**Supplementary materials and methods**

**1. Additional methods**

***a) Stimuli validation and FACS codes***

Photographs were calibrated in Adobe Photoshop CS6 for standard white backgrounds and conformity in colour and lighting. Each photo shows models looking directly at the camera with exposed teeth. Facial actions are described using action units (AUs) from the Facial Action Coding System (FACS) [1]. A certified FACS coder (JW) ensured that essential AUs associated with the relevant facial expressions were present in the stimuli. These include: happy, AU12 (lip corner puller) and AU6 (cheek raiser); and angry, AU4 (brow lowerer), AU7 (lid tightener), and AU10 (lip raiser) [2]. Finally, expressions were rated by independent parties blind to the experiment using an open-ended question: “What emotion do you believe is being expressed?”. This yielded 100% agreement on the relevant valences (*N*=8). The marginally super-normal size of the faces served to engage horses’ attention and enhance visual acuity.

***b) Photograph of experimental setup***

******

*Figure I. Depiction of the experimental setup, taken as a still from the video clip of a positive stimulus presentation. Includes distance between the horse’s nose and the stimulus, and the positions of experimenters.*

***c) Behavioural coding scheme***

*Table I. Definitions of behaviourally coded variables*

|  |  |
| --- | --- |
| Behaviour | Coding scheme definition |
| Looking durations |  |
| *Binocular look* | Horse’s head is directed centrally towards the stimulus. If the trial begins whilst the horse is facing away from the centre (which was not a common occurrence), no looking behaviour is coded until the horse deliberately moves its head into a particular orientation. |
| *Monocular look* | The horse is attentive to the stimulus with its head turned to the left or right respectively. Attentiveness is determined by the horse having at least one ear and/or eye focused on the stimulus. |
| Approach and avoid |  |
| *Approach* | Any extension of the horse’s head, or movement of the body, towards the stimulus from the horse’s original position. |
| *Avoid* | Any increase of distance from stimulus combined with one or more concurrent stress-related behaviours (e.g. nostril dilation, head bobbing). |
| Stress-related behaviours |  |
| *Nostril dilation* | The skin above the nostrils is inflated as the air is blown outwards; generally driven by strong exhalation (blowing). |
| *Tail swish* | Horse moves tail several times to the left and right with visible ‘swishing’ movement. |
| *Head bob* | Horse moves nose and head up and down in tight, rapid movements whilst paying attention to the stimulus (attention determined by ear and eye directions). |
| *Lick and chew* | Horse chews and protrudes tongue with no external stimulus as a cause (e.g. not chewing hay or biting wood). |
| *Increased eye whites* | Horse’s eye widens to show additional white sclera compared with their resting state. |

**2. Additional heart rate information**

***a) Heart rate data: Subjects and exclusion criteria***

Heart rate beat corrections were deemed necessary due to errors commonly found in data from Polar Equine heart rate monitors. Once corrected however, Polar HR readings correlate well with ECG recordings [3]. Type 1 errors were most commonly observed – anomalous single-beat spikes in the waveform. Data containing more than 5% heartbeat errors were excluded from analysis (*N*=10). Beats were corrected using Kubios HRV© version 2.2 software (Biosignal Analysis and Medical Imaging Group, 2014); errors were corrected using an Artefact Correction level of 0.3 and smoothed with a Lambda value of 500 as advised by previous research (e.g. [4]). Following correction, one further horse was excluded due to an irregular heartbeat, resulting in 17 subjects in the main HR analyses (13 geldings, 4 mares; ages 7-23 years, *M*=16.24, *SD*=4.55). For heart rate recovery calculations only, a final two horses were removed as their HRs did not return to baseline within 4 minutes after the test ended, leaving 15 subjects (11 geldings, 4 mares; ages 17-23 years, *M*=16.07, *SD*=4.29).

***b) Heart rate recovery calculations***

The heart rate recovery measure represents the time taken for the HR to return to baseline (bpm) following the peak HR observed during the test. The heart rate was considered to have returned to baseline when there were five consecutive beats equal to the baseline mode. Within these five beats, one beat was allowed to exceed the mode by one unit only.

**3. Additional analyses and results**

***a) Reliability analyses***

10 out of 56 videos were double-coded by AS and KG and found to be reliable at or above *r* = .85, *p* ≤ .002 using Spearman’s rho correlations. First monocular look, *r* = 1.0; binocular looking time, *r* = .95; gaze-left, *r* = .96; gaze-right, *r* = .97; approach, *r* = 1.0; avoid, *r* = .85).

***b) Comparisons between blind and not-blind trials***

In 9 out of 56 trials (18%) experimenters were not blind to the stimulus being presented. All such trials occurred in round 1, so Mann-Whitney U tests investigated potential differences between responses to positive and negative stimuli in blind and not-blind trials within trial 1. No significant differences between the blind and not-blind trials were found (see Table II below). N.B. Due to the very low incidence of not-blind trials, some statistical comparisons have low power.

