Horses form cross-modal representations of adults and children

Plotine Jardat, Monamie Ringhofer, Shinya Yamamoto, Chloé Gouyet, Rachel Degrande, Céline Parias, Fabrice Reigner, Ludovic Calandreau, Léa Lansade

Accepted Manuscript

Horses form cross-modal representations of adults and children

Plotine Jardat[1,*], Monamie Ringhofer[2], Shinya Yamamoto[3,4], Chloé Gouyet[1], Rachel Degrande[1], Céline Parias[1], Fabrice Reigner[5], Ludovic Calandreau[1], Léa Lansade[1,*]

[1] CNRS, IFCE, INRAE, Université de Tours, PRC, F-37380, Nouzilly, France

[2] Department of Animal Science, Teikyo University of Science

[3] Institute for Advanced Study, Kyoto University, Kyoto, Japan

[4] Wildlife Research Center, Kyoto University, Kyoto, Japan

[5] UEPAO, INRAE F-37380, Nouzilly, France

* corresponding authors – lea.lansade@inrae.fr ; plotine.jardat@gmail.com

ORCID IDs: P Jardat 0000-0003-0374-5588 ; M Ringhofer 0000-0002-4813-939X ; S Yamamoto 0000-0001-7938-1106 ; C Gouyet 0000-0001-5548-053X ; R Degrande 0000-0003-4976-1073 ; L Calandreau 0000-0002-7535-5733 ; L Lansade 0000-0002-4185-9714

Abstract

Recently, research on domestic mammals' sociocognitive skills toward humans has been prolific, allowing us to better understand the human-animal relationship. For example, horses have been shown to distinguish human beings on the basis of photographs and voices and to have cross-modal mental representations of individual humans and human emotions. This leads to questions such as the extent to which horses can differentiate human attributes such as age. Here, we tested whether horses discriminate human adults from children. In a cross-modal paradigm, we presented 31 female horses with two simultaneous muted videos of a child and an adult saying the same neutral sentence, accompanied by the sound of an adult's or child's voice speaking the sentence. The horses looked significantly longer at the videos that were incongruent with the heard voice than at the congruent videos. We conclude that horses can match adults' and children's faces and voices cross-modally. Moreover, their heart rates increased during children's vocalizations but not during adults'. This suggests that in addition to having mental representations of adults and children, horses have a stronger emotional response to children's voices than adults' voices.

Keywords: human-animal relationship, emotion, cross-modal recognition, social cognition, *Equus caballus*

Statements and declarations

The authors declare no competing interests. This study was funded by Institut Français du Cheval et de l'Equitation (IFCE), grant number 32 000809-Cognition Equine.

<u>Author contributions</u>: All authors devised the protocol. P.J., C.G., C.P., F.R., L.L. implemented the protocol. P.J., C.G., L.L. coded the videos and analyzed the data from heart rate monitoring and behavior coding. P.J., M.R, S.Y., C.G., R.D., L.C. and L.L. revised the analysis and report.

Acknowledgments

We thank Adèle, Amande, Coline and the other volunteers who were filmed to make up the stimuli of this experiment; Estel Blasi for drawing Figure 1; and Springer Nature Authors Services for English language and text flow editing.

Introduction

Recently, research on domestic mammals' sociocognitive skills toward humans has been prolific, allowing us to better understand the human-animal relationship (Jardat and Lansade 2021). Horses, for example, can distinguish human beings on the basis of photographs (Lansade et al. 2020a, b) as well as voices (Smith et al. 2018). They can also match the body and voice of an individual human (Proops and Mccomb 2012; Lampe and Andre 2012), and they are sensitive to the type of speech used to address them (Lansade et al. 2021). Furthermore, horses have been observed to behave differently toward a human depending on what they know about the human, such as whether the human knows where food is hidden (Ringhofer and Yamamoto 2017; Trösch et al. 2019b; Ringhofer et al. 2021), what emotions this human has expressed (Proops et al. 2018), or how this human has behaved toward another horse (Trösch et al. 2020b). These findings suggest questions such as whether horses behave differently toward humans according to their age, and a first question that can be raised is whether they can discriminate human adults from children.

Few studies have explored this question in animals. Dolphins seem to swim closer to children than to adults in swim-with-dolphin programs (Brensing and Linke 2003). In dogs, behavioral synchronization and attachment styles seem to differ between the relationship of a dog and adult owner and that of a dog and child owner (Wanser et al. 2020, 2021). It is also possible that horses react differently to children compared with adults, as has commonly been reported by horsemen and horsewomen. For example, in a survey of French-speaking horse owners posted on social media, 76% of the respondents declared that their horses behaved differently toward children compared with adults (see Online Resource 1).

In humans, age is an essential attribute used to categorize conspecifics and can be assessed by both facial and vocal cues (Moyse 2014). Indeed, there are facial and vocal differences between people of different age groups. Facial characteristics (e.g., nose width to mouth length and face width proportions) evolve throughout the life span (Farkas et al. 1985). Children's and adults' voices differ in various acoustic parameters, such as fundamental frequency, speech rate and formants (Glaze et al. 1988; Sussman and Sapienza 1994; Lee et al. 1999; Stathopoulos et al. 2011; Dilley et al. 2013). We hypothesized that horses could differentiate human adults from children too, and that they could use these differences in facial and vocal features to do so as they are known to use these types of cues to recognize individual humans or human emotions (Proops and Mccomb 2012; Lampe and Andre 2012; Trösch et al. 2019a).

