

Cross-modal individual recognition in domestic horses (*Equus caballus*) extends to familiar humans

Leanne Proops* and Karen McComb

Mammal Vocal Communication and Cognition Research, School of Psychology, University of Sussex, Brighton BN1 9QH, UK

It has recently been shown that some non-human animals can cross-modally recognize members of their own taxon. What is unclear is just how plastic this recognition system can be. In this study, we investigate whether an animal, the domestic horse, is capable of spontaneous cross-modal recognition of individuals from a morphologically very different species. We also provide the first insights into how cross-modal identity information is processed by examining whether there are hemispheric biases in this important social skill. In our preferential looking paradigm, subjects were presented with two people and playbacks of their voices to determine whether they were able to match the voice with the person. When presented with familiar handlers subjects could match the specific familiar person with the correct familiar voice. Horses were significantly better at performing the matching task when the congruent person was standing on their right, indicating marked hemispheric specialization (left hemisphere bias) in this ability. These results are the first to demonstrate that cross-modal recognition in animals can extend to individuals from phylogenetically very distant species. They also indicate that processes governed by the left hemisphere are central to the cross-modal matching of visual and auditory information from familiar individuals in a naturalistic setting.

Keywords: *Equus caballus*; individual recognition; lateralization; interspecific communication; animal cognition; social behaviour

1. INTRODUCTION

Being able to identify individual social partners is central to developing sophisticated social relationships. Cross-modal matching was once thought to be unique to humans [1], and although individual recognition is believed to be widespread [2], it has been hard to prove conclusively in animals because this requires a demonstration not only that discrimination between identity cues occurs at the level of the individual, but also that current sensory cues match stored information about that specific individual. In addition, if animals are capable of combining cues to an individual's identity cross-modally, then this suggests that these cues are ultimately stored in some form of higher-order representation that is independent of modality. We recently used an expectancy violation paradigm to demonstrate that domestic horses are capable of spontaneous cross-modal individual recognition of conspecifics [3]. The ability to match the sight of familiar individuals with their voice, without explicit training, has also since been demonstrated in non-human primates (*Lophocebus albigena*, *Macaca mulatta*) and crows (*Corvus macrorhynchos*), indicating that the ability is likely to have a long phylogenetic history [4–6].

What remains unclear is the extent to which this recognition system is flexible in the types of individual encoded.

It is well known that humans (and other animals) become specialized in processing identity information from members of their own species during infancy. Young human infants are just as capable of categorizing heterospecific auditory, visual and cross-modal stimuli as human stimuli but between 6 and 10 months of age these discriminatory abilities begin to become specialized towards the processing of the language and the sight of individuals from their own species, and more precisely from their own socio-ethnic group [7]. This perceptual narrowing can, however, be slowed with exposure to other species [8]. For domestic animals, humans also represent significant social partners and often become more familiar than conspecifics. It would therefore be highly adaptive for animals that develop strong bonds with members of another species, such as domestic horses, to be able to cross-modally recognize these significant individuals. Domestic dogs have been shown to match the image of a single familiar person with their voice, but this was in a design where matching of familiar stimuli rather than activation of a representation of a specific individual might have occurred [9]. Domestic horses are capable of discriminating between different human faces, and have been shown to discriminate between familiar and unfamiliar people [10,11], but genuine cross-modal individual recognition of heterospecifics has yet to be demonstrated outside the primate taxa [4]. The prevalence of this ability across different species therefore remains unknown. It also remains possible that the relatively similar physical features of humans and primates in some way facilitate recognition of humans by primates [12]. Thus, one aim of this study is to assess

* Author for correspondence (leanne.proops@gmail.com).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2012.0626> or via <http://rspb.royalsocietypublishing.org>.

whether a species that is morphologically and phylogenetically very different from humans, yet forms a close relationship with them, is capable of cross-modal recognition of a range of familiar people. This would suggest not only that cross-modal individual recognition is widespread among mammals but also that it is highly plastic in the morphology of individuals that it can code.

Moreover, in our study, we also provide the first investigation into cerebral functional asymmetries associated with the ability to spontaneously recognize social partners cross-modally. It is now clear that hemispheric specialization has its origins in pre-mammalian brains before 500 Ma [13]. A general explanation of the roles of the different hemispheres suggests that the left hemisphere provides focused attention and is involved in top-down processing, controlling well-established patterns of behaviour, approach responses, pro-social behaviour and the categorization of familiar stimuli, and as such may have a positive cognitive bias. The right hemisphere is more under bottom-up control and is central to the processing of novel, potentially threatening stimuli and producing corresponding escape behaviour [13,14]. There is some debate as to whether the right hemisphere controls the reaction to all stimuli producing strong emotional reactions (the right hemisphere hypothesis) or whether the right hemisphere controls the reaction to stimuli with negative affect and the left controls responses to those with a positive affect (the valence hypothesis) [15–18].

The few studies that have assessed cerebral asymmetries in the processing of cross-modal information have mainly been concerned with multisensory integration of voice–face information in humans and have involved presenting subjects with unimodal visual or auditory identity information to determine which brain areas are activated by stimuli from both modalities [19,20]. Results tend to show a synchronization of activation between face and voice selective areas, often with activation of additional cortical areas that may be the location of semantic information about particular individuals (so-called ‘person identity nodes’) [21]. Specific right hemispheric activation of these areas has been reported by some studies [22–24]. Two studies to date have investigated hemispheric specialization in animals during learnt audiovisual matching tasks. Rhesus macaques that learnt to associate six non-biological sounds with six visual patterns were significantly impaired when lesions were made to the left hemisphere, but were unimpaired when lesions were performed on the right [25]. The second task is particularly relevant here: bottlenose dolphins (*Tursiops truncatus*) were trained to associate audio stimuli (including known signature whistles, human voices and tones) with visual objects (including videos of dolphins and people). Here, subjects showed no strong eye preference but were significantly better at matching the audio-visual stimuli across all categories if objects were viewed by the right eye [26]. This study also suggested that the type of audio-visual stimuli (conspecific, heterospecific, non-biological) did not affect lateralization.

