final survey to assess how well humans can recognise horse behavioural responses to human interactions. Categories of affective state were divided into positive, negative and neutral, with overt and subtle scenarios for positive and negative situations. Media explanations describe what the horse is doing in the scenario in response to human interactions

Categorisation	Description	Media type	Video length (s)
Overt positive	Horse following target for food reward	Video	8.0
Overt positive	Foal sniffing human crouched down	Video	7.5
Overt positive	Horse being embraced by human	Image	—
Overt positive	Horse lowering head towards human	Image	—
Overt positive	Horse drinking water from hose held by human	Video	8.0
Overt positive	Allogrooming during human petting	Video	9.0
Overt positive	Horse following target for food reward	Video	8.0
Subtle positive	Horse reaching toward humans over fence	Image	—
Subtle positive	Olfactory investigation of human	Image	—
Subtle positive	Horse reaching neck over stall door towards human	Image	—
Subtle positive	Horse reaching neck over stall door towards human	Image	—
Subtle positive	Horse angling head and neck towards human	Image	—
Subtle positive	Horse resting head in human's arms	Image	—
Neutral	Horse having ear bonnet put on	Video	5.0
Neutral	Horse standing for human to mount	Video	8.0
Neutral	Horse standing in crossties	Video	8.0
Neutral	Horse having hind foot picked out	Video	5.0
Neutral	Horse being bridled	Video	9.0
Neutral	Horse being cross-tied	Video	11.0
Neutral	Horse being bridled	Video	8.0
Subtle negative	Horse being ridden through gate	Video	6.0
Subtle negative	Horse being ridden through water	Video	4.0
Subtle negative	Horse being caught in field	Video	7.0
Overt negative	Horse being asked to cross tarp on ground	Video	10.0
Overt negative	Horse being loaded onto trailer	Video	10.0
Overt negative	Horse being brushed	Video	10.0
Overt negative	Horse having bit size measured	Video	10.0
Overt negative	Horse being petted over stall door	Video	7.0
Overt negative	Horse avoiding being caught in paddock	Video	20.0
Overt negative	Horse raising head high away from human	Image	—
Overt negative	Horse being sprayed with fly spray	Video	9.0

behavioural responses to humans in a variety of human-horse interactions. Media were selected to portray both overt and subtle responses from the horse. The overt category contained obvious affiliative, agonistic or avoidant behaviours, such as a horse trying to bite a human. Conversely, the subtle category contained less noticeable physical responses from the horse such as a horse backing away while the human is trying to place a halter on their head. To account for potential dampened expression of positive or negative behavioural responses due to the horse's training, and at the suggestion of the experts, instances of passive coping were grouped as neutral responses to human-horse interactions.

An initial media bank ($n = 77$ pieces) was distributed to seven external equine behaviour experts for validation to confirm the media categorisations. External experts comprised equine

behaviour researchers who had an average of 50 peer-reviewed articles focused on horse behaviour. For media categorisations to be considered validated, at least 85% of the experts (i.e. six of the seven external validators) must have agreed with the categorisation. The final survey contained 31 pieces of media (Table 1) which included 22 videos and nine images.

The videos ranged from 4–20 s in length, averaging 8.5 s each, with a total running time of 187.5 s (approximately 3 min). All human faces or identifying features were blurred to respect the privacy of those depicted in the media. In some instances, other horses that were not interacting with the humans in frame were also blurred to avoid any confusion for participants regarding which horse they should be assessing. To focus on the human ability to read solely the physical responses of horses, only visual cues were

explored, and all audio was removed from videos to avoid influence from horse vocalisations (e.g. whinny, nicker, snort) or human commentary.

Data collection

The survey was created using Qualtrics (Qualtrics, Provo, UT, USA), a web-based survey platform. The survey was divided into four main sections: (1) Informed consent; (2) Participant demographics; (3) Categorisation of the 31 media pieces; and (4) the Multidimensional Assessment of Interoceptive Awareness (Version 2) (MAIA-2) (for survey questions, see S1 in the Supplementary material).

Participants were recruited using snowball sampling across relevant social media platforms (e.g. Facebook, Instagram and LinkedIn). Recruitment letters were distributed to university faculty and students via e-mail lists and advertisements were placed in equine newsletters (e.g. Equine Guelph). The survey was accessible either by a URL link or QR code. The survey was open to any participant aged 18 or older and previous experience with horses was not required. The survey was available from June 30 to July 27 2022 (total of 29 days).

After informed consent, the survey collected demographic data from participants including age, gender, education, country of residence and experience with horses. The survey then moved into the media categorisations, however prior to starting this section, participants were provided with the following definitions of positive, neutral and negative human-horse interactions:

Positive interaction: The horse exhibits a positive response to human interaction. The horse may seem attentive, engaged or appear to enjoy the interaction with the human.

Neutral interaction: The horse does not exhibit a noticeable response to human interaction. The horse may seem disengaged while still following human direction or participating in the interaction.

Negative interaction: The horse exhibits a negative response to human interaction. The horse may seem fearful, avoidant or attempt to stop the interaction.

The 31 pieces of media were presented to participants in randomised order. All nine images were uploaded directly onto Qualtrics and all 22 videos were uploaded to YouTube (www.youtube.com) then embedded in Qualtrics using HTML code. YouTube controls were left on, allowing the participants to replay the video as often as required before selecting their answer. Each piece of media was displayed independently from the others and once an answer had been submitted participants were unable to return to previous questions to change their answers. For each media piece, participants were asked to categorise the experience from the horse's point of view as Positive, Likely Positive, Neutral, Likely Negative or Negative. To avoid biasing participant categorisation of media, the adjective 'overt' was omitted, and the adjective subtle was replaced with 'likely'.

Two attention check questions were included within the media analysis section, roughly splitting it into three parts. The attention checks served to monitor the participant's attention to detail and flag potential instances of the participants selecting answers at random before moving on to the next question.

Each media question collected data pertaining to the participant's time budget while completing the survey which included the time of their first and last click per question, total click counts, and the time of page submission. At the end of the media section, an open-ended question asked participants to describe what

behavioural clues from the horse they used to classify both positive and negative human-horse interactions (e.g. ear position).

The final part of the survey consisted of the Multidimensional Assessment of Interoceptive Awareness (Version 2) (MAIA-2) to assess participants' interoception (Mehling *et al.* 2018). The MAIA-2 consists of 37 items and the participant selected how often they identified with each item on a scale of 0 to 5, 0 being never and 5 being always. For example, participants were presented with the following prompt: "When I feel pain or discomfort, I try to power through it", then asked to rate the degree to which they relate to the prompt. This section was purposely placed last to avoid any 'emotional warm-up' the participants could have experienced, as participant self-reflection of their own feelings and emotions prior to media assessment could better prepare them to evaluate the horses' reactions to human interaction. Of the 37 questions within the MAIA-2, each belonged to a subset of one to seven questions that paired with the one following eight factors: noticing, not-distracting, not-worrying, attention regulation, emotional awareness, self-regulation, body listening and trust. The MAIA-2 questions were presented in random order to each participant and can be seen in S2 in the Supplementary material.