We investigated the capacity of horses to differentiate human adults and children based on facial and vocal cues using a cross-modal paradigm. Cross-modal paradigms aim to determine whether animals have mental representations of subjects or objects that combine different modalities, which implies discrimination between these subjects or objects. These paradigms consist in presenting animals with stimuli from different modalities, either two pictures accompanied by a sound that corresponds to one of the pictures (preference looking paradigms) or different combinations of a picture and a sound presented together that may or may not correspond to each other (expectancy violation paradigms). For example, horses, dogs, and cats were presented with vocal and visual stimuli from individual humans (e.g., they could see two people while one of these people's voices was broadcast). They looked for a different amount of time at the person whose voice they were hearing compared to the other person. This showed that they had mental representations of these individual humans that combined their vocal and visual characteristics (Adachi et al. 2007; Proops and Mccomb 2012; Takagi et al. 2019). Similarly, horses, dogs and cats reacted differently to images of humans expressing anger and joy, according to their congruence with (i.e., whether they matched) the vocalization that was played (Albuquerque et al. 2016; Nakamura et al. 2018; Trösch et al. 2019a; Quaranta et al. 2020). For example, horses looked longer at the images that were incongruent with the sound, suggesting that they were surprised by the contradictory information (Trösch et al. 2019a). This showed that these animals have multimodal mental representations of human emotions. In the present study, we implemented this type of procedure to examine whether horses form multimodal mental representations of human adults and children by showing horses images of adults' and children's faces while playing adults' or children's voices.

We measured the horses' heart rates as a potential indicator of change in arousal or emotional valence felt by the horses (Jardat et al. 2022). Indeed, a rise in heart rate has been observed in horses in response to human voices expressing anger (Trösch et al. 2019a) or using pet-directed speech (Jardat et al. 2022), in fearful or stressful situations (Lansade et al. 2008; Munsters et al. 2012), and in a positively arousing situation (Briefer et al. 2015a). Similar reactions have been observed in other mammals (Forkman et al. 2007), with for example dogs' heart rates fluctuating in response to different human facial expressions (Siniscalchi et al. 2018), and goats' heart rates increasing in a positively arousing situation (Briefer et al. 2015b).

Horses were simultaneously shown images of two human faces (one of an adult and one of a child) while the voice of an adult or a child was played over a speaker. We hypothesized that the horses would look at the videos for different amounts of time according to their congruency with the voice heard and would possibly spend longer looking at the incongruent video, as reported for a similar protocol in horses (Trösch et al. 2019a). Moreover, we hypothesized that the horses would react differently to child than to adult stimuli (as suggested in other species - Brensing and Linke 2003; Wanser et al. 2020, 2021) and that consequently, their heart rates would vary differently in response to children's voices than in response to adults' voices.

Materials and Methods

Ethics statement

Our experiment was approved by the Val de Loire Ethical Committee (CEEA VdL, Nouzilly, France, authorization number CE19-2022-1503-1). Animal care and experimental treatments complied with the French and European guidelines for housing and care of animals used for scientific purposes (European Union Directive 2010/63/EU) and were performed under the authorization and supervision of official veterinary services (agreement number F371752 delivered to the UEPAO animal facility by the veterinary service of the Département d'Indre et Loire, France). The animals lived in groups, were not food deprived during the experiment and did not undergo any invasive procedures. All methods were carried out in accordance with the relevant guidelines and regulations for direct human involvement in the study. All the people participating in the study provided informed consent.

Subjects

The study involved 31 Welsh mares (*Equus caballus*) aged 8.9 ± 2.4 years (mean ± sd) reared at the Animal Physiology Experimental Unit PAO (UEPAO, 37380 Nouzilly, France, DOI: 10.15454/1.55738963217 28955E12), INRAE. They lived in groups in indoor stalls on straw and had free access to an outdoor paddock. Hay and water were available ad libitum. These horses are used only for research purposes and are handled daily by human adults. They interact with children on rare occasions (approximately twice a month for less than an hour).

Experimental setup

The experimental setup was based on that used to investigate cross-modal recognition of human emotions in horses (Trösch et al. 2019a). The experiment took place in a large familiar stall (3.5 m large, 4.5 m long, Fig. 1). Before the beginning of the experiment, the experimenters made sure that no external sound (horse neighs, human voices or machine noises) was audible from the inside of the stall. The horses were placed in the middle of the stall and attached by two loose ropes. The videos

