



HAL
open science

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

► **To cite this version:**

Plotine Jardat, Monamie Ringhofer, Shinya Yamamoto, Chloé Gouyet, Rachel Degrande, et al.. Horses form cross-modal representations of adults and children. *Animal Cognition*, 2023, 26, pp.369-377. 10.1007/s10071-022-01667-9 . hal-03751952

HAL Id: hal-03751952

<https://hal.inrae.fr/hal-03751952>

Submitted on 16 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

2 Horses form cross-modal representations of 3 adults and children

4 Plotine Jardat^{1,*}, Monamie Ringhofer², Shinya Yamamoto^{3,4}, Chloé Gouyet¹, Rachel Degrande¹, Céline
5 Parias¹, Fabrice Reigner⁵, Ludovic Calandreau¹, Léa Lansade^{1,*}

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

7 ² Department of Animal Science, Teikyo University of Science

8 ³ Institute for Advanced Study, Kyoto University, Kyoto, Japan

9 ⁴ Wildlife Research Center, Kyoto University, Kyoto, Japan

10 ⁵ UEPAO, INRAE F-37380, Nouzilly, France

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

12 ORCID IDs: P Jardat 0000-0003-0374-5588 ; M Ringhofer 0000-0002-4813-939X ; S Yamamoto 0000-
13 0001-7938-1106 ; C Gouyet 0000-0001-5548-053X ; R Degrande 0000-0003-4976-1073 ; L Calandreau
14 0000-0002-7535-5733 ; L Lansade 0000-0002-4185-9714

15

16 **Abstract**

17 Recently, research on domestic mammals' sociocognitive skills toward humans has been prolific,
18 allowing us to better understand the human-animal relationship. For example, horses have been
19 shown to distinguish human beings on the basis of photographs and voices and to have cross-modal
20 mental representations of individual humans and human emotions. This leads to questions such as
21 the extent to which horses can differentiate human attributes such as age. Here, we tested whether
22 horses discriminate human adults from children. In a cross-modal paradigm, we presented 31 female
23 horses with two simultaneous muted videos of a child and an adult saying the same neutral
24 sentence, accompanied by the sound of an adult's or child's voice speaking the sentence. The horses

25 looked significantly longer at the videos that were incongruent with the heard voice than at the
26 congruent videos. We conclude that horses can match adults' and children's faces and voices cross-
27 modally. Moreover, their heart rates increased during children's vocalizations but not during adults'.
28 This suggests that in addition to having mental representations of adults and children, horses have a
29 stronger emotional response to children's voices than adults' voices.

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

32 **Statements and declarations**

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

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

38 **Acknowledgments**

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

42

43 Introduction

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

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

64 In humans, age is an essential attribute used to categorize conspecifics and can be assessed
65 by both facial and vocal cues (Moyses 2014). Indeed, there are facial and vocal differences between
66 people of different age groups. Facial characteristics (e.g., nose width to mouth length and face
67 width proportions) evolve throughout the life span (Farkas et al. 1985). Children's and adults' voices

68 differ in various acoustic parameters, such as fundamental frequency, speech rate and formants
69 (Glaze et al. 1988; Sussman and Sapienza 1994; Lee et al. 1999; Stathopoulos et al. 2011; Dilley et al.
70 2013). We hypothesized that horses could differentiate human adults from children too, and that
71 they could use these differences in facial and vocal features to do so as they are known to use these
72 types of cues to recognize individual humans or human emotions (Proops and McComb 2012; Lampe
73 and Andre 2012; Trösch et al. 2019a).

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

94 children by showing horses images of adults' and children's faces while playing adults' or children's
95 voices.

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

104 Horses were simultaneously shown images of two human faces (one of an adult and one of a
105 child) while the voice of an adult or a child was played over a speaker. We hypothesized that the
106 horses would look at the videos for different amounts of time according to their congruency with the
107 voice heard and would possibly spend longer looking at the incongruent video, as reported for a
108 similar protocol in horses (Trösch et al. 2019a). Moreover, we hypothesized that the horses would
109 react differently to child than to adult stimuli (as suggested in other species - Brensing and Linke
110 2003; Wanser et al. 2020, 2021) and that consequently, their heart rates would vary differently in
111 response to children's voices than in response to adults' voices.

