2024-05-29

Measuring affect-related attention bias to emotionally valenced visual stimuli in horses

Sarah Kappel *University of Plymouth*

Marco A.Ramirez Montes De Oca *University of Bristol*

Sarah Collins *School of Biological and Marine Sciences*

Katherine Herborn *School of Biological and Marine Sciences*

Michael Mendl *University of Bristol*

et al. *See next page for additional authors*

Measuring affect-related attention bias to emotionally valenced visual stimuli in horses

Kappel, Sarah; Oca, Marco A.Ramirez Montes De; Collins, Sarah; Herborn, Katherine; Mendl, Michael; Fureix, Carole

Published in:
Applied Animal Behaviour Science

DOI:
10.1016/j.applanim.2024.106303

Publication date:
2024

Document version:
Publisher's PDF, also known as Version of record

Applied Animal Behaviour Science 275 (2024) 106303

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

journal homepage: www.elsevier.com/locate/applanim

Measuring affect-related attention bias to emotionally valenced visual stimuli in horses

Sarah Kappel [a,*,1], Marco A.Ramirez Montes De Oca [b], Sarah Collins [a], Katherine Herborn [a], Michael Mendl [b], Carole Fureix [a,b]

[a] University of Plymouth, School of Biological and Marine Sciences, Portland Square, Drake Circus, Plymouth PL4 8AA, UK
[b] University of Bristol, Bristol Veterinary School, Langford House, Langford BS40 5DU, UK

ARTICLE INFO

Keywords:
Emotional state
Affective measures
Visual attention
Horse behaviour
Animal welfare

ABSTRACT

Negative affect appears to mediate animals' attention to competing emotional stimuli (e.g., threatening vs. non-threatening conspecific face images), similarly to anxiety-related enhanced attention to social threat reported in humans. To investigate this 'attention bias' (AB, i.e., the differential attention allocation to certain types of information over others) in horses, we developed a visual AB test assessing horses' attention towards image pairs showing unfamiliar conspecifics' facial expressions indicating, a) negative (social threat), b) more neutral (at rest), and c) positive (food anticipation) situations. We predicted that horses exhibit greater attention to negative compared to neutral or positive face images (as a normal adaptive response), and that horses in negative affective states (inferred from validated welfare indices comprising direct (health, behaviour) and indirect (housing, management) measures summarised as individual welfare scores and subscores reflecting health, social and environmental aspects) show greater AB to negative face stimuli and all images overall. Comparing AB to positive *versus* neutral social stimuli is rarely considered in AB studies, we therefore explored horses' AB responses without *a priori* predictions. Over six trials, 44 horses from three facilities were shown stimulus pairs (negative/neutral, negative/positive, positive/neutral) presented simultaneously on two projector screens. Attention was assessed as absolute attention duration to each image, the proportion of time the negative/positive stimulus was attended to relative to the other stimulus, and overall attention (i.e., duration of head turns towards both stimuli combined). AB to stimulus type, side, effects of facility and individual characteristics (welfare and subscores, age) was analysed using linear and generalised mixed-effect models. Against our predictions, horses attended to the images within the three stimulus pairs for similar lengths of time (negative-neutral: W=1870.5, p=0.2572; negative-positive: W=2542.5, p=0.9296; positive-neutral: W=1762.5, p=0.1019). Due to Covid-19 interruptions, our sample size was lower than our estimated required number (N=113). Still, lower welfare (X^2_1=4.71, p=0.03) and health scores (X^2_1=4.13, p=0.04) significantly predicted shorter attention to the negative face stimuli, possibly reflecting threat avoidance previously reported in other animals. We found significant facility effects on overall attention to the stimuli (X^2_2=77.42, p<0.001), likely due to varying yard-specific conditions (e.g., lighting, noise). This highlights that external influences on visual attention require consideration when conducting cognitive tests at different testing sites. Further methodological investigation (e.g., test cue suitability, perceptual processing of computer-generated images; test stimuli familiarity; individual differences) is needed to evaluate the potential of AB as an indicator of affective valence in horses.

1. Introduction

Measuring changes in cognitive processing (e.g., biases in attention, memory, and decision-making) is an important approach to studying non-human animal emotions (Mendl et al., 2009; Paul et al., 2005), since affective valence (i.e., pleasant or unpleasantness of the emotional state) can be inferred from the direction of cognitive biases (Mendl et al., 2010; Paul et al., 2005). In humans, negative affect is associated with a greater expectation of negative outcomes under ambiguity (i.e., 'pessimism'), heightened attention allocation toward negative information,

* Corresponding author.
E-mail address: sarahkappel09@gmail.com (S. Kappel).
1 Present address: University of British Columbia, Faculty of Land and Food Systems, Animal Welfare Program, Vancouver, BC V6T 1Z4, Canada.

https://doi.org/10.1016/j.applanim.2024.106303
Received 13 February 2024; Received in revised form 24 May 2024; Accepted 28 May 2024
Available online 29 May 2024

and greater likelihood of negative memory recall (reviewed in e.g., Blanchette and Richards, 2010; Paul et al., 2005). That underlying negative mood (i.e., longer-lasting states arisen from the accumulation of short-lasting emotions in response to specific stimuli) induces pessimistic judgement also in animals has been demonstrated in various mammal, bird, and insect species (reviewed in e.g., Gygax, 2014; Lagisz et al., 2020). However, applying judgment bias tasks (JBTs) in animals has significant limitations (Bethell, 2015; Crump et al., 2018; Roelofs et al., 2016). These can include *i)* labour-intensive training phases (e.g., discrimination training of positive/negative cues before animals' responses to intermediate cues can be tested), *ii)* sampling biases towards 'learners' (and exclusion of non-learning animals, which possibly have more negative affect since severe stress can impair cognitive abilities, Mendl, 1999), *iii)* ambiguity could be perceived as novel rather than intermediate (mixed positive/negative) stimuli (Bateson and Nettle, 2015; Doyle et al., 2010; Roelofs et al., 2016), and/or *iv)* discrimination training could act as cognitive enrichment for animals kept in barren environments (hence influencing affect and responses during testing; Crump et al., 2018). Developing new cognitive measures free from these caveats is therefore desirable (Crump et al., 2018; Roelofs et al., 2016).

Assessing affect-driven attentional biases (ADABs), i.e., the emotion-induced differential allocation of attention to certain information compared to others, has been promoted as an alternative, and potentially more practical, cognitive marker of animal affect (Crump et al., 2018). Unlike JBTs, attention bias tests (ABTs) rely on innate tendencies to cognitively process sensory information, such as facilitated attention to biologically relevant visual stimuli (e.g., social threat indicated through facial expressions, tested e.g., in Bethell et al., 2012; Howarth et al., 2021). In addition, attention allocation to these stimuli can be measured with no, or relatively little, animal training (e.g., in Howarth et al., 2021 macaques were only trained to sit by a target for AB testing). Attention allocation can be behaviourally assessed via e.g., measuring eye gaze or head movements toward visual cues, whereby the location of target fixation and attention duration (i.e., looking time) can be quantified (Winters et al., 2015).

In humans, trait (i.e., personality) and state anxiety are associated with modulation of attention towards threatening information (reviewed in e.g., Bar-Haim et al., 2007; MacLeod et al., 1986; Yiend, 2010). The majority of AB studies report anxiety-induced AB *towards* threat (i.e., faster detection of and/or difficulties in disengaging from threat cues), although opposite AB effects (i.e., attention shifted *away from* threat cues) have been found in some studies (Bar-Haim et al., 2007). Furthermore, the relevance of stimulus type to the subject's concern or situation influences AB in humans. For instance, social phobia induces stronger AB to social threat stimuli than non-social threat stimuli (Williams et al., 1996) and AB to pain-related information is found in chronic pain patients (reviewed in Schoth et al., 2012).

Attention to threat also appears to be modulated by affective state in non-human animals (reviewed Crump et al., 2018). Rhesus macaques show AB to conspecific threat faces paired with comparatively more neutral face images, with attentional avoidance of threat faces found following negative affect manipulation (veterinary handling, Bethell et al., 2012). In sheep (Lee et al., 2016) and cattle (Lee et al., 2018), the pharmacological induction of anxiety resulted in increased looking duration towards the location of a threat (a live dog presented shortly before its removal for the remaining test duration), while administering an anxiolytic lowered sheep's attention to the threat location.[2]

The aim of this study was to develop a test to assess whether horses show AB to emotionally-valenced social stimuli, and to investigate whether AB responses are reflective of affective valence determined *via* individual welfare level quantified as a score derived from validated equine welfare indices. In humans, psychological (e.g., anxiety, Bar-Haim et al., 2007) and physical stress (e.g., chronic pain, Schoth et al., 2011) are associated with AB. We therefore tested how different types of stressors influence horses' attention by subdividing welfare indices into social, health and environmental factors. Developing measures of horse affect is important as, despite the identification of housing and management-related issues that compromise equine well-being (e.g., Dalla Costa et al., 2016; Hockenhull and Whay, 2014; Lesimple et al., 2010; Minero and Canali, 2009), the identification and validation of reliable measures of affective valence, and ultimately welfare, are relatively scarce in horses.

Here, we investigate whether horses show AB towards unfamiliar conspecifics' valenced facial expressions associated with negative (social threat), positive (anticipation of food), or comparatively more neutral (resting) contexts. Horses were previously reported to discriminate between printed photographs of conspecifics with valenced facial expressions. Indeed, they showed differential behavioural (i.e., approach positive faces, avoid approaching agonistic faces) and physiological responses (i.e., decreased heart rate viewing positive faces, increased heart rate viewing agonistic faces) to these stimuli (Wathan et al., 2016). We predicted that horses would attend to negative face images for longer than neutral or positive faces (because attention to threat is a normal adaptive response, Ohman and Mineka, 2001), and that this effect would be stronger in horses with poorer welfare scores (and lower subscores) as negative affect further increases AB to threat (e.g., Lee et al., 2016).

We also explored horses' attention to the positive relative to the neutral face, since a bias towards positive stimuli could be relevant when assessing quality of life, since good welfare requires the presence of positive experiences (Boissy et al., 2007). ABs favouring positive information are understudied in humans and other animals (Crump et al., 2018), making predictions of the direction and strength of such a bias difficult. Finally, we expected that horses with lower welfare scores (and lower subscores) would show greater overall attention to both stimuli since negative affect heightens general vigilance (i.e., increased monitoring of the environment for potential threat) in other ungulates (Lee et al., 2016; Monk et al., 2018, 2019, 2020).