*Table II. Comparison of blind vs. not-blind trials in behavioural measures*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Emotion | Variable | Condition | Mdn | ±95% CI | *W* | *z* | *p* |
| Happy | Laterality: Duration | Blind | 0.06 | 0.24 | 19 | -0.14 | .89 |
|  |  | Not-blind | 0.03 | 0.18 |  |  |  |
|  | Avoid | Blind | 1.14 | 2.1 | 65.5 | -1.40 | .16 |
|  |  | Not-blind | 6.47 | 4.43 |  |  |  |
|  | Approach | Blind | 2.30 | 3.56 | 24 | -0.87 | .39 |
|  |  | Not-blind | 0.22 | 3.68 |  |  |  |
| Angry | Laterality: Duration | Blind | 0.21 | 0.31 | 15 | -1.00 | .32 |
|  |  | Not-blind | 0.20 | 0.44 |  |  |  |
|  | Avoid | Blind | 2.84 | 2.99 | 57 | -1.41 | .16 |
|  |  | Not-blind | 5.04 | 2.38 |  |  |  |
|  | Approach | Blind | 0.00 | 2.67 | 32 | -0.86 | .39 |
|  |  | Not-blind | 0.00 | 1.57 |  |  |  |

*Happy: blind n=10, not-blind n=4; angry: blind n=9, not-blind n=5*

*Table III. Comparison of blind vs. not-blind trials in heart rate measures*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Emotion | Variable | Condition | Mdn | ±95% CI | *W* | *z* | *p* |
| Happy | Start to peak (s) | Blind | 33.00 | 3.68 | 6 | -1.00 | .32 |
|  |  | Not-blind | 28.40 | 7.25 |  |  |  |
|  | HR change (bpm) | Blind | 0.85 | 2.63 | 7 | -0.67 | .51 |
|  |  | Not-blind | -0.91 | 1.71 |  |  |  |
|  | HR recovery (s) | Blind | 6.05 | 16.63 | 26 | -0.34 | .74 |
|  |  | Not-blind | 16.85 | 26.75 |  |  |  |
|  | HR max (bpm) | Blind | 42.00 | 9.02 | -24.5 | -0.84 | .40 |
|  |  | Not-blind | 57.00 | 33.32 |  |  |  |
| Angry | Start to peak (s) | Blind | 1.90 | 14.72 | 21 | -0.98 | .33 |
|  |  | Not-blind | 25.1 | 11.47 |  |  |  |
|  | HR change (bpm) | Blind | 1.08 | 7.41 | 21 | -0.98 | .33 |
|  |  | Not-blind | 2.70 | 2.96 |  |  |  |
|  | HR recovery (s) | Blind | 51.00 | 28.83 | 23 | -0.49 | .62 |
|  |  | Not-blind | 41.20 | 52.28 |  |  |  |
|  | HR max (bpm) | Blind | 35.00 | 17.38 | 20 | -1.23 | .22 |
|  |  | Not-blind | 53.00 | 10.23 |  |  |  |

*Happy: blind n=6, not-blind n=2; angry: blind n=5, not-blind n=5*

***c) Descriptive statistics for heart rate data***

*Table IV. Heart rate data medians and 95% confidence intervals*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Positive |  | Negative |  |
|  | **Mdn** | **±95% CI** | **Mdn** | **±95% CI** |
| Latency to reach max HR (s) | 32.1 | 4.01 | 19.4 | 6.50 |
| HR change (baseline to test) (bpm) | 0.34 | 3.23 | 1.08 | 2.35 |
| Maximum HR during test (bpm) | 41.0 | 6.44 | 41.0 | 5.71 |
| Recovery time (return to mode) (s) | 7.4 | 26.19 | 24.5 | 30.85 |

***d) Analysis of stress-related behaviours***

Figure II shows the number of horses who performed each stress-related behaviour towards each emotional expression. Apart from increased white visible sclera, behaviours are performed more often to the negative stimuli, though no individual behaviour is significantly different (two-tailed binomial tests) between positive and negative stimuli: nostril dilate, *p* = .36; eye whites, *p* = 1.0, tail swish, *p* = 1.0; head bob, *p* = .75; lick and chew, *p* = 1.0. The showing of eye whites may not have been a reliable measure of stress as horses were observed to show eye whites not only when avoiding stimuli, but also when they extended and raised their nose to look, touch, or otherwise explore stimuli.

*Figure II. Number of horses performing each stress-related behaviour by expression*

**Supplementary references**

1. Ekman P, Friesen WV. 1978 *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto: Consulting Psychologists Press.

2. Waller BM, Cray JJ, Burrows AM. 2008 Selection for universal facial emotion. *Emotion* **8**, 435–439. (doi:10.1037/1528-3542.8.3.435)

3. Marchant-Forde RM, Marlin DJ, Marchant-Forde JN. 2004 Validation of a cardiac monitor for measuring heart rate variability in adult female pigs: Accuracy, artefacts and editing. *Physiol. Behav.* **80**, 449–458. (doi:10.1016/j.physbeh.2003.09.007)

4. Schmid A, Biau S, Möstl E, Becker-Birck M, Morillon B, Aurich J, Faure JM, Aurich C. 2010 Changes in cortisol release and heart rate variability in sport horses during long-distance road transport. *Domest. Anim. Endocrinol.* **38**, 179–189. (doi:10.1016/j.domaniend.2009.10.002)