Applied Animal Behaviour Science 300 (2026) 106998

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

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

Face to FACE® – Investigating horses' perception of facial expressions performed by an android robot

Paolo Baragli [a,b], Lorenzo Frassineti [c], Martina Felici [d], Claire Ricci-Bonot [e], Alice Galotti [f], Micaela Sgorbini [a], Elisabetta Palagi [f], Lorenzo Cominelli [g], Enzo Pasquale Scilingo [b,g], Chiara Scopa [a,*], Antonio Lanatà [c]

[a] *Department of Veterinary Sciences, University of Pisa, Viale delle Piagge 2, Pisa 56124, Italy*
[b] *Research Center "E. Piaggio", University of Pisa, Largo L. Lazzarino 1, Pisa 56122, Italy*
[c] *Department of Information Engineering, Università degli Studi di Firenze, Via Santa Marta 3, Florence 50139, Italy*
[d] *Department of Agricultural and Food Sciences, University of Bologna, Viale G. Fanin 46, Bologna 40127, Italy*
[e] *Animal Behaviour, Cognition and Welfare Group, Dept of Life Sciences, University of Lincoln, Lincolnshire LN6 7DL, United Kingdom*
[f] *Unit of Ethology, Department of Biology, University of Pisa, Via Volta 6, Pisa 56126, Italy*
[g] *Department of Information Engineering, University of Pisa, Largo L. Lazzarino 1, Pisa 56122, Italy*

ARTICLE INFO

Keywords:
Human-animal interaction
Heart Rate Variability
EquiFACS
Emotional communication
Humanoid robot
Facial expression recognition

ABSTRACT

The ability to perceive and respond to emotional cues across species plays a central role in human–animal interactions and has implications for animal welfare, cognition, and social communication. Horses are believed to be able to understand the valence of human facial emotions presented in photographs, videos, and live interactions. However, the extent to which horses rely on purely visual facial cues, independently of human presence and other sensory modalities, remains underrated. To address this issue, the present study investigated whether horses can distinguish human-like emotional facial expressions displayed by an android robot under highly controlled conditions. Twelve adult mares were individually exposed to four facial expressions (neutral, happy, angry, surprise) performed by the robot FACE (Facial Automaton for Conveying Emotions®), with no humans present during testing. Heart rate variability (HRV), behavioural responses, and facial movements coded using the Equine Facial Action Coding System (EquiFACS) were recorded. The experimental design included baseline periods before and after exposure to the robot, allowing assessment of autonomic recovery and potential carry-over effects. Exposure to the robot induced significant autonomic changes compared to baseline, as reflected by shifts in HRV parameters across all conditions. Emotional expressions elicited concurrent sympathetic and parasympathetic modulations, indicating increased arousal. However, differences among emotional expressions were limited and inconsistent. Behavioural responses were non-specific, with increased locomotor activity observed across both neutral and emotional conditions. EquiFACS analysis revealed no significant variation in Action Units' recruitment depending on the expressions, suggesting that horses displayed a common facial response pattern regardless of the robot's display. To further explore multivariate response patterns, a machine-learning approach combining physiological and behavioural features was applied. Classification accuracy was moderate, with the neutral condition being identified more reliably than emotional expressions, which showed substantial overlap. Together, these findings indicate that while the android robot elicited measurable autonomic and behavioural activation, horses did not show evidence of discriminating the emotional valence of human-like facial expressions based on visual cues alone. Overall, the results suggest that horses' responses reflected a general affective arousal rather than emotion-specific recognition. These findings challenge assumptions about equine sensitivity to isolated human facial expressions and highlight the importance of multimodal cues (i.e. vocal, postural, or contextual information) in interspecies emotional communication. The

* Corresponding author.
E-mail address: chiara.scopa@vet.unipi.it (C. Scopa).

https://doi.org/10.1016/j.applanim.2026.106998
Received 17 February 2026; Received in revised form 16 March 2026; Accepted 27 March 2026
Available online 2 April 2026

use of humanoid robots provides a valuable methodological tool for isolating visual sensory channels and controlling for unintentional human influence, offering new perspectives for research on animal emotions, cognition, and welfare.

1. Introduction

Social animals engage in emotional exchanges, within and across species. Emotions serve communicative functions through mechanisms such as emotional contagion, whereby one individual's emotional state influences that of another (Pérez-Manrique and Gomila, 2022). Given the importance of human–animal relationships, increasing attention has focused on how domestic animals perceive and respond to human emotions (Baciadonna et al., 2018; Benz-Schwarzburg et al., 2020; Jardat and Lansade, 2022). Research suggests that some domestic mammals recognize human emotions and modify their behavior accordingly (i.e. interspecific emotional contagion) (Galvan and Vonk, 2016; Huber et al., 2017; Trösch et al., 2020). Horses are a versatile species exhibiting advanced interspecific cognition that extends to complex communication with humans (Scopa et al., 2019). They adapt their behavior to human facial and vocal cues, whether presented live (Lansade et al., 2020; Proops et al., 2018), via images, or in videos (Nakamura et al., 2018; Smith et al., 2016) demonstrating interspecies emotional contagion (Jardat et al., 2025). Notably, horses show lateralized behavioral responses to variations in emotional valence and arousal, such as joy or fear. Emotions modulate autonomic, endocrine, and motor responses to environmental challenges (Panksepp, 1986; Venkatraman et al., 2017; Waller et al., 2020). Facial expressions, arising from coordinated muscular, neural, and autonomic activity, play a key role in emotional contagion in humans (Herrando and Constantinides, 2021) and non-human animals (Jardat and Lansade, 2022; Tate et al., 2006). Although individual and experiential factors can obscure direct physiology–behaviour links, physiological measures remain valuable indicators of affective states (McCraty and Zayas, 2014). Heart rate variability (HRV) reflects autonomic nervous system regulation and beat-to-beat fluctuations (Task Force, 1996; Von Borell et al., 2007; Karemaker, 2022) and is widely used as a marker of arousal in humans and other animals (McCraty and Shaffer, 2015; Shi et al., 2017). In horses, HRV can be useful tool in assessing arousal along emotional and stress-related responses (Kuwahara et al., 1996; Visser et al., 2002).

Research on humans' emotional contagion increasingly employs android robots capable of reproducing human facial expressions. Social robotics integrates neuroscientific insights into machines designed for human-like interaction (Cross et al., 2019) and provides a controlled tool to investigate social cognition (Henschel et al., 2020). Androids that closely resemble humans (Ishiguro and Nishio, 2007) enable standardized, repeatable, and ecologically valid presentations of emotional stimuli (Sato et al., 2022). Based on the Facial Action Coding System (FACS; Ekman and Friesen, 1978), they generate defined facial configurations (Action Units, AUs) with high precision. Furthermore, their three-dimensional human-like presence enhances realism compared with static images or virtual agents (Li, 2015).

Following this framework, the present study examined whether horses discriminate the valence of emotional expressions displayed by an android robot (Facial Automaton for Conveying Emotions®, FACE): if horses are sensitive to visual emotional cues, distinct physiological and behavioural responses should emerge across robot expressions. To explore this possibility, twelve mares were individually tested in a stall while facing the robot and wearing validated textile electrodes to record electrocardiogram (ECG) and compute HRV (Felici et al., 2021; Guidi et al., 2017).

Previous studies showed that horses recognize human emotional valence across photographs, videos, and live interactions (Baba et al., 2019; Jardat et al., 2023; Jardat et al., 2024; Merkies et al., 2022; Proops et al., 2018; Smith et al., 2016). We therefore predict differentiation of robot-displayed expressions at physiological and behavioural levels. Angry faces (negative valence) should increase sympathetic and reduce vagal activity, whereas happy faces (positive valence) should enhance vagal activation (Prediction 1). Behaviourally, horses are expected to show a left-eye bias toward angry faces and a right-eye bias toward happy faces, with greater attention to emotional than neutral expressions (Prediction 2). Surprise expressions should elicit intermediate responses, such as longer looking times than neutral faces. Finally, horse facial expressions are predicted to vary in response to emotional, but not neutral, robot displays (Prediction 3), although interpretation should remain cautious given the limited application of EquiFACS in equine research (Lundblad et al., 2021; Phelipon et al., 2024; Rashid et al., 2020; Tomberg et al., 2023).

2. Materials & methods

2.1. Animals and husbandry

Twelve healthy Standardbred mares (mean age 8.8 ± 1.1 years), housed at the Department of Veterinary Sciences of the University of Pisa (Italy), were selected for this study. The mares were kept in paddocks (75×75 m) and fed concentrate, hay, and water *ad libitum*. All had previously been used as recipients in an embryo transfer program, and none were pregnant during the tests.