112 Materials and Methods

113 **Ethics statement**

114 Our experiment was approved by the Val de Loire Ethical Committee (CEEA VdL, Nouzilly, France,
115 authorization number CE19-2022-1503-1). Animal care and experimental treatments complied with
116 the French and European guidelines for housing and care of animals used for scientific purposes
117 (European Union Directive 2010/63/EU) and were performed under the authorization and
118 supervision of official veterinary services (agreement number F371752 delivered to the UEPAO

119 animal facility by the veterinary service of the Département d'Indre et Loire, France). The animals
120 lived in groups, were not food deprived during the experiment and did not undergo any invasive
121 procedures. All methods were carried out in accordance with the relevant guidelines and regulations
122 for direct human involvement in the study. All the people participating in the study provided
123 informed consent.

124

125 **Subjects**

126 The study involved 31 Welsh mares (*Equus caballus*) aged 8.9 ± 2.4 years (mean \pm sd) reared at the
127 Animal Physiology Experimental Unit PAO (UEPAO, 37380 Nouzilly, France, DOI:
128 10.15454/1.55738963217 28955E12), INRAE. They lived in groups in indoor stalls on straw and had
129 free access to an outdoor paddock. Hay and water were available ad libitum. These horses are used
130 only for research purposes and are handled daily by human adults. They interact with children on
131 rare occasions (approximately twice a month for less than an hour).