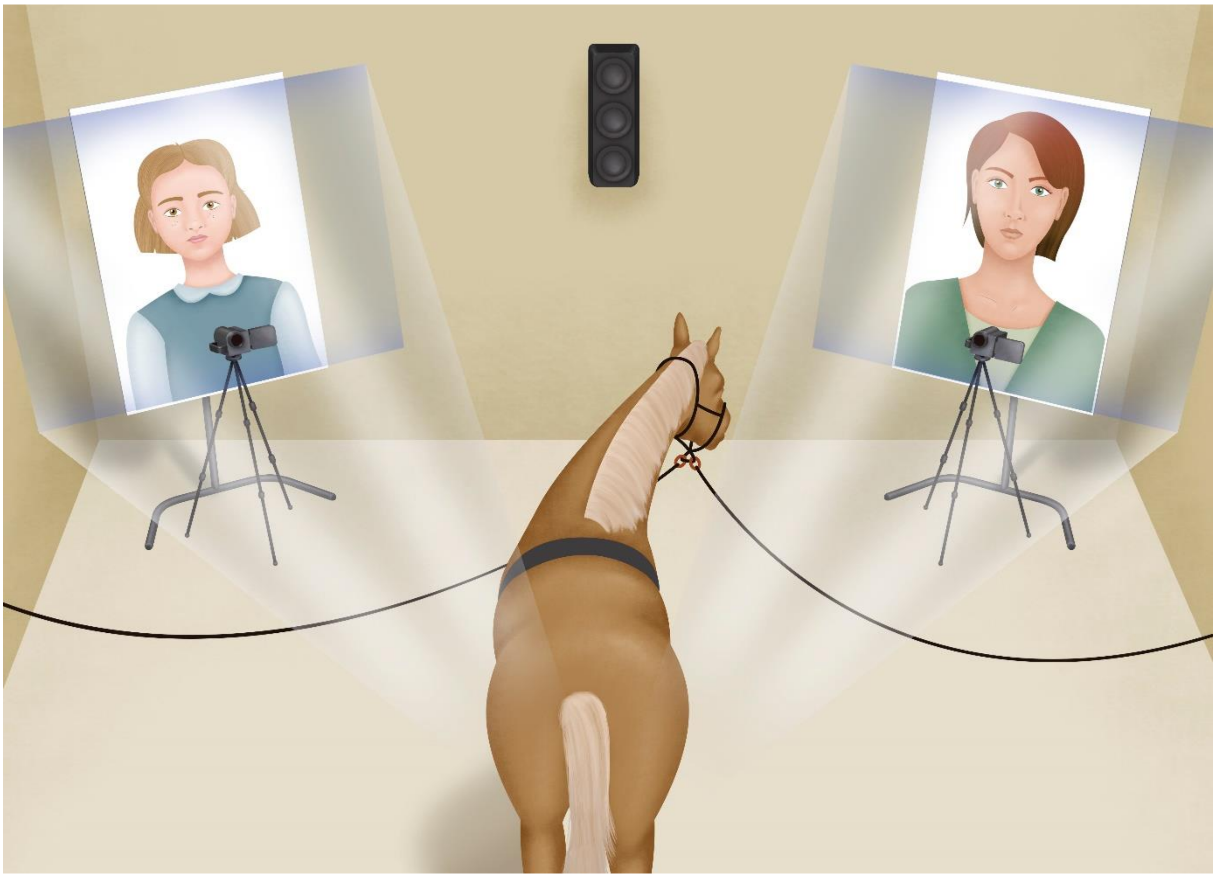

Fig. 1 Schematic representation of the experimental setup. Two simultaneous muted videos of a child and an adult saying the same neutral sentence were projected on the screens, while the sound of an adult's or child's voice speaking this sentence was played over a loudspeaker.

were projected on two projection screens (1x2 m) placed vertically on both sides of the horse (Fig. 1). For safety reasons, an assistant stayed with the horse to ensure that it did not panic or become entangled in the ropes. They stood along the wall at the level of the horse; the side of the horse on which they stood was counterbalanced among the horses. The assistant never interacted with the horse during the tests but remained still with their head down, looking neither at the screen nor at the horse. The assistant was instructed to stop the test if the horse panicked, but that never happened. The experiment was filmed by three cameras (Sony HDR CX450), one in front of each screen (Fig. 1) and one at the back, above the horse. An overview camera (GoPro Hero Black) allowed the experimenter to follow the running of the experiment from outside the stall and control the projection accordingly. The horses were equipped with a heart monitor system composed of a captor placed on the horse and a watch showing and recording real-time heart rate values (Polar Equine RS800CX Science, Polar Oy, Finland).

Stimuli

During the test sessions, a muted video of a child and a muted video of an adult were played simultaneously, one on each projection screen. At the same time, the voice of another person (child or adult) was played over a loudspeaker placed between the projection screens (Fig. 1). The voice played was not the original voice of either of the two projected videos to prevent the horses from matching the heard voice to the mouth movements of any of the videos. The faces were projected at a height of approximately 160 cm and were approximately twice the size of a real human's face (as in Trösch et al. 2019a). The sound was broadcast from a loudspeaker placed between the screens and in front of the horse (approximately 60 dB from where the horse's head was positioned).

Since horses may differentiate male and female faces and voices of humans, we used only female stimuli in this experiment. Four young girls aged 6 to 9 years and four women aged 46 to 58 years who the horses had never seen or heard were filmed. They were recorded in similar conditions, against a white background and with the face centered and the shoulders visible. They said a predefined neutral sentence that was identical for all participants ("J'aime la glace à la pistache; j'aime la glace au chocolat", meaning "I like pistachio ice cream; I like chocolate ice cream"), their facial expression was neutral. The acoustic characteristics of the sounds were obtained using Praat V6.1.16 (http://www.fon.hum.uva.nl/praat/); with a mean pitch of 269±36 Hz for children's voices and 227±15 Hz for adults' voices, and a pitch range of 392±18 Hz for children's voices and 416±1 Hz for adults' voices.

The videos shown during the tests lasted 76 seconds and were divided into four sections, two with a child's voice and two with an adult's voice (Fig. 2). Each section was composed of the sentence being repeated 4 times in a row and lasted 16 seconds. Four different video stimuli were made up, each with the faces of the four women and the four young girls, two voices of each type, twice the adult face on the right, and twice the congruent face on the right (Fig 2). The horses were randomly assigned one of the stimuli, so that the order in which the people were seen and the side on which

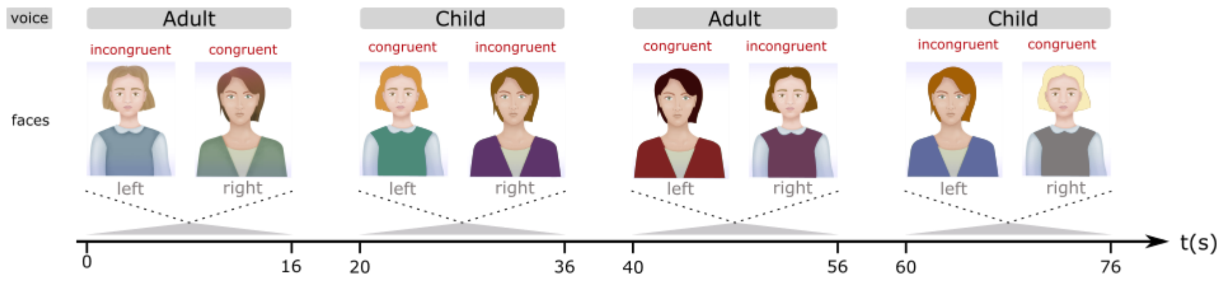

Fig. 2 Example of the composition of the stimuli shown to a horse. Eight different faces were shown, and four different voices were heard.

the congruent face was presented with each voice were counterbalanced among the horses. The videos were validated by 20 persons with 100% accuracy in categorizing the age group.