To minimize potential emotional interference, the mares underwent a 30-day acclimation period to the experimental procedures. During this time, they became accustomed to staying alone in a box for about one hour without displaying behaviours indicative of an altered emotional state. The mares were also accustomed to wearing the chest belt and to the mild noises produced by the robot's electromechanical equipment during operation, in order to prevent motor sounds from representing a novel stimulus during the experimental sessions.

2.2. Design of the experiment

The mare was led into the test box by a handler using a halter. Once released, it was left alone inside the box for 5 min, with the door open, to record basal ECG signals and behaviours (Baseline1). Then the door was closed, and the robot was placed in front of the test box door, sitting on a chair, approximately 2 m away from the horse. After that, the stall door was open again, and the mare could see the robot in front of her (see Fig. 1 for the setup of the experiment). A recording camera (HDC-SD9, Panasonic, Osaka, Japan) was positioned behind the robot, elevated above its head, to capture the horse's behaviour. The robot was positioned, according to the peculiarities of this species' vision (Rørvang et al., 2020). Two wooden crossbars were placed to prevent the animal from leaving the stall and avoid any physical contact with the robot. After opening the box door, the individuals who had positioned the robot moved to the adjacent box. During the test, all personnel involved remained in the adjoining box and stood there in silence. This was necessary to avoid the "Clever Hans effect" (Samhita and Gross, 2013). In the stall adjacent to the horse's, a console was installed to remotely control the robot's facial expressions. For the first 2 min, the robot maintained a neutral facial expression (EXP0). After this period, the robot's expression was changed to one of the pre-selected emotional faces using a remote control. This expression was maintained for an additional 2 min (EXP1). Then, the robot's expression returned to the neutral face for 30 s (neutral). Next, the second emotional face was displayed for another 2-min block (EXP2). The same sequence was repeated for the last emotional face (EXP3). The robot displayed a total of three emotional faces: 'surprise' (S) (Fig. 1(a)), 'happy' (H) (Fig. 1(b)), and 'angry' (A) (Fig. 1(c)) (Nakamura et al., 2018; Smith et al., 2016) (see FACE facial expressions video in Supplementary Materials). The sequence of the facial expressions was randomized across horses, whereas the 'neutral' expression (N) (Fig. 1(d)) was always presented first and served as a control (see Table S.1 for the exact sequence). After the last trial, the stall door was closed, and the robot was moved out of the mare's sight. A baseline data collection was then repeated using the same methodology for an additional 5 min after the end of testing (Baseline2). The sequence of the experimental design is reported in Figure S.1. The transitions between the robot's expressions were synchronized using a stopwatch and the video recording. The timing of each phase was recorded on paper, allowing precise determination of the start and end times of each phase from the videos for behavioural analysis (Fig. 2).

2.3. FACE android robot

The robot implemented in this study is a humanoid called FACE (Facial Automaton for Conveying Emotions®), characterized by hyper-realistic adult female features and specifically designed for social robotics (Pioggia et al., 2004). It is composed of a passive body on the top of which a Hanson Robotics' head has been mounted. The head is designed to accommodate 32 servomotors that control the robot's neck, eyes, mouth, and facial expressions. The face of the robot is made of Frubber1, a registered material with skin-like mechanical and aesthetic features. This hardware is controlled by SEAI (Social Emotional Artificial Intelligence), a distributed control architecture comprising perception, cognitive, and actuation systems, which endow the robot with expressive and communicative capabilities (Cominelli et al., 2018). The facial expressions implemented on FACE are based on configurations derived from the Facial Action Coding System (FACS; Ekman and Friesen, 1978), which describes facial expressions in terms of anatomically grounded facial muscle activations called Action Units (AUs). Through coordinated control of the facial servomotors, the robot can reproduce prototypical emotional facial configurations in a controlled and repeatable manner. This actuation architecture was adopted to provide highly standardised visual stimuli while minimising variability in the presentation of facial cues across trials. The distributed control architecture enables coordinated activation of multiple facial regions, including the eyes, mouth, and neck, allowing reproducible expression patterns across trials. This configuration makes the platform particularly suitable for experimental paradigms requiring strict control of stimulus presentation. FACE has been previously evaluated as a social robot platform in therapeutic/HRI contexts, and SEAI has been implemented and tested on the robot, supporting its use as a controlled source of human-like facial cues in experiments (Mazzei et al., 2011; Cominelli et al., 2018).

2.4. Data collection

2.4.1. Heart Rate Variability

Across all the experimental conditions, ECG traces were obtained from each horse to extrapolate HRV. A textile system developed and validated by the University of Pisa, consisting of a belt placed around the horse's chest containing two textile electrodes (size 13 ×5 cm, developed by Smartex Srl, Navacchio, Italy) placed in a modified base-apex configuration, was used to acquire the ECG tracing (Felici et al., 2021; Guidi et al., 2017). This system appears to be non-invasive to the animal as it does not require shaving of the coat or the use of adhesive electrodes and the horses were accustomed to wearing a saddle or similar equipment. The system was equipped with a recording device (Biopac, Goleta, USA) featuring Bluetooth Low Energy (BLE) connectivity, allowing for real-time monitoring of ECG signals via a mobile device and a dedicated application. ECG recordings were made at a sampling rate of 250 Hz. The belt was also equipped with a piezoelectric sensor to detect respiratory dynamics, which was necessary to eliminate artefacts due to sinus arrhythmia. Before starting the initial phase of the test, the horses were habituated to wear the belt system. It was necessary to let the horses familiarize themselves with the belt and to check the system's functionality remotely.

The first phase of the ECG trace analysis involved identifying the R-peaks of the QRS complexes using the QRS detector in Kubios software, based on the Pan-Tompkins algorithm. Each RR sequence obtained was then subjected to a series of preprocessing steps: 1) band-pass filtering

Fig. 1. Experimental setup and robot facial expressions: (a) 'surprise' (S), (b) 'happy' (H), (c) 'angry' (A), (d) 'neutral' (N).