132

133 **Experimental setup**

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

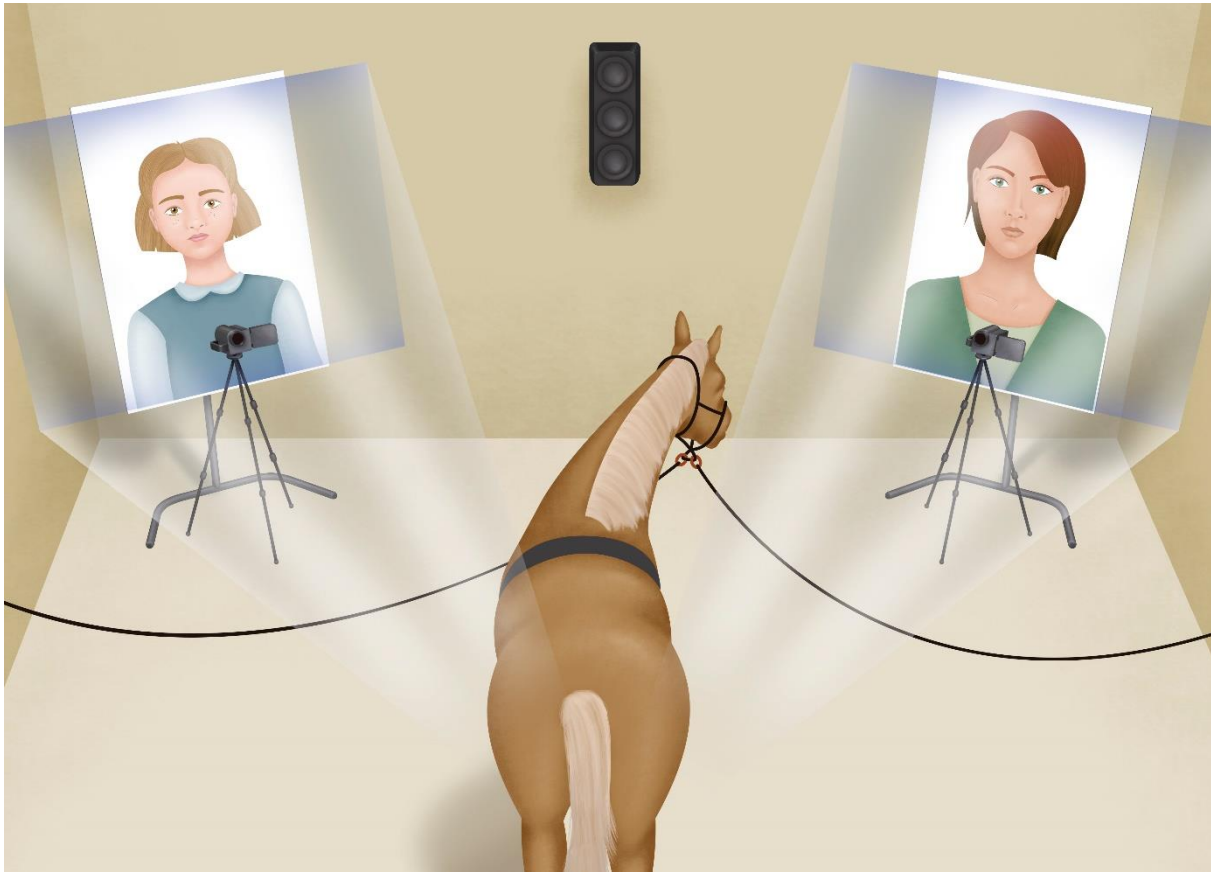


Fig. 1 Schematic representation of the experimental setup. Two simultaneous muted videos of a child and an adult saying the same neutral sentence were projected on the screens, while the sound of an adult's or child's voice speaking this sentence was played over a loudspeaker.

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

151

152 **Stimuli**

153 During the test sessions, a muted video of a child and a muted video of an adult were played
154 simultaneously, one on each projection screen. At the same time, the voice of another person (child
155 or adult) was played over a loudspeaker placed between the projection screens (Fig. 1). The voice
156 played was not the original voice of either of the two projected videos to prevent the horses from
157 matching the heard voice to the mouth movements of any of the videos. The faces were projected at
158 a height of approximately 160 cm and were approximately twice the size of a real human's face (as in
159 Trösch et al. 2019a). The sound was broadcast from a loudspeaker placed between the screens and
160 in front of the horse (approximately 60 dB from where the horse's head was positioned).
161 Since horses may differentiate male and female faces and voices of humans, we used only female
162 stimuli in this experiment. Four young girls aged 6 to 9 years and four women aged 46 to 58 years
163 who the horses had never seen or heard were filmed. They were recorded in similar conditions,
164 against a white background and with the face centered and the shoulders visible. They said a
165 predefined neutral sentence that was identical for all participants ("J'aime la glace à la pistache;
166 j'aime la glace au chocolat", meaning "I like pistachio ice cream; I like chocolate ice cream"), their
167 facial expression was neutral. The acoustic characteristics of the sounds were obtained using Praat
168 V6.1.16 (<http://www.fon.hum.uva.nl/praat/>); with a mean pitch of 269 ± 36 Hz for children's voices
169 and 227 ± 15 Hz for adults' voices, and a pitch range of 392 ± 18 Hz for children's voices and 416 ± 1 Hz
170 for adults' voices.

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

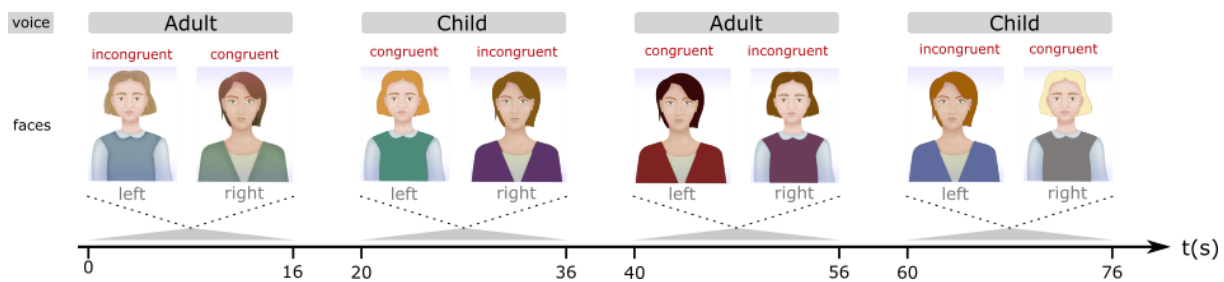


Fig. 2 Example of the composition of the stimuli shown to a horse. Eight different faces were shown, and four different voices were heard.

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

179

180 **Habituation**

181 In the habituation phase, the horse was led to the middle of the stall facing the screens. It was then
 182 attached with two loose ropes, the assistant took their place, and the habituation phase began.

183 Identical scenes of nature accompanied by birdsongs were projected on both screens, while the
 184 assistant watched the horse's heart rate on the Polar watch. In this phase, the assistant could
 185 reposition the horse if it had its back to the screen. Once the horse was calm (not neighing, pulling on

186 the ropes with her head, or trying to turn around or leave) and her heart rate had stayed below 100
 187 bpm for two consecutive minutes, the test phase began immediately. If this criterion was not met

188 after five minutes, the session ended, and a new one was scheduled for the next day. All the horses
 189 but three met the criterion on day 1 and therefore continued with the test session on the same day.