2. Methods

2.1. Ethical approval

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

2.2. Animals and housing

The sample size required to detect an effect was estimated through power analysis (80% power, 2-tailed, $p \leq 0.05$). We used data from 15 pilot horses (not included in this study) tested to collect preliminary data on differences in attention duration to negative, positive, and neutral conspecific face images to calculate the N needed to demonstrate significant discrimination between these stimuli. The resulting estimated sample size was 113 animals. We therefore aimed to test animals from six horse facilities (yards; approached *via* web search and word-of-mouth) with approximatively 20 animals per yard. During piloting, we were unable to collect pilot data on horse welfare (due to time limitations). We therefore planned to use welfare data collected at the first three facilities to conduct further power analysis. However, our data collection period coincided with COVID-19 restrictions, so in total 47 horses from three different yards were finally enrolled in this study. Three horses were excluded from data collection or analysis (one due to

[2] although see Monk et al., (2018, 2019) for attentional avoidance of the threat location in sheep given an anxiogenic.

camera failure, one because of disruptions during testing, and one horse was reluctant to approach the test equipment), resulting in 44 animals used for data analysis (yard 1: N=11; yard 2: N=16; yard 3: N=17).

All horses belonged to private owners (all animals from yard 1 had different owners, horses from the other yards belonged to two riding schools), were 4–28 years old (mean age ± SD, 13.75 ± 5.83 years), of mixed breeds, and 17 (39%) were females (see Table 1 for each individual's details). Housing conditions varied between facilities as we aimed to recruit horses from environments with either comparatively 'naturalistic' housing conditions (i.e., conditions aligned with horses' species-specific needs to live in stable groups (Fureix et al., 2012; Waring, 2003) with free access to roughage and exercise (Krueger et al., 2021), or comparatively more 'restrictive' environments, i.e., limited social contact (e.g., Lansade et al., 2008; Yarnell et al., 2015), restricted outdoor access (Henry et al., 2017; Löckener et al., 2016; McGreevy et al., 1995) or limited access to roughage (Goodwin et al., 2002; Ninomiya et al., 2004), all known to cause behavioural, physiological and psychological problems in horses (reviewed in Krueger et al., 2021; Lesimple, 2020; Minero and Canali, 2009).

At yard 1, horses were kept in a relatively naturalistically pasture/track system all-year around and in a stable herd consisting of the same animals housed together for more than three months prior to our study. Grass and hay were available *ad libitum*. All horses received supplement feed (feed brands varied between owners) once per day. Yards 2 and 3 were riding schools that housed their horses comparatively more restrictively in single stalls (at yard 3 six horses were tie-stalled) with limited opportunity for social contact (i.e., some animals only had visual but no physical contact with conspecifics), and bedding consisted of either straw or wood shavings (individual housing conditions presented in Table S1 in the supplementary material). Horses from these yards had access to pasture with the same group of conspecifics (ranging between 3 and 10 animals) during parts of the day and/or at night (pasture time depended on weather conditions and riding lessons). Hay availability was restricted (i.e., facilities adjusted the amount of hay for each horse based on its body weight) and provided 2–3 times per day. At yard 2, horses did not receive any supplemental feed (horses were on a 'diet' due to lower workload during COVID-19 restrictions). At yard 3, all horses received an adjusted diet (Thunderbrook Equestrian™) once per day. Water was freely accessible through automatic troughs at yard 1 and 2 and provided in water buckets at yard 3.

2.3. Welfare assessment

We evaluated horses' individual welfare using a holistic approach comprising multiple assessment steps. We graded environmental, management (indirect) and animal-based (direct) measures in a welfare assessment protocol based on the AWIN protocol (Minero et al., 2015) and Dalla Costa et al. (2014). Briefly, a single welfare score was calculated per horse derived from the sum of scores assessed for each environmental, management, and social resource provided to this individual, as well as the horse's health (i.e., sign of injuries, nostril discharge) and behavioural status (presence/absence of abnormal behaviour or abnormal repetitive behaviour, reaction towards an unfamiliar human). Table 2 provides details of our assessment protocol and calculation of individuals' welfare scores. Some health-based indicators recommended in the literature (AWIN protocol, Minero et al., 2015) were excluded from our protocol for practical reasons, i.e., eye discharge (horses at yard 1 wore fly masks; coat condition – parts of the coat had been clipped in some horses) and body condition scoring (BCS). BCS was originally assessed by two observers (ICC 0.73) after Henneke et al. (1983), but scores showed little variation from normal (between 4 and 8, mean 5.82 ± 0.8; 5 is considered normal). The maximum score of our welfare assessment protocol was 20, with higher scores putatively reflecting more positive welfare and a more positive affective state. Affective consequences associated with different types of stressors might influence attention differently, which is why we summarised social,

Table 1
Horse information and welfare scores. See Table 2 for details on calculation of the welfare score.

ID	Yard	Sex	Age	Breed	Welfare score
1	yard 1	gelding	17	Thoroughbred	15.92
2	yard 1	gelding	7	Cob	17.42
3	yard 1	gelding	24	New Forest Pony	16.75
4	yard 1	gelding	19	Welsh Pony	17.00
5	yard 1	mare	22	Connemara Pony	16.75
6	yard 1	gelding	7	Irish draft x	17.58
7	yard 1	mare	9	Appaloosa	17.42
8	yard 1	gelding	14	Connemara Pony x Camargue Pony	16.75
9	yard 1	gelding	13	Thoroughbred	19.00
10	yard 1	gelding	20	Thoroughbred	15.58
11	yard 1	mare	10	Thoroughbred	16.42
12	yard 2	gelding	11	Cob	12.92
13	yard 2	gelding	15	Irish Cob	14.33
14	yard 2	gelding	7	Irish Sport Horse	12.42
15	yard 2	mare	9	Cob	15.58
16	yard 2	gelding	7	Cob	15.92
17	yard 2	mare	13	Irish Sport Horse	11.33
18	yard 2	gelding	14	Cob	14.17
19	yard 2	gelding	9	Welsh Cob X	12.92
20	yard 2	mare	10	Cob	15.25
21	yard 2	mare	9	Cob	14.25
22	yard 2	mare	11	Cob	15.25
23	yard 2	mare	7	Fell	15.42
24	yard 2	gelding	13	Irish Sport Horse	13.25
25	yard 2	mare	9	TB X Cob	13.33
26	yard 2	gelding	9	Irish Sport Horse	15.50
27	yard 2	gelding	12	Welsh	13.67
28	yard 3	gelding	9	Cob	16.75
29	yard 3	mare	10	Cob x Thoroughbred	12.75
30	yard 3	mare	18	Welsh pony	13.92
31	yard 3	gelding	9	New Forest Pony	17.75
32	yard 3	gelding	26	Welsh Pony	15.25
33	yard 3	gelding	17	Connemara Grey	15.41
34	yard 3	gelding	28	New Forest Pony	14.75
35	yard 3	gelding	19	New Forest Pony	12.75
36	yard 3	gelding	18	Shire Cross Irish Pony	16.25

(continued on next page)

Table 1 (*continued*)

ID	Yard	Sex	Age	Breed	Welfare score
37	yard 3	mare	15	Welsh pony	16.91
38	yard 3	mare	16	Exmoor Pony	14.75
39	yard 3	mare	14	Percheron Cross	15.75
40	yard 3	gelding	15	Welsh pony	16.75
41	yard 3	gelding	14	Welsh x Anglo Arab	17.25
42	yard 3	mare	4	Shire x Thoroughbred	15.75
43	yard 3	mare	24	New Forest Pony	14.25
44	yard 3	gelding	26	Welsh pony	13.25

health and environmental factors in separate subscores (see Table 2 for details). We assessed the effects of social aspects on horses' attention as our test stimuli were social stimuli, and human literature shows that the relevance of stimulus type to the subject's concern or situation influences AB findings (e.g., social phobia induces stronger AB to social threat stimuli than non-social threat stimuli; Williams et al., 1996). We also assessed the effect of health indicators as scientific evidence suggests that chronic pain influences AB in humans (e.g., reviewed in Schoth et al., 2011). The environmental factors we assessed captivate species-specific needs not reflected in the other two scores.

2.4. Assessment of affect-related attention bias (ADAB)

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

2.4.1. Stimuli preparation (prior to testing)

Four stimulus horses (ID A, B, C and D) unfamiliar to the test animals were photographed during each of the three situations (negative, positive and neutral) using a NIKON D3200 camera (focal length 30.00 mm). Images were edited so that only the horses' heads on a white background were visible (see Table 4 for example images and supplementary Table S2 for all test stimuli). Mean luminosity values were extracted for each image via the histogram function in Adobe Photoshop. To compare the mean luminosity across stimulus types (negative/positive/neutral) and horses, we used a two-way ANOVA with Tukey post hoc pairwise comparison tests. There was no effect of stimulus type ($F=0.65$, $p=0.55$), but there was an effect of stimulus horse, the latter indicating significant differences between images of horses A-B, A-C, A-D and B-C (all p-values <0.05), but not B-D ($p=0.11$) or C-D ($p=0.71$). We therefore controlled for stimulus horse ID in the later analysis.

Pictures were subsequently arranged so that during testing, each test horse was exposed to one Microsoft PowerPoint presentation containing a total of six stimulus slides, each proceeded by a blank inter-trial slide. Each stimulus slide showed two side-by-side images of the same stimulus horse displaying differing facial expressions (i.e., negative-positive, positive-neutral, or negative-neutral). Two stimulus horses were randomly chosen from the four available to create each presentation (i.e., in total six stimulus slides showing the two selected stimulus horses in each stimulus pair combination). The order of stimulus slides was pseudo-randomised so that if the same type of stimulus pairing (e.g., positive *vs.* neutral) happened to be shown on two consecutive trials, the type of facial expression was counterbalanced across sides (e.g., positive (left) – neutral (right) followed by neutral (left) – positive (right)). Four presentations, each containing a different combination of stimulus horses, were created in total. This 'set' of four presentations allowed to test four horses (one presentation per horse), to be repeated until the 44 horses were be tested. However, prior to the set use by SK when testing the horses, the four presentations were re-named (i.e., anonymised to SK), and their order of presentation within each set was randomly allocated by another experimenter (CF). This ensured the experimenter (SK) conducting the tests was blinded to the content of the presentations (since each PowerPoint started with a blank slide, and SK did not look at the screens while testing, see section below).