Our study therefore not only provides the first assessment of cross-modal individual recognition of social partners with physical characteristics that are very different to the subject species, but also provides the first insights into the neural processes at work during the natural assimilation and activation of cross-modal social information in

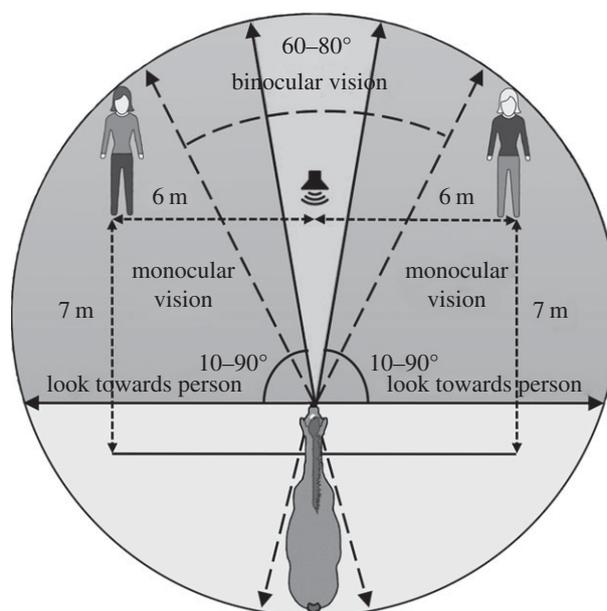


Figure 1. Diagram showing the experimental set-up. The behavioural coding of subject head orientations defined as looks towards each person ($10\text{--}90^\circ$ from centre), the speaker (less than 10° from centre) and elsewhere (greater than 90° from centre) are also shown in relation to the binocular ($60\text{--}80^\circ$) and monocular visual fields of the horse.

animals. The domestic horse is an ideal animal model for this research because it has a complex social organization and close relationship to man, making individual recognition of humans a highly functional ability. We employed a preferential looking paradigm to determine whether horses could match playbacks of human voices with the sight of those specific people. Two people stood either side of a loudspeaker from which the voices of each person (giving two renditions of the subject's name) were played with an interval of 15 s between presentations of the different voices (figure 1). Horses stood facing the people on the centre line, and their responses to the congruent and incongruent person were recorded. In experiment 1, we tested 32 horses for the ability to use cross-modal information in a basic task involving discrimination of known individuals from strangers. In experiment 2, 40 horses were tested for whether they could perform the more complex task of cross-modal individual recognition by virtue of presenting them with a choice between a number of different familiar humans (10 pairs of handlers were presented to four subjects each). To ensure that the humans were not inadvertently cuing the horses, a control was included in which handlers listened to white noise on small headphones to prevent them hearing the playback itself. If horses are capable of cross-modal discrimination of familiar from unfamiliar people and recognition of familiar handlers, we predicted they would be quicker to look, and look for longer at the person that matched the voice just heard [27,28]. Analyses of side preferences and success at discrimination in relation to the side on which the matched person was standing were conducted to examine whether horses showed any behavioural asymmetry in orienting response and in recognition/discrimination ability. Visual and auditory control trials were also conducted to test for hemispheric biases and possible handler/stranger preferences to the presentation of unimodal stimuli.