190 The three other horses needed a second habituation session; they met the habituation criterion in
 191 this second session and continued with the test session on the same day.

192

193 **Tests**

194 Immediately after the horse met the habituation criterion, two videos of adults' and children's faces
 195 accompanied by adults' and children's voices (see **Stimuli** section above) were projected. The
 196 conditions were the same as those during habituation, but the assistant never intervened. The

197 assistant was unaware of the side on which each video appeared (adult or child, congruent or
198 incongruent). At the end of the test, the horse was led directly back to its stable.

199

200 **Behavioral and physiological analyses**

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

213 The total time spent looking at the stimuli during a section was highly variable among horses (from 1
214 to 16 seconds over the 16-second sections). To take this variability into account, we calculated, for
215 each horse, a preference index measuring the propensity to look more at the incongruent video in
216 each section relative to the total time spent looking at the videos. The index was defined as $(INC -$
217 $CON)/(INC + CON)$ with INC being the time spent looking at the incongruent video and CON the time
218 spent looking at the congruent video. Then, for each horse, the mean of this index over the four
219 sections was calculated. The index varied from -1 to 1, with positive values indicating that a horse
220 looked more at the incongruent videos and negative values the opposite.

221 Some horses tried to keep the exit door, situated behind them, in their field of vision
222 throughout the duration of the test (probably because they were not comfortable being alone in the

223 stall). Consequently, they almost never looked at either the right or the left screen, regardless of the
224 congruence of the video and the voice heard. For this reason, these horses were excluded: fifteen
225 horses which looked at one side for less than two seconds on average per trial were not taken into
226 account in the analysis (a duration of 2 s had previously been used in the literature – Gácsi et al.
227 2004; Pattison et al. 2010; Somppi et al. 2012; Trösch et al. 2020a).

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

236

237 **Statistical analyses**

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

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

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

250

251 Results

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

260 Discussion

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

265 In line with our hypothesis, the horses looked more at the video that was incongruent with
266 the vocalization played. In other words, they differentiated between human adults and children on
267 the videos based on their congruence or incongruence with the voice they heard at the same time,
268 which shows that horses can differentiate human adults and children. This finding is consistent with
269 previous studies on dolphins and dogs (Brensing and Linke 2003; Wanser et al. 2020, 2021).

270 Moreover, these results show that horses can match a face and voice when they come from the
271 same age group (adults or children), which suggests that they bear mental representations of adults

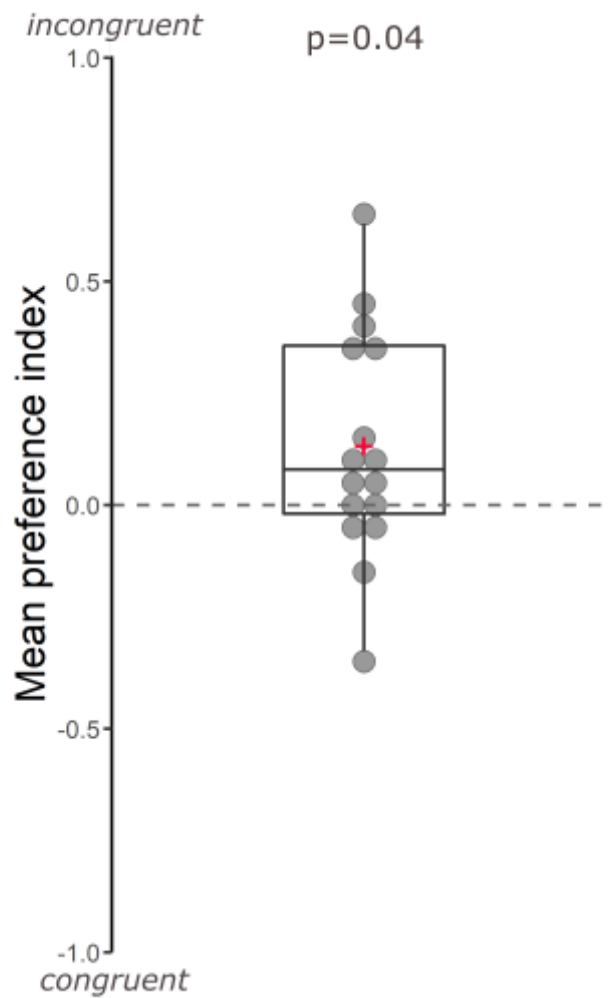


Fig. 3 Mean preference index. $(INC - CON) / (INC + CON)$, with INC being the time spent looking at the incongruent video and CON the time spent looking at the congruent video. Boxplot: median, 1st and 3rd quartiles. Gray dots: individual indexes. Red cross: mean group index. p-value: comparison to 0,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