2.4.2. Experimental set-up and procedure

The test area was a familiar arena equipped with two back projection screens (HOIN, 140 cm H x 212 cm W, 254 cm in diameter, placed 1 m above ground), and two LCD projectors (HITACHI CP-X303WN) connected to a computer (Lenovo Thinkpad 13) and positioned at 50 cm height in 2.5 m distance behind the screens (Fig. 1). At yard 1, the projectors were always placed in front of the screen for practical reasons of space limitation. At yard 2, back projection seemed unsuitable (i.e., images barely visible with human eyes) for eight horses (IDs 12, 13, 14, 15, 17, 24, 25, 27 in Table 1), and therefore images were front projected for these horses. The projected horse face images were the size of a medium-sized horse head (approx. 55 cm long, 25 cm wide). Two cameras (GoPro Hero 3, 3+) were used for recording horse behaviour from the front (i.e., horse is visible from the front, stimuli not visible) and the back (i.e., horse visible from behind, both screens in full view). The back view videos were subsequently edited (VSDC free video editor software) for blinding purposes, by placing white squares over the screens so that only the stimulus horses' muzzle was visible. This allowed the observer to record the start and the end of a trial when extracting the tested horse's behaviour from footage, whilst remaining blind to the type of stimuli shown on screen.

Each horse was habituated to the test equipment (screen turned off), first from an approx. 10 m distance, then from the test position standing as close as possible to the first ground pole (5 in Fig. 1) and with head and body aligned with the second ground pole (6 in Fig. 1). Once the horse stood quietly in front of the screens with head in a relaxed position (i.e., with muzzle at approx. chest height), the experimenter (hereafter E, SK) stood next to the horses' shoulder with her head directed to the horse's withers (i.e., E standing at a 90° angle to the horses' shoulder). Whilst glancing towards the horse's neck to judge its alignment with the second ground pole, E turned on both screens using a remote control (Kensington wireless presenter) to show the first stimulus slide (i.e., started the first trial) at the moment the horses' head was aligned with the ground pole. The images were shown for at least 20 s regardless of whether the horse looked at the screens or not (see ethogram Table 3). If the horse was still looking at the screens after 20 s, the images were shown until the horse no longer appeared to look at the stimuli.[3] When done, E showed a blank slide on the screen, switched her position to stand at the horse's opposite shoulder, corrected the horse's position if necessary, and started a new trial as soon as the horse's head aligned with the ground pole. E monitored the horses' head/neck movements but never looked directly at the screens during a trial (ensuring blinding to the stimuli projected). All six trials were conducted this way. Disturbances (e.g., noise, person unexpectedly entering the test area) were signalled by E by directly looking into the back camera and mouthing the word "disturbance" to identify trials to be excluded during video analysis.

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

Table 2
Overview of indicators selected to assess individuals' welfare level and description of scores. Each of the 20 indicators weighted equally into the overall score. For each assessment point, a score between 0 and 1 was given with the higher score reflecting better welfare conditions. When more than two outcomes (i.e., presence or absence of resource) were possible (see e.g., confinement: always, part-time, never), fractional values were assigned to each possible condition. Subscores reflecting environment, social and health conditions were calculated based on the factors listed under the corresponding table heading. References to literature addressing specific welfare indicators are given in brackets.

#	Welfare indicator [References]	Answer option	Description of score.	Score
	ENVIRONMENTAL FACTORS			
1	Confinement (Heleski et al., 2002; Henry et al., 2017; Löckener et al., 2016; McGreevy et al., 1995)	Always	The horse has no outdoor access or is outside for less than 12 h per week.	0
		Part-time	The horse has outdoor access for at least 12 h per week (excluding work, husbandry procedures).	0.5
		Never	The horse is always kept outdoors (e.g., on paddock or on pasture) (excluding work, husbandry procedures).	1
2	Stall space sufficiency*	No	Stall dimensions are smaller than required according to the British Horse Society (BHS) guidelines.	0
		Yes	Stall dimensions are within or bigger than required according to the BHS guidelines. Horses always kept outdoors were given a score of 1.	1
3	Water availability*	No	Water is not always freely accessible or manually filled (e.g., buckets).	0
		Yes	Water is always freely accessible (e.g., via an automatic system).	1
4	Ventilation (Clarke, 1987; Holcombe et al., 2001)	No	No ventilation is provided.	0
		Partially	Ventilation is provided during parts of the day but not always (e.g., barn doors are shut at night preventing constant airflow).	0.5
		Always	Ventilation is always provided (e.g., via air-condition system or horses are kept outside).	1
5	Forage opportunity (Ninomiya et al., 2004)	Never	The horse has no opportunity to forage.	0
		Restricted	The horse has restricted (time/amount) forage opportunity through provision of forage at least twice per day.	0.5
		Always	The horse has always the opportunity to forage.	1
6	Forage diversity (Goodwin et al., 2002; Ninomiya et al., 2004)	No	No forage diversity is available.	0
		Yes, part-time	The horse was opportunity to choose between types forages for parts of the day (e.g., grass and hay available on pasture) but not full-time (e.g., indoors only hay was provided).	0.5
		Yes	The horse has a choice in forages through provision of at least two forage options (e.g., hay available on grass pasture).	1
7	Bedding type (McGreevy et al., 1995; Nazarenko et al., 2018; Pedersen et al., 2004)	No bedding/ Rubber mat or wood shavings	No bedding is provided.	0
		Straw	Straw is provided as bedding.	1
8	Visual scope (Ninomiya et al., 2007; Ruet et al., 2019)	Never	The horse has no opportunity to broaden its visual scope (e.g., no window, closed stable doors)	0
		Part-time	The horse has ability to boarded visual scope (e.g., via half-closed stable door, windows, access to pasture) for parts of the day.	0.5
		Always	The horse always has the ability to broaden its visual scope (e.g., via windows, half-open stable door, full-time on pasture).	1
	SOCIAL INDICATORS			
9	Social contact (daily duration and level of physical contact) (Lansade et al., 2008; Lesimple et al., 2019; Yarnell et al., 2015)	Full-time no social contact	The horse has never social contact.	0
		Part-time no social contact and part-time restricted	The horse has no social contact at times (e.g., in stable) but has restricted (e.g., only visual) contact to other horses during parts of the day (e.g., top-half of stable door is opened during the day allowing to see other horses).	0.25
		Full-time restricted	The horse has opportunity of sniff and/or view other horses e.g., through stable bars but cannot physically interact further than sniffing with other horses.	0.5
		Part-time restricted and part-time full social contact	The horse has restricted opportunity to interact other horses (e.g., in stable) but full social contact (full-body) to others for parts of the day (e.g., on pasture).	0.75
		Full-time social	The horse always has full social contact (full-body) to other horses.	1
10	Social stability (opportunity to maintain social relationships) (Fureix et al., 2012; Waring, 2003)	Never	The horse has no social contact or social partners are changing (e.g., horses are allocated to neighbouring stables at random, mixing neighbours frequently).	0
		Partially	The horse has the same social partner(s) inside stable but social partner (s) outside vary frequently (or *vice versa*).	0.5
		Full-time	The horse is kept with the same social partner(s) continuously.	1
	*ANIMAL-BASED INDICATORS**			
11	Signs of practice induced injuries	Yes	Injuries (e.g., fresh skin lesions, healed scars, white hairs, and bald skin patches) likely caused by management procedures or ill-fitting tack (e.g., lesions mouth corners, white hair on withers) are present.	0

(*continued on next page*)

Table 2 (*continued*)

#	Welfare indicator [References]	Answer option	Description of score.	Score
		No	The horse has no signs of injuries linked to management procedures.	1
12	Signs of other injuries	Yes	Injuries (e.g., fresh skin lesions, healed scars, white hairs, bald skin patches) bigger than 1x2cm^2 or longer than 4 cm not likely to be caused by management procedure (e.g., bite marks on neck) are present.	0
		No	No injuries are present.	1
13	Nasal discharge	Yes	Flow of discharge (watery, transparent or yellow/green) is present from one or both nostrils.	0
		No	Nostrils have no discharge.	1
14	Lameness	Present	The horse's gait is altered with uneven weight bearing (limp) whilst standing or walking. If the horse received treatment for lameness during the time of the study it was scored as lame regardless whether it presented as lame.	0
		Not present	The horse is walking/standing with even weight bearing on all four feet and was not treated for lameness during the study.	1
15	Swollen joints	Present	Swelling of joints evident in increased diameter of one or more limps.	0
		Not present	No swelling of joints present.	1
16	Manure consistency	Abnormal	Manure consistency is different from the normal shape or texture (e.g., too watery (cow dung like with no shape or too dry, i.e., very hard).	0
		Normal	The horse's manure has normal consistency.	1
17	Abnormal hoof condition/ shape	Present	One or more hooves are overgrown, cracked or misshaped.	0
		Absent	No signs of hoof neglect present.	1
OTHER INDICATORS				
18	Abnormal behaviour (assessed during behaviour observation during 5x1h sessions) [modified from AWIN (Minero et al., 2015)]	Present	The horse displayed a form of abnormal behaviour (e.g., teeth grinding, cribbing, head tossing, abnormal posture during resting, i.e., withdrawn state) during observation periods.	0
		Absent	The horse was not observed displaying abnormal behaviours during observation periods.	1
19	Documented horse health history†•	Acute or chronic condition present	The horse had an acute or chronic condition possibly compromising welfare at the time of the study.	0
		Documented as healthy	The horse had no known underlying health conditions.	1
20	Human interaction score (voluntary (VAT) and forced approach test (FAT) results combined)	A voluntary approach test (VAT) was conducted once while the horses were feeding hay (head down) in their stalls. The experimenter (E) entered the stall, stood motionless by the door facing the horse at approximatively 1.5 m distance and remained there for 5 min or less if the horses approached E earlier. The forced approach test (FAT) was conducted in the similar manner but upon entry, E directly approached the horse whilst lifting her right arm and pointing it towards the horses' shoulder and moving closer to the horse to touch its withers for 3 s. The FAT was carried out three times on three different days. All tests were video recorded and later scored by a single observer (SK) following the ethogram described below. The VAT and mean FAT score were averaged and scaled to a range between 0 and 1 as the human interaction score.		
		Positive response	The horse stops feeding, lifts its head up and approaches the experimenter with forward pointing ears or sniffs the experimenter during FAT for more than 2 s	1
		Attentive	The horse is continuing to feed but moves its ears forwards when experimenter appears, the head is lifted up or/and its head is rotated towards the experimenter. Quick sniff (less than 2 seconds).	0.5
		Feeding – inattentive (to experimenter)	The horse is feeding (with its head down), no obvious change in behaviour when experimenter appears.	0
		Negative - threat	The horse rotates its ears backwards, head possibility lifting or moving towards the experimenter or remains in same position as before the experimenter entered.	-0.5
		Negative - avoidant	The horse lifts its head, possibly startles on entry of the experimenter, ears backwards and horse actively moves away from experimenter.	-0.5
		Negative - aggressive	The horse lifts its head, and approaches the experimenter (walking or with head moving towards experimenter) while displaying a threatening attitude (ears back and teeth showing).	-1

† We were unable to schedule a veterinary health check for every horse due to monetary constraints and instead relied on horses' documented health history.
• Health history was also included as a factor in the health sub-score.
* all animal based indicators derived from the AWIN (2015) horse welfare assessment form (Minero et al., 2015).