Habituation

In the habituation phase, the horse was led to the middle of the stall facing the screens. It was then attached with two loose ropes, the assistant took their place, and the habituation phase began. Identical scenes of nature accompanied by birdsongs were projected on both screens, while the assistant watched the horse's heart rate on the Polar watch. In this phase, the assistant could reposition the horse if it had its back to the screen. Once the horse was calm (not neighing, pulling on the ropes with her head, or trying to turn around or leave) and her heart rate had stayed below 100 bpm for two consecutive minutes, the test phase began immediately. If this criterion was not met after five minutes, the session ended, and a new one was scheduled for the next day. All the horses but three met the criterion on day 1 and therefore continued with the test session on the same day. The three other horses needed a second habituation session; they met the habituation criterion in this second session and continued with the test session on the same day.

Tests

Immediately after the horse met the habituation criterion, two videos of adults' and children's faces accompanied by adults' and children's voices (see **Stimuli** section above) were projected. The conditions were the same as those during habituation, but the assistant never intervened. The assistant was unaware of the side on which each video appeared (adult or child, congruent or incongruent). At the end of the test, the horse was led directly back to its stable.

Behavioral and physiological analyses

The videos of the tests from the front cameras were analyzed by the same coder without sound and with the screens not visible in the videos so that the coder did not know which type of voice was being broadcast (adult or child) or the side on which the congruent video appeared. All occurrences of defecations, head shakes, scratching the ground and rearing were noted. The total time spent looking at each video (on the right or left of the horse) was quantified. The horse was considered to be looking at the right screen when her left eye was visible from the camera placed under the right screen, and vice versa. Then, according to the type of video (congruent or incongruent) that was playing at each moment of the test, the total time spent looking at each type of video was quantified. A second observer gauged the time spent looking at each screen in 20% of the videos. The interobserver reliability was evaluated by an interclass correlation coefficient (ICC). The lower bound was 0.95, and the estimate was 0.97; thus, interobserver reliability was considered excellent (Koo and Li 2016).

The total time spent looking at the stimuli during a section was highly variable among horses (from 1 to 16 seconds over the 16-second sections). To take this variability into account, we calculated, for each horse, a preference index measuring the propensity to look more at the incongruent video in each section relative to the total time spent looking at the videos. The index was defined as (INC – CON)/(INC + CON) with INC being the time spent looking at the incongruent video and CON the time spent looking at the congruent video. Then, for each horse, the mean of this index over the four sections was calculated. The index varied from -1 to 1, with positive values indicating that a horse looked more at the incongruent videos and negative values the opposite.

Some horses tried to keep the exit door, situated behind them, in their field of vision throughout the duration of the test (probably because they were not comfortable being alone in the stall). Consequently, they almost never looked at either the right or the left screen, regardless of the congruence of the video and the voice heard. For this reason, these horses were excluded: fifteen horses which looked at one side for less than two seconds on average per trial were not taken into account in the analysis (a duration of 2 s had previously been used in the literature – Gácsi et al. 2004; Pattison et al. 2010; Somppi et al. 2012; Trösch et al. 2020a).

Heart rate data were extracted from the Polar recordings. A visual correction was applied to eliminate artifactual beats (as recommended by von Borell et al., 2007). The difference in heart rate (beats per minute - bpm) between the last and first three seconds of each section was calculated. Then, for each horse, the mean of this difference over the two sections of each voice type (adult or child) was calculated. Three horses neighed during the tests, with their abdominal and thoracic movements interfering with the heart rate signal, so these individuals were excluded from this analysis. Data were missing for two other individuals due to technical issues with the heart rate monitor. Therefore, the heart rate data analysis concerned 11 individuals.

Statistical analyses

All statistical analyses were performed using R 4.1.1 (R Core Team, 2013). Due to the sample size, nonparametric tests were used. The significance threshold was fixed at $\alpha \leq 0.05$.

To test whether the preference index was different from 0 at the group level, which would indicate that the horses looked more at one type of video (congruent or incongruent) than at the other, we used a two-tailed Wilcoxon test (n=16, *wilcox.test* function). To test whether the mean differences in heart rate over the child or adult voice sections were different from 0, we used two-tailed Wilcoxon tests (n=11, *wilcox.test* function). To check for an effect of the age of the horses, we tested whether the preference index was correlated to the age of the horses, using a Spearman correlation test (n=16, *spearman_test* function).

Defecations, head shakes, scratching the ground, rearing and vocalizations were expressed by too few horses to allow a statistical analysis (defecations, n = 2; head shakes, n = 2; scratching the ground, n = 0; rearing, n = 1; vocalizations, n = 3).

Results

The mean preference index was significantly different from 0 (Wilcoxon test, V=107, p=0.04, Fig. 3), meaning that the horses looked more at the incongruent video than at the congruent video relative to the total time spent looking at the videos. The mean heart rate variation during the adult vocalization sections was not different from 0 (Wilcoxon test, V=57, p=0.52, Fig. 4), and it was different from 0 during the child vocalization sections (Wilcoxon test, V=79, p=0.05, Fig. 4), indicating that the horses' heart rates increased during the children's vocalizations but not during the adults' vocalizations. There was no correlation between the preference index and the age of the horses (Spearman test, Z=0.18, p=0.86).