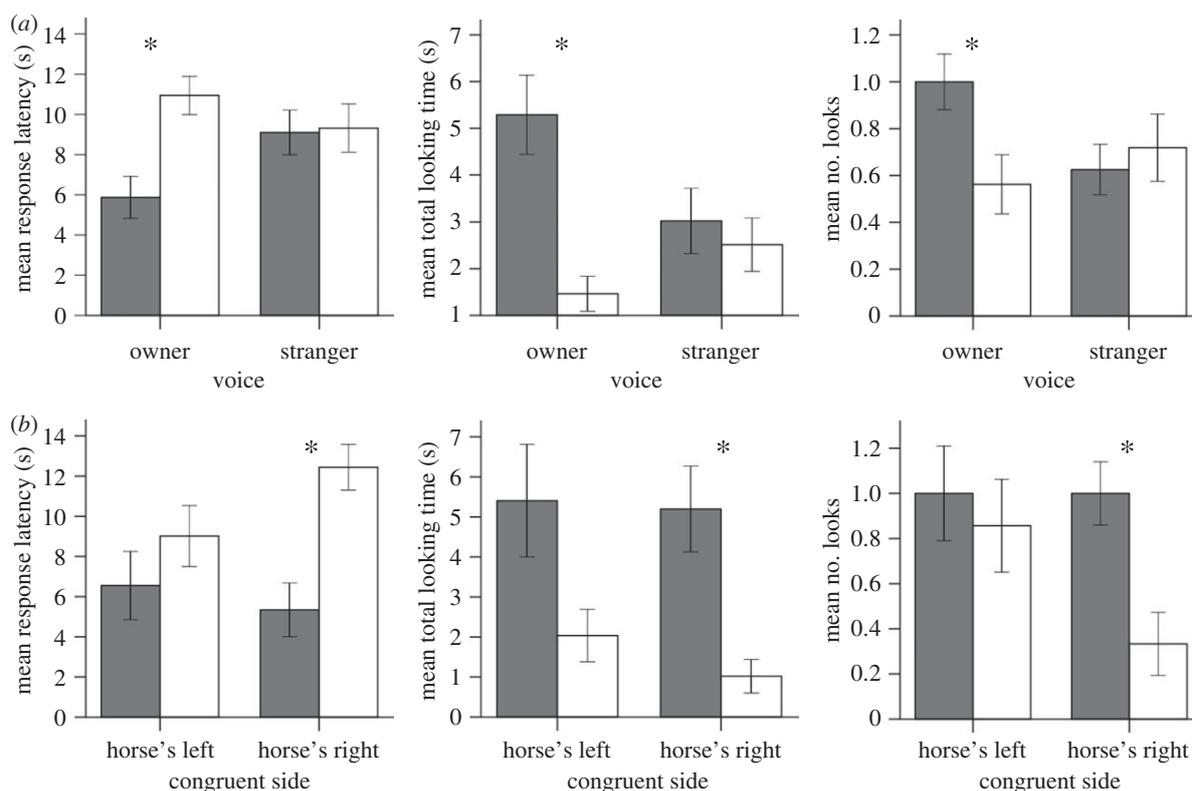


Figure 2. Experiment 1. Cross-modal discrimination of familiar human handlers and strangers. (a) Mean \pm s.e.m. for the behavioural responses of subjects to the matched (congruent person: grey bars) versus the mismatched (incongruent person: white bars) person during the trials in which the owner and the stranger were heard ($*p < 0.05$). (b) Mean \pm s.e.m. for the behavioural responses of subjects to the owner's voice during the trials in which the owner was on the right side of the horse versus when they were on the left ($*p < 0.05$).

2. RESULTS

(a) Experiment 1. Cross-modal discrimination of familiar human handlers from strangers

As predicted, horses were faster to look and spent more time overall looking at the congruent compared with the incongruent person (response latency: $F_{1,31} = 5.72$, $p = 0.023$; looking time: $F_{1,31} = 10.82$, $p = 0.003$). However, for total looking time, there was also a significant interaction between congruency and whether the voice heard was the owner or the stranger, suggesting that the time spent looking at the congruent and incongruent person was different depending on whether the voice was the owner's or the stranger's (looking time: $F_{1,31} = 5.27$, $p = 0.029$). This was also supported by a parallel trend in response latency that bordered on significance (response latency: $F_{1,31} = 4.058$, $p = 0.053$).

When the results are broken down further, it emerges that although the horses were faster to look and spent significantly more time looking at the owner when they heard the owner's voice (response latency: $t_{31} = -3.10$, $p < 0.01$; looking time: $t_{31} = 3.77$, $p < 0.005$), they did not respond faster and look for longer at the stranger when they heard the stranger's voice (response latency: $t_{31} = -0.13$, n.s.; looking time: $t_{31} = 0.62$, n.s.). In addition, horses tended to look more often at the owner compared with the stranger when they heard the owner's voice ($z = 1.94$, $p = 0.052$) but did not look more often at the stranger compared with the owner when they heard the stranger's voice ($z = 0.66$, $p = 0.51$). We therefore excluded from further analysis the trials in which the stranger's voice was played (figure 2a).

Within the responses to the owners' calls, we found that subjects were faster to respond, looked for longer and looked more often at the owner compared with the stranger when the owner was standing on the right side of the subject (response latency: $t_{17} = -3.27$, $p = 0.005$; looking time: $t_{17} = 3.19$, $p = 0.005$; number of looks: $z = 2.63$, $p = 0.023$). However, when the owner was standing on the left side of the subject, they did not look more quickly, for longer or more often at the owner compared with the stranger, although the difference in total looking time did border on significance (response latency: $t_{17} = -1.03$, $p = 0.32$; looking time: $t_{17} = 2.05$, $p = 0.061$; number of looks: $z = 0.37$, $p = 0.71$), indicating that horses were considerably poorer at making the match when the owner was not standing on the right side (figure 2b).

There were no significant differences in the direction of the first look (owner: 19 right turns, 10 left turns and 3 no responses; $n = 29$, $K = 19$, $p = 0.14$; stranger: 12 right turns, 14 left turns and 6 no responses; $n = 26$, $K = 12$, $p = 0.85$), or in total number of looks given in each direction (owner: 30 right turns, 20 left turns; $n = 50$, $K = 30$, $p = 0.20$; stranger: 21 right turns, 21 left turns; $n = 42$, $K = 21$, $p > 0.99$). The results of the unimodal visual control trials were indicative of a tendency to look to the right when presented with the visual stimuli ($t = -1.74$, $p = 0.096$) but, crucially, this was not stronger when the owner was on the right compared with the stranger (owner: $t_{11} = 1.30$, $p = 0.22$; stranger: $t_{11} = 1.14$, $p = 0.28$). Overall, the subjects did not look more at their owner than the stranger ($t_{23} = -0.094$, $p = 0.93$). In the auditory control trials, there was no overall difference in