2.4.3. Behavioural measures to assess attention

The horse behaviour was scored from video by SK (intra-observer reliability assessed using the icc function in the psych R package: 0.932) using Noldus Observer XT 14. A second observer not involved in the study independently scored videos of 10 horses (22.7% of samples) and inter-observer reliability was 0.87. The ethogram is shown in Table 3. The lateral position of the eyes and their limited movement within the eye socket makes it difficult to assess what horses are looking at by looking at the eyes themselves. We thus used proxies of overt visual attention as previously described in horses (Waring, 2003; Wathan et al., 2016), i.e., head and ear positions. Head and ear positions were scored continuously whilst images were shown. A horse was assessed as attending to a stimulus if it turned its head either to the right or the left with both or one ear forward suggesting engagement of the visual (binocular) system to look at the screen ahead of it. If a horse kept its head turned towards one screen whilst moving both its ears from a forward to a lateral or backward position for more than 1 s, it was no longer considered attending the screen. Only time periods >1 s

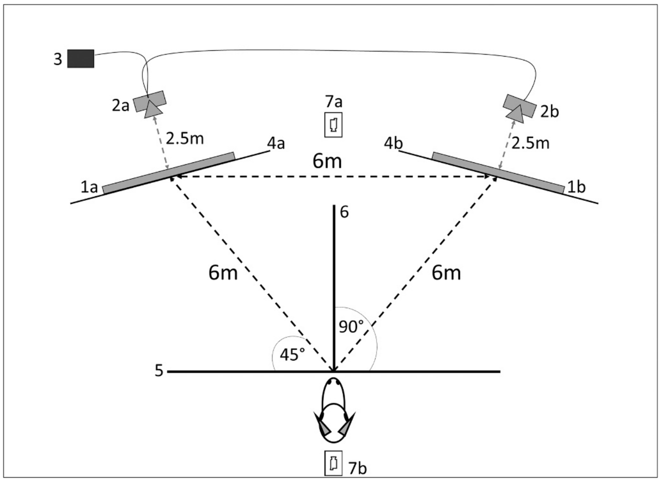

Fig. 1. Experimental setup. The test stimuli were projected onto two screens (1a, b) via two LCD projectors (2a, b) receiving input from a laptop (3). The midpoint of two ground poles (3 m long, Ø 98 mm) placed in front of the screens (4a, b) marked the screens' centre point. A ground pole (5) ensured consistent distance between animals and the screens. Horses were positioned with their front feet as close as possible to the midpoint of the pole. Perpendicular to this location, a second ground pole (6) was placed which marked the midline between both screens as a visual reference for horses' head position during testing and post hoc behaviour analysis. Cameras recorded the horses from the front (7a) and behind (7b).

(threshold chosen by us to facilitate differentiation between continuously moving and static ear positions) where the horse was attending to either screen were included in the final data analysis. Although horses possibly looked at the stimuli whilst carrying their head straight forward (i.e., both eyes at equal distance to the stimuli, see Table 3), we did not include these time points in the final analysis as it was impossible to determine which stimuli they were attending to without clear indication of the side they were looking at.

2.5. Data analysis

Data were analysed in R v. 4.0.3 (R Core team, 2021). For each horse, four indices of attention were calculated per trial. Firstly, all time periods (>1 s) a horse was observed looking at one of the two stimuli were combined to assess the duration of **absolute attention to the target stimulus** (i.e., negative or positive). Likewise, duration of attention to the **other** stimulus was assessed by combining all time points the horses were recorded looking at the opposite image (e.g., in negative stimulus pairs, the opposite stimulus was either positive or neutral, in positive stimulus pairs the opposite stimulus was neutral). **Total attention duration** was assessed as the sum of durations attending to the target stimulus and the other stimulus. Lastly, the **proportion of time looking at the target stimulus** was calculated relative to the total duration of attending both stimuli (i.e., attention to stimulus divided by total attention duration). Assessing both absolute attention duration and proportional values when measuring attention biases might be important because absolute differences give a direct measure of the difference in attention duration to the two competing stimuli, but variation in looking durations can influence this AB measure because of differences in absolute looking times between individuals (Lewis et al., 2021). We therefore also analysed proportional attention duration values.

A value of 0.001 was added to all absolute values so that trials with zero attention durations for one of the two stimuli thereby enabling inclusion in the statistical analysis of trials where horses looked at one stimulus only. We included trials where horses only looked at one stimulus as our test design ensured that horses were able to visually perceive both images at trial onset, hence excluding these trials did not

Table 3
Ethogram of horse behaviours scored during the attention bias test. Behaviours indicating that the horses were no longer looking at the stimuli once the first 20 s of stimulus presentation elapsed were categorised as behavioural events terminating the trial.

Behaviour	Description
Behaviours indicating attention towards the stimuli	
Head position (side scored from the horse's perspective)	
Head turn right	Head is held at, or higher than, chest height, with the centre of the forehead rotated away from the midline towards the left corner of the front view video footage (head direction scored from the horse's perspective hence contralateral side when seen from front view). The left eye is more visible on front view footage than the right eye (which might be partially visible or not at all).
Head turn left	Head is held at, or higher than, chest height, with the centre of the forehead rotated away from the midline towards the right corner of the video frame. The right eye is more visible on front view footage than the left eye (which might be partially visible or not at all).
Ear position	
Ears both forward	Head is held at, or higher than, chest height, and both ears are rotated forwards so that the insides of the ears are fully visible from frontal view footage while the outer parts of the ears (pinnae), apart from the tips, are not.
Ears both back	Head is held at, or higher than, chest height, the ears are rotated backwards so that the insides of the ears are not visible from frontal view footage, but the pinnae become more visible, and the tip of the ears are pointing backwards, or are possibly not visible if ears are pressed flat against head.
Ears lateral	Head is held at, or higher than, chest height, the ears are rotated outwards so that the inside of both ears only partly visible, the pinnae become more visible at the base of the head, and tips of the ears point outwards.
Right ear forward	The right ear is rotated forwards so that the inside of the right ear is fully visible while the left ear is in lateral or backward position.
Left ear forward	The left ear is rotated forwards so that the inside of the left ear is fully visible while the right ear is in lateral or backward position.
Behaviours indicating attention away from stimuli terminating a trial	
Head turn away right	The horses' neck is bend to its right-hand side with its head at least in profile view with the forehead no longer visible from frontal view footage.
Head turn away left	The horses' neck is bend to its left-hand side with its head at least in profile view with the forehead no longer visible from frontal view footage.
Horse interacting with the handler	Physical contact between any part of the horse's head and the handler (e.g., sniffing, head rubbing).
Exploring the ground	The horse is sniffing the ground/poles.

seem appropriate. Trials during which the horse did not attend to either of the stimuli, as well as trials with external disturbances (e.g., loud banging, dog barking) were excluded. This yielded an unbalanced data set with both data points per stimulus pair not being available for all horses. Consequently, we only considered the first trial per stimulus pair a horse attended to at least one of the stimuli to assess attentional biases at the group level (i.e., data of repeated trials (same stimulus pair) were excluded from this analysis). Since absolute duration data (attention to the stimulus, attention to the opposite stimulus and total attention duration) were not normally distributed (Shapiro-Wilk test, $p<0.001$), central tendency measures are reported as median and Q1-Q3 range. Wilcoxon signed rank tests were used to compare attention duration and proportion of time horses looked at the images (i.e., stimulus and opposite stimulus) presented as pairs. The number of trials per yard used for this analysis is shown in Table 4.

To investigate affect-related AB, our predictions of how horses would attend the emotional stimuli in relation to their welfare score were as follows:

1) **Attention bias to negative stimuli** is reflected in (a) longer absolute attention duration and (b) greater proportion of time attending

Table 4
Example image of stimuli and number of trials per stimulus pair used for group-level AB testing when horses first attended to the images.

Stimulus pair	Yard	N trials
Negative-neutral	1	10
	2	15
	3	15
Negative-positive	1	11
	2	14
	3	17
Positive-neutral	1	8
	2	15
	3	17

to the negative horse face in comparison to putatively less negative stimuli (i.e., a comparatively more neutral or positive horse face). We expected that horses with lower welfare scores would attend the negative stimuli for longer and for a greater proportion of time than horses with comparatively higher welfare scores.

2) **Attention bias to positive stimuli**, measured as (a) absolute attention duration and (b) proportion of time attending the positive horse face relative to the comparatively more neutral stimulus, is modulated by welfare score (exploratory analysis, no direction of ADAB predicted due to the lack testing ADAB towards positive stimuli in current literature).
3) **Heightened overall attention** is reflected in a significant negative relationship between total attention duration and welfare score, with horses scoring lower on welfare showing more attention to the stimuli compared to horses with comparatively higher welfare scores.