Data curation

The attention checks, in conjunction with the total time participants spent viewing the media, served as a means of response exclusion. If the participant spent less than the total video run time on the entire media section (approximately 3 min) it was assumed that the media were not viewed in their entirety and random answers were selected. In this case the participant's data were removed. If the participant failed one or both attention checks, their data were flagged and compared to the total time spent in the section. The data from any participant that failed one or both attention checks and had a view time less than the media run time for that section were excluded. Some data-sets were incomplete, which represented participants who skipped questions and did not complete the entire survey prior to submission. If a participant's data-set was only missing responses within the MAIA-2 section, their data were still used for analysis within the media section but excluded during subsequent analyses directly related to the MAIA-2. Incomplete questionnaires did not represent withdrawal of consent, as participants who wished to withdraw were instructed not to submit the survey and their data were not recorded. Withdrawal after submission was not possible as no identifying information was collected to link a response to a particular person.

Data analysis

All analyses were conducted using SAS (v9.4, SAS Institute, Cary, NC, USA) with significance levels specified at $\alpha = 0.05$. Expert categorisations of all media pieces were considered to be 100% correct. The ability of participants to accurately characterise the human-horse interactions compared to the expert categorisations was analysed using two criteria. The first criterion required the participant to exactly match the expert categorisation of any media piece (expert exact match). The second criterion accepted participant responses that were close to expert categorisations (expert close match): for example, if the experts categorised a media piece as a 'positive' interaction, participant responses of both 'positive' or 'likely positive' were accepted as correct. Similarly, if the experts categorised a media piece as 'negative', participant responses of both 'negative' or 'likely negative' were accepted as correct. If the

experts categorised a media piece as 'neutral', participant responses of 'neutral', 'likely negative' or 'likely positive' were all accepted as correct. The percent agreement between the expert and participant categorisations for both criteria were separately analysed for each media piece using a Pearson's Chi-squared test. A General Linear Mixed Model with participant as a random factor examined the main factors of participant age, gender, education and level of experience with horses on their percent agreement with expert categorisations under both criteria. Tukey-Kramer *post hoc* multiple comparisons further examined significant results between the percent agreement of participant responses relative to the experts and the main factors.

The MAIA-2 scores were calculated as described by Mehling *et al.* (2018) and as available in the public domain (www.osher.ucsf.edu/maia). Total scores represent the average of a six-point scale across all 37 questions and all scores ranged from zero to five. Each of the eight factors (e.g. Noticing) that contributed to the total MAIA-2 score also had a factor-specific score. Factor-specific scores were determined by taking the average response value across all factor-specific questions (e.g. Q1–Q4 are questions specific to the 'Noticing' factor). Higher total scores in the MAIA-2 represent a higher degree of interoception, and higher factor-specific scores represent a higher aptitude for each factor. Pearson's product-moment correlations were run individually to determine any relationship with total and factor-specific MAIA-2 scores on the percent agreement between the participants' and the experts' categorisations of the media for both criteria.

Responses to the open-ended questions describing what behavioural clues from the horse participants used to classify both positive and negative human-horse interactions were analysed qualitatively using content analyses. Each response was coded using *a priori* themes of horse body parts (eyes, ears, face, head and neck, body, tail), engagement with or avoidance of the human, and general demeanour describing either positive or negative affective states. The frequency of the mention of each theme was calculated as the number of mentions divided by the total number of responses.

Results

A total of 618 surveys were received, with 84 surveys excluded as being incomplete or failing the time or attention tests, resulting in 534 useable surveys. Not every participant answered every question as all questions were optional.

Participant demographics

Participants ($n = 534$) were mostly females ($n = 482$; 90.3%) residing in Canada ($n = 395$; 75.2%). Participant age was fairly evenly distributed across categories (18–25 years: $n = 111$; 20.8%; 26–35 years: $n = 122$; 22.8%; 36–45 years: $n = 84$; 15.7%; 46–55 years: $n = 77$; 14.4%; > 55 years: $n = 139$; 26.0%). The majority of participants indicated they had advanced ($n = 174$; 32.6%) or expert ($n = 199$; 37.3%) level horse experience. Most participants were well-educated having completed a college or university degree ($n = 196$; 36.7%) or pursued further qualifications ($n = 222$; 41.6%), with 21.8% ($n = 117$) having no college or university education.

Percent agreement of participants to experts' categorisations

Criterion 1

Expert categorisations were considered 100% correct and under Criterion 1 participants had to match expert categorisations exactly.

Participant categorisations ranged from 32.5–87.5% agreement and did not reflect expert exact matches for any of the 31 media pieces ($\chi^2 = 1,392.8$, $df = 4$; $P < 0.0001$). Overall expert exact matches by participants occurred 52.5% of the time. However, participants performed better for certain scenarios relative to others. On average, participants were better at identifying overt positive and overt negative human-horse interaction scenarios relative to subtle positive, neutral and subtle negative categorisations (Figure 1). Participant categorisations in only four questions exceeded 80% agreement to expert categorisations and these were all in the overt negative category (Q 27, 29, 30 and 31).

Criterion 2

Expert categorisations were considered 100% correct and under Criterion 2 participants had to closely match expert categorisations. For example, if the participant selected likely positive but the correct answer was overt positive, it would still be considered correct. Participant percent agreement was higher under Criterion 2 compared to Criterion 1 ($P < 0.0001$ for all categories) but still did not closely match expert categorisation for any question ($\chi^2 = 470.0$, $df = 3$; $P < 0.0001$). On the whole, participants closely matched expert categorisations 78.5% of the time with a range of 15.3–98.5%. Participants were somewhat better at correctly identifying positive scenarios (80.4% agreement) compared to neutral (75.4% agreement) or negative (78.2% agreement) scenarios.

Influence of participant age

Criterion 1

Participants who were older than 55 years had lower percent agreement to expert exact matches of neutral scenarios ($F_{4,518} = 5.24$; $P = 0.0004$) compared to participants under 46 years. Participant age did not influence percent agreement in any other category (all $P > 0.1000$).