Discussion

Our results show that at the group level, the horses which looked at both screens during the experiment differentiated the videos of children and adults based on their congruence with the voice they were hearing. Moreover, their heart rates increased when they heard children's voices but not when they heard adults' voices.

In line with our hypothesis, the horses looked more at the video that was incongruent with the vocalization played. In other words, they differentiated between human adults and children on the videos based on their congruence or incongruence with the voice they heard at the same time, which shows that horses can differentiate human adults and children. This finding is consistent with previous studies on dolphins and dogs (Brensing and Linke 2003; Wanser et al. 2020, 2021). Moreover, these results show that horses can match a face and voice when they come from the same age group (adults or children), which suggests that they bear mental representations of adults

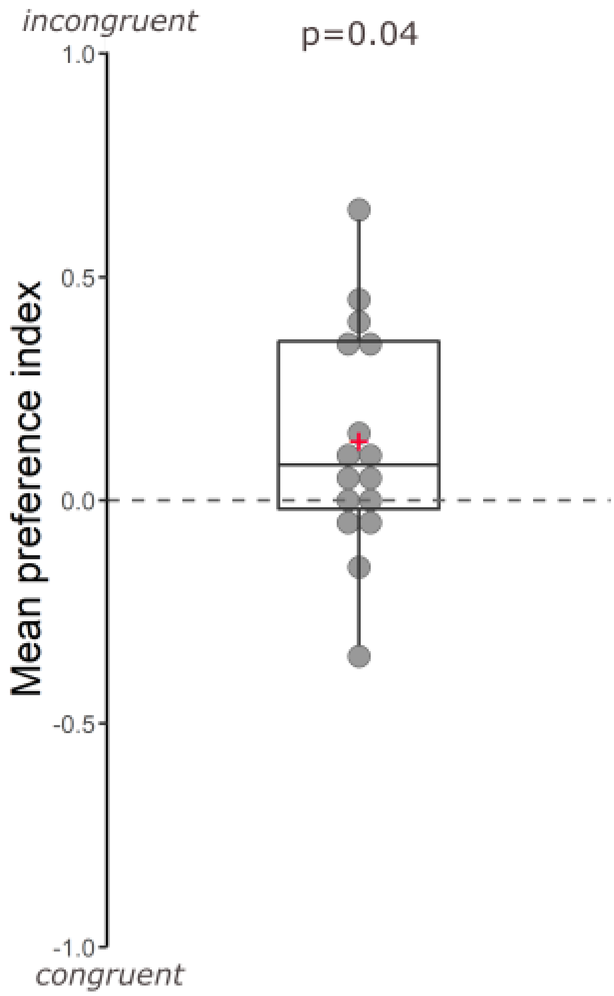

Fig. 3 Mean preference index. (INC – CON) / (INC + CON), with INC being the time spent looking at the incongruent video and CON the time spent looking at the congruent video. Boxplot: median, 1st and 3rd quartiles. Gray dots: individual indexes. Red cross: mean group index. p-value: comparison to 0, Wilcoxon test.

and children in which vocal and facial features are associated (Albuquerque et al. 2016; Quaranta et al. 2020). The horses were not trained prior to the study, and the videos and vocalizations were from people they had never seen before; therefore, the horses could not have recognized particular faces or voices presented in the experiment. Moreover, as the voice that was broadcast was that of a different person from the ones in the videos, there was no correlation between the timing of the facial movements observed and the voice heard. Eight different people whom the horses had never seen were used as stimuli, suggesting that this ability is not restricted to a specific person (a bias discussed in Trösch et al. 2019a). However, as only women and young girls were used, further study