To test our hypotheses, we fitted separate linear mixed-effect models (lmer function in the lme4 package, Bates et al., 2015) with attention duration to the target stimulus (i.e., negative or positive; square root transformed for better model fit), and total attention duration (log--transformed) as the dependent variables. Facility (yard 1, 2, 3), horses' age, sex, and welfare score, as well as opposite stimulus type (i.e., neutral or positive), stimulus side (left, right), and stimulus horse ID (horse A, B, C, D, see Table S2) were set as fixed factors. Collinearity of fixed factors was checked with the vif function (from the lme4 package) as factors with vif >5 would have indicated a problematic amount of collinearity between predictors (James et al., 2014), but this was not the case. We included all trials in which a horse attended to at least one of the images in this analysis and therefore included horse ID as random factor in the models. The drop1 function (base R) was used for all models to exclude all non-significant factors (except welfare score, which was included in all models) to reach the minimum adequate model (model selection based on the smaller Akaike Information Criterion (AIC)). The ANOVA function with type III sum of squares implemented from the 'car' package (Fox and Weisberg, 2019) was used to generate p-values for individual factors within each model and post-hoc analysis was done with Tukey pairwise comparisons (multcomp package, Hothorn et al., 2008). Results were reported as significant at a threshold of $p \leq 0.05$ and results between 0.05 and 0.1 were reported as 'trends' given that these might convey meaningful biological variation (Stoehr, 1999). For model validation, model residual diagnostics were done using the 'DHARMa' package (Hartig and Lohse, 2022).

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

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

3. Results

3.1. Does the valence of the stimuli overall influence horses' attention?

Considering the absolute attention duration (i.e., all time periods (>1 s) a horse was observed looking at the target or other stimulus), we found no statistically significant evidence for AB as horses viewed the images within the three stimulus pairs for similar lengths of time (negative-neutral: W=1870.5, p=0.2572; negative-positive: W=2542.5, p=0.9296; positive-neutral: W=1762.5, p=0.1019; Fig. 2A). Likewise, the proportion of time horses looked at the stimuli did not significantly differ within the stimulus pairs (negative-neutral: W=1863.5, p=.2472; negative-positive: W= 2626, p=0.6684; positive-neutral: W=1639.5, p=0.0278; Fig. 2B).

3.2. Testing hypothesis 1: is attention to negative stimuli influenced by welfare score?

Overall welfare score significantly predicted the absolute duration horses looked at the negative stimuli (X^2_1=4.71, p=0.03), but in contrast to our prediction, a lower welfare score (i.e., greater likelihood for negative affective state) predicted shorter absolute attention duration to the negative stimuli. Moreover, facility had a significant effect on attention duration (X^2_2=22.93, p<0.001). Yard 2 horses looked at the negative stimulus for significantly longer than horses from yard 1 (z=4.68, p<0.001) and yard 3 (z=-3.45, p=0.001), and horses from yard 3 attended to the negative images significantly longer than yard 1 horses (z=2.6, p=0.02). Mean overall welfare scores significantly differed between yards (mean±SD, yard 1: 17.1±0.9; yard 2: 14.0 ±1.28; yard 3: 15.3±1.57; all p<0.001), but there was no significant interaction between yard and welfare score on attention measures (X^2_2=0.16, p=0.92). In addition, the type of opposite stimulus significantly influenced the duration horses attended to the negative stimulus (X^2_1=4.1, p=0.04). When paired with the positive stimulus, horses looked at the negative stimulus for significantly longer than when it was presented with the neutral stimulus (z=2.01, p=0.04).

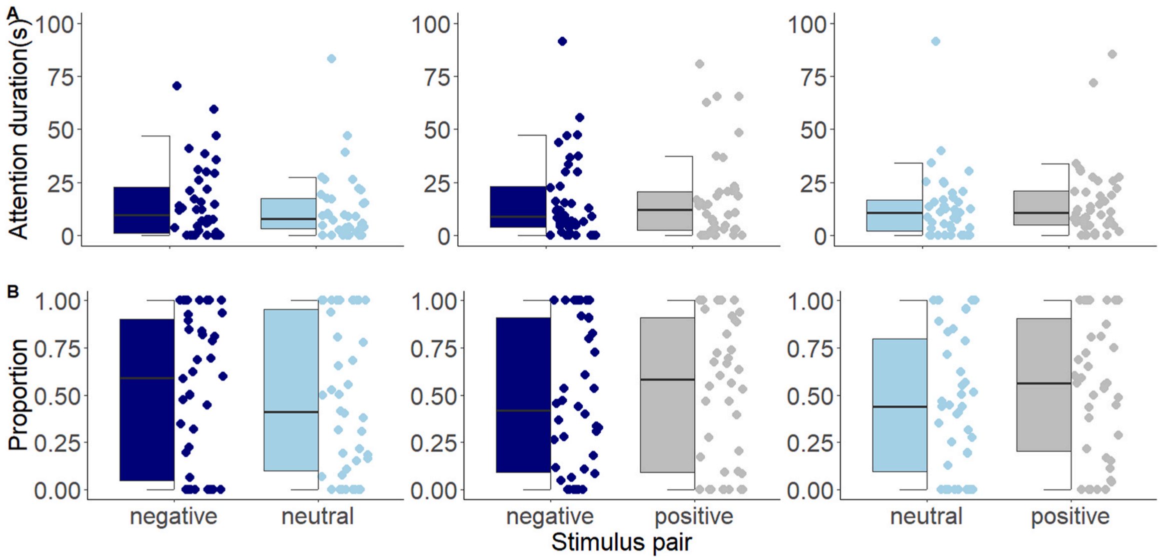

Fig. 2. Absolute duration (A) and proportion of time (B) horses attended to stimulus pairs for the first time. Boxplots show data at group level and dots show individual data points. Wilcoxon signed rank test, p>0.05 for all stimulus pairs.

Considering horses' health scores (sum of animal-based indices directly reflecting animal health), this score also had a significant effect (X_1^2=4.13, p=0.04) on horses' attention to the negative stimuli. Horses with better health scores looked at the negative stimuli for longer than horses with poorer health scores. We found no significant effects of social (X_1^2=1.4, p=0.22) or environment score (X_1^2=0.01, p=0.97).

Considering the proportion of time horses attended to the negative stimulus, we found no significant effect of overall welfare score (X_1^2=1.83, p=0.17), or any of the other factors we investigated (yard: X_1^2 =1.92, p=0.38; opposite stimulus: X_1^2 =0.68, p=0.4; sex: X_1^2 =1.71, p=0.19; age: X_1^2 =0.02, p=0.89; stimulus horse ID: X_3^2 = 0.1, p=0.99; stimulus side; X_1^2 =0.6, p=0.43; welfare score: X_1^2 =1.83, p=0.17; environment score X_1^2 =0.12, p=0.71). However, higher levels of both health (X_1^2 =3.41, p=0.06) and social scores (X_1^2 =3.61, p=0.057) showed non-significant tendencies to predict higher proportions of time attending to negative stimuli.

3.3. Testing hypothesis 2: is attention to positive stimuli influenced by welfare score?

Neither the absolute duration of attention to the positive stimulus (X_1^2=1.32, p=0.25), nor the proportion of time horses looked at the positive stimulus (X_1^2=0.96, p=0.32) were significantly predicted by overall welfare score or any of the other factors we explored (duration: yard: X_1^2 =0.91, p=0.63; sex: X_1^2 =0.44, p=0.5; age: X_1^2 =1.21, p=0.27; stimulus horse ID: X_3^2 = 3.2, p=0.34; stimulus side; X_1^2 =1.22, p=0.26; welfare score: X_1^2 =1.32, p=0.25; social score: X_1^2 =0.86, p=0.35; health score: X_1^2 =0.99, p=0.31, environment score: X_1^2 =3.11, p=0.57; proportion of time horses looked: yard: X_1^2 =0.26, p=0.87; sex: X_1^2 =1.53, p=0.21; age: X_1^2 =0.23, p=0.62; stimulus horse ID: X_3^2 = 2.4, p=0.49; stimulus side; X_1^2 =0.13, p=0.71; welfare score: X_1^2 =0.96, p=0.32; social score: X_1^2 =0.1, p=0.74; health score: X_1^2 =1.18, p=0.27; environment score: X_1^2 =0.75, p=0.38).

3.4. Testing hypothesis 3: is overall attention to the stimulus pairs predicted by welfare score?

Overall attention, assessed as the total duration horses attended to both images within the stimulus pairs, was not significantly predicted by overall welfare (X_1^2=0.72, p=0.39), social (X_1^2=1.6, p=0.19), health (X_1^2=0.26, p=0.6) or environment score (X_1^2=1.48, p=0.22). However, facility had a significant effect on attention duration (X_2^2=77.42, p<0.001). Horses from yard 2 attended to both stimuli for significantly longer than horses from yard 1 (z=8.73, p<0.001) and yard 3 (z=-2.93, p=0.009), and horses from yard 3 spent significantly more time looking at the images than horses from yard 1 (z=6.29, p<0.001). Moreover, ID of the stimulus horse had a significant effect on overall attention (X_3^2=11.45, p=0.009) with stimulus horse B being attended to significantly longer than stimulus horse D (z=-3.22, p=0.006), but not significant differences between the other stimulus horse pairs (A-B: z=2.35, p=0.08; A-C: z=0.38, p=0.98; A-D: z=-0.58, p=0.93; B-C: z=-1.91, p=0.21; C-D: z=-1.05, p=0.71).

4. Discussion

We tested whether horses exhibit a bias in attention (AB) when presented with a combination of two digital images of unfamiliar conspecifics' faces showing either negative, positive, or comparatively neutral facial expressions. We expected that negative stimuli would be attended for longer when presented alongside comparatively more positive or neutral stimuli, and that this bias would be exacerbated in horses with poorer welfare scores (proxy for poorer underlying affective state) as a result of negative affect modulation on attention to threat as reported in humans (Bar-Haim et al., 2007) and other animals (Crump et al., 2018). In contrast to our predictions, we found no significant evidence of AB to specific types of valenced face images as horses attended to all stimuli for similar lengths of time. However, we did observe a modulating effect of overall welfare score and associated health score on AB to negative stimuli but in the opposite direction to our predictions; horses with lower welfare or health scores spent less time attending to the negative images. Several reasons could explain our results/predictions discrepancy, which we discuss below.