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

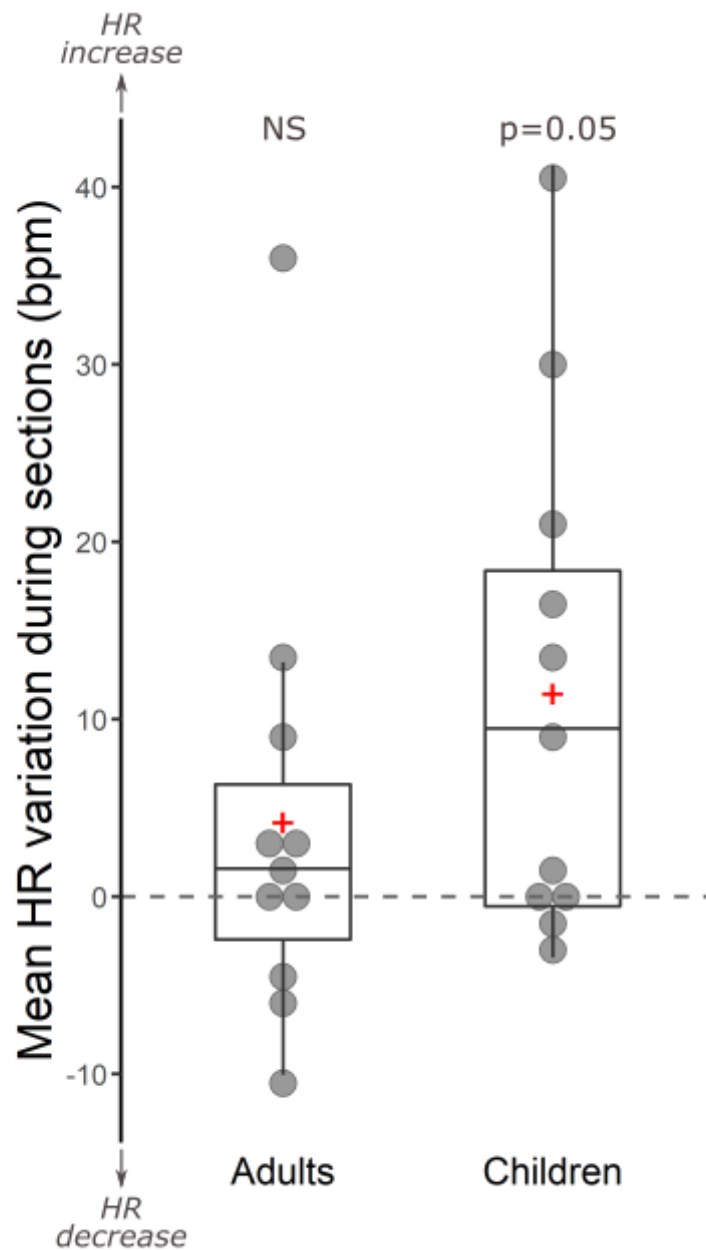


Fig. 4 Mean heart rate variation during adults' and children's vocalizations. Boxplot: median, 1st and 3rd quartiles. Gray dots: individual means. Red cross: mean group variation. p-value: comparison to 0, Wilcoxon test.

280 would be necessary to test whether these results can be generalized to men and young boys.

281 Similarly, only mares were tested in this experiment, and it would be interesting to test whether the

282 results can be generalized to stallions and geldings.