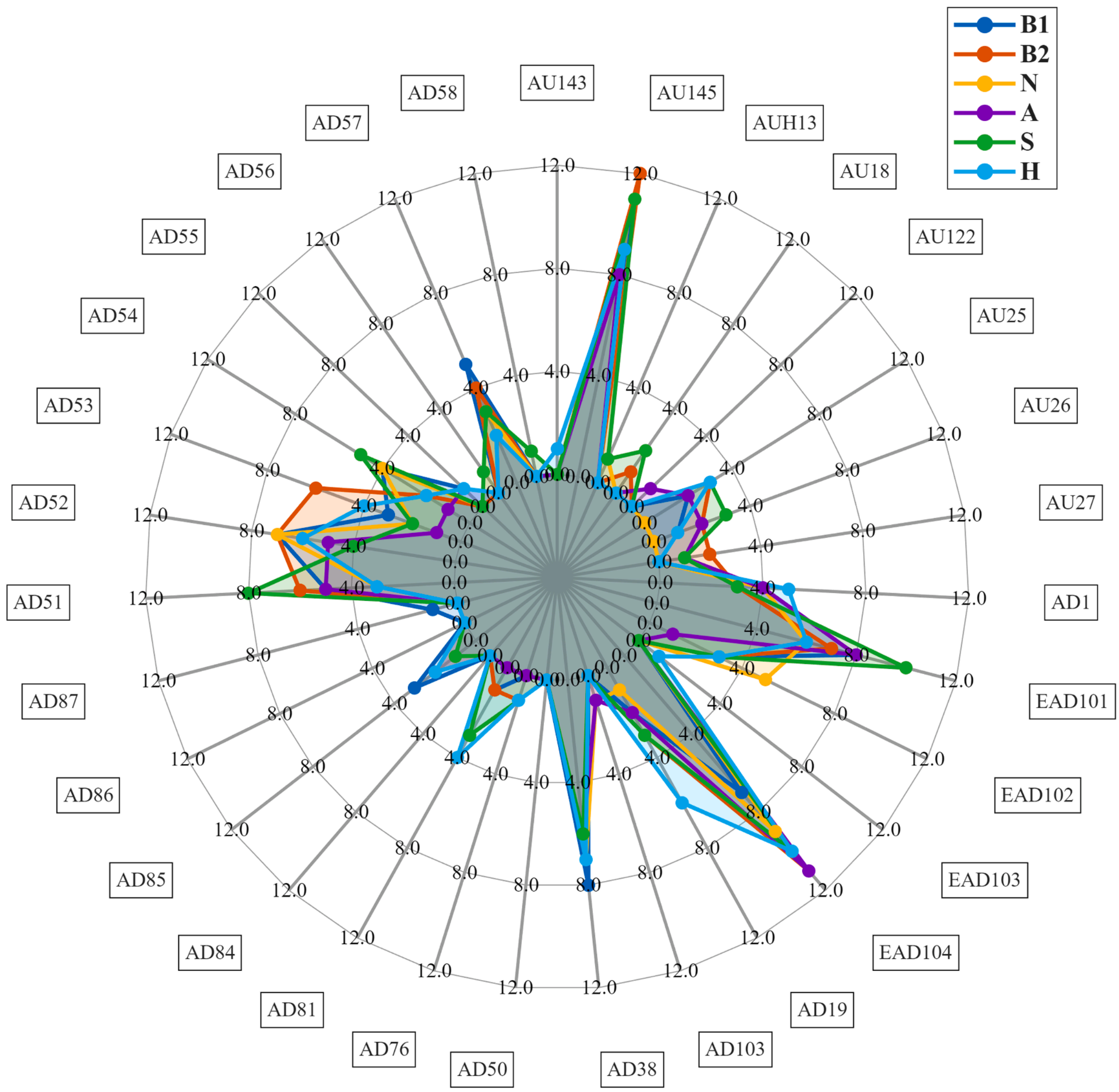

Fig. 2. Spider plot representing the EquiFACS distribution among FACE stimuli and conditions. B1 = baseline 1, B2 = baseline 2, N = 'neutral', A = 'angry', S = 'surprise', H = 'happy'.

(0.05–40 Hz) to minimise noise, 2) squaring of the data to emphasise R-peaks, and 3) application of a moving average filter (with a 150 ms window) to define adjacent peaks. Preprocessing decision rules included amplitude thresholding and comparison with expected values between consecutive R-waves. Any ectopic R peaks were rectified using an algorithm specialised in the removal of motion artefacts for equine ECG recordings (Lanatà et al., 2015).

The RR intervals obtained were then used to calculate HRV in time- and frequency-domains, as well as nonlinear parameters (Acharya, 2006; Task Force, 1996). Regarding the frequency domain, autoregressive modeling was used to estimate HRV spectra and define two primary spectral bands: low frequency (LF, 0.01–0.07 Hz) and high frequency (HF, 0.07–0.6 Hz), as reported in the literature for horses (Cottin et al., 2005; Kuwahara et al., 1996; Stucke et al., 2015). All procedures for physiological signal analysis were performed blindly, and the operator was unaware of the robot's emotional phases. The HRV parameters extracted, along with their brief descriptions, can be found in Supplementary Materials, Table S.2.

2.4.2. Behaviours

Two students analysed the collected videos, unaware of the test, and specifically trained in behaviour analysis, using the focal animal sampling method. Placing the camera behind the robot's head enabled blind behaviour analysis with the two students who independently scored the behaviours without knowing the sequence of emotional expressions displayed by the robot to the horse. Regarding the behavioural features, we considered a total of 10 variables: attention to the left, attention to the right, exploration to the left, exploration to the right, total looking time, walking behaviours, standing behaviours, and instances when the subject was out-of-view (see Supplementary Materials, Table S.3, for a detailed description of the behaviours considered).

2.4.3. Equine Facial Action Coding System

For each video collected, a random 10-second sample was extracted using Shotcut software (version 21.03.21) for a total of 6 clips for each horse (one sample for each phase, Baseline1, neutral face, emotional face 1, emotional face 2, emotional face 3, Baseline2). Those videos were analysed using EquiFACS to detect horses' emotional facial expressions. The starting point of the clip was randomly selected by using the 'sample' function in R software (version 4.1.1). A video sample was considered valid when the horse was visible for at least 10 s. A total of 72 video samples were extracted, and 68 of them were considered valid for the

EquiFACS analysis. The videos were coded by two certified EquiFACS coders (C.R.B. and A.G.). Each coder scored 27 videos independently, and 14 videos were randomly selected to be scored by both coders for the reliability check. Facial movements were coded according to the EquiFACS manual (https://animalfacs.com/equifacs_new), with Action Units (AUs) recorded as present or absent using a one–zero sampling method (Bateson and Martin, 2021). Inter–observer reliability was assessed by treating each AU as a binary variable (present/absent) per video. For each video, AU vectors from each coder were extracted, concatenated across the dataset, and a single Cohen's Kappa (Landis and Koch, 1977) coefficient was computed ($\kappa = (P_o - P_e)/(1 - P_e)$). The resulting reliability was $\kappa = 0.78$, indicating "substantial agreement". Overall, coders showed consistent agreement across the majority of AUs: more than half of the AUs reached moderate to substantial reliability, with several units achieving perfect agreement, while only a small subset of ears movement AUs (EAD101 and EAD 104) displayed lower reliability.

2.5. Statistical analysis

Regarding the HRV features, the first analysis performed was a paired test between the first baseline period (Baseline1) and each subsequent period for every extracted HRV feature. The Shapiro-Wilk test was used to assess the normality assumption (level of significance $p < 0.05$). When the assumption of normality was met, the Paired t-Test was used; otherwise, the non-parametric Wilcoxon Signed Rank Test was applied. Moreover, the Bonferroni correction for multiple hypotheses was applied. Then, a group analysis among the conditions with the presence of the robot was performed (i.e., considering the HRV features obtained for the following FACE stimuli: 'neutral', 'happy', 'angry', and 'surprise'). For each horse, HRV features were normalized according to the corresponding first Baseline1 value, following this formula:

$$HRV_{FACE\_NORM} = HRV_{FACE}/HRV_{Baseline1}$$

where HRV_{FACE} refers to any of the FACE stimuli considered, other than Baseline1. Considering that no differences were found in the previous paired test on Baseline2, this condition was excluded from this analysis (for further details, please see Section Results). When the assumption of normality was met, the ANOVA test was performed; otherwise, the Friedman test was used (level of significance: $p < 0.05$). Moreover, pairwise comparisons were performed among the different stimuli administered, with the Bonferroni correction applied for multiple comparisons. For the behavioural analysis, the same statistical analysis conducted for HRV features was considered. The only difference was in the normalization of behavioural features (hereafter BEHA) for the group analysis, which in this case followed the formula: $BEHA_{FACE\_NORM} = BEHA_{FACE} - BEHA_{Baseline1}$.

Finally, for the EquiFACS analysis, since the data were categorical (0 =absent, 1 =present), a Generalized Linear Model (GLM) was built for each pairwise comparison between FACE stimuli. For each GLM, the binomial distribution and the logit link function were considered. For each AU, the level of significance was set to 0.05.

2.6. Machine-learning analysis

To identify potential relationships across the feature domains considered (mainly HRV and BEHA) and assess whether a multivariate characterization of the horse's status during FACE stimuli was feasible, a preliminary Machine Learning (ML) experiment was conducted. The ML workflow consisted of the following points:

- -from the original dataset and according to the experiment performed, BEHA normalized features and HRV normalized features were imported in Matlab (MATLAB version R2023b, Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com), as well as the classes for each observation (e.g., 'neutral' – 'happy' – 'angry' – 'surprise': N-H-A-S). Since the normalized versions of the features were used, and considering the results obtained during the statistical analysis, the baseline conditions (B1 and B2) were excluded from all the ML experiments. This was done to characterize the horses' status in the presence of the robot.
- the leave-One-Subject-Out (LOSO) validation framework was used following a Bayesian optimization scheme. In this work, the ensemble supervised models Bag and AdaBoosting were considered. For more details regarding the validation framework and hyperparameter optimization, please refer to the Supplementary Materials, Machine Learning Validation Framework.

In summary, the first ML experiments performed considered all the FACE stimuli (i.e., 'neutral' – 'happy' – 'angry' – 'surprise': N-H-S-A) and all possible combinations among feature domains (i.e., HRV, BEHA, and HRV+BEHA). Then, as a form of ablation study, each FACE stimulus was selectively removed in various combinations, ensuring that at least three different conditions remained in each case. This allowed us to understand how the performance of the models differs according to the FACE stimuli included in the ML experiments and whether the feature domain (i.e., HRV and BEHA) is necessary or redundant for some stimuli. All these results are reported and summarized in the Results Section.

3. Results

Due to the complexity of the analysis, the results are presented in sections corresponding to the analysed parameters.