While innate preferential attentional processing of emotional faces has been well described in humans, knowledge of what type of information other (non-primate) animals extract from faces is still limited (Leopold and Rhodes, 2010). Wathan et al., (2016) found that horses were more likely to approach the printed images of unknown conspecifics' positive faces than negative faces, which is an appropriate response to avoid social conflict, possibly reflecting an adaptive behaviour in this social-living species. However, approach/avoidance response behaviours are controlled by stages of cognitive processing that follow initial attentional processing and stimulus appraisal (i.e., stimulus relevance to the individual) before an appropriate response is determined (Scherer, 1999). Hence, our findings are not necessarily comparable with the previous horse study, as we tested responses in horses held by an experimenter and putatively linked to different stages of stimulus processing (i.e., looking here *versus* approaching/avoiding in Wathan et al., 2016).

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

In addition to familiarity, other facial characteristics (e.g., coat colour, head shape, facial markings) might contribute to preferential attention unrelated to emotional state. We used four different stimulus horses and found that stimulus horse ID had significant effects on total attention duration (considering both stimuli combined). Horses looked longer at stimulus horse B (dark brown coloured horse face with two small white areas on its forehead and muzzle) than stimulus horse D (chestnut-coloured horse face with a wide white vertical stripe covering (blaze) most of its head). Mean luminosity values did not significantly differ between the images of these horses. It is possible that horse D was more difficult to identify as a horse face compared to horse B given its wide strip of white merging with the white background. Bellegarde et al., (2017) also observed that face images of one goat were more attended than those of the other three stimulus goats (regardless of emotional information the faces conveyed), although instead of facial characteristics, social rank might have influenced goats' attention in this study as the stimulus goats were familiar conspecifics. It could also be that slight differences in the head angle, and the fact that horse faces were shown in absence of the rest of the horse's body further influenced horses' perception of these stimuli. Interestingly, two independent groups of raters (n=5 and n=6, with varying degrees of expertise in horse behaviour) showed relatively low interobserver reliability when independently rating blind the stimuli based on the descriptions of valenced facial expressions described in Wathan et al., (2016) (Krippendorff's alpha coefficients: α=0.55 and α=0.49). This suggests that photographic displays may not convey all the intended information.

It is also possible that, beyond the putative valence of the facial expression, the horses did not recognise the images as conspecific faces overall (Kappel et al., 2023). Our test stimuli were computer-generated projections, i.e., artificial images made for the human eye while equine vision differs from ours (Rørvang et al., 2020). Several studies have reported that horses can successfully discriminate between artificial images such as printed photos of objects (Hanggi, 2001), horses (Wathan et al., 2016) and humans (Proops et al., 2018; Smith et al., 2016), screen images of humans (Lansade et al., 2020), and videos of human-horse interactions (Trösch et al., 2020). Existing evidence of image recognition in horses should nonetheless be approached with caution because previous findings might have alternative explanations (e.g., rapid learning and generalisation in (Hanggi, 2001), lack of control conditions and statistical weaknesses in Proops et al., 2018 addressed by Amici 2019, discussed in further detail in Kappel et al., 2023) and do not fully explain whether animals learn to discriminate between visual stimuli or perceive images as representations of real individuals. A recent study in goats found no spontaneous preference for approaching images of familiar over unfamiliar herd members suggesting that goats did not perceive the images as representations of real conspecifics (Langbein et al., 2023). Furthermore, goats showed no difference in learning performance when trained to discriminate between familiar and unfamiliar conspecific images in a 4-choice discrimination task indicating that they used the images solely as visual stimuli with no reference to the real conspecifics (Langbein et al., 2023). Similarly, in a follow-up study using some of the same horses, we found that after learning to discriminate between two real-life objects of different colours and shapes, only one horse spontaneously approached the rewarded object when presented with on-screen images of these objects (Kappel et al., 2023). However, that the horses were not able to reliably transfer from inanimate objects to images in our follow-up study does not eliminate the possibility that horses might recognise biologically more important cues (such as conspecific faces) from images.

The finding that absolute attention duration to the negative stimuli was modulated by overall welfare and the health score might support this argument. Contrary to our predictions, horses with lower welfare/health scores attended to negative face images for shorter durations compared to horses with higher welfare scores. Avoidance of threat cues has been described in other animals. Macaques direct their attention away from social threat stimuli (conspecific faces) following a stressful handling procedure, although they show sustained attention to aggressive conspecific faces during enriched housing (Bethell et al., 2012). In sheep, the pharmacological induction of anxiety-like states decreased attention to threat (live dog) compared to controls (Monk et al., 2018), although Lee et al., (2016) found the opposite, and (Monk et al., 2019) found no significant effects. In humans affected by social phobias, avoidance of anxiety-eliciting stimuli (e.g., social situations) is a form of coping (Chen et al., 2020). Hence, horses with putatively poorer welfare might have disengaged from the negative faces faster to avert further negative experiences. In order to test attentional avoidance of certain types of stimuli more accurately, a similar approach to the operant conditioning conducted by Raoult and Gygax, (2018), who trained sheep to turn on/off videos to avoid seeing valenced social stimuli, could be helpful. Alternatively, measuring eye preference in response to valenced stimuli (i.e., visual lateralisation) might be useful when interpreting animals' affective responses to emotional stimuli (reviewed in e.g., Leliveld et al., 2013), although monocular vision is difficult to assess in animals with laterally positioned eyes and a wide visual field (Raoult and Gygax, 2018).

We also assessed whether horses showed AB to the positive stimulus in relation to the neutral stimulus, as prioritised attending to positive information over neutral information might also be of biological significance (Gupta, 2019; Pool et al., 2016). However, horses attended to both types of stimuli equally and none of the welfare scores modulated attention to positive images. Nevertheless, when paired with the positive, stimulus horses' attention to the negative stimulus was significantly longer than when paired with the neutral stimulus. This observation might suggest that attentional processing for potential threat could be influenced by the type of competing emotional information. Maybe horses recognised the emotional valence of positive faces more rapidly, and hence attended to the negative stimulus for longer, compared to when viewing the neutral faces that might be perceived as more ambiguous in terms of the emotional information conveyed, therefore may have required longer attentional processing leaving less time to attend to the paired negative stimulus. We are not aware of any other animal studies testing the varying effects of positive and comparatively more neutral stimuli on animals' attention to threat. However, Bellegarde et al., (2017) tested goats with morphed images (merging negative and positive goat faces at different gradations, i.e., 25%, 50%, or 75% positive) putatively presenting intermediate cues to positive and negative goat faces and found that the merged stimuli did not significantly differ in the time they were attended compared to the original faces.

Our inference of horses' affective valence was necessarily indirect and we should acknowledge that inferring affect from our overall welfare score, and its subscores (social, health and environment scores), were only approximations of the likelihood that animals were experiencing negative/positive affective states. The presence (or absence) of certain resources may not determine animals' experience of their environment to the same extent and furthermore variations in individuals' coping abilities may influence their perception of their environmental conditions, and ultimately their emotional state. However, we did find that better health scores, in line with better results from the overall welfare score, significantly predicted longer attention to negative stimuli whereas environment and social scores had no significant effects. Horses did not differ much in the latter two scores since most animals from the same facility were kept in similar conditions, so lack of variation in these measures may have limited the likelihood of finding significant relationships. Although we planned to use manipulations likely to directly influence affective state (e.g., exposing horses to sudden novel stimuli directly before AB test (negative manipulation), or rewarding stimuli such as food (positive manipulation) as done in previous AB studies, our experimental plans were significantly affected by Covid-19 restrictions. COVID-19 also prevented us from achieving our estimated sample size. Nonetheless, our final sample size was bigger than that of Wathan et al., (2016) reporting significant behavioural differences (approach/avoidance) to valenced stimuli when testing a maximum of 13 horses. We therefore expected to see significant effects of overall attention to the valenced stimuli. It is possible that our analysis of welfare-mediated AB effects might have been underpowered. Thus, testing a larger number of animals might have led to different findings.

We also found that, facility, but not any of the welfare scores, had a significant effect on how long both images were attended. Although welfare scores differed significantly between facilities, no interaction between both factors was found. This implies that non-welfare-related factors varying between facilities, such as environmental conditions during test (e.g., lighting, noise), disparities in the test setup, or other yard differences that we could not measure, may have induced variations in horses' attention to the stimuli (see e.g., Rosenberger et al., 2021 showing that goats' cognitive performance can be influenced by research site). Furthermore, the presence of the experimenter could have altered horses' response to the stimuli since horses' might modify their behavioural response in emotional situations while being handled (Squibb et al., 2018).

We did not apply a standardised stimulus duration (which is the general procedure for human and other animal AB studies), but instead allowed the horses to attend to the images until they directed their attention away from the screens. Pilot horses varied greatly (few seconds to more than 2 min) in attending to the images and setting an arbitrary stimulus cut-off time could have meant that the images disappeared while the horses were still attending to the stimuli, potentially resulting in confounded AB results. If eye tracking was available in horses, more precise measures such as initial attention capture (i.e., which stimulus is attended first) or latency to disengage from stimuli, a more detailed assessment of attentional processing could be conducted.

5. Conclusion

The current study is the first to investigate affect-related AB in horses. Overall, we found no significant evidence of attentional biases when horses viewed image pairs of unfamiliar conspecifics' showing negative, positive, or comparatively more neutral facial expressions. However, attention to the negative emotional stimuli was modulated by overall welfare score and health score, with horses assessed as having putatively lower welfare and poorer health attending to the negative stimuli significantly shorter than horses with putatively better welfare and health. In addition, emotion-unrelated stimulus characteristics as well as experimental circumstances associated with testing location may have influenced their response to visual stimuli. Our findings highlight refinements to experimental designs needed to test attention bias in horses, including a continuing need to further understand horse perceptual abilities in order to allow development of AB tests with better species-appropriate stimuli.

Funding

This work was funded by the University of Plymouth (GD105241-106).

CRediT authorship contribution statement

Sarah Kappel: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marco A Ramirez Montes De Oca:** Writing – review & editing, Investigation. **Sarah Collins:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Katherine Herborn:** Writing – review & editing, Supervision, Formal analysis. **Michael Mendl:** Writing – review & editing, Supervision, Conceptualization. **Carole Fureix:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

No forms of generative artificial intelligence were used in the process of writing this manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the riding schools for the use of their horses and facilities. We thank Sharyn Bistre Dabbah for her help with video analysis for inter-observer reliability testing. We are grateful to the ABW groups at Bristol and Plymouth University for their time categorising the face images for stimulus selection reliability testing.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.applanim.2024.106303.