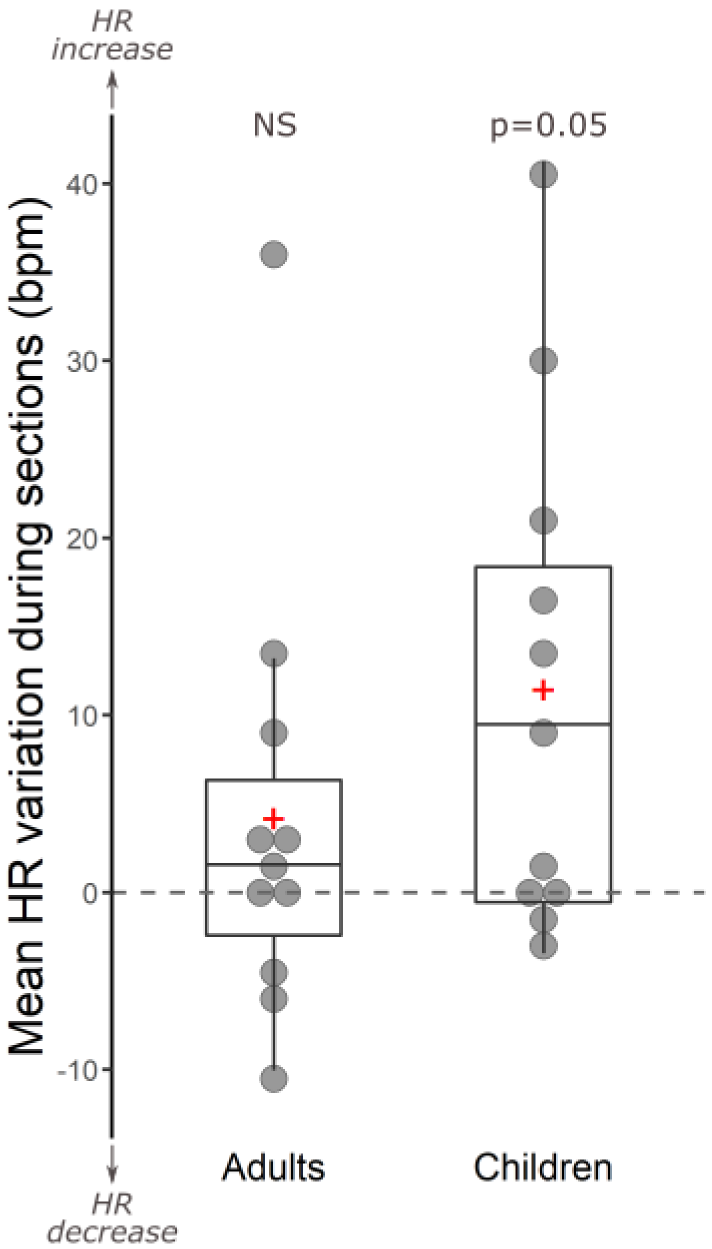

Fig. 4 Mean heart rate variation during adults' and children's vocalizations. Boxplot: median, 1st and 3rd quartiles. Gray dots: individual means. Red cross: mean group variation. p-value: comparison to 0, Wilcoxon test.

would be necessary to test whether these results can be generalized to men and young boys. Similarly, only mares were tested in this experiment, and it would be interesting to test whether the results can be generalized to stallions and geldings.

In this study, the horses looked longer at the videos that did not match the vocalization they heard. This is in line with a previous experiment using the preference looking paradigm in horses, in which they looked more at incongruent than congruent stimuli when presented with emotional faces and voices of humans (Trösch et al. 2019a). Similarly, in expectancy violation paradigms presenting human voices and faces horses, cats and dogs looked more at images which were incongruent with the sound (Adachi et al. 2007; Lampe and Andre 2012; Nakamura et al. 2018; Takagi et al. 2019). These longer looks toward incongruent stimuli can be explained by the animals being intrigued by contradictory information (Trösch et al. 2019a). However, in other experiments on domestic mammals' cross-modal recognition of human individuals and human emotions using the preference looking paradigm, the animals looked more at the congruent stimuli (Proops and Mccomb 2012; Albuquerque et al. 2016; Quaranta et al. 2020). These differences may be explained by variations in experimental conditions, for example, differences in the emotions or level of stress provoked by the presented stimuli.

In this study, the horses' heart rates increased when they heard children's voices but not when they heard adults' voices. This difference in physiological reaction to the two types of voices reinforces the idea that horses differentiate children's voices from adults' voices and shows that the children's voices induced a stronger emotional reaction. This is in line with previous findings suggesting that some species react differently to children than to adults: dolphins and dogs seem to interact differently with these two age groups (Brensing and Linke 2003; Wanser et al. 2020, 2021). An increase in heart rate can indicate either increased arousal or a change in the valence of emotions felt by horses (Jardat et al. 2022), and several hypotheses can explain such reactions to children's voices. First, the horses may have reacted to the novelty of this type of voice. Indeed, the horses in this experiment were less familiar with children's voices, as they seldom interact with children but are handled daily by adults. Thus, they could have been surprised by this type of voice, explaining a change in arousal level (Russell and Barrett 1999). Second, the horses may have reacted to the intrinsic characteristics of children's voices. Indeed, children's voices are notably characterized by differences in pitch and other prosodic features, such as speech rate (Dilley et al. 2013). These characteristics could elicit arousal and consequently a rise in heart rate, similar to horses' reaction to pet-directed speech, a type of speech that has specific acoustic characteristics such as a high pitch and a slow rate of speech (Jardat et al. 2022). Moreover, a change in arousal level or emotional valence in response to children's voices may be explained by an attraction to juvenile traits. Attraction to juvenile traits is well known in humans toward both conspecifics and pets (Archer and Monton 2011; Little 2012). It is thought to be related to parental care (Archer 1997) and could be shared by other mammal species, such as horses.

In conjunction with the first hypothesis (the horses may have reacted to the novelty of children's voices), it would be interesting to investigate the influence of horses' familiarity with children on their physiological reaction to this type of voice and on their reaction to congruent and incongruent presentations of faces and voices of children and adults. In addition, as we demonstrated that horses differentiate human adults from children, it would be interesting to investigate whether they behave differently toward these two age groups, as is commonly reported by horse owners. For example, in our survey posted on social media, 82% of the respondents who said they observed differences in their horses' behavior when interacting with children said that the animal was gentler or calmer toward children than toward adults (see Supplementary Information). Moreover, other types of categorization among humans could be investigated, such as differentiating between men and women.

It should be noted that in our experiment, 15 out of the 31 horses that were initially involved had to be excluded from the analysis, because they looked almost exclusively at one projection screen, right or left. This behavior can be explained by the horses trying to keep the exit door which was behind them in their visual field, by turning their head to one side throughout the experiment. Indeed, the horses used in this study live in groups permanently and are highly social, which can cause mild discomfort during social isolation. In future experiments using the same paradigm with highly social horses, a longer adaptation to the experimental set-up may be needed to avoid this issue. Moreover, since this study investigated only females, further research could also test male horses, without specifically hypothesizing on sex-related differences in the results.

Conclusion

In this study, horses associated the voices and faces of human adults and children in a cross-modal paradigm. This shows that horses are capable of differentiating human children and adults on the basis of both visual and vocal cues and that they form cross-modal representations of adults and children. Moreover, horses' emotional response to children's voices was stronger than that to adults' voices, which should be further investigated. Whether horses perceive and interact with human adults or children differently can be an important issue for equestrian practice, particularly in riding schools. These results also allow us to better understand the relationship between humans and horses, which has direct implications for horse management and welfare.