Criterion 2

Participants who were older than 55 years had lower percent agreement to expert close matches of neutral ($F_{4,518} = 3.73$; $P = 0.0053$) and subtle negative ($F_{4,518} = 3.74$; $P = 0.0052$) scenarios compared to

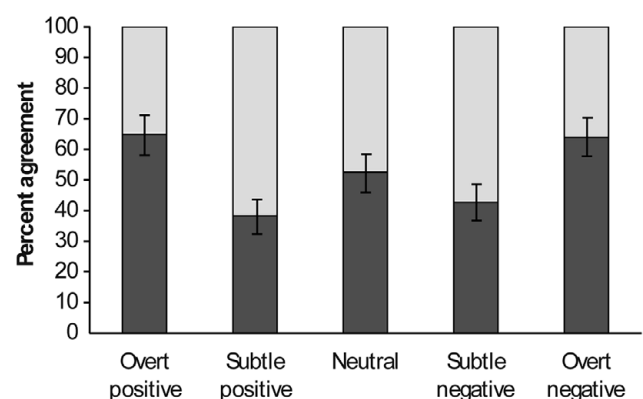


Figure 1. Mean percentage (dark grey bars [\pm SEM]) of human participants ($n = 534$) exactly matching categorisations by experts (considered to be 100% correct; light grey bars) of various positive, negative and neutral scenarios of human-horse interactions. Responses did not match expert categorisations in any of the categories ($P < 0.0001$).

participants aged 26–35 years. Participant age did not influence percent agreement in any other category (all $P > 0.2100$).

Influence of participant gender

Criterion 1

Women outperformed men in exactly matching expert categorisations of overt positive ($F_{2,518} = 3.83$; $P = 0.0223$) human-horse interaction scenarios. Women tended to outperform men in exactly matching expert categorisations of subtle negative ($F_{2,518} = 2.35$; $P = 0.0961$) and overt negative ($F_{2,518} = 3.35$; $P = 0.0963$). Participant gender did not influence percent agreement in any other category (all $P > 0.2200$).

Criterion 2

Women outperformed men in closely matching expert categorisations of overt positive ($F_{2,518} = 3.75$; $P = 0.0242$) and subtle negative ($F_{2,518} = 5.92$; $P = 0.0029$) human-horse interaction scenarios. Women tended to outperform men in closely matching expert categorisations of overt negative ($F_{2,518} = 2.87$; $P = 0.0577$) scenarios. Participant gender did not influence percent agreement in any other category (all $P > 0.1200$).

Influence of participant level of education

Criterion 1

Education did not influence participants' exact expert matches of overt positive ($F_{2,518} = 0.73$; $P = 0.4801$) and subtle negative ($F_{2,518} = 1.04$; $P = 0.3542$) human-horse interaction scenarios compared to expert categorisations. However, participants with more education did better at exactly matching expert categorisations of subtle positive ($F_{2,518} = 3.52$; $P = 0.0304$), neutral ($F_{2,518} = 11.60$; $P < 0.0001$) and overt negative ($F_{2,518} = 4.52$; $P = 0.0114$) scenarios than those with less education (Figure 2).

Criterion 2

Participants with more education did better at closely matching expert categorisations of neutral ($F_{2,518} = 13.18$; $P < 0.0001$), subtle negative ($F_{2,518} = 4.34$; $P = 0.0135$) and overt negative ($F_{2,518} = 3.73$; $P = 0.0246$) scenarios than those with less education. Participant

education did not influence percent agreement in any other category (all $P > 0.1000$).

Influence of horse experience

Criterion 1

Participants with more experience with horses performed better at exactly matching the experts than those with less horse experience when categorising both overt positive ($F_{4,518} = 2.72$; $P = 0.0290$) and overt negative ($F_{4,518} = 4.38$; $P = 0.0017$) human-horse interaction scenarios (Figure 3). Horse experience did not affect the percentage of exact matches of subtle positive, neutral or subtle negative scenarios (all $P > 0.2770$).

Criterion 2

Participants with more experience with horses performed better at closely matching the experts than those with less horse experience when categorising subtle negative ($F_{4,518} = 6.87$; $P < 0.0001$) human-horse interaction scenarios. Participants with more experience with horses tended to perform better at closely matching the experts than those with less horse experience when categorising subtle positive ($F_{4,518} = 2.19$; $P = 0.0689$) human-horse interaction scenarios. Horse experience did not affect the percentage of closely matched categorisations of overt positive, neutral or overt negative scenarios (all $P > 0.5735$).

Influence of MAIA-2 scores

Participant responses to the MAIA-2 are found in Table 2. Participant scores for any of the MAIA-2 factor-scales were not correlated to the percent agreement of categorisations of various scenarios of human-horse interactions compared to expert categorisations for either exact matches (Criterion 1; $r = 0.087$, $n = 455$; $P = 0.0631$) or close matches (Criterion 2; $r = -0.062$, $n = 455$; $P = 0.1989$).

Qualitative analysis

Participants were asked to describe what clues from the horse's behaviour they used to categorise both positive and negative human-horse interactions. Responses involving the ears, eyes, face, head and neck, body, tail and movement appeared in both the

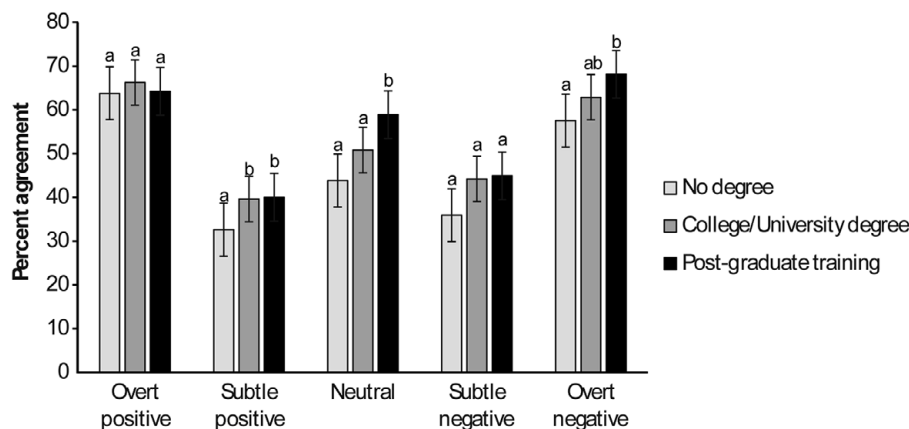


Figure 2. Mean (\pm SEM) percentages of human participants ($n = 534$) exactly matching expert categorisations (considered to be 100% correct) of various positive, negative and neutral scenarios of human-horse interactions according to their level of education. Within a category, bars with different superscripts (a,b) differ significantly ($P < 0.031$).