References

Amici, F., 2019. Memories of emotional expressions in horses. Learn. Behav. *47* (3), 191–192. https://doi.org/10.3758/s13420-018-0363-9.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M.J., van IJzendoorn, M. H., 2007. Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. Psychol. Bull. *133* (1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw.

Bateson, M., Nettle, D., 2015. Development of a cognitive bias methodology for measuring low mood in chimpanzees. PeerJ 3 (11), e998. https://doi.org/10.7717/peerj.998.

Bellegarde, L.G.A., Haskell, M.J., Duvaux-Ponter, C., Weiss, A., Boissy, A., Erhard, H.W., 2017. Face-based perception of emotions in dairy goats. Appl. Anim. Behav. Sci. *193*, 51–59. https://doi.org/10.1016/j.applanim.2017.03.014.

Bethell, E.J., 2015. A "How-To" guide for designing judgment bias studies to assess captive animal welfare. J. Appl. Anim. Welf. Sci. *18* (sup1), S18–S42. https://doi.org/10.1080/10888705.2015.1075833.

Bethell, E.J., Holmes, A., MacLarnon, A., Semple, S., 2012. Evidence that emotion mediates social attention in rhesus macaques. PLoS ONE *7* (8), e44387. https://doi.org/10.1371/journal.pone.0044387.

Blanchette, I., Richards, A., 2010. The influence of affect on higher level cognition: a review of research on interpretation, judgement, decision making and reasoning. Cogn. Emot. *24* (4), 561–595. https://doi.org/10.1080/02699930903132496.

Boissy, A., Manteuffel, G., Jensen, M.B., Moe, R.O., Spruijt, B., Keeling, L.J., Winckler, C., Forkman, B., Dimitrov, I., Langbein, J., Bakken, M., Veissier, I., Aubert, A., 2007. Assessment of positive emotions in animals to improve their welfare. Physiol. Behav. *92* (3), 375–397. https://doi.org/10.1016/j.physbeh.2007.02.003.

Brooks, M., Kristiansen, K., van Benthem, K., Magnusson, A., Berg, C., Nielsen, A., Skaug, H., Maechler, M., Bolker, B., 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R. J. *9* (2), 378–400.

Bruce, V., & Young, A. (2012). Face perception (pp. xi, 481). Psychology Press. ⟨https://doi.org/10.4324/9780203721254⟩.

Burke, D., Sulikowski, D., 2013. The evolution of holistic processing of faces. Front. Psychol. *4* https://doi.org/10.3389/fpsyg.2013.00011.

Chen, J., van den Bos, E., Westenberg, P.M., 2020. A systematic review of visual avoidance of faces in socially anxious individuals: influence of severity, type of social situation, and development. J. Anxiety Disord. *70*, 102193 https://doi.org/10.1016/j.janxdis.2020.102193.

Clarke, A.F., 1987. A review of environmental and host factors in relation to equine respiratory disease. Equine Vet. J. *19* (5), 435–441. https://doi.org/10.1111/j.2042-3306.1987.tb02638.x.

Crump, A., Arnott, G., Bethell, E., 2018. Affect-driven attention biases as animal welfare indicators: review and methods. Animals *8* (8), 136. https://doi.org/10.3390/ani8080136.

Dalla Costa, E., Dai, F., Lebelt, D., Scholz, P., Barbieri, S., Canali, E., Zanella, A., Minero, M., 2016. Welfare assessment of horses: the AWIN approach. Anim. Welf. *25* (4), 481–488. https://doi.org/10.7120/09627286.25.4.481.

Dalla Costa, E., Murray, L., Dai, F., Canali, E., Minero, M., 2014. Equine on-farm welfare assessment: a review of animal-based indicators. Anim. Welf. *23* (3), 323–341. https://doi.org/10.7120/09627286.23.3.323.

Douma, J.C., Weedon, J.T., 2019. Analysing continuous proportions in ecology and evolution: a practical introduction to beta and Dirichlet regression. Methods Ecol. Evol. *10* (9), 1412–1430. https://doi.org/10.1111/2041-210X.13234.

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. Behav. Process. *83* (3), 349–352. https://doi.org/10.1016/j.beproc.2010.01.019.

Fox, J., & Weisberg, S. (2019). An R Companion to Applied Regression (3rd ed.). Sage, Thousand Oaks CA. ⟨https://socialsciences.mcmaster.ca/jfox/Books/Companion/⟩.

Fureix, C., Bourjade, M., Henry, S., Sankey, C., Hausberger, M., 2012. Exploring aggression regulation in managed groups of horses Equus caballus. Appl. Anim. Behav. Sci. *138* (3–4), 216–228. https://doi.org/10.1016/j.applanim.2012.02.009.

Goodwin, D., Davidson, H.P.B., Harris, P., 2002. Foraging enrichment for stabled horses: effects on behaviour and selection. Equine Vet. J. *34* (7), 686–691. https://doi.org/10.2746/042516402776250450.

Gupta, R., 2019. Positive emotions have a unique capacity to capture attention. In: In *Progress in Brain Research*, Vol. 247. Elsevier, pp. 23–46. https://doi.org/10.1016/bs.pbr.2019.02.001.

Gygax, L., 2014. The A to Z of statistics for testing cognitive judgement bias. Anim. Behav. *95*, 59–69. https://doi.org/10.1016/j.anbehav.2014.06.013.

Hanggi, E., 2001. Can horses recognize pictures? Proc. Third Int. Conf. Cogn. Sci., Beijing, China *52–56*.

Hartig, F., & Lohse, L. (2022). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models (0.4.6). ⟨https://CRAN.R-project.org/package=DHARMa⟩.

Heleski, C.R., Shelle, A.C., Nielsen, B.D., Zanella, A.J., 2002. Influence of housing on weanling horse behavior and subsequent welfare. Appl. Anim. Behav. Sci. *78* (2), 291–302. https://doi.org/10.1016/S0168-1591(02)00108-9.

Henneke, D.R., Potter, G.D., Kreider, J.L., Yeates, B.F., 1983. Relationship between condition score, physical measurements and body fat percentage in mares. Equine Vet. J. *15* (4), 371–372. https://doi.org/10.1111/j.2042-3306.1983.tb01826.x.

Henry, S., Fureix, C., Rowberry, R., Bateson, M., Hausberger, M., 2017. Do horses with poor welfare show 'pessimistic' cognitive biases? Sci. Nat. *104* (1–2), 8. https://doi.org/10.1007/s00114-016-1429-1.

Hockenhull, J., Whay, H.R., 2014. A review of approaches to assessing equine welfare: a review of approaches to assessing equine welfare. Equine Vet. Educ. *26* (3), 159–166. https://doi.org/10.1111/eve.12129.

Holcombe, S.J., Jackson, C., Gerber, V., Jefcoat, A., Berney, C., Eberhardt, S., Robinson, N.E., 2001. Stabling is associated with airway inflammation in young Arabian horses. Equine Vet. J. *33* (3), 244–249. https://doi.org/10.2746/042516401776249606.

Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous Inference in General Parametric Models. Biom. J. *50* (3), 346–363.

Howarth, E.R.I., Kemp, C., Thatcher, H.R., Szott, I.D., Farningham, D., Witham, C.L., Holmes, A., Semple, S., Bethell, E.J., 2021. Developing and validating attention bias tools for assessing trait and state affect in animals: a worked example with Macaca mulatta. Appl. Anim. Behav. Sci. *234*, 105198 https://doi.org/10.1016/j.applanim.2020.105198.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014). An Introduction to Statistical Learning with Applications in R. Springer.

Kappel, S., Ramirez Montes De Oca, M.A., Collins, S., Herborn, K., Mendl, M., Fureix, C., 2023. Do you see what I see? Testing horses' ability to recognise real-life objects from 2D computer projections. Anim. Cogn. *26* (4), 1147–1159. https://doi.org/10.1007/s10071-023-01761-6.

Krueger, K., Esch, L., Farmer, K., Marr, I., 2021. Basic Needs in Horses?—A Literature Review. Animals *11* (6), 1798. https://doi.org/10.3390/ani11061798.

Lagisz, M., Zidar, J., Nakagawa, S., Neville, V., Sorato, E., Paul, E.S., Bateson, M., Mendl, M., Løvlie, H., 2020. Optimism, pessimism and judgement bias in animals: a systematic review and meta-analysis. Neurosci. Biobehav. Rev. *118*, 3–17. https://doi.org/10.1016/j.neubiorev.2020.07.012.

Langbein, J., Moreno-Zambrano, M., Siebert, K., 2023. How do goats "read" 2D-images of familiar and unfamiliar conspecifics? Frontiers in Psychology 14. https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1089566.

Lansade, L., Bouissou, M.-F., Erhard, H.W., 2008. Reactivity to isolation and association with conspecifics: a temperament trait stable across time and situations. Appl. Anim. Behav. Sci. *109* (2–4), 355–373. https://doi.org/10.1016/j.applanim.2007.03.003.

Lansade, L., Colson, V., Parias, C., Trösch, M., Reigner, F., Calandreau, L., 2020. Female horses spontaneously identify a photograph of their keeper, last seen six months previously. Sci. Rep. *10* (1), 6302. https://doi.org/10.1038/s41598-020-62940-w.

Lee, C., Cafe, L.M., Robinson, S.L., Doyle, R.E., Lea, J.M., Small, A.H., Colditz, I.G., 2018. Anxiety influences attention bias but not flight speed and crush score in beef cattle. Appl. Anim. Behav. Sci. *205*, 210–215. https://doi.org/10.1016/j.applanim.2017.11.003.

Lee, C., Verbeek, E., Doyle, R., Bateson, M., 2016. Attention bias to threat indicates anxiety differences in sheep. Biol. Lett. *12* (6), 20150977. https://doi.org/10.1098/rsbl.2015.0977.

Leliveld, L.M.C., Langbein, J., Puppe, B., 2013. The emergence of emotional lateralization: evidence in non-human vertebrates and implications for farm animals. Appl. Anim. Behav. Sci. *145*, 1–14. https://doi.org/10.1016/j.applanim.2013.02.002.

Leopold, D.A., Rhodes, G., 2010. A comparative view of face perception. J. Comp. Psychol. *124* (3), 233–251. https://doi.org/10.1037/a0019460.

Lesimple, C., 2020. Indicators of horse welfare: state-of-the-art. Animals *10* (2), 294. https://doi.org/10.3390/ani10020294.