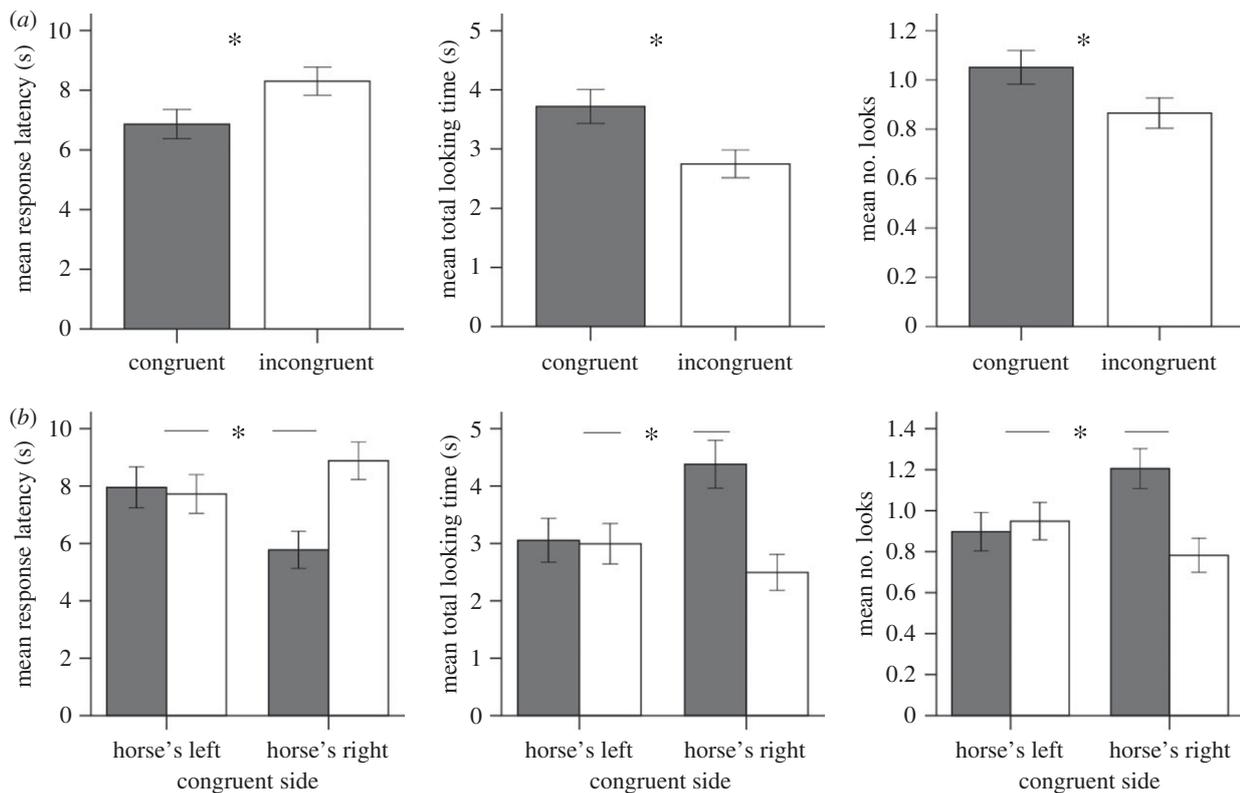


Figure 3. Experiment 2. Cross-modal individual recognition of familiar human handlers. (a) Mean \pm s.e.m. for the behavioural responses of subjects to the matched (congruent person: grey bars) versus the mismatched (incongruent person: white bars) person ($*p < 0.05$). (b) Mean \pm s.e.m. for the behavioural responses of subjects during trials in which the congruent person was on the right side of the horse versus when they were on the left ($*p < 0.05$).

the amount of time subjects spent looking to the left versus the right ($t_{23} = 1.57$, $p = 0.13$), nor was there any lateralized response to the stranger's voice alone ($t_{11} = -0.41$, $p = 0.69$). However, horses spent more time looking to the left than the right when they heard their owner's voice ($t_{11} = 2.77$, $p = 0.018$).

(b) Experiment 2. Cross-modal individual recognition of familiar human handlers

Subjects were faster to look, looked more often and looked for longer at the familiar person whose voice they had just heard compared with the familiar person whose voice they had not heard (response latency: $F_{1,312} = 6.815$, $p = 0.009$; looking time: $F_{1,312} = 11.164$, $p = 0.001$; number of looks: $F_{1,234} = 7.801$, $p = 0.006$; figure 3a), clearly indicating cross-modal recognition of specific individual handlers.

Across the three behavioural measures (response latency, looking time and number of looks), models containing the predictor variables sex and side on which the congruent person stood best explained the data (figure 3b). Specifically, horses found it easier to correctly match the handler with their voice if they were standing on their right side, and female subjects were better at recognizing their handlers (response latency: congruency \times side, $F_{1,312} = 6.593$, $p = 0.011$; congruency \times sex, $F_{1,312} = 3.942$, $p = 0.048$; looking time: congruency \times side, $F_{1,312} = 6.446$, $p = 0.012$; congruency \times sex, $F_{1,312} = 9.348$, $p = 0.002$; number of looks: congruency \times side, $F_{1,234} = 7.937$, $p = 0.005$; congruency \times sex, $F_{1,234} = 7.801$, $p = 0.006$). Additionally, there was a main effect

of age on response latency, with horses from the youngest age group (0–5 years) responding faster to the playbacks ($F_{4,312} = 3.408$, $p = 0.010$). There was also a main effect of trial part/playback on number of looks, with subjects giving more looks overall in response to the first voice heard in a trial compared with the second voice ($F_{1,234} = 4.699$; $p = 0.031$). Measures of familiarity, trial order and number of handlers did not significantly explain variance in recognition ability and were not included in the final models (see electronic supplementary material, §§ESM3–ESM5 for comparison of hypothesized models).

There were no significant differences in the direction the subjects first looked when they heard a familiar handler's voice (81 right turns, 61 left turns and 14 no responses; $n = 142$, $K = 81$, $p = 0.11$). There was, however, a significant difference in the total number of looks given in each direction, with subjects looking right more often than left (168 right turns, 131 left turns; $n = 299$, $K = 168$, $p = 0.037$).

3. DISCUSSION

In our first experiment, subjects looked faster, for longer and more often at their owner when they heard their owner's voice than when they heard a stranger's voice. In contrast, they did not look for longer at the stranger when they heard the unfamiliar voice. Thus, subjects were able to match a familiar voice with a familiar person but did not match an unfamiliar voice with an unfamiliar person. The matching of the owner with their voice does not reflect a spontaneous preference for looking towards the owner (as confirmed by the results of the visual control

trials); instead subjects actively associate the audio and visual stimuli. Whether the failure to match the stranger with the unknown voice reflects an inability to infer that an unknown voice comes from an unknown individual or they are not motivated to respond to a stranger calling their name is unclear. Crucially, in the second experiment, subjects proved able to match a specific familiar voice with a specific familiar human handler. This indicates that the sight of the handler activated a multimodal memory of that specific individual, allowing subjects to match the sight of that particular person with the sound of their voice. Furthermore, the ecologically valid methodology and the large number of handlers pairs presented suggest that horses use this recognition strategy naturally to identify numerous individual people in their day-to-day lives.