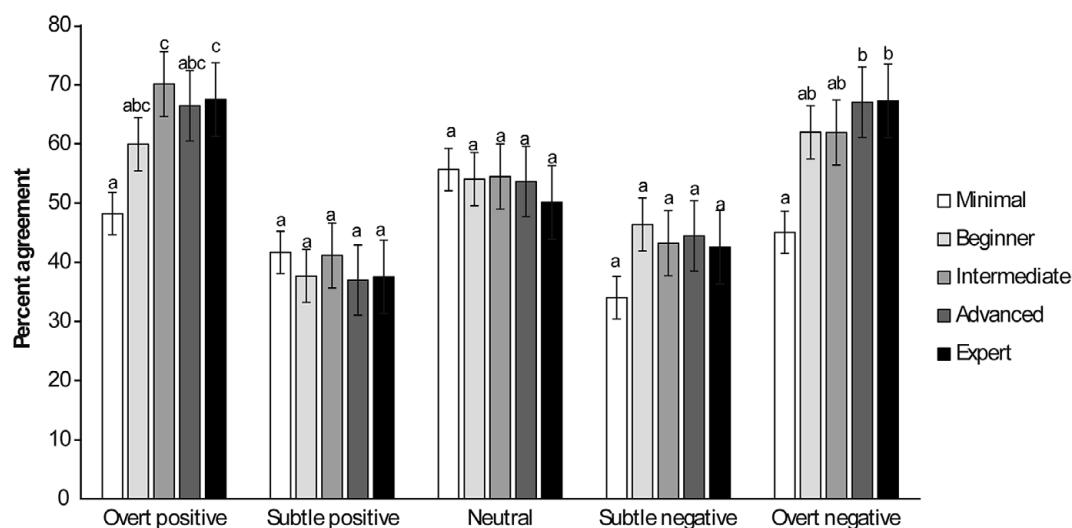


Figure 3. Mean (\pm SEM) percentages of human participants ($n = 534$) exactly matching categorisations by experts (considered to be 100% correct) of varying depictions of positive, negative and neutral human-horse interactions relative to their level of horse experience. Within a category, bars with different superscripts (a,b,c) differ significantly ($P < 0.03$).

Table 2. Average human participant ($n = 534$) scores for each factor scale answered within the Multidimensional Assessment of Interoceptive Awareness questionnaire (MAIA-2). Scores ranged from 0–5 with higher scores indicating higher interoception for that scale

MAIA factor scale	Mean	SD
Total score	4.08	0.617
Noticing	4.64	0.888
Not distracting	3.21	1.044
Not worrying	3.78	0.947
Attention regulation	4.13	0.811
Emotional awareness	4.36	0.816
Self regulation	3.76	1.027
Body listening	4.16	1.008
Trusting	4.56	0.864

positive and negative descriptors provided by participants (Table 3).

Along with body parts, participants also referred to whether the horse was engaged with the human ($n = 271$; 59.3%) during positive interactions or avoided ($n = 229$; 50.2%) the human during negative interactions. Engagement was described as seeking human attention, voluntarily approaching, leaning in or staying close to the human, initiating contact, nuzzling, doing what the human asked of them and their ability to leave if they chose. Seventeen respondents (3.7%) specifically referred to the horse as ‘happy’ while the majority of respondents ($n = 286$; 62.6%) described the demeanour of the horse in other ways, including curious, inquisitive, calm and relaxed. A small number of respondents ($n = 11$; 2.4%) explicitly mentioned the actions of the human contributing to their assessment of the scenario, such as maintaining a loose lead rope connection, being relaxed themselves or providing treats or release cues to the horse.

Avoidance was described as being evasive, looking away, pulling back, trying to get away from human, displaying defensive reactions or being uncooperative to the human’s request. The demeanour of the horse included descriptors such as fear, tension, disinterest,

jaded, aggression, stress, anxious, inhibited, reactive, unsure, unhappy, irritated, apprehensive and distant. Fifteen respondents (3.3%) indicated the actions of the human being responsible for the horse’s response, either by the human pulling on the reins or using a lot of tension on the lead rope, forcing the horse to do something they did not want to do or using tools to restrict motion, not ensuring the horse understood their request, not being aware of the horse’s reactions or being rushed.

Discussion

The results presented here corroborate findings from other researchers in that generally humans are not very successful in distinguishing the affective state of horses during their interactions with humans (Bell *et al.* 2019; Rogers & Bell 2022; Luke *et al.* 2023). Participants did better at correctly characterising the valence of the interactions as positive or negative but may have lacked the finesse required to observe the more subtle signs. In particular, neutral responses to human interaction were challenging for the participants to identify in horses. Age, gender, education and horse experience all affected the participants’ ability to categorise the horses’ affective states. Older participants struggled more to correctly categorise neutral scenarios, women generally outperformed men, and those with more education and more horse experience were better able to categorise certain scenarios. How well a participant was in touch with their own inner bodily sensations did not impact their categorisations of the affective states of the horses. Participants provided many descriptors of horse behaviour to qualify their categorisations, mostly relying upon the ears and eyes of the horse and how engaged they were with the human. While participants were able to define various behaviours the horse might portray, there seemed to be a disconnect between recognising the behaviours and linking them to affective states.

In this study, the expert categorisations were taken as 100% correct, although others have shown that experts do not always agree. Pearson *et al.* (2021) showed a poor agreement among equine behaviour experts and veterinarians when assessing behavioural indicators of stress in horses undergoing veterinary treatment. Bell *et al.* (2019) showed imperfect consensus of horse affective state among six equine behaviour experts rating videos. However,

Table 3. Responses provided by survey participants (n = 457) for both positive and negative indicators of horse affective state after viewing media pieces (n = 31). Frequencies represent the number of participants (n; %) who indicated that they utilised these physical horse traits to categorise human-horse interactions during the media analysis. Qualifiers are examples of the qualitative descriptions respondents wrote

Comment	Positive frequency	Positive qualifiers	Negative frequency	Negative qualifiers
Ears	n = 308; 67.0%	Forward, relaxed, soft, toward person, up, floppy, straight,	n = 272; 60.0%	Back, pinned, fixed, pressed
Eyes	n = 147; 32.2%	Soft, relaxed, calm, curious, on human, closing, interested, gentle, kind, quiet, sleepy	n = 191; 41.9%	Whale eye, open wide, bulging, alert, empty, tension, hard, rolling, anxious, no eye contact
Face	n = 131; 28.7%	Lick and chew, relaxed facial muscles, loose jaw/lips, facial expression, soft	n = 90; 19.7%	Nostrils flared, curling lips, mouth clamped, tight lips, facial tension, pain face, grimace
Head & neck	n = 110; 24.1%	Neutral position, head low, relaxed, neck stretched toward person	n = 143; 31.4%	Head tossing, raised head, head shy, tension in neck, fussy with head
Body	n = 112; 24.5%	Relaxed posture, engaged, inviting, no visible tension, soft, body language	n = 93; 20.4%	Tense posture, stiff, increased muscle tone, body language, body leaning away
Tail	n = 14; 3.1%	Still, quiet, swaying, wagging, relaxed	n = 39; 9.0%	Flicking, raised, tense, swishing, whipping, clamped
Movement	n = 42; 9.2%	Standing still, movement toward person, standing near person, moves calmly, relaxed movement	n = 261; 57.2%	Escape movements, frozen, jerky, agitated, flight response, skittish, frantic, sudden, pulling away, quick, jumpy

Pannewitz and Loftus (2023) reached consensus among 30 equine behaviour experts on behavioural indicators of frustration in horses. Young *et al.* (2012) found the ratings of thirteen equine professionals on the behavioural responses of horses in various stressful situations (e.g. clipping, fireworks) using a simple behaviour scale ranging from no stress (1) to extreme stress (10) to reliably correspond to both heart rate and cortisol measures from those horses. Since at least six of our seven (> 85%) experts had to agree on the behavioural categorisation of the media, and those media that did not reach that benchmark were not included in the survey, we can be fairly confident that the included media depicted the assigned affective state. However, this resulted in our media not being divided evenly among the various categories, with fewer media pieces in slightly negative scenarios. This could imply that even our experts had difficulty interpreting these scenarios since there were fewer media pieces agreed upon in this category.