Lesimple, C., Fureix, C., Menguy, H., Hausberger, M., 2010. Human direct actions may alter animal welfare, a study on horses (Equus caballus). PLoS ONE *5* (4), e10257. https://doi.org/10.1371/journal.pone.0010257.

Lesimple, C., Gautier, E., Benhajali, H., Rochais, C., Lunel, C., Bensaïd, S., Khalloufi, A., Henry, S., Hausberger, M., 2019. Stall architecture influences horses' behaviour and the prevalence and type of stereotypies. Appl. Anim. Behav. Sci. *219*, 104833 https://doi.org/10.1016/j.applanim.2019.104833.

Lewis, L.S., Kano, F., Stevens, J.M.G., DuBois, J.G., Call, J., Krupenye, C., 2021. Bonobos and chimpanzees preferentially attend to familiar members of the dominant sex. Anim. Behav. *177*, 193–206. https://doi.org/10.1016/j.anbehav.2021.04.027.

Löckener, S., Reese, S., Erhard, M., Wöhr, A.-C., 2016. Pasturing in herds after housing in horseboxes induces a positive cognitive bias in horses. J. Vet. Behav. *11*, 50–55. https://doi.org/10.1016/j.jveb.2015.11.005.

MacLeod, C., Mathews, A., Tata, P., 1986. Attentional bias in emotional disorders. J. Abnorm. Psychol. *95* (1), 15–20. https://doi.org/10.1037/0021-843X.95.1.15.

McGreevy, P., French, N., Nicol, C., 1995. The prevalence of abnormal behaviours in dressage, eventing and endurance horses in relation to stabling. Vet. Rec. *137*, 36–37. https://doi.org/10.1136/vr.137.2.36.

Mendl, M., 1999. Performing under pressure: stress and cognitive function. Appl. Anim. Behav. Sci. *65* (3), 221–244. https://doi.org/10.1016/S0168-1591(99)00088-X.

Mendl, M., Burman, O.H.P., Paul, E.S., 2010. An integrative and functional framework for the study of animal emotion and mood. Proc. R. Soc. B: Biol. Sci. *277* (1696), 2895–2904. https://doi.org/10.1098/rspb.2010.0303.

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. Appl. Anim. Behav. Sci. *118* (3–4), 161–181. https://doi.org/10.1016/j.applanim.2009.02.023.

Minero, M., Canali, E., 2009. Welfare issues of horses: an overview and practical recommendations. Ital. J. Anim. Sci. *8* (sup1), 219–230. https://doi.org/10.4081/ijas.2009.s1.219.

Minero, M., Dalla Costa, E., Dai, F., Lebelt, D., Scholz, P., 2015. AWIN Welf. Assess. Protoc. Horses. https://doi.org/10.13130/AWIN_HORSES_2015.

Monk, J.E., Belson, S., Colditz, I.G., Lee, C., 2018. Attention bias test differentiates anxiety and depression in sheep. Front. Behav. Neurosci. *12*, 246. https://doi.org/10.3389/fnbeh.2018.00246.

Monk, J.E., Lee, C., Belson, S., Colditz, I.G., Campbell, D.L.M., 2019. The influence of pharmacologically-induced affective states on attention bias in sheep. PeerJ *7*, e7033. https://doi.org/10.7717/peerj.7033.

Monk, J.E., Lee, C., Dickson, E., Campbell, D.L.M., 2020. Attention bias test measures negative but not positive affect in sheep: a replication study. Animals *10* (8), 1314. https://doi.org/10.3390/ani10081314.

Nazarenko, Y., Westendorf, M.L., Williams, C.A., Mainelis, G., 2018. The effects of bedding type in stalls and activity of horses on stall air quality. J. Equine Vet. Sci. *67*, 91–98. https://doi.org/10.1016/j.jevs.2018.03.014.

Ninomiya, S., Kusunose, R., Sato, S., Terada, M., Sugawara, K., 2004. Effects of feeding methods on eating frustration in stabled horses. Anim. Sci. J. *75* (5), 465–469. https://doi.org/10.1111/j.1740-0929.2004.00214.x.

Ninomiya, S., Sato, S., Sugawara, K., 2007. Weaving in stabled horses and its relationship to other behavioural traits. Appl. Anim. Behav. Sci. *106* (1–3), 134–143. https://doi.org/10.1016/j.applanim.2006.06.014.

Ogle, D., Doll, J., Wheeler, P., & Dinno, A. (2022). FSA:Fisheries Stock Analysis [R package version 0.9.3].

Ohman, A., Mineka, S., 2001. Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. Psychol. Rev. *108* (3), 483–522. https://doi.org/10.1037/0033-295x.108.3.483.

Paul, E.S., Harding, E.J., Mendl, M., 2005. Measuring emotional processes in animals: the utility of a cognitive approach. Neurosci. Biobehav. Rev. *29* (3), 469–491. https://doi.org/10.1016/j.neubiorev.2005.01.002.

Pedersen, G.R., Søndergaard, E., Ladewig, J., 2004. The influence of bedding on the time horses spend recumbent. J. Equine Vet. Sci. *24* (4), 153–158. https://doi.org/10.1016/j.jevs.2004.03.013.

Pool, E., Brosch, T., Delplanque, S., Sander, D., 2016. Attentional bias for positive emotional stimuli: a meta-analytic investigation. Psychol. Bull. *142* (1), 79–106. https://doi.org/10.1037/bul0000026.

Proops, L., Grounds, K., Smith, A.V., McComb, K., 2018. Animals remember previous facial expressions that specific humans have exhibited. Curr. Biol. *28* (9), 1428–1432.e4. https://doi.org/10.1016/j.cub.2018.03.035.

R Core team. (2021). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. ⟨https://www.R-project.org/⟩.

Raoult, C.M.C., Gygax, L., 2018. Valence and intensity of video stimuli of dogs and conspecifics in sheep: approach-avoidance, operant response, and attention. Article 7. Animals *8* (7) https://doi.org/10.3390/ani8070121.

Roelofs, S., Boleij, H., Nordquist, R.E., van der Staay, F.J., 2016. Making decisions under ambiguity: judgment bias tasks for assessing emotional state in animals. Front. Behav. Neurosci. *10* https://doi.org/10.3389/fnbeh.2016.00119.

Rørvang, M.V., Nielsen, B.L., McLean, A.N., 2020. Sensory abilities of horses and their importance for equitation science. Front. Vet. Sci. *7*, 633. https://doi.org/10.3389/fvets.2020.00633.

Rosenberger, K., Simmler, M., Langbein, J., Keil, N., Nawroth, C., 2021. Performance of goats in a detour and a problem-solving test following long-term cognitive test exposure. R. Soc. Open Sci. *8* (10), 210656 https://doi.org/10.1098/rsos.210656.

Ruet, A., Lemarchand, J., Parias, C., Mach, N., Moisan, M.-P., Foury, A., Briant, C., Lansade, L., 2019. Housing horses in individual boxes is a challenge with regard to welfare. Animals *9* (9), 621. https://doi.org/10.3390/ani9090621.

Scherer, K.R., 1999. Appraisal theory. In Handbook of cognition and emotion. John Wiley & Sons Ltd, pp. 637–663. https://doi.org/10.1002/0470013494.ch30.

Schoth, D.E., Nunes, V.D., Liossi, C., 2012. Attentional bias towards pain-related information in chronic pain; a meta-analysis of visual-probe investigations. Clin. Psychol. Rev. *32* (1), 13–25. https://doi.org/10.1016/j.cpr.2011.09.004.

Sert, N.P., du, Hurst, V., Ahluwalia, A., Alam, S., Avey, M.T., Baker, M., Browne, W.J., Clark, A., Cuthill, I.C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S.T., Howells, D.W., Karp, N.A., Lazic, S.E., Lidster, K., MacCallum, C.J., Macleod, M., Würbel, H., 2020. The ARRIVE guidelines 2.0: updated guidelines for reporting animal research. PLOS Biol. *18* (7), e3000410 https://doi.org/10.1371/journal.pbio.3000410.

Smith, A.V., Proops, L., Grounds, K., Wathan, J., McComb, K., 2016. Functionally relevant responses to human facial expressions of emotion in the domestic horse (*Equus caballus*. Biol. Lett. *12* (2), 20150907. https://doi.org/10.1098/rsbl.2015.0907.

Squibb, K., Griffin, K., Favier, R., Ijichi, C., 2018. Poker face: discrepancies in behaviour and affective states in horses during stressful handling procedures. Appl. Anim. Behav. Sci. *202*, 34–38. https://doi.org/10.1016/j.applanim.2018.02.003.

Stoehr, A.M., 1999. Are significance thresholds appropriate for the study of animal behaviour? Anim. Behav. *57* (5), F22–F25. https://doi.org/10.1006/anbe.1998.1016.

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. Anim. Cogn. *23* (4), 643–653. https://doi.org/10.1007/s10071-020-01369-0.

Vögeli, S., Wolf, M., Wechsler, B., Gygax, L., 2015. Housing conditions influence cortical and behavioural reactions of sheep in response to videos showing social interactions of different valence. Behav. Brain Res. *284*, 69–76. https://doi.org/10.1016/j.bbr.2015.02.007.

Waring, G.H. (2003). Horse Behaviour (2nd ed.). Noyes Publications.

Wathan, J., Proops, L., Grounds, K., McComb, K., 2016. Horses discriminate between facial expressions of conspecifics. Sci. Rep. *6* (1), 38322 https://doi.org/10.1038/srep38322.

Williams, J., Mathews, A., MacLeod, C., 1996. The emotional Stroop task and psychopathology. Psychol. Bull. *120* (1), 3–24. https://doi.org/10.1037/0033-2909.120.1.3.

Winters, S., Dubuc, C., Higham, J.P., 2015. Perspectives: the looking time experimental paradigm in studies of animal visual perception and cognition. Ethology *121* (7), 625–640. https://doi.org/10.1111/eth.12378.

Yarnell, K., Hall, C., Royle, C., Walker, S.L., 2015. Domesticated horses differ in their behavioural and physiological responses to isolated and group housing. Physiol. Behav. *143*, 51–57. https://doi.org/10.1016/j.physbeh.2015.02.040.

Yiend, J., 2010. The effects of emotion on attention: a review of attentional processing of emotional information. Cogn. Emot. *24* (1), 3–47. https://doi.org/10.1080/02699930903205698.