283 In this study, the horses looked longer at the videos that did not match the vocalization they

284 heard. This is in line with a previous experiment using the preference looking paradigm in horses, in

285 which they looked more at incongruent than congruent stimuli when presented with emotional faces

286 and voices of humans (Trösch et al. 2019a). Similarly, in expectancy violation paradigms presenting
287 human voices and faces horses, cats and dogs looked more at images which were incongruent with
288 the sound (Adachi et al. 2007; Lampe and Andre 2012; Nakamura et al. 2018; Takagi et al. 2019).
289 These longer looks toward incongruent stimuli can be explained by the animals being intrigued by
290 contradictory information (Trösch et al. 2019a). However, in other experiments on domestic
291 mammals' cross-modal recognition of human individuals and human emotions using the preference
292 looking paradigm, the animals looked more at the congruent stimuli (Proops and McComb 2012;
293 Albuquerque et al. 2016; Quaranta et al. 2020). These differences may be explained by variations in
294 experimental conditions, for example, differences in the emotions or level of stress provoked by the
295 presented stimuli.

296 In this study, the horses' heart rates increased when they heard children's voices but not
297 when they heard adults' voices. This difference in physiological reaction to the two types of voices
298 reinforces the idea that horses differentiate children's voices from adults' voices and shows that the
299 children's voices induced a stronger emotional reaction. This is in line with previous findings
300 suggesting that some species react differently to children than to adults: dolphins and dogs seem to
301 interact differently with these two age groups (Brensing and Linke 2003; Wanser et al. 2020, 2021).
302 An increase in heart rate can indicate either increased arousal or a change in the valence of emotions
303 felt by horses (Jardat et al. 2022), and several hypotheses can explain such reactions to children's
304 voices. First, the horses may have reacted to the novelty of this type of voice. Indeed, the horses in
305 this experiment were less familiar with children's voices, as they seldom interact with children but
306 are handled daily by adults. Thus, they could have been surprised by this type of voice, explaining a
307 change in arousal level (Russell and Barrett 1999). Second, the horses may have reacted to the
308 intrinsic characteristics of children's voices. Indeed, children's voices are notably characterized by
309 differences in pitch and other prosodic features, such as speech rate (Dilley et al. 2013). These
310 characteristics could elicit arousal and consequently a rise in heart rate, similar to horses' reaction to
311 pet-directed speech, a type of speech that has specific acoustic characteristics such as a high pitch

312 and a slow rate of speech (Jardat et al. 2022). Moreover, a change in arousal level or emotional
313 valence in response to children's voices may be explained by an attraction to juvenile traits.
314 Attraction to juvenile traits is well known in humans toward both conspecifics and pets (Archer and
315 Monton 2011; Little 2012). It is thought to be related to parental care (Archer 1997) and could be
316 shared by other mammal species, such as horses.

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

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

337 Conclusion

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

346 Declarations

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

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

358 Funding: This study was funded by Institut Français du Cheval et de l'Équitation (IFCE), grant number
359 32 000809-Cognition Equine.

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

363 References

- 364 Adachi I, Kuwahata H, Fujita K (2007) Dogs recall their owner's face upon hearing the owner's voice.
365 Anim Cogn 10:17–21. <https://doi.org/10.1007/s10071-006-0025-8>
- 366 Albuquerque N, Guo K, Wilkinson A, et al (2016) Dogs recognize dog and human emotions. Biol Lett
367 12:20150883. <https://doi.org/10.1098/rsbl.2015.0883>
- 368 Archer J (1997) Why do people love their pets? Evol Hum Behav 18:237–259.
369 [https://doi.org/10.1016/S0162-3095\(99\)80001-4](https://doi.org/10.1016/S0162-3095(99)80001-4)
- 370 Archer J, Monton S (2011) Preferences for Infant Facial Features in Pet Dogs and Cats. Ethology
371 117:217–226. <https://doi.org/10.1111/j.1439-0310.2010.01863.x>
- 372 Breusing K, Linke K (2003) Behavior of dolphins towards adults and children during swim-with-
373 dolphin programs and towards children with disabilities during therapy sessions. Anthrozoos
374 16:315–331. <https://doi.org/10.2752/089279303786992035>
- 375 Briefer EF, Maigrot A-L, Mandel R, et al (2015a) Segregation of information about emotional arousal
376 and valence in horse whinnies. Sci Rep 5:9989. <https://doi.org/10.1038/srep09989>
- 377 Briefer EF, Tettamanti F, McElligott AG (2015b) Emotions in goats: Mapping physiological,
378 behavioural and vocal profiles. Anim Behav 99:131–143.
379 <https://doi.org/10.1016/j.anbehav.2014.11.002>
- 380 Dilley LC, Wieland EA, Gamache JL, et al (2013) Age-Related Changes to Spectral Voice Characteristics
381 Affect Judgments of Prosodic, Segmental, and Talker Attributes for Child and Adult Speech. J
382 Speech, Lang Hear Res 56:159–177. [https://doi.org/10.1044/1092-4388\(2012/11-0199\)](https://doi.org/10.1044/1092-4388(2012/11-0199))
- 383 Farkas LG, Hreczko TA, Kolar JC, Munro IR (1985) Vertical and Horizontal Proportions of the Face in
384 Young Adult North American Caucasians. Plast Reconstr Surg 75:328–337.
385 <https://doi.org/10.1097/00006534-198503000-00005>