On average, our survey participants matched expert categorisations of the affective states of the horses depicted in the media only 52.5% of the time, which does not differ much from chance. Similar to Bell *et al.* (2019), participants were better at recognising overt positive and negative behaviours the horses displayed but were less discerning of the subtle behaviours. This lack of awareness of subtle signals can lead to escalation of dangerous behaviours and compromise both human and horse safety (Bell *et al.* 2019; Rogers & Bell 2022).

When analysed under broader standards (i.e. Criterion 2), our survey participants' responses improved although they still did not match the expert responses. Under this less stringent analysis, it was clear the participants recognised the valence of the horses' affective states as positive or negative even if they missed the more subtle signals. Participants struggled most with identifying neutral situations where the horse did not appear responsive to or engaged with the human, highlighting how difficult it can be to accurately identify both subtle and neutral horse behaviours. A horse's response to human interaction is a combination of their own temperament, previous experiences with humans, individual coping styles and training (Wechsler 1995; Seaman *et al.* 2002; McGreevy *et al.* 2009; Bell *et al.* 2019; Hausberger *et al.* 2021). These factors may lead to minimal behavioural responses that may mask the horse's affective

state. It could be that participants misidentified neutral human-horse interactions to be more positively valenced (Lesimple & Hausberger 2014; Bell *et al.* 2019). If the horse depicted in the neutral media was a passive copier, this may improperly translate to humans as the horse being calm or relaxed, which are desirable traits (Bell *et al.* 2019). Recent research suggests that relying solely upon physical behaviours may not be a reliable indicator of horse affect during stressful situations or interactions as physiological measures may still indicate a stress response (Squibb *et al.* 2018). For the purposes of this study, human-horse interactions depicting a horse that was disengaged from or minimally participating in the interaction with a human was assigned a neutral categorisation by the experts, but this highlights that despite expert consensus, it may still be an inaccurate categorisation of the true valence of affect being experienced by the horse.

In general, age did not influence the participants' ability to match expert categorisation of human-horse interactions with the exception of those older than 55 years performing poorer than younger participants when categorising neutral (Criteria 1 and 2) and subtle negative (Criterion 2) scenarios. Older participants viewing photographs of cats were similarly less successful in identifying positive or negative affect than younger participants (Dawson *et al.* 2019). Others have shown that the ability to recognise emotions in human facial expressions decreased with age, most particularly when regarding images of neutral or negative emotions (Malykhin *et al.* 2023).

In this study, women outperformed men in correctly categorising obvious positive and negative affective states in horses and tended to outperform men in correctly categorising subtle indicators of affective state. The participants in this study represented more females than males, which is typical of the equine industry in general (Dashper 2016; Fenner *et al.* 2019). More women than men believe that horses can feel emotions (Hötzel *et al.* 2019) and this, together with the fact that women are more empathic toward animals than men (e.g. Christov-Moore *et al.* 2014), may account for this result. Similarly, women were better at decoding the valence of horse whinnies than men (Merkie *et al.* 2021).

Survey participants with a higher level of education were better able to categorise some horse affective states than those with less

education. However, Pannewitz and Loftus (2023) showed no difference among equine behaviour professionals with or without a PhD in identifying behavioural indicators of frustration in horses. Similarly, education was not found to be a significant factor in human ability to distinguish affective states in cats (Dawson *et al.* 2019). Formal education in general is believed to increase human cognitive abilities and was shown to be a stronger factor than experience in recognition of the importance of ecosystems (Lima & Bastos 2019). However, the influence of formal education on recognition of horse behaviours remains inconclusive.

Within the overt positive and negative categories, participants with more horse experience did perform better than those with little horse experience, although this pattern did not appear in the subtle positive, neutral or subtle negative categories. Horse experience has been related to the ability to distinguish horse behaviours with conflicting outcomes. Fox (2023) showed that lay people performed just as well as practitioners of equine-assisted services when describing the affective states of horses in videos even though they reported a lower level of horse experience. Contrarily, Braun *et al.* (2024) showed that those with more horse experience did better at characterising the affective state of horses in photographs compared to inexperienced people, although they still only correctly identified affective state 50% of the time. Fourth year veterinary students were better at characterising horse behaviour than first year students with less self-reported horse experience (Guinnefollau *et al.* 2019). Participants with more experience performed better when assessing the videos depicting in-hand dressage and behavioural rehabilitation but not when assessing ridden horse videos (Bell *et al.* 2019). It could be that experience interplays with cognitive bias – we process information and make decisions based upon our own perceptions and memories, which are led by our experiences (Azzopardi 2021). Those who spend extended periods of time with horses tend to overconfidently assess their horses' affective state as positive and misinterpret indicators of negative affect (Lesimple & Hausberger 2014; Bell *et al.* 2019; Bornmann *et al.* 2021) and those involved in the daily care of horses underestimate the expression of negative affect as they are likely desensitised to behaviours that, in their opinion, do not affect a horse's overall well-being (Lesimple & Hausberger 2014).

Empathy is the ability to recognise the affective state of another and to respond with appropriate emotion. Despite there being a close relationship between empathy seen between humans and interoception (Ernst *et al.* 2013; Hübner *et al.* 2021) results from this study showed no correlation between participant interoception as assessed by the MAIA-2 and their ability to correctly categorise the affective states of horses interacting with humans. The MAIA-2 is a commonly used measure of interoception (Desmedt *et al.* 2022; Vig *et al.* 2022) and has demonstrated relationships to interpreting facial expressions of others (Hübner *et al.* 2021). To the authors' knowledge, the MAIA-2 has not been used to evaluate interpretation of affective states of animals. The lack of correlation between participants' interoception and their ability to correctly categorise human-horse interactions could indicate that the MAIA-2 is not transferrable across species, as even experts struggle to reach a consensus regarding which behavioural markers suggest a horse may be stressed (Pearson *et al.* 2021). An alternate explanation is that people's perceptions of their own thoughts and emotions are more subjective than their perception of others, as observed within the human medical field where practitioners tend to underestimate their patients' pain (Marquié *et al.* 2003).