3.1. Heart Rate Variability

HRV time domain parameters. Compared to Baseline1, the statistical analysis revealed a significant increase in RR distance when the horses faced the robot's 'happy' and ‘surprise’ expressions. Since RR distance is an inverse measure of heart rate (HR), the HR of horses decreased during the robot's 'happy' and ‘surprise’ expressions compared to the baseline condition (Baseline1). Furthermore, stdRR, TINN, and stdDER2, other features of the time domain, decreased during the robot's ‘angry’ expression, indicating that this expression reduces HRV in horses (see Table 1 for statistical details).

HRV frequency domain parameters. The analysis revealed an increase in HF when the robot displayed 'angry' and 'surprised' expressions. Among the nonlinear parameters, only SD2RR showed significant differences compared to Baseline1, decreasing during both 'angry' and 'surprise' expressions, indicating a reduced autonomic modulation (see Table 1 for statistical details). Interestingly, no differences were found between Baseline1 and Baseline2, i.e., when the robot was not present. These data indicate that the horses' sympathetic-parasympathetic balance quickly returned to baseline levels, regardless of the changes induced by the robot's expressions, suggesting that the horses did not

Table 1
Statistical results for the paired tests between Baseline1 conditions and the other FACE stimuli (only the statistically significant features are reported). Both the p-value obtained, the median direction of the HRV features (↑ = greater in the other FACE conditions, ↓ = greater in the Baseline1) and the statistically significant comparisons are reported. H= ’happy’, A= ’angry’, S= ’surprise’.

HRV feature	Significant pairwise comparison FACE conditions (p-value, direction)
meanRR	H (e-04, ↑), S (e-04, ↑)
stdRR	A (e-04, ↓)
TINN	A (e-04, ↓)
stdDER2	A (e-04, ↓)
HF prc pow (Welch)	A (e-03, ↑), S (e-03, ↑)
SD2RR	A (e-03, ↓), S (e-03, ↓)

experience any long-term effects from observing them (see Table 1 for statistical details).

The different facial expressions performed by the robot were compared to the ‘neutral’ one (excluding the two baselines), in terms of horses' cardiac correlates. Compared to the ‘neutral’ expression, the RR distance was higher when the horse faced the S expression, suggesting that the horses had a lower average HR during the robot's 'surprise' expression than during the ‘neutral’ expression. Furthermore, our results showed that cardiac entropy was lower (i.e., the signal is more predictable and less complex) during the ‘neutral’ expression than during the ‘happy’ and ‘angry’ expressions.

Meanwhile, significant statistically difference in the sympathetic–vagal regulation of HR was described by the Hurst parameter, when the horses were subjected to the A expression.

On the other hand, the SD12_LPP parameters (1, 2, 6, and 7) indicated modulation of the slow dynamics of HRV at different temporal scales during the emotional faces (‘happy’, ‘angry’ and ‘surprise’) compared to the ‘neutral’ face.

Therefore, significant differences emerged when FACE expressed an emotion compared to when it remained neutral in front of the horses (*Prediction 1* partially supported; see Table 2 for statistical details).

3.2. Behaviours

Among these, the analysis focused on walking (*Walk_F*) and standing (*Stand_F*), as they were the only features yielding meaningful data (*Prediction 2* not supported). After the post-hoc correction, only the comparison between Baseline1 and the 'surprise' and ‘angry’ expressions for the behaviours walk and stand resulted statistically significant ($p = 0.004$ and $p = 0.002$ for stand; $p = e\text{-}05$ and $p = 0.001$ for walk, respectively). Specifically, horses walked less during the S condition (expressed as % of time); however, a high degree of inter-individual variability was observed. As for HRV analysis, no differences were found between Baseline1 and Baseline2 for all the behavioural features considered. For the group analysis of the FACE stimuli, excluding the baseline periods, only few difference where found between ‘angry’ with ‘neutral’ and ‘happy’ ($p = 0.04$ and $p = 0.01$ respectively, see Table 3 for statistical details).

3.3. EquiFACS

From the EquiFACS analysis, no differences emerged between the facial expressions displayed by the robot (*Prediction 3* not supported).

During the statistical analysis, the EquiFACS features did not obtain any statistically significant results, showing a poor level of discrimination between the FACE stimuli. For these reasons, they were excluded from the machine-learning analysis and are only partially reported in this section.

Table 2
Statistical results for the analysis performed on HRV features between all the FACE stimuli. Both the p-value related to differences between groups (level of significance 0.05) and the statistically significant comparison are reported. N = ’neutral’, H= ’happy’, A= ’angry’, S= ’surprise’. ↓ denotes a median lower value of N conditions than the other one tested.

HRV feature	p-value	Significant pairwise comparison with N (p-value, direction)
meanRR	0.032	N vs. S (0.04, ↓)
N bis sq Entr	0.007	N vs. H (0.01, ↓), N vs. A (0.03, ↓)
Hurst	0.047	N vs. A (0.04, ↓)
SD12_LPP_1	0.017	N vs. H (0.04, ↓) N vs. S (0.03, ↓)
SD12_LPP_2	0.029	N vs. H (0.03, ↓)
SD12_LPP_6	0.029	N vs. A (0.03, ↓)
SD12_LPP_7	0.032	N vs. A (0.03, ↓)

3.4. Machine Learning

In Fig. 3, the confusion matrix for the ML experiment, considering all FACE stimuli, is reported. In this case, an overall accuracy of 59% was obtained. The best model was an AdaBoosting with the following hyperparameters: learn rate= 0.0027, number learning cycles= 248, minimun leaf size = 8, mamixum number of split = 40, number of features = 9, as determined by the Relief algorithm (k = 5). The 9 features used in the best model were a combination of HRV and BEHA domains. The following features were considered: NN50; DET_RQA (determinism of recurrence plot); RECT1_CRQA (Recurrence Time of 1st Type for the cross-recurrence plot), Displ_F; Std_P; DFA; meanDER1; FractDim; meanDER2. It is noteworthy that the FACE's 'neutral' conditions achieved a True Positive Rate (TPR) of nearly 90%. At the same time, the highest classification error occurred in discriminating among the other FACE stimuli (i.e., ‘happy’, ‘angry’, ‘surprise’), with a mean False Negative Rate of 51%. For a complete overview of the model's performance, in addition to the confusion matrix, in Fig. 3, the following metrics are reported for each FACE stimulus: True Positive Rate (TPR), False Negative Rate (FNR), Positive Predictive Value (PPV), and the Negative Predictive Value (NPV).

Furthermore, Table 4 reports the other ML results obtained by excluding each FACE stimulus iteratively. It is noteworthy that the performance of the models increased compared to the 4 classes experiments, spanning from 69% to 78% in accuracy. Interestingly, it appears that the N conditions resulted in an increase in model complexity, at least in terms of the number of features, as well as the inclusion of both HRV and BEHA features. Moreover, the inclusion of HRV features allowed a high TPR only for N conditions (but a lower TPR for the other FACE stimuli for the case ‘neutral’-’happy’-’angry’ N-H-A and ‘neutral’-’happy’-’surprise’ N-H-S in Table 4). Instead, when stimuli other than N were considered (e.g., H-S-A), the best model was obtained using only the BEHA features. Moreover, in the case of N-S-A, considering the best model with only the BEHA features, the 'neutral' condition was the worst in terms of TPR (50%).

4. Discussion

The results indicate that horses' arousal levels changed in response to the robot's 'happy', 'angry', 'surprise', and 'neutral' facial expressions when compared to baseline. However, the data did not provide evidence that horses could discriminate these expressions in terms of emotional valence. Therefore, visual cues emitted by the android robot, even under highly standardised conditions, appeared insufficient to demonstrate that horses perceived the emotional meaning of human-like facial expressions. In addition, heart rate variability (HRV) changed during exposure to the robot but returned to baseline thereafter, suggesting an adaptive autonomic response without signs of sustained stress (Koolhaas et al., 2011), indicating that horses did not perceive the robot as an aversive stimulus and that the experimental protocol did not compromise their welfare.

The ability of animals to perceive and respond to human emotions has long been a topic of scientific interest (Albuquerque et al., 2016; Benz-Schwarzburg et al., 2020; Nawroth et al., 2018). The well-known Clever Hans case, in which a horse responded to subtle and unconscious human cues (Pfungst, 1911), remains a paradigmatic example of the often-underestimated influence of human presence on animal behaviour, also in research contexts (Samhita and Gross, 103; Trestman, 2015). A growing evidence suggests that horses can perceive human emotions: they discriminate between happy and angry expressions in photographs (Smith et al., 2016) and videos (Jardat et al., 2023), and respond differently to people previously seen displaying such emotions, indicating that they recall and utilize affective information in later interactions (Proops et al., 2018).