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

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

392 Glaze LE, Bless DM, Milenkovic P, Susser RD (1988) Acoustic characteristics of children’s voice. *J Voice*
393 2:312–319. [https://doi.org/10.1016/S0892-1997\(88\)80023-7](https://doi.org/10.1016/S0892-1997(88)80023-7)

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

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

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

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

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

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

408 Lansade L, Colson V, Parias C, et al (2020b) Female horses spontaneously identify a photograph of

409 their keeper, last seen six months previously. *Sci Rep* 10:6302. [https://doi.org/10.1038/s41598-](https://doi.org/10.1038/s41598-020-62940-w)
410 020-62940-w

411 Lansade L, Trösch M, Parias C, et al (2021) Horses are sensitive to baby talk : Pet-directed speech
412 facilitates communication with humans in a pointing task and during grooming. *Anim Cogn*
413 5:999–1006. <https://doi.org/10.1007/s10071-021-01487-3>

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

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

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

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

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

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

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

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

436 Quaranta A, D’Ingeo S, Amoruso R, Siniscalchi M (2020) Emotion recognition in cats. *Animals*
437 10:1107. <https://doi.org/10.3390/ani10071107>

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

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

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

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

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

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

454 Stathopoulos ET, Huber JE, Sussman JE (2011) Changes in acoustic characteristics of the voice across
455 the life span: Measures from individuals 4–93 years of age. *J Speech, Lang Hear Res* 54:1011–
456 1021. [https://doi.org/10.1044/1092-4388\(2010/10-0036\)](https://doi.org/10.1044/1092-4388(2010/10-0036))

457 Sussman JE, Sapienza C (1994) Articulatory, developmental, and gender effects on measures of
458 fundamental frequency and jitter. *J Voice* 8:145–156. <https://doi.org/10.1016/S0892->
459 1997(05)80306-6

460 Takagi S, Arahori M, Chijiwa H, et al (2019) Cats match voice and face: cross-modal representation of
461 humans in cats (*Felis catus*). *Anim Cogn* 22:901–906. <https://doi.org/10.1007/s10071-019->
462 01265-2

463 Trösch M, Bertin E, Calandreau L, et al (2020a) Unwilling or willing but unable: can horses interpret
464 human actions as goal directed? *Anim Cogn* 23:1035–1040. <https://doi.org/10.1007/s10071->
465 020-01396-x

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

469 Trösch M, Pellon S, Cuzol F, et al (2020b) Horses feel emotions when they watch positive and
470 negative horse–human interactions in a video and transpose what they saw to real life. *Anim*
471 *Cogn* 23:643–653. <https://doi.org/10.1007/s10071-020-01369-0>

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

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

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

481 Wanser SH, Simpson AC, MacDonald M, Udell MAR (2020) Considering Family Dog Attachment
482 Bonds: Do Dog-Parent Attachments Predict Dog-Child Attachment Outcomes in Animal-Assisted
483 Interventions? Front Psychol 11:2293. <https://doi.org/10.3389/FPSYG.2020.566910/BIBTEX>
484

485 **FIGURE CAPTIONS**

486

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

490

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

493

494 **Fig. 3** Mean preference index. $(INC - CON) / (INC + CON)$, with INC being the time spent looking at
495 the incongruent video and CON the time spent looking at the congruent video. Boxplot: median, 1st
496 and 3rd quartiles. Gray dots: individual indexes. Red cross: mean group index. p-value: comparison to
497 0, Wilcoxon test.

498 **Fig. 4** Mean heart rate variation during adults' and children's vocalizations. Boxplot: median, 1st and
499 3rd quartiles. Gray dots: individual means. Red cross: mean group variation. p-value: comparison to 0,
500 Wilcoxon test.