Typically, those actively involved in equestrian activities pride themselves on being able to read their horse, often attributing

anthropomorphic states to their horse's expression of emotion (Hötzel *et al.* 2019). Our results did not show much evidence of anthropomorphism in the qualitative responses provided by participants. Despite performing poorly in recognising subtle positive and negative categories, some survey participants were still able to describe subtle indicators of affect (e.g. muscular tension). However, most qualitative descriptions were either excessively broad (e.g. ears, eyes) or referred to more overt indicators of behaviour (e.g. biting). Similarly, Bell *et al.* (2019) concluded that participants were more likely to list overt indicators of behaviour although some would be able to identify more subtle cues, like changes in eye aperture (Bell *et al.* 2019). Survey participants gauged horse affective state mainly in terms of expression of specific body parts. In particular, the eyes and ears were referred to most often as indicative of how a horse was feeling. More body descriptors were suggested to evaluate negative affective state (e.g. head and neck position) than positive affective state, and more comments in general were provided for negative states. This supports the idea that identifying positive affective states in horses is much more challenging than identifying negative affective states (Zeitler-Feicht *et al.* 2024). Participants also voiced more concern over human actions creating or exacerbating the scenarios leading to negative affective states in the horse than scenarios leading to positive affective states. It is reassuring that survey participants were able to recognise the harmful actions of humans in the scenarios but disappointing that overall they were still unable to characterise when horses were experiencing negative affective states. A similar study evaluating human ability to recognise distress in horses also found that some participants would be comfortable with their horse being subjected to the same treatment or situation that could result in negative affect (Bell *et al.* 2019).

Rapid advances in technology make it now increasingly possible to decipher horse behaviour using artificial intelligence. Simple detection of key body points and their change over time allows for gross determination of animal activity as an indicator of health and well-being (Kleanthous *et al.* 2022). More sophisticated approaches utilise deep learning to determine patterns and features specific to affective states reflecting potential pain in horses based on more frequent changes in posture (Martin-Cirera *et al.* 2024) and the Horse Grimace Scale (Lencioni *et al.* 2021). While these techniques are very promising for behaviour recognition and would drastically reduce time commitment for assessment in a non-invasive manner, the methodology still remains in its infancy (Rohan *et al.* 2024).

Animal welfare implications

To safeguard both human and horse welfare during human-horse interactions, it is imperative that humans recognise positive human-horse interactions to promote their reoccurrence, and negative human-horse interactions to avoid subjecting horses to unnecessary stress. However, similar to other studies, our survey participants were unable to successfully deduce horse affective state in various scenarios compared to expert categorisations. Participants were more successful at identifying overt indicators of affect over subtle indicators and could generally infer the positive or negative valence of a scenario. Human literature suggests that those who are more aware of their own internal body sensations may have more empathy for others. To our knowledge, this is the first study that has attempted to relate human interoception to identification of horse affective states. However, our results showed participants'

interoception determined by the MAIA-2 questionnaire did not impact their ability to match expert categorisations of human-horse interactions. Despite participants being able to qualitatively describe negative affect in horses, this did not translate to their ability to characterise affective state, underscoring the subconscious disconnect between behavioural signals and affective states. Future research would benefit from focusing on subtle indicators of horse affect, to identify if participants misconstrue subtle indicators of negative affect as neutral, or even positive, as many involved in horse care inaccurately identify or ignore negative indicators of affect. Although participants with more self-reported experience with horses performed better at identify affective states, continued improvement in education and awareness of horse behaviour will ensure continued improvement for horse welfare.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/awf.2024.55>.

Competing interest. None.