Emotions are complex phenomena involving coordinated physiological, behavioural, and cognitive changes (Barrett and Bliss-Moreau,

Table 3
Median (interquartile range, IQR) of behavioural variables recorded across experimental conditions. B1 and B2 indicate baseline phases preceding and following exposure, respectively. Values are presented as median (IQR). Non-parametric analyses were applied due to non-normal data distribution.

	ATTdx_F	ATTsx_F	EXPdx_F	EXPsx_F	selATT_F	selEXP_F	Displ_F	Stand_F	Walk_F	Out_F
B1	6.75 (8)	6.00 (4.25)	3.00 (2.50)	1.00 (2.25)	5.50 (5.75)	5.50 (6.50)	5.12 (10.12)	8.50 (5.75)	8.75 (5.00)	9.25 (7.25)
Neutral	5.00 (4.50)	4.25 (9.75)	3.25 (4.25)	2.75 (2.75)	3.75 (6.25)	7.00 (3.50)	4.75 (5.50)	8.25 (3.50)	7.25 (3.50)	7.00 (6.75)
Happy	5.75 (10.00)	6.25 (6.75)	2.50 (5.25)	2.5 (4.00)	2.50 (1.75)	5.25 (4.00)	5.60 (6.35)	7.00 (4.75)	6.75 (4.50)	7.00 (7.50)
Surprise	4.25 (6.75)	8.25 (9.00)	3.25 (4.50)	1.25 (3.75)	2.00 (3.75)	5.50 (3.25)	4.10 (7.25)	4.75 (3.50)	5.00 (6.00)	5.75 (7.50)
Angry	6.50 (15.50)	4.50 (6.25)	2.25 (5.25)	1.00 (1.25)	2.50 (4.75)	3.50 (3.25)	2.25 (5.35)	5.50 (5.50)	4.75 (5.25)	5.50 (8.25)
B2	10 (8.75)	6.00 (9.00)	2.25 (3.00)	1.50 (3.25)	2.50 (4.75)	5.25 (3.25)	5.87 (6.87)	7.50 (3.75)	5.75 (3.50)	10.25 (10.50)

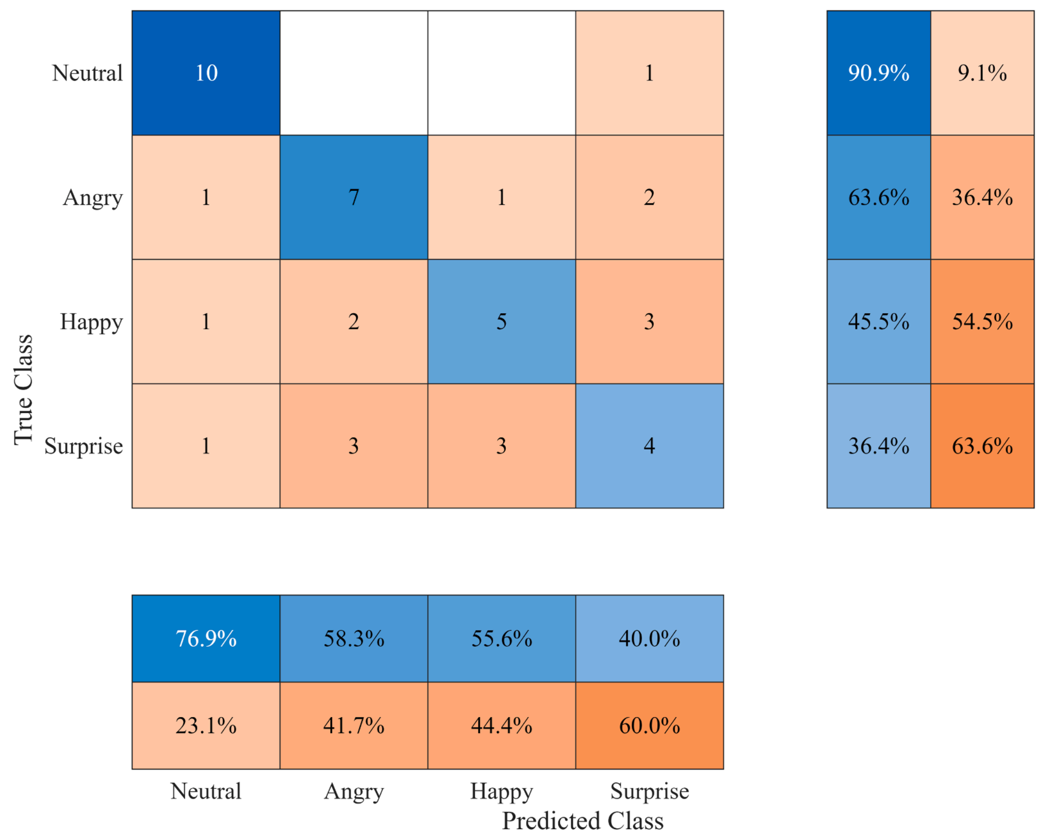

Fig. 3. Confusion matrix for the ML experiment ('neutral'-'angry'-'happy'-'surprise' N–A–H–S), overall accuracy 59%. TPR and FNR for each class are reported in the right columns; PPV and NPV values are shown below.

Table 4
Results obtained in the ML experiments excluding iteratively each FACE stimuli covering all possible triplets. Both the overall accuracy of the models obtained during LOSO validation and the TPR for each class considered are reported. TPR = True Positive Rate; H-S-A = 'happy'-'surprise'-'angry' class; N-H-A = 'neutral'-'happy'.

ML Experiment	Model	Accuracy (TPR for each class)	Number of HRV features	Number of BEHA features
H-S-A	AdaBoosting with ReliefF	77% (83%,75%,75%)	Not present in the best ML model	4
N-H-A	AdaBoosting with ReliefF	78% (90%,63%,81%)	32	8
N-S-A	AdaBoosting with ReliefF	69% (50%,75%,83%)	Not present in the best ML model	8
N-H-S	AdaBoosting with ReliefF	75% (90%, 45%,90%)	35	9

2009; LeDoux, 2021; Mendl et al., 2010). In this study, HRV analysis revealed that horses' Autonomic Nervous System activity was affected by both the robot's presence and its facial expressions.

No significant differences emerged between Baseline1 and the 'neutral' face; however, both the 'happy' and 'surprise' faces increased RR interval, resulting in a reduced heart rate. The 'surprise' expression also raised HF, indicating enhanced vagal tone, and reduced SD2RR, suggesting decreased cardiac complexity. In contrast, exposure to the 'angry' face elicited mixed responses, with fluctuating markers of both sympathetic and vagal activation, suggesting unstable autonomic regulation. Additional changes in long-term HRV parameters (SD12_LPP; (Nardelli et al., 2017) across all faces further indicate that the robot's faces influenced slower cardiac regulation, though their meaning in horses remains unclear.

HRV linear parameters showed mixed results when emotional faces were compared with baseline and the 'neutral' face. Nonlinear parameters revealed lower entropy for 'neutral' face, reflecting more monotonic cardiac activity, whereas 'happy' and 'angry' increased entropy, suggesting more complex and healthier autonomic regulation.

However, rises in SD12 components for both 'happy' and 'angry' might indicate an altered balance. The 'surprise' face increased RR interval (lower HR) and elevated SD12, suggesting partial adaptability of cardiac dynamics to external stimuli. These findings highlight the challenges of interpreting HRV, particularly in species such as horses that exhibit high vagal tone, where HR variations up to about 120 bpm mainly reflect vagal fluctuations (Clément and Barrey, 1995; Coote and White, 2015; Physick-Sheard et al., 2000). Overall, differences between neutral and emotional conditions were small and did not indicate dominance of either sympathetic or parasympathetic activity.

The only behaviour that differed was locomotor activity (walking), which tended to increase when the robot was positioned in front of the horses, regardless of its expression. Together with HRV data, this suggests general arousal, manifested as greater movement (Briefer et al., 2015), but provides no evidence of valence discrimination.

The Circumplex Model of Affect (CMA) provides a useful framework for interpreting these findings by conceptualising emotions along two primary dimensions: valence (positive to negative) and arousal (low to high) (Mendl et al., 2010; Posner et al., 2005). Discrete emotions are typically short-lived, and event-focused, whereas broader affective states encompass more enduring valenced sensations and motivational conditions (Barrett, 2006; Lindquist et al., 2012). While discrete emotions are often associated with relatively consistent physiological and behavioural patterns (Berridge, 2000; Caeiro et al., 2013; Dolensek et al., 2020), affective states are better described as continuous valence–arousal dimensions that can modulate perception and cognition even in the absence of a specific eliciting event (Barret, 2017a, 2017b; Lindquist et al., 2012). In animal research, this distinction is particularly relevant, as many experimental paradigms are more likely to capture affective states rather than discrete emotions. Accordingly, physiological responses such as HRV fluctuations should primarily be interpreted as indicators of arousal, given that both positive and negative affective states may be characterised by high arousal (Phelipon et al., 2024). Inferences about emotional valence therefore require the integration of physiological measures with behavioural and cognitive indicators (Kuppens et al., 2010).