Declarations

Data availability: The datasets generated and analyzed during the current study are available in the INRAE data repository from the following link: https://doi.org/10.15454/SKHGKZ

Compliance with ethical standards: The authors declare no conflicts of interest. All the people participating in the study provided informed consent. Animal care and experimental treatments complied with the French and European guidelines for housing and care of animals used for scientific purposes (European Union Directive 2010/63/EU) and were performed under the authorization and supervision of official veterinary services (agreement number F371752 delivered to the UEPAO animal facility by the veterinary service of the Département d'Indre et Loire, France). This experiment experiment was approved by the Val de Loire Ethical Committee (CEEA VdL, Nouzilly, France, authorization number CE19-2022-1503-1). This study was reported in accordance with ARRIVE guidelines.

Funding: This study was funded by Institut Français du Cheval et de l'Equitation (IFCE), grant number 32 000809-Cognition Equine.

Author contributions: All authors devised the protocol. P.J., C.G., C.P., F.R., L.L. implemented the protocol. P.J., C.G., L.L. coded the videos and analyzed the data from heart rate monitoring and behavior coding. P.J., M.R, S.Y., C.G., R.D., L.C. and L.L. revised the analysis and report.

References

Adachi I, Kuwahata H, Fujita K (2007) Dogs recall their owner’s face upon hearing the owner’s voice. Anim Cogn 10:17–21. https://doi.org/10.1007/s10071-006-0025-8

Albuquerque N, Guo K, Wilkinson A, et al (2016) Dogs recognize dog and human emotions. Biol Lett 12:20150883. https://doi.org/10.1098/rsbl.2015.0883

Archer J (1997) Why do people love their pets? Evol Hum Behav 18:237–259. https://doi.org/10.1016/S0162-3095(99)80001-4

Archer J, Monton S (2011) Preferences for Infant Facial Features in Pet Dogs and Cats. Ethology 117:217–226. https://doi.org/10.1111/j.1439-0310.2010.01863.x

Brensing K, Linke K (2003) Behavior of dolphins towards adults and children during swim-with-dolphin programs and towards children with disabilities during therapy sessions. Anthrozoos 16:315–331. https://doi.org/10.2752/089279303786992035

Briefer EF, Maigrot A-L, Mandel R, et al (2015a) Segregation of information about emotional arousal and valence in horse whinnies. Sci Rep 5:9989. https://doi.org/10.1038/srep09989

Briefer EF, Tettamanti F, McElligott AG (2015b) Emotions in goats: Mapping physiological, behavioural and vocal profiles. Anim Behav 99:131–143. https://doi.org/10.1016/j.anbehav.2014.11.002

Dilley LC, Wieland EA, Gamache JL, et al (2013) Age-Related Changes to Spectral Voice Characteristics Affect Judgments of Prosodic, Segmental, and Talker Attributes for Child and Adult Speech. J Speech, Lang Hear Res 56:159–177. https://doi.org/10.1044/1092-4388(2012/11-0199)

Farkas LG, Hreczko TA, Kolar JC, Munro IR (1985) Vertical and Horizontal Proportions of the Face in Young Adult North American Caucasians. Plast Reconstr Surg 75:328–337. https://doi.org/10.1097/00006534-198503000-00005

Forkman B, Boissy A, Meunier-Salaün MC, et al (2007) A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. Physiol Behav 92:340–374. https://doi.org/10.1016/j.physbeh.2007.03.016

Gácsi M, Miklód Á, Varga O, et al (2004) Are readers of our face readers of our minds? Dogs (Canis familiaris) show situation-dependent recognition of human’s attention. Anim Cogn 24:144–153. https://doi.org/10.1007/s10071-003-0205-8

Glaze LE, Bless DM, Milenkovic P, Susser RD (1988) Acoustic characteristics of children’s voice. J Voice 2:312–319. https://doi.org/10.1016/S0892-1997(88)80023-7

Jardat P, Calandreau L, Ferreira V, et al (2022) Pet-directed speech improves horses’ attention toward humans. Sci Rep 12:4297. https://doi.org/10.1038/s41598-022-08109-z

Jardat P, Lansade L (2021) Cognition and the human–animal relationship: a review of the sociocognitive skills of domestic mammals toward humans. Anim Cogn. https://doi.org/10.1007/s10071-021-01557-6

Koo TK, Li MY (2016) A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 15:155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lampe JF, Andre J (2012) Cross-modal recognition of human individuals in domestic horses (Equus caballus). Anim Cogn 15:623–630. https://doi.org/10.1007/s10071-012-0490-1

Lansade L, Bouissou MF, Erhard HW (2008) Fearfulness in horses: A temperament trait stable across time and situations. Appl Anim Behav Sci 115:182–200. https://doi.org/10.1016/j.applanim.2008.06.011

Lansade L, Colson V, Parias C, et al (2020a) Human Face Recognition in Horses: Data in Favor of a Holistic Process. Front Psychol 11:2311. https://doi.org/10.3389/fpsyg.2020.575808

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

Lansade L, Trösch M, Parias C, et al (2021) Horses are sensitive to baby talk : Pet-directed speech facilitates communication with humans in a pointing task and during grooming. Anim Cogn 5:999–1006. https://doi.org/10.1007/s10071-021-01487-3

Lee S, Potamianos A, Narayanan S (1999) Acoustics of children’s speech: Developmental changes of temporal and spectral parameters. J Acoust Soc Am 105:1455–1468. https://doi.org/10.1121/1.426686

Little AC (2012) Manipulation of Infant-Like Traits Affects Perceived Cuteness of Infant, Adult and Cat Faces. Ethology 118:775–782. https://doi.org/10.1111/J.1439-0310.2012.02068.X