References

- Arrazola A and Merkies K** 2020 Effect of human attachment style on horse behaviour and physiology during equine-assisted activities—a pilot study. *Animals* **10**(7): 1–13. <https://doi.org/10.3390/ani10071156>
- Azzopardi L** 2021 Cognitive biases in search: A review and reflection of cognitive biases in information retrieval. *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval (CHIIR '21)* pp 27–37. Association for Computing Machinery: New York, NY, USA. <https://doi.org/10.1145/3406522.3446023>
- Bell C, Rogers S, Taylor J and Busby D** 2019 Improving the recognition of equine affective states. *Animals* **9**(12): 1124. <https://doi.org/10.3390/ani9121124>
- Boissy A, Manteuffel G, Jensen MB, Moe RO, Spruijt B, Keeling LJ, Winckler C, Forkman B, Dimitrov I, Langbein J, Bakken M, Veissier I and Aubert A** 2007 Assessment of positive emotions in animals to improve their welfare. *Physiology and Behavior* **92**(3): 375–397. <https://doi.org/10.1016/j.physbeh.2007.02.003>
- Bornmann T, Randle H and 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>
- Bradshaw J and Casey R** 2007 Anthropomorphism and anthropocentrism as influences in the quality of life of companion animals. *Animal Welfare* **16** (S1): 149–154. <https://doi.org/10.1017/S0962728600031869>
- Braun MN, Müller-Klein A, Sopp MA, Michael T, Link-Dorner U and Lass-Hennemann J** 2024 The human ability to interpret affective states in horses' body language: The role of emotion recognition ability and previous experience with horses. *Applied Animal Behaviour Science* 106171. <https://doi.org/10.1016/j.applanim.2024.106171>
- Briefer Freymond S, Briefer EF, von Niederhäusern R and Bachmann I** 2013 Pattern of social interactions after group integration: A possibility to keep stallions in group. *PLoS One* **8**(1): e54688. <https://doi.org/10.1371/journal.pone.0054688>
- Budzyńska M** 2014 Stress reactivity and coping in horse adaptation to environment. *Journal of Equine Veterinary Science* **34**: 935–941. <https://doi.org/10.1016/j.jevs.2014.05.010>
- Burla JB, Ostertag A, Patt A, Bachmann I and Hillmann E** 2016 Effects of feeding management and group composition on agonistic behaviour of group-housed horses. *Applied Animal Behaviour Science* **176**: 32–42. <https://doi.org/10.1016/j.applanim.2016.01.011>
- Christov-Moore L, Simpson EA, Coudé G, Grigaityte K, Iacoboni M and Ferrari PF** 2014 Empathy: gender effects in brain and behavior. *Neuroscience and Biobehavioral Reviews* **46**(4): 604–627. <https://doi.org/10.1016/j.neubiorev.2014.09.001>
- Craig AD** 2016 Interoception and emotion: A neuroanatomical perspective. In: Barrett LF, Lewis M and Haviland-Jones JM (eds) *Handbook of Emotions, Fourth Edition*. Guilford Press: New York, NY, USA.
- Dashper K** 2016 Human-Animal Relationships in Equestrian Sport and Leisure pp 194. Routledge: London, UK. <https://doi.org/10.4324/9781315678139>
- Dawson LC, Cheal J, Niel L and Mason G** 2019 Humans can identify cats' affective states from subtle facial expressions. *Animal Welfare* **28**(4): 519–531. <https://doi.org/10.7120/09627286.28.4.519>
- de Oliveira AR, Gozalo-Marcilla M, Ringer SK, Schauvliege S, Fonseca MW, Trindade PHE, Filho JNPP and Luna SPL** 2021 Development and validation of the facial scale (FaceSed) to evaluate sedation in horses. *PLoS One* **16** (6): e0251909. <https://doi.org/10.1371/journal.pone.0251909>
- de Santis M, Contalbrigo L, Borgi M, Cirulli F, Luzi F, Redaelli V, Stefani A, Toson M, Odore R, Vercelli C, Valle E and Farina L** 2017 Equine assisted interventions (EAIs): Methodological considerations for stress assessment in horses. *Veterinary Sciences* **4**(3): 44. <https://doi.org/10.3390/vetsci4030044>
- Desmedt O, Heeren A, Corneille O and Luminet O** 2022 What do measures of self-report interoception measure? Insights from a systematic review, latent factor analysis, and network approach. *Biological Psychology* **169**: 108289. <https://doi.org/10.1016/j.biopsycho.2022.108289>
- Ekhholm Fry N** 2021 Welfare considerations for horses in therapy and education services. *The Welfare of Animals in Animal-Assisted Interventions* pp 219–242. Springer International Publishing: London, UK. https://doi.org/10.1007/978-3-030-69587-3_9
- Ernst J, Northoff G, Böker H, Seifritz E and Grimm S** 2013 Interoceptive awareness enhances neural activity during empathy. *Human Brain Mapping* **34**(7): 1615–1624. <https://doi.org/10.1002/hbm.22014>
- Fenner K, Caspar G, Hyde M, Henshall C, Dhand N, Probyn-Rapsey F, Dashper, K, McLean A and McGreevy P** 2019 It's all about the sex, or is it? Humans, horses and temperament. *PLoS One* **14**(5): e0216699. <https://doi.org/10.1371/journal.pone.0216699>
- Fox S** 2023 Accuracy of horse affect assessments: a comparison of equine assisted mental health professionals, non-equine assisted mental health professionals, and laypeople. Master's Thesis, Mount Royal University, Lethbridge, AB, Canada.
- Goodwin D** 1999 The importance of ethology in understanding the behaviour of the horse. *Equine Veterinary Journal* **31**(28): 15–19. <https://doi.org/10.1111/j.2042-3306.1999.tb05150.x>
- Green TC and Mellor DJ** 2011 Extending ideas about animal welfare assessment to include “quality of life” and related concepts. *New Zealand Veterinary Journal* **59**(6): 263–271. <https://doi.org/10.1080/00480169.2011.610283>
- Guinefolou L, Gee EK, Bolwell CF, Norman EJ and Rogers CW** 2019 Benefits of animal exposure on veterinary students' understanding of equine behaviour and self-assessed equine handling skills. *Animals* **9**: 620. <https://doi.org/10.3390/ani9090620>
- Hausberger M, Lesimple C and Henry S** 2021 Detecting welfare in a non-verbal species: Social/cultural biases and difficulties in horse welfare assessment. *Animals* **11**(8): 2249. <https://doi.org/10.3390/ani11082249>
- Hausberger M and Muller C** 2002 A brief note on some possible factors involved in the reactions of horses to humans. *Applied Animal Behaviour Science* **76**: 339–344. [https://doi.org/10.1016/S0168-1591\(02\)00016-3](https://doi.org/10.1016/S0168-1591(02)00016-3)
- Hausberger M, Roche H, Henry S and Visser EK** 2008 A review of the human-horse relationship. *Applied Animal Behaviour Science* **109**(1): 1–24. <https://doi.org/10.1016/j.applanim.2007.04.015>
- Hötzel MJ, Vieira MC and Leme DP** 2019 Exploring horse owners' and caretakers' perceptions of emotions and associated behaviors in horses. *Journal of Veterinary Behavior* **29**: 18–24. <https://doi.org/10.1016/j.jveb.2018.10.002>
- Hübner AM, Trempler I, Gietmann C and Schubotz RI** 2021 Interoceptive sensibility predicts the ability to infer others' emotional states. *PLoS One* **16** (10): e0258089. <https://doi.org/10.1371/journal.pone.0258089>
- Kleantous N, Hussain AJ, Khan W, Sneddon J, Al-Shamma'a A and Liatsis P** 2022 A survey of machine learning approaches in animal behaviour. *Neurocomputing* **491**: 442–463. <https://doi.org/10.1016/j.neucom.2021.10.126>
- Ladewig J** 2019 Body language: Its importance for communication with horses. *Journal of Veterinary Behavior* **29**: 108–110. <https://doi.org/10.1016/j.jveb.2018.06.042>
- Lencioni GC, de Sousa RV, de Souza Sardinha EJ, Corrêa RR and Zanella AJ** 2021 Pain assessment in horses using automatic facial expression recognition through deep learning-based modeling. *PLoS One* **16**(10): e0258672. <https://doi.org/10.1371/journal.pone.0258672>