To deepen these relationships, a Machine Learning (ML) approach was applied to classify physiological and behavioural responses to the robot's expressions. The model achieved a moderate accuracy of 59% (Fig. 3). The 'neutral' condition was identified with high reliability (True Positive Rate ≈ 90%), indicating that it produced the most distinctive pattern; whereas the emotional expressions overlapped considerably, with an average False Negative Rate of 51%, showing that the model struggled to distinguish between them

In three-class models (e.g., N-H-A, N-H-S), classification relied on both HRV and behavioural parameters. Accuracy remained high for 'neutral' (≈90%) and 'angry' or 'surprise' (81–90%), but lower for 'happy' (45–63%). Across all models, including 'neutral', the HRV parameter NN50 consistently contributed to accuracy, while 'happy' remained the most difficult expression to classify. In models excluding 'neutral' face, classification depended mainly on behavioural parameters, suggesting that physiological measures were less informative for distinguishing emotional responses. Horses appeared to employ a wide range of attention- and exploration-related behaviours, with no behaviour uniquely associated with any emotion. Rather than relying on a single discriminative response, horses used a combination of exploratory and attentional behaviours, suggesting a cognitive strategy rather than a purely emotional reaction.

The robot provided only visual facial cues, lacking vocal, olfactory, or movement-based information that horses are likely to use to interpret emotion. Previous studies have shown that horses integrate facial and vocal cues when processing human emotions (Nakamura et al., 2018; Proops and McComb, 2012; Trösch et al., 2020), emphasizing the importance of multimodal communication in perceiving emotional valence. The absence of such cues or possibly the low conspicuity of the visual one, may have limited the ability to discriminate between positive and negative expressions. This limitation also raises questions about the robustness of studies claiming that horses (and other animals) recognize human emotions solely from photographs or videos, where multimodal information is absent.

Identifying reliable indicators of positive emotion in horses remains challenging, as distinguishing between discrete emotions and broader affective states is inherently difficult (Humphreys et al., 2025). Positive affect involves dynamic interplay between the sympathetic and parasympathetic systems, making physiological biomarkers highly context-dependent (Pawluski et al., 2017; Phelipon et al., 2024; Ralph and Tilbrook, 2016). Moreover, individual differences in temperament and experience, along with social bonds shaped by domestication (Steklis et al., 2025), further complicate the interpretation of emotional valence in horse–human interactions (Mendl et al., 2010; Visser et al., 2002). Consequently, much of the existing literature probably reflects affective rather than discrete emotional states (Barrett and Bliss-Moreau, 2009; Mendl et al., 2009), whose expression is strongly influenced by context and individual variability (Humphreys et al., 2025).

EquiFACS analysis revealed no significant differences across the robot's facial expressions, consistent with ongoing debates on the limits of facial metrics alone. In particular, the EquiFACS analysis revealed a similar distribution of AUs across all the robot's conditions, indicating that the horses recruited the same facial AUs regardless of the facial expression perceived. This is reflected in the overlap among conditions (Fig. 2). The similarity in AUs distributions suggests that horses may employ a core set of facial actions when confronted with ambiguous or unfamiliar stimuli, regardless of the intended expression. The limited discriminative power of EquiFACS features may therefore reflect the constraints of a unimodal approach, reinforcing the need for integrated, multimodal analyses to assess emotion more reliably (Jardat et al., 2025; Nakamura et al., 2018; Proops and McComb, 2012; Trösch et al., 2020).

One additional aspect that may influence responses to artificial social stimuli is the highly human-like yet artificial nature of the robot's face. In human perception, such stimuli may sometimes evoke the so-called 'uncanny valley' effect. Although evaluating such phenomenon in non-human animals remains speculative, it cannot be excluded that horses interpret artificial human-like faces differently from natural human expressions. Future studies could therefore explore whether the inclusion of additional cues, such as motion dynamics or vocal signals, influences horses' interpretation of robotic emotional displays.

Despite limitations such as the small sample size and the inclusion of a novel 'surprise' expression, which may have acted as a distractor, this study represents the first attempt to combine HRV and behavioural data to analyse horses' responses to android robot facial expressions. This integrative approach suggests that HRV measures may capture stimulus-related variation more effectively than behaviour alone, offering a promising direction for future research.

5. Conclusion

The robot's emotional expressions elicited a state of arousal; however, no evidence emerged to support the horses' ability to discriminate the valence of these expressions. The observed behavioral responses were broad and variable, suggesting cognitive activation driven by a general motivation to interpret the stimulus rather than discrete emotion recognition. This pattern is more consistent with a general positive affective state. Overall, the neutral expression elicited more coherent and readily classifiable physiological and behavioral responses, whereas emotional expressions produced more variable and overlapping patterns. These findings indicate a largely similar modulation of autonomic nervous system activity across emotional expressions, associated with a shared repertoire of attention–exploration behaviors. Android robots provide highly standardized and controllable stimuli, allowing the exclusion of direct human interference and the non-conscious human-derived cues. As such, they represent a promising tool for investigating animal emotions, cognition, and welfare, supporting more ethical and evidence-based approaches to human–animal interaction.

CRediT authorship contribution statement

Martina Felici: Writing – original draft, Methodology, Investigation, Data curation. **Lorenzo Frassineti:** Writing – original draft, Software, Formal analysis, Data curation. **Paolo Baragli:** Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Antonio Lanatà:** Writing – review & editing, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Chiara Scopa:** Writing – original draft, Investigation, Data curation. **Enzo Pasquale Scilingo:** Writing – review & editing, Supervision, Software, Project administration, Methodology, Conceptualization. **Lorenzo Cominelli:** Validation, Software, Methodology, Investigation. **Elisabetta Palagi:** Writing – review & editing, Supervision, Project administration. **Micaela Sgorbini:** Writing – review & editing, Project administration, Investigation. **Alice Galotti:** Investigation, Formal analysis. **Claire Ricci-Bonot:** Investigation, Formal analysis.

Ethics

The Ethics Committee on Animal Experimentation of the University of Pisa approved the experimental protocol (reference No. 49/2020). All methods are in accordance with the proposed ARRIVE guidelines.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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.

Acknowledgments

The authors wish to thank the Department of Veterinary Sciences, University of Pisa, for their support in the management and care of the horses involved in this study.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.applanim.2026.106998.

References

Acharya, R.U., Paul Joseph, K., Kannathal, N., Lim, C.M., Suri, J.S., 2006. Heart rate variability: a review. MBEC Med. Biol. Eng. Comput. 44 (12), 1031–1051. https://doi.org/10.1007/s11517-006-0119-0.

Albuquerque, N., Guo, K., Wilkinson, A., Savalli, C., Otta, E., Mills, D., 2016. Dogs recognize dog and human emotions. Biol. Lett. 12 (1), 20150883. https://doi.org/10.1098/rsbl.2015.0883.

Baba, C., Kawai, M., Takimoto-Inose, A., 2019. Are horses (*Equus caballus*) sensitive to human emotional cues? Animals 9 (9), 630. https://doi.org/10.3390/ani9090630.

Baciadonna, L., Duepjan, S., Briefer, E.F., Padilla de la Torre, M., Nawroth, C., 2018. Looking on the bright side of livestock emotions—the potential of their transmission to promote positive welfare. Front. Vet. Sci. 5, 218. https://doi.org/10.3389/fvets.2018.00218.

Barrett, L.F., 2006. Are emotions natural kinds? Persp. Phsychol. Sci. 1 (1), 28–58. https://doi.org/10.1111/j.1745-6916.2006.0000.

Barrett, L.F., 2017a. How emotions are made: the secret life of the brain. Houghton Mifflin Harcourt, Boston, MA.

Barrett, L.F., 2017b. The theory of constructed emotion: an active inference account of interoception and categorization. Soc. Cogn. Aff. Neurosci. 12 (1), 1–23. https://doi.org/10.1093/scan/nsw154.

Barrett, L.F., Bliss-Moreau, E., 2009. She's emotional. He's having a bad day: Attributional explanations for emotion stereotypes. Emotion 9 (5), 649.

Bateson, M., Martin, P., 2021. Measuring Behaviour: an Introductory Guide, fourth ed. Cambridge University Press, Cambridge.