Our results indicate that individual recognition abilities in animals can be highly versatile, encoding individuals that are morphologically very different from the species itself. In humans, auditory, visual and cross-modal perceptual narrowing occurs over the course of the first year of life [7]. Animals also show a specialized ability to discriminate conspecific faces [29] but, similar to humans, animals with significant exposure to heterospecifics do demonstrate improved discriminatory abilities [30–32] and, in some cases, appear to process heterospecific identity information similarly to that of conspecifics [33,34]. Thus, it is clear that familiarity with heterospecifics can, to some extent, enable identity information from other species to be processed in similar ways to that of conspecific individuals. The current research shows that this adaptability not only extends to the most complex of social recognition tasks—that of cross-modal individual recognition—but also that such mechanisms can be employed in the recognition of familiar individuals in a highly divergent taxon.

Overall horses were significantly better at matching the familiar person with the sound of their voice when the correct person stood on their right side, with this asymmetry being more pronounced in the second experiment. Subjects also showed a trend towards preferring to orient to the right when the playbacks were heard. Although horses were free to look at the people with both eyes prior to and during the playback of vocalizations, the position of the people lay in their monocular field when subjects were facing forwards (figure 1). We therefore interpret the above-mentioned findings as the horses having been better at identifying the person initially with the right eye. In the visual control trials, there was no strongly lateralized response to viewing the stranger and the owner. This suggests that the hemispheric specialization for recognition reported here does not reflect a preference for observing people in general, or familiar people in particular, when they are on their right side. Moreover, when subjects heard the voices of their owners without the owner being present, subjects actually tended to look more to left, indicating a right hemisphere bias. Thus, it would appear that the left hemisphere dominance observed in the experimental trials is due to specific features of cross-modal recognition rather than of the presentation of the visual or auditory stimuli *per se*. There are a number of possible explanations as to why the left hemisphere could be central to cross-modal individual recognition in horses.

Although visual and cross-modal studies matching voice–face stimuli in humans and primates have tended to reveal an overall right hemisphere bias in processing, right hemispheric activation is followed by subsequent increases in left hemispheric activation [21,35–38]. It has been suggested that the right hemisphere is responsible for the initial processing of identity information and the assessment of novelty versus familiarity, whereas the left hemisphere may be involved in the more top-down retrieval of memories and details associated with specific individuals [39]. Crucially, our study involves the retrieval of information about genuine social partners associated with long-term relationships rather than responding to arbitrary voice–face pairs learnt in laboratory-based studies. Moreover, although the right hemisphere is involved in the processing of faces, it is widely accepted that the left hemisphere is involved in the categorization of other familiar objects and in matching with sample tasks [14,36,40]. In our study, horses were presented with the sight of the whole body of their owners, and thus may not have been using the face to assign identity. Such a task would require the use of mental templates based on previous experience and the implementation of established strategies of behaviour that are governed by the left hemisphere. This is consistent with previous research demonstrating a left hemisphere dominance in audio-visual matching tasks where stimuli are learnt during experimental trials [25,26]. What our results indicate is that processes governed by the left hemisphere are also central to the cross-modal matching of visual and auditory identity information from familiar individuals in a naturalistic setting. In addition, the valence hypothesis suggests that the left hemisphere governs pro-social behaviour and responses to stimuli regarded as positive, thus emotional responses to the stimuli may also have contributed to the left hemisphere bias [14,16,18]. It is entirely possible that the sight and sound of the handler produced a positive emotion in subjects causing them to process the information more efficiently with the left hemisphere. In addition, the cerebral asymmetry may reflect a desire for the horses to approach the familiar person when their name was called; indeed, a few of the horses did try to approach the handler upon hearing the familiar voice.

We also found a sex bias in recognition ability, with female horses being significantly better at recognizing particular handlers, or possibly more motivated to respond differentially to particular handlers when one of their voices was heard. Female horses are also better able to gauge the attentional state of humans, suggesting that female horses may generally be better able to interact with humans compared with male horses [41]. However, the authors are not aware of any research that indicates that mares bond more closely to people than gelded males, and a comparison of personality traits and trainability in young mares and geldings rated the sexes as equally ‘affable’ [42]. In the wild, horse social organization has been described as matriarchal and females form close social bonds, the quality of which has been shown to directly impact on the fitness of individuals [43]. Thus, the importance of forming good social relationships and of recognition of offspring may lead to enhanced socio-cognitive skills in females, although why this was not evident in previous research into the cross-modal recognition of conspecifics is unclear [3]. Sex differences in

motor laterality have also been found in horses, with males having a left side (right hemisphere) bias and females preferring to turn to the right side (left hemisphere) [44]. Thus, if cross-modal recognition requires mechanisms found primarily in the left hemisphere, then females may be predisposed to be more skilled at this task. Further research is required to address this finding.

4. CONCLUSIONS

These results demonstrate that domestic horses are capable of cross-modal individual recognition of familiar human handlers. This is the first evidence that an animal is capable of cross-modal recognition outside its own taxon and suggests that such an ability is likely to be widespread. It further demonstrates the adaptability of recognition systems—particularly because humans and horses are phylogenetically and morphologically so different. In addition, we also provide the first report of strong cortical asymmetry during the processing of cross-modal identity information in a naturalistic setting. By determining the prevalence and plasticity of individual recognition systems and providing insights into the hemispheric specializations on which these mechanisms are based across species, we broaden our understanding of the evolutionary history and neural bases of conceptual knowledge and social cognition.