Moyse E (2014) Age estimation from faces and voices: A review. In: Psychologica Belgica. Ubiquity Press Ltd, pp 255–265

Munsters CCBM, Visser KEK, van den Broek J, Sloet van Oldruitenborgh-Oosterbaan MM (2012) The influence of challenging objects and horse-rider matching on heart rate, heart rate variability and behavioural score in riding horses. Vet J 192:75–80. https://doi.org/10.1016/j.tvjl.2011.04.011

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

Pattison KF, Miller HC, Rayburn-Reeves R, Zentall T (2010) The case of the disappearing bone: Dogs’ understanding of the physical properties of objects. Behav Processes 85:278–282. https://doi.org/10.1016/J.BEPROC.2010.06.016

Proops L, Grounds K, Smith AV, McComb K (2018) Animals Remember Previous Facial Expressions that Specific Humans Have Exhibited. Curr Biol 28:1428-1432.e4. https://doi.org/10.1016/j.cub.2018.03.035

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

Quaranta A, D’ingeo S, Amoruso R, Siniscalchi M (2020) Emotion recognition in cats. Animals 10:1107. https://doi.org/10.3390/ani10071107

Ringhofer M, Trösch M, Lansade L, Yamamoto S (2021) Horses with sustained attention follow the pointing of a human who knows where food is hidden. Sci Reports 2021 111 11:1–9. https://doi.org/10.1038/s41598-021-95727-8

Ringhofer M, Yamamoto S (2017) Domestic horses send signals to humans when they face with an unsolvable task. Anim Cogn 20:397–405. https://doi.org/10.1007/s10071-016-1056-4

Russell JA, Barrett LF (1999) Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. J Pers Soc Psychol 76:805–819. https://doi.org/10.1037/0022-3514.76.5.805

Siniscalchi M, D’Ingeo S, Quaranta A (2018) Orienting asymmetries and physiological reactivity in dogs’ response to human emotional faces. Learn Behav 46:574–585. https://doi.org/10.3758/s13420-018-0325-2

Smith AV, Proops L, Grounds K, et al (2018) Domestic horses (Equus caballus) discriminate between negative and positive human nonverbal vocalisations. Sci Rep 8:13052. https://doi.org/10.1038/s41598-018-30777-z

Somppi S, Törnqvist H, Hänninen L, et al (2012) Dogs do look at images: eye tracking in canine cognition research. Anim Cogn 15:163–174. https://doi.org/10.1007/s10071-011-0442-1

Stathopoulos ET, Huber JE, Sussman JE (2011) Changes in acoustic characteristics of the voice across the life span: Measures from individuals 4-93 years of age. J Speech, Lang Hear Res 54:1011–1021. https://doi.org/10.1044/1092-4388(2010/10-0036)

Sussman JE, Sapienza C (1994) Articulatory, developmental, and gender effects on measures of fundamental frequency and jitter. J Voice 8:145–156. https://doi.org/10.1016/S0892-1997(05)80306-6

Takagi S, Arahori M, Chijiiwa H, et al (2019) Cats match voice and face: cross-modal representation of humans in cats (Felis catus). Anim Cogn 22:901–906. https://doi.org/10.1007/s10071-019-01265-2

Trösch M, Bertin E, Calandreau L, et al (2020a) Unwilling or willing but unable: can horses interpret human actions as goal directed? Anim Cogn 23:1035–1040. https://doi.org/10.1007/s10071-020-01396-x

Trösch M, Cuzol F, Parias C, et al (2019a) Horses categorize human emotions cross-modally based on facial expression and non-verbal vocalizations. Animals 9:862. https://doi.org/10.3390/ani9110862

Trösch M, Pellon S, Cuzol F, et al (2020b) 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 23:643–653. https://doi.org/10.1007/s10071-020-01369-0

Trösch M, Ringhofer M, Yamamoto S, et al (2019b) Horses prefer to solicit a person who previously observed a food-hiding process to access this food: A possible indication of attentional state attribution. Behav Processes 166:103906. https://doi.org/10.1016/j.beproc.2019.103906

von Borell E, Langbein J, Després G, et al (2007) Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals - A review. Physiol Behav 92:293–316. https://doi.org/10.1016/j.physbeh.2007.01.007

Wanser SH, MacDonald M, Udell MAR (2021) Dog-human behavioral synchronization: family dogs synchronize their behavior with child family members. Anim Cogn 24:747–752. https://doi.org/10.1007/s10071-020-01454-4

Wanser SH, Simpson AC, MacDonald M, Udell MAR (2020) Considering Family Dog Attachment Bonds: Do Dog-Parent Attachments Predict Dog-Child Attachment Outcomes in Animal-Assisted Interventions? Front Psychol 11:2293. https://doi.org/10.3389/FPSYG.2020.566910/BIBTEX

FIGURE CAPTIONS

Fig. 1 Schematic representation of the experimental setup. Two simultaneous muted videos of a child and an adult saying the same neutral sentence were projected on the screens, while the sound of an adult's or child's voice speaking this sentence was played over a loudspeaker.

Fig. 2 Example of the composition of the stimuli shown to a horse. Eight different faces were shown, and four different voices were heard.

Fig. 3 Mean preference index. (INC – CON) / (INC + CON), with INC being the time spent looking at the incongruent video and CON the time spent looking at the congruent video. Boxplot: median, 1^{st} and 3^{rd} quartiles. Gray dots: individual indexes. Red cross: mean group index. p-value: comparison to 0, Wilcoxon test.

Fig. 4 Mean heart rate variation during adults' and children's vocalizations. Boxplot: median, 1^{st} and 3^{rd} quartiles. Gray dots: individual means. Red cross: mean group variation. p-value: comparison to 0, Wilcoxon test.