Benz-Schwarzburg, J., Monsó, S., Huber, L., 2020. How dogs perceive humans and how humans should treat their pet dogs: linking cognition with ethics. Front. Psychol. 11, 3587. https://doi.org/10.3389/fpsyg.2020.584037.

Berridge, K.C., 2000. Measuring hedonic impact in animals and infants: microstructure of affective taste reactivity patterns. Neurosci. Biobehav. Rev. 24, 173–198. https://doi.org/10.1016/S0149-7634(99)00072-X.

Briefer, E., Maigrot, A.L., Mandel, R., Briefer Freymond, S., Bachmann, I., Hilmann, E., 2015. Segregation of information about emotional arousal and valence in horse whinnies. Sci. Rep. 5, 9989. https://doi.org/10.1038/srep09989.

Caeiro, C.C., Waller, B.M., Zimmermann, E., Burrows, A.M., Davila-Ross, M., 2013. OrangFACS: A muscle-based facial movement coding system for orangutans (*Pongo spp.*). Int. J. Prima 34, 115–129. https://doi.org/10.1007/s10764-012-9652-x.

Clément, F., Barrey, E., 1995. Heart rate fluctuations in the horse at rest: (1) investigation of heart rate changes by spectrum analysis. C. R. Acad. Sci. III Sci. Vie 318, 859–865.

Cominelli, L., Mazzei, D., De Rossi, D.E., 2018. SEAI: Social emotional artificial intelligence based on Damasio's theory of mind. Front. Robot. Ai. 5, 6. https://doi.org/10.3389/frobt.2018.00006.

Coote, J.H., White, M.J., 2015. CrossTalk proposal: bradycardia in the trained athlete is attributable to high vagal tone. J. Physi 593 (8), 1745. https://doi.org/10.1113/jphysiol.2014.284364.

Cottin, F., Médigue, C., Lopes, P., Petit, E., Papelier, Y., Billat, V.L., 2005. Effect of exercise intensity and repetition on heart rate variability during training in elite trotting horse. Int. J. Sports Med. 26 (10), 859–867. https://doi.org/10.1055/s-2005-837462.

Cross, E.S., Hortensius, R., Wykowska, A., 2019. From social brains to social robots: applying neurocognitive insights to human-robot interaction. Philos. Trans. R. Soc. B Biol. Sci. 374 (1771), 20180024. https://doi.org/10.1098/rstb.2018.0024.

Dolensek, N., Gehrlach, D.A., Klein, A.S., Gogolla, N., 2020. Facial expressions of emotion states and their neuronal correlates in mice. Science 368 (6486), 89–94. https://doi.org/10.1126/science.aaz9468.

Ekman, P., Friesen, W.V., 1978. Facial Action Coding System. Consulting Psychologists Press, Palo Alto, CA. https://doi.org/10.1037/t27734-000.

Felici, M., Nardelli, M., Lanatà, A., Sgorbini, M., Scilingo, E.P., Baragli, P., 2021. Smart textiles biotechnology for electrocardiogram monitoring in horses during exercise on treadmill: Validation tests. Equine Vet. J. 53 (2), 373–378. https://doi.org/10.1111/evj.13296.

Galvan, M., Vonk, J., 2016. Man's other best friend: domestic cats (*F. silvestris catus*) and their discrimination of human emotion cues. Anim. Cogn. 19, 193–205. https://doi.org/10.1007/s10071-015-0927-4.

Guidi, A., Greco, A., Felici, F., Leo, A., Ricciardi, E., Bianchi, M., Bicchi, A., Valenza, G., Scilingo, E.P., 2017. Heart rate variability analysis during muscle fatigue due to prolonged isometric contraction. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 1324–1327. https://doi.org/10.1109/EMBC.2017.8037076.

Henschel, A., Hortensius, R., Cross, E.S., 2020. Social cognition in the age of human–robot interaction. Trends Neurosci. 43 (6), 373–384. https://doi.org/10.1016/j.tins.2020.03.013.

Herrando, C., Constantinides, E., 2021. Emotional contagion: a brief overview and future directions. Front. Psychol. 12, 712606. https://doi.org/10.3389/fpsyg.2021.712606.

Huber, A., Barber, A.L., Faragó, T., Müller, C.A., Huber, L., 2017. Investigating emotional contagion in dogs (*Canis familiaris*) to emotional sounds of humans and conspecifics. Anim. Cogn. 20 (4), 703–715. https://doi.org/10.1007/s10071-017-1092-8.

Humphreys, S., Freire, R., Waran, N., Randle, H., 2025. The challenge of identifying positive emotional indicators in equids: a scoping review to evaluate if this approach is putting the cart before the horse (*Equus caballus*). Appl. Anim. Behav. Sci., 106760 https://doi.org/10.1016/j.applanim.2025.106760.

Ishiguro, H., Nishio, S., 2007. Building artificial humans to understand humans. J. Artific. Organs 10, 133–142. https://doi.org/10.1007/978-981-10-8702-8_2.

Jardat, P., Lansade, L., 2022. Cognition and the human–animal relationship: a review of the sociocognitive skills of domestic mammals toward humans. Anim. Cogn. 25, 369–384. https://doi.org/10.1007/s10071-021-01557-6.

Jardat, P., Liehrmann, O., Reigner, F., Parias, C., Calandreau, L., Lansade, L., 2023. Horses discriminate between human facial and vocal expressions of sadness and joy. Anim. Cogn. 26, 1733–1742. https://doi.org/10.1007/s10071-023-01817-7.

Jardat, P., Menard, Z., Parias, C., Reigner, F., Calandreau, L., Lansade, L., 2024. Horses can learn to identify joy and sadness against other basic emotions from human facial expressions. Behav. Proc. 220, 105081. https://doi.org/10.1016/j.beproc.2024.105081.

Jardat, P., Yamamoto, S., Ringhofer, M., Tanguy-Guillo, N., Parias, C., Reigner, F., Calandreau, L., Lansade, L., 2025. Emotional contagion of fear and joy from humans to horses using a combination of facial and vocal cues. Sci. Rep. 15, 17689. https://doi.org/10.1038/s41598-025-98794-3.

Karemaker, J.M., 2022. The multibranched nerve: vagal function beyond heart rate variability. Biol. Psychol. 172, 108378. https://doi.org/10.1016/j.biopsycho.2022.108378.

Koolhaas, J.M., Bartolomucci, A., Buwalda, B., de Boer, S.F., Flügge, G., Korte, S.M., Meerlo, P., Murison, R., Olivier, B., Palanza, P., Richter-Levin, G., Sgoifo, A., Steimer, T., Stiedl, O., van Dijk, G., Wöhr, M., Fuchs, E., 2011. Stress revisited: a critical evaluation of the stress concept. Neurosci. Biobehav. Rev. 35 (5), 1291–1301. https://doi.org/10.1016/j.neubiorev.2011.02.003.

Kuppens, P., Oravecz, Z., Tuerlinckx, F., 2010. Feelings change: accounting for individual differences in the temporal dynamics of affect. J. Pers. Soc. Psychol. 99 (6), 1042. https://doi.org/10.1037/a0020962.

Kuwahara, M., Hashimoto, S.I., Ishii, K., Yagi, Y., Hada, T., Hiraga, A., Kai, M., Kubo, K., Oki, H., Tsubone, H., et al., 1996. Assessment of autonomic nervous function by power spectral analysis of heart rate variability in the horse. J. Auton. Nerv. Syst. 60, 43–48. https://doi.org/10.1016/0165-1838(96)00028-8.

Lanatà, A., Guidi, A., Baragli, P., Valenza, G., Scilingo, E.P., 2015. A novel algorithm for movement artifact removal in ecg signals acquired from wearable systems applied to horses. PloS One 10 (10), e0140783. https://doi.org/10.1371/journal.pone.0140783.

Landis, J.R., Koch, G.G., 1977. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 363–374. https://doi.org/10.2307/2529786.

Lansade, L., Colson, V., Parias, C., Reigner, F., Bertin, A., Calandreau, L., 2020. Human face recognition in horses: data in favor of a holistic process. Front. Psychol. 11, 575808. https://doi.org/10.3389/fpsyg.2020.575808.

LeDoux, J.E., 2021. What emotions might be like in other animals. Curr. Biol. 31 (13), R824–R829. https://doi.org/10.1016/j.cub.2021.05.005.

Li, J., 2015. The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Intern. J. Hum. Comput. Stud. 77, 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001.

Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F., 2012. The brain basis of emotion: a meta-analytic review. Behav. Brain Sci. 35 (3), 121–143. https://doi.org/10.1017/S0140525X11000446.