- Lesimple C and Hausberger M 2014 How accurate are we at assessing others' well-being? The example of welfare assessment in horses. *Frontiers in Psychology* 5. <https://doi.org/10.3389/fpsyg.2014.00021>
- Li X, Peng C, Qin F, Luo Q, Ren Z, Wang X, Feng Q, Liu C, Li Y, Wei D and Qiu J 2024 Basolateral amygdala functional connectivity in alexithymia: Linking interoceptive sensibility and cognitive empathy. *Neuroscience* 539: 12–20. <https://doi.org/10.1016/j.neuroscience.2023.12.014>
- Lima FP and Bastos RP 2019 Perceiving the invisible: Formal education affects the perception of ecosystem services provided by native areas. *Ecosystem Services* 40: 101029. <https://doi.org/10.1016/j.ecoser.2019.101029>
- Luke KL, McAdie T, Smith BP and Warren-Smith AK 2022 New insights into ridden horse behaviour, horse welfare and horse-related safety. *Applied Animal Behaviour Science* 246: 105539. <https://doi.org/10.1016/j.applanim.2021.105539>
- Luke KL, Rawluk A, McAdie T, Smith BP and Warren-Smith AK 2023 How equestrians conceptualise horse welfare: Does it facilitate or hinder change? *Animal Welfare* 32: e59. <https://doi.org/10.1017/awf.2023.79>
- Malykhin N, Pietrasik W, Aghamohammadi-Sereshki A, Ngan Hoang K, Fujiwara E and Olsen F 2023 Emotional recognition across the adult lifespan: Effects of age, sex, cognitive empathy, alexithymia traits, and amygdala subnuclei volumes. *Journal of Neuroscience Research* 101(3): 367–383. <https://doi.org/10.1002/jnr.25152>
- Marquié L, Raufaste E, Lauque D, Mariné C, Ecoiffier M and Sorum P 2003 Pain rating by patients and physicians: evidence of systematic pain miscalibration. *Pain* 102(3): 289–296. [https://doi.org/10.1016/S0304-3959\(02\)00402-5](https://doi.org/10.1016/S0304-3959(02)00402-5)
- Marson F, Revital N-Z, Paoletti P, Glicksohn J, Harris T, Elliott MA, Carducci F and Ben-Soussan TD 2024 When the body fosters empathy: The interconnectivity between bodily reactivity, meditation, and embodied abstract concepts. *Progress in Brain Research*. <https://doi.org/10.1016/bs.pbr.2024.05.004>
- Martin-Cirera A, Nowak M, Norton T, Auer U and Oczak M 2024 Comparison of Transformers with LSTM for classification of the behavioural time budget in horses based on video data. *Biosystems Engineering* 242: 154–168. <https://doi.org/10.1016/j.biosystemseng.2024.04.014>
- McDonnell SM and Haviland JCS 1995 Agonistic ethogram of the equid bachelor band. *Applied Animal Behaviour Science* 43(3): 147–188. [https://doi.org/10.1016/0168-1591\(94\)00550-X](https://doi.org/10.1016/0168-1591(94)00550-X)
- McGreevy P 2012 *Equine Behavior: A Guide for Veterinarians and Equine Scientists, Second Edition* pp 378. Elsevier Health Sciences: Amsterdam, The Netherlands.
- McGreevy PD, Oddie C, Burton FL and McLean AN 2009 The horse-human dyad: Can we align horse training and handling activities with the equid social ethogram? *Veterinary Journal* 181(1): 12–18. <https://doi.org/10.1016/j.tvjl.2009.03.005>
- Mehling WE, Acree M, Stewart A, Silas J and Jones A 2018 The multidimensional assessment of interoceptive awareness, version 2 (MAIA-2). *PLoS One* 13(12): e0208034. <https://doi.org/10.1371/journal.pone.0208034>
- Merkies K, Crouchman E and Belliveau H 2021 Human ability to determine affective states in domestic horse whinnies. *Anthrozoös* 35(3): 483–494. <https://doi.org/10.1080/08927936.2021.1999605>
- Merkies K and Franzin O 2021 Enhanced understanding of horse–human interactions to optimize welfare. *Animals* 11: 1–14. <https://doi.org/10.3390/ani11051347>
- Merkies K, Hayman B and Ijichi CL 2024 Examining the human-horse bond from the human perspective. *Anthrozoös* 37(2): 231–244. <https://doi.org/10.1080/08927936.2024.2308403>
- Pannewitz L and Loftus L 2023 Frustration in horses: Investigating expert opinion on behavioural indicators and causes using a Delphi consultation. *Applied Animal Behaviour Science* 258: 105818. <https://doi.org/10.1016/j.applanim.2022.105818>
- Pearson G, Waran N, Reardon RJM, Keen J and 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: 105291. <https://doi.org/10.1016/j.applanim.2021.105291>
- Pierard M, McGreevy P and Geers R 2019 Effect of density and relative aggressiveness on agonistic and affiliative interactions in a newly formed group of horses. *Journal of Veterinary Behavior* 29: 61–69. <https://doi.org/10.1016/j.jveb.2018.03.008>
- Rogers S and Bell C 2022 Perceptions of fear and anxiety in horses as reported in interviews with equine behaviourists. *Animals* 12: 2904. <https://doi.org/10.3390/ani12212904>
- Rohan A, Saad Rifaq M, Junayed Hasan Md, Asghar F, Kashif Bashir A and Dottorini T 2024 Application of deep learning for livestock behaviour recognition: A systematic literature review. *Computers and Electronics in Agriculture* 224: 109115. <https://doi.org/10.1016/j.compag.2024.109115>
- Rørvang MV, Nielsen BL and McLean AN 2020 Sensory abilities of horses and their importance for equitation science. *Frontiers in Veterinary Science* 7: 633. <https://doi.org/10.3389/fvets.2020.00633>
- Russell JA 2003 Core affect and the psychological construction of emotion. *Psychological Review* 110(1): 145–172. <https://doi.org/10.1037/0033-295X.110.1.145>
- Seaman SC, Davidson HPB and Waran NK 2002 How reliable is temperament assessment in the domestic horse (*Equus caballus*)? *Applied Animal Behaviour Science* 78(2): 175–191. [https://doi.org/10.1016/S0168-1591\(02\)00095-3](https://doi.org/10.1016/S0168-1591(02)00095-3)
- Squibb K, Griffin K, Favier R and Ijichi C 2018 Poker face: Discrepancies in behaviour and affective states in horses during stressful handling procedures. *Applied Animal Behaviour Science* 202: 34–38. <https://doi.org/10.1016/j.applanim.2018.02.003>
- Thompson K and Clarkson L 2019 How owners determine if the social and behavioral needs of their horses are being met: Findings from an Australian online survey. *Journal of Veterinary Behavior* 29: 128–133. <https://doi.org/10.1016/j.jveb.2018.12.001>
- Vanutelli ME and Balconi M 2015 Empathy and prosocial behaviours. Insights from intra- and inter-species interactions. *Rivista Internazionale di Filosofia e Psicologia* 6(1): 88–109. <https://doi.org/10.4453/rifp.2015.0007>
- Vig L, Köteles F and Ferentzi E 2022 Questionnaires of interoception do not assess the same construct. *PLoS One* 17(8): e0273299. <https://doi.org/10.1371/journal.pone.0273299>
- Waiblinger S, Boivin X, Pedersen V, Tosi MV, Janczak AM, Visser EK and Jones RB 2006 Assessing the human-animal relationship in farmed species: A critical review. *Applied Animal Behaviour Science* 101: 185–242. <https://doi.org/10.1016/j.applanim.2006.02.001>
- Wechsler B 1995 Coping and coping strategies: a behavioural view. *Applied Animal Behaviour Science* 43(2): 123–134. [https://doi.org/10.1016/0168-1591\(95\)00557-9](https://doi.org/10.1016/0168-1591(95)00557-9)
- Yeates JW 2011 Is “a life worth living” a concept worth having? *Animal Welfare* 20(3): 397–406. <https://doi.org/10.1017/S0962728600002955>
- Young T, Creighton E, Smith T and Hosie C 2012 A novel scale of behavioural indicators of stress for use with domestic horses. *Applied Animal Behaviour Science* 140: 33–43. <https://doi.org/10.1016/j.applanim.2012.05.008>
- Zeitler-Feicht MH, Hartmann E, Erhard MH and Baumgartner M 2024 Which affiliative behaviour can be used as a valid, reliable and feasible indicator of positive welfare in horse husbandry? *Applied Animal Behaviour Science* 273: 106236. <https://doi.org/10.1016/j.applanim.2024.106236>