5. METHODS

(a) *Study animals*

Subjects were naive to the experimental set-up and participated in only one of the following experiments (see electronic supplementary material, §ESM1 for subject details). In experiment 1, a total of 32 horses were recruited from five private livery yards in Norfolk and Sussex, UK. Ages ranged from 1.6 to 31 years ($x \pm \text{s.e.} = 12.83 \pm 1.47$). The horses comprised 14 gelded males and 18 mares. In experiment 2, a total of 40 horses were recruited from nine private livery yards and riding schools in Kent and Sussex, UK. Ten pairs of handlers were recruited, each sharing the care of four of the subjects. One horse was sold during the study, leaving a total of 39 subjects in the final analysis. Ages ranged from 2 to 25 years ($x \pm \text{s.e.} = 13.97 \pm 0.94$). The horses comprised 23 gelded males, 1 stallion and 15 mares.

(b) *Call acquisition*

Handlers were recorded calling the names of their horses in a stern voice. Handlers participating in experiment 1 were also recorded calling the name of an unfamiliar subject from another livery yard. In this way, the voices of handlers were used both as the familiar voice for their own horse and as the voice of a stranger for a different subject. See electronic supplementary material, §ESM2 for additional details.

(c) *Playback procedure*

A preferential looking paradigm was employed (see figure 1 for details of experimental set-up). Experiments were carried out in a familiar paddock or school during February 2009–July 2010. In experiment 1, the two people were a familiar handler and a person unknown to the subject. In experiment 2, the people were two highly familiar handlers. Each test consisted of a playback of one person saying the name of the subject twice, with a 1 s interval between the calls, followed by 15 s of silence (in which the response was monitored),

then the playback of the other person calling the subject's name twice, followed by another 15 s of silence. The speaker was disguised by either vegetation or showjumping wings. The order of the voices and the side that the people stood on were counterbalanced across trials. The person holding the subjects remained still, looking forward and did not interact with the horses. In experiment 1, subjects were given one trial in which the horses heard the two calls of each person once. In experiment 2, subjects were given two trials separated by at least one week, presenting the two calls from each person in a counterbalanced order across trials. The order in which the voices was played and the sides on which the handlers stood were counterbalanced across these trials. In order to ensure that the people participating in the study were not giving any unintentional cues when they heard the playback of their voices, 16 subjects in experiment 2 received one trial in which the handlers wore small earpiece headphones and listened to white noise from handheld MP3 players and one trial without headphones. This white noise masked the sound of the playbacks. The responses of the horses to the trials with and without the headphones were compared to ensure that the horses' recognition ability was not significantly improved when the handlers could also hear the playbacks. See electronic supplementary material, §ESM3 for further details of procedure and equipment.

Subsequent to the experimental trials, additional unimodal control trials were given to 12 subjects from experiment 1 to determine whether there were any lateralized responses to the visual and auditory stimuli when they were presented separately. The experimental set-up was the same as in the experimental trials. Subjects were given two visual control trials: one in which the owner stood on the right of the subject and the stranger on the left, and one in which the stranger stood on the right and the owner on the left. No playback was heard and their responses were recorded for 15 s. During the auditory control trials, subjects heard the same sequence of calls as during the experimental trials (one person calling their name twice followed by 15 s of silence followed by the other person calling their name twice) but there were no people standing either side of the speaker. The order of control trials was counterbalanced across subjects.

(d) *Behavioural and statistical analysis*

The total amount of time spent looking at each of the people, the speaker and elsewhere was recorded. Horses have laterally placed eyes with a small (60–80°) binocular field of vision and almost complete (80–90%) decussation of the optic nerves, suggesting that behavioural asymmetries reflect asymmetries in hemispheric activation [45,46]. A look was thus defined as being at either of the people if the horse's nose was between 10° and 90° from the centre line and was recorded as being in the direction of the speaker if the horse's nose was facing a point within 10° to the left or the right of the speaker; a look was recorded as 'elsewhere' if it was over 90° from the centre point (figure 1). See electronic supplementary material, §ESM4 for additional details.

Response latencies, the total amount of looking time and the number of looks given to the congruent and incongruent person were recorded as dependent variables (DVs). For subjects that did not look at one (or both) of the people during a playback, a maximum time of 15 s was assigned as the response latency towards that particular individual. The videos were coded blind in random order by L.P. For

experiment 1, the total looking time and number of looks for 22/32 subjects (69%) were scored by a second coder providing inter-observer reliabilities of 0.704 ($p < 0.0001$) and 0.703 ($p < 0.0001$), respectively. For experiment 2, the videos of 23/39 subjects (59%) were second coded, providing an inter-observer reliability of 0.674 ($p < 0.0001$) for response latency, 0.723 ($p < 0.0001$) for total looking time and 0.627 ($p < 0.0001$) for number of looks (measured by Spearman's rho correlation).

(i) Experiment 1

For the DVs of response latency and total looking time, initial 2×2 repeated-measures ANOVAs were conducted with congruency (congruent/incongruent) and person heard (owner/stranger) as within subject factors. Additional repeated-measures t -tests were performed to assess the differences in latency and looking time towards the incongruent and congruent person for the owner's voice and the stranger's voice separately (with the significance value adjusted using the Bonferroni correction $0.05/n$, where n is the number of hypotheses tested). The DV number of looks was not considered to be a continuous variable owing to the limited distribution of responses (range 0–2); thus, non-parametric Wilcoxon signed-ranks tests were used to compare the differences in number of looks given to the incongruent and congruent person for the owner's voice and the stranger's voice separately. The responses to the owner's voice were then divided into those trials where the owners stood on the left of the subject and those in which they stood on the right. Individual t -tests were then performed on the response latency and total time data for the two groups to assess the effects of side on performance. The effect of side on number of looks given was also analysed using Wilcoxon signed-ranks tests. Repeated-measures t -tests were performed to assess the differences in looking time towards the left and right for the four control conditions (visual control with owner on the right, visual control with the stranger on the right, auditory control with owner's voice, auditory control with stranger's voice). A t -test was also performed to determine whether the subjects showed any preference for looking at their owner or the stranger across the visual control trials.