Lundblad, J., Rashid, M., Rhodin, M., Haubro Andersen, P., 2021. Effect of transportation and social isolation on facial expressions of healthy horses. PLoS ONE 16, e0241532. https://doi.org/10.1371/journal.pone.0241532.

Mazzei, D., Lazzeri, N., Billeci, L., Igliozzi, R., Mancini, A., Ahluwalia, A., Muratori, F., De Rossi, D., 2011. Development and evaluation of a social robot platform for therapy in autism. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 4515–4518. https://doi.org/10.1109/IEMBS.2011.6091119.

McCraty, R., Shaffer, F., 2015. Heart rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. GAHMJ 4 (1), 46–61. https://doi.org/10.7453/gahmj.2014.07.

McCraty, R., Zayas, M., 2014. Intuitive intelligence, self-regulation, and lifting consciousness. GAHMJ 3 (2), 56–65. https://doi.org/10.7453/gahmj.2014.013.

Mendl, M., Burman, O.H., Parker, R.M., 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.

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

Merkies, K., Sudarenko, Y., Hodder, A.J., 2022. Can ponies (*Equus caballus*) distinguish human facial expressions? Animals 12 (18), 2331. https://doi.org/10.3390/ani12182331.

Nakamura, K., Takimoto-Inose, A., Hasegawa, T., 2018. Cross-modal perception of human emotion in domestic horses (*Equus caballus*). Sci. Rep. 8 (1), 8660. https://doi.org/10.1038/s41598-018-26892-6.

Nardelli, M., Greco, A., Bolea, J., Valenza, G., Scilingo, E.P., Bailón, R., 2017. Investigation of lagged Poincaré plot reliability in ultra-short synthetic and experimental heart rate variability series. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 2329–2332. https://doi.org/10.1109/EMBC.2017.8037322.

Nawroth, C., Albuquerque, N., Savalli, C., Single, M.S., McElligott, A.G., 2018. Goats prefer positive human emotional facial expressions. R. Soc. Open Sci. 5 (8), 180491. https://doi.org/10.1098/rsos.180491.

Panksepp, J., 1986. The anatomy of emotions. In: Biological foundations of emotion. Academic Press, San Diego, CA, pp. 91–124. https://doi.org/10.1016/B978-0-12-558703-7.50010-3.

Pawluski, J., Jego, P., Henry, S., Bruchet, A., Palme, R., Coste, C., Hausberger, M., 2017. Low plasma cortisol and fecal cortisol metabolite measures as indicators of compromised welfare in domestic horses (*Equus caballus*). Plos One 12 (9), e0182257. https://doi.org/10.1371/journal.pone.0182257.

Pérez-Manrique, A., Gomila, A., 2022. Emotional contagion in nonhuman animals: a review. Wiley Interdiscip. Rev. Cogn. Sci. 13, e1560. https://doi.org/10.1002/wcs.1560.

Pfungst, O., 1911. Clever Hans (the horse of Mr. von Osten): a contribution to experimental animal and human psychology. Holt, Rinehart and Winston, New York, NY.

Phelipon, R., Ascione, A., Ruet, A., Bertin, A., Biau, S., Arnould, C., Boichot, L., Lansade, L., 2024. Differences in behaviour, facial expressions and locomotion between positive anticipation and frustration in horses. Appl. Anim. Behav. Sci. 274, 106262. https://doi.org/10.1016/j.applanim.2024.106262.

Physick-Sheard, P.W., Marlin, D.J., Thornhill, R., Schroter, R.C., 2000. Frequency domain analysis of heart rate variability in horses at rest and during exercise. Equ. Vet. J. 32 (3), 253–262. https://doi.org/10.2746/042516400776563572.

Pioggia, G., Ahluwalia, A., Carpi, F., Marchetti, A., Ferro, M., Rocchia, W., Rossi, D.D., 2004. FACE: facial automaton for conveying emotions. Appl. Bionics Biomech. 1 (2), 91–100. https://doi.org/10.1533/abib.2004.1.2.91.

Posner, J., Russell, J.A., Peterson, B.S., 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev. Psychopath. 17 (3), 715–734. https://doi.org/10.1017/S0954579405050340.

Proops, L., McComb, K., 2012. Cross-modal individual recognition in domestic horses (*Equus caballus*) extends to familiar humans. Proc. Biol. Sci. 279 (1741), 3131–3138. https://doi.org/10.1098/rspb.2012.0626.

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. https://doi.org/10.1016/j.cub.2018.03.035.

Ralph, C.R., Tilbrook, A.J., 2016. The usefulness of measuring glucocorticoids for assessing animal welfare. J. Anim. Sci. 94 (2), 457–470. https://doi.org/10.2527/jas.2015-9645.

Rashid, M., Silventoinen, A., Gleerup, K.B., Andersen, P.H., 2020. Equine Facial Action Coding System for determination of pain-related facial responses in videos of horses. Plos One 15 (11), e0231608. https://doi.org/10.1371/journal.pone.0231608.

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.

Samhita, L., Gross, H.J., 2013. The "clever Hans phenomenon" revisited. Comm. Integr. Biol. 6 (6), e27122. https://doi.org/10.4161/cib.27122.

Sato, W., Namba, S., Yang, D., Nishida, S.Y., Ishi, C., Minato, T., 2022. An android for emotional interaction: Spatiotemporal validation of its facial expressions. Front. Psychol. 12, 800657. https://doi.org/10.3389/fpsyg.2021.800657.

Scopa, C., Contalbrigo, L., Greco, A., Lanatà, A., Scilingo, E.P., Baragli, P., 2019. Emotional transfer in human–horse interaction: new perspectives on equine assisted interventions. Animals 9 (12), 1030. https://doi.org/10.3390/ani9121030.

Shi, H., Yang, L., Zhao, L., Su, Z., Mao, X., Zhang, L., Liu, C., 2017. Differences of heart rate variability between happiness and sadness emotion states: a pilot study. J. Med. Biol. Eng. 37, 527–539. https://doi.org/10.1007/s40846-017-0238-0.

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.

Steklis, N.G., Peñaherrera-Aguirre, M., Steklis, H.D., 2025. Nay to Prey: challenging the view of horses as a "Prey" species. Animals 15 (5), 641. https://doi.org/10.3390/ani15050641.

Stucke, D., Ruse, M.G., Lebelt, D., 2015. Measuring heart rate variability in horses to investigate the autonomic nervous system activity–Pros and cons of different methods. Appl. Anim. Behav. Sci. 166, 1–10. https://doi.org/10.1016/j.applanim.2015.02.007.

Task Force, 1996. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Circulation 93, 1043–1065.

Tate, A.J., Fischer, H., Leigh, A.E., Kendrick, K.M., 2006. Behavioural and neurophysiological evidence for face identity and face emotion processing in animals. Philos. Trans. R. Soc. B Biol. Sci. 361 (1476), 2155–2172. https://doi.org/10.1098/rstb.2006.1937.

Tomberg, C., Petagna, M., de Selliers de Moranville, L.A., 2023. Horses (*Equus caballus*) facial micro-expressions: insight into discreet social information. Sci. Rep. 13, 8625. https://doi.org/10.1038/s41598-023-35807-z.

Trestman, M., 2015. Clever Hans, Alex the Parrot, and Kanzi: what can exceptional animal learning teach us about human cognitive evolution? Biol. Theory 10, 86–99. https://doi.org/10.1007/s13752-014-0199-2.

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.

Venkatraman, A., Edlow, B.L., Immordino-Yang, M.H., 2017. The brainstem in emotion: a review. Front. Neuroanat. 11, 15. https://doi.org/10.3389/fnana.2017.00015.

Visser, E.K., Van Reenen, C.G., Van der Werf, J.T.N., Schilder, M.B.H., Knaap, J.H., Barneveld, A., Bolkhuis, H.J., 2002. Heart rate and heart rate variability during a novel object test and a handling test in young horses. Physiol. Behav. 76, 289–296. https://doi.org/10.1016/S0031-9384(02)00698-4.

Von Borell, E., Dobson, H., Prunier, A., 2007. Stress, behaviour and reproductive performance in female cattle and pigs. Horm. Behav. 52 (1), 130–138. https://doi.org/10.1016/j.yhbeh.2007.03.014.

Waller, B.M., Julle-Daniere, E., Micheletta, J., 2020. Measuring the evolution of facial 'expression' using multi-species FACS. Neurosci. Biobehav. Rev. 113, 1–11. https://doi.org/10.1016/j.neubiorev.2020.02.031.