(ii) Experiment 2

Owing to the larger sample size and multiple trial protocol employed in experiment 2, each DV could be analysed in an individual linear-mixed model (with a scaled identity covariance structure, using a maximum-likelihood estimation). The fit of potential models was determined using Akaike's information criterion corrected for small samples (AICc) and ranked using Δ AICc to determine the best-fit model. All factors listed later were included in a global model, and factors with little or no predictive value were systematically removed to produce the final, best model (see electronic supplementary material, §ESM5 for tables of the global models and top eight hypothesized models).

In these models, handler's voice was nested within subject as a random factor. The main effect assessed was congruency (i.e. whether there were any significant differences in the response latency, number of looks and total time spent looking at the matched versus mismatched person). The following potential predictor variables were also included as fixed factors: horse sex, age (grouped variable, GV), side on which handler stood, number of handlers, number of years the horse had known each handler (GV), estimated number of hours a day spent with each handler (GV) and total exposure time (GV; calculated

by the number of hours a day spent with the person \times years known). Order effects were assessed by the repeated measure of trial order (1 and 2) with playback (1 and 2) nested within trial. To assess the effect of these additional potential predictor variables on the ability to distinguish between congruent and incongruent people, each factor was included as an interaction variable with congruency (factor \times congruency).

Within both experiments, the direction of the initial look and the total number of times subjects looked in each direction when they heard the two voices were compared using two-tailed binomial probability tests to determine whether there was a group-level orienting asymmetry when hearing familiar or unfamiliar voices. To ensure that handlers were not unintentionally cuing the horses during the trials, the recognition ability of the subjects during experiment 2 when handlers could hear the playbacks and when they could not were compared using a $2 \times 2 \times 2$ repeated-measures ANOVA with playback (voice 1/voice 2), congruency (congruent/incongruent) and trial type (with headphones/without headphones) as within-subject factors. No significant differences between the trials in which handlers could hear the playbacks and those in which they could not were found (response latency: $F_{1,15} = 0.663$, $p = 0.43$; total looking time: $F_{1,15} = 2.391$, $p = 0.14$; number of looks: $F_{1,15} = 1.901$, $p = 0.19$), and the data were therefore combined in all analyses.

We are indebted to Emily Bacon and Jessica Hooker for their contribution to data collection and second coding, and to the owners and handlers of the horses for their support and willingness to facilitate this project. We are also grateful to Prof. Lesley Rogers and Prof. Richard Andrew for their helpful discussions. This study complied with the University of Sussex regulations on the use of animals and was approved by the School of Psychology ethics committee. This work is supported by a quota studentship from the BBSRC (to L.P., supervised by K.M.).

REFERENCES

- Ettlinger, G. 1967 Analysis of cross-modal effects and their relationship to language. In *Brain mechanisms underlying speech* (ed. F. L. Darley), pp. 53–60. New York, NY: Grune & Stratton.
- Seyfarth, R. M. & Cheney, D. L. 2009 Seeing who we hear and hearing who we see. *Proc. Natl Acad. Sci. USA* **106**, 669–670. (doi:10.1073/pnas.0811894106)
- Proops, L., McComb, K. & Reby, D. 2009 Cross-modal individual recognition in domestic horses (*Equus caballus*). *Proc. Natl Acad. Sci. USA* **106**, 947–951. (doi:10.1073/pnas.0809127105)
- Sliwa, J., Duhamel, J. R., Pascalis, O. & Wirth, S. 2011 Spontaneous voice–face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proc. Natl Acad. Sci. USA* **108**, 1735–1740. (doi:10.1073/pnas.1008169108)
- Bovet, D. & Deputte, B. L. 2009 Matching vocalizations to faces of familiar conspecifics in grey-cheeked Mangabeys (*Lophocebus albigena*). *Folia Primatol.* **80**, 220–232. (doi:10.1159/000235688)
- Kondo, N., Izawa, E.-I. & Watanabe, S. 2012 Crows cross-modally recognize group members but not non-group members. *Proc. R. Soc. B* **279**, 1937–1942. (doi:10.1098/rspb.2011.2419)
- Scott, L. S., Pascalis, O. & Nelson, C. A. 2007 A domain-general theory of the development of perceptual discrimination. *Curr. Dir. Psychol.* **16**, 197–201. (doi:10.1111/j.1467-8721.2007.00503.x)
- Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M. & Nelson, C. A. 2005

- Plasticity of face processing in infancy. *Proc. Natl Acad. Sci. USA* **102**, 5297–5300. (doi:10.1073/pnas.0406627102).
- 9 Adachi, I., Kuwahata, H. & Fujita, K. 2007 Dogs recall their owner's face upon hearing the owner's voice. *Anim. Cogn.* **10**, 17–21. (doi:10.1007/s10071-006-0025-8)
 - 10 Stone, S. M. 2009 Human facial discrimination in horses: can they tell us apart? *Anim. Cogn.* **13**, 51–61. (doi:10.1007/s10071-009-0244-x)
 - 11 Krueger, K., Flauger, B., Farmer, K. & Maros, K. 2011 Horses (*Equus caballus*) use human local enhancement cues and adjust to human attention. *Anim. Cogn.* **14**, 187–201. (doi:10.1007/s10071-010-0352-7)
 - 12 Taubert, J. 2009 Chimpanzee faces are 'special' to humans. *Perception* **38**, 343–356. (doi:10.1068/p6254)
 - 13 MacNeilage, P. F., Rogers, L. J. & Vallortigara, G. 2009 Origins of the left and right brain. *Sci. Am.* **301**, 60–67. (doi:10.1038/scientificamerican0709-60)
 - 14 Rogers, L. J. 2010 Relevance of brain and behavioural lateralization to animal welfare. *Appl. Anim. Behav. Sci.* **127**, 1–11. (doi:10.1016/j.applanim.2010.06.008)
 - 15 Rogers, L. J. 2002 Lateralization in vertebrates: its early evolution, general pattern, and development. *Adv. Study Behav.* **31**, 107–161. (doi:10.1016/s0065-3454(02)80007-9)
 - 16 Davidson, R. J. 1992 Emotion and affective style: hemispheric substrates. *Psychol. Sci.* **3**, 39–43. (doi:10.1111/j.1467-9280.1992.tb00254.x)
 - 17 Ley, R. G. & Bryden, M. P. 1979 Hemispheric differences in processing emotions and faces. *Brain Lang.* **7**, 127–138. (doi:10.1016/0093-934X(79)90010-5)
 - 18 Workman, L., Peters, S. & Taylor, S. 2000 Lateralisation of perceptual processing of pro- and anti-social emotions displayed in chimeric faces. *Laterality* **5**, 237–249. (doi:10.1080/713754378)
 - 19 Daisley, J. N., Mascalonzi, E., Rosa-Salva, O., Rugani, R. & Regolin, L. 2009 Lateralization of social cognition in the domestic chicken (*Gallus gallus*). *Phil. Trans. R. Soc. B* **364**, 965–981. (doi:10.1098/rstb.2008.0229)
 - 20 Farmer, K., Krueger, K. & Byrne, R. W. 2010 Visual laterality in the domestic horse (*Equus caballus*) interacting with humans. *Anim. Cogn.* **13**, 229–238. (doi:10.1007/s10071-009-0260-x)
 - 21 Campanella, S. & Belin, P. 2007 Integrating face and voice in person perception. *Trends Cogn. Sci.* **11**, 535–543. (doi:10.1016/j.tics.2007.10.001)
 - 22 von Kriegstein, K. & Giraud, A. L. 2006 Implicit multisensory associations influence voice recognition. *PLoS Biol.* **4**, 1809–1820. (doi:10.1371/journal.pbio.0040326)
 - 23 Joassin, F., Pesenti, M., Maurage, P., Verreckett, E., Bruyer, R. & Campanella, S. 2009 Cross-modal interactions between human faces and voices involved in person recognition. *Cortex* **47**, 367–376. (doi:10.1016/j.cortex.2010.03.003)
 - 24 Hertrich, I., Dietrich, S. & Ackermann, H. 2011 Cross-modal interactions during perception of audiovisual speech and nonspeech signals: an fMRI study. *J. Cogn. Neurosci.* **23**, 221–237. (doi:10.1162/jocn.2010.21421)
 - 25 Gaffan, D. & Harrison, S. 1991 Auditory–visual associations, hemispheric specialization and temporal–frontal interaction in the rhesus monkey. *Brain* **114**, 2133–2144. (doi:10.1093/brain/114.5.2133)
 - 26 Delfour, F. & Marten, K. 2006 Lateralized visual behavior in bottlenose dolphins (*Tursiops truncatus*) performing audio-visual tasks: the right visual field advantage. *Behav. Processes* **71**, 41–50. (doi:10.1016/j.beproc.2005.09.005)
 - 27 Spelke, E. S. 1985 Preferential looking methods as tools for the study of cognition in infancy. In *Measurement of audition and vision in the first year of postnatal life* (eds G. Gottlieb & N. Krasnegor), pp. 323–363. Norwood, NJ: Ablex.
 - 28 Ghazanfar, A. A., Tureson, H. K., Maier, J. X., van Dinther, R., Patterson, R. D. & Logothetis, N. K. 2007 Vocal-tract resonances as indexical cues in rhesus monkeys. *Curr. Biol.* **17**, 425–430. (doi:10.1016/j.cub.2007.01.029)
 - 29 Leopold, D. A. & Rhodes, G. 2010 A comparative view of face perception. *J. Comp. Psychol.* **124**, 233–251. (doi:10.1037/a0019460)
 - 30 Adachi, I., Kuwahata, H., Fujita, K., Tomonaga, M. & Matsuzawa, T. 2009 Plasticity of ability to form cross-modal representations in infant Japanese macaques. *Dev. Sci.* **12**, 446–452. (doi:10.1111/j.1467-7687.2008.00780.x)
 - 31 Racca, A., Amadei, E., Ligout, S., Guo, K., Meints, K. & Mills, D. 2010 Discrimination of human and dog faces and inversion responses in domestic dogs (*Canis familiaris*). *Anim. Cogn.* **13**, 525–533. (doi:10.1007/s10071-009-0303-3)
 - 32 Sugita, Y. 2008 Face perception in monkeys reared with no exposure to faces. *Proc. Natl Acad. Sci. USA* **105**, 394–398. (doi:10.1073/pnas.0706079105)
 - 33 Kendrick, K. M. 2006 Brain asymmetries for face recognition and emotion control in sheep. *Cortex* **42**, 96–98. (doi:10.1016/S0010-9452(08)70328-9)
 - 34 Guo, K., Meints, K., Hall, C., Hall, S. & Mills, D. 2009 Left gaze bias in humans, rhesus monkeys and domestic dogs. *Anim. Cogn.* **12**, 409–418. (doi:10.1007/s10071-008-0199-3)
 - 35 Belin, P. 2006 Voice processing in human and non-human primates. *Phil. Trans. R. Soc. B* **361**, 2091–2107. (doi:10.1098/rstb.2006.1933)
 - 36 Brancucci, A., Lucci, G., Mazzatenta, A. & Tommasi, L. 2009 Asymmetries of the human social brain in the visual, auditory and chemical modalities. *Phil. Trans. R. Soc. B* **364**, 895–914. (doi:10.1098/rstb.2008.0279)
 - 37 Peirce, J. W. & Kendrick, K. M. 2002 Functional asymmetry in sheep temporal cortex. *Neuroreport* **13**, 2395–2399. (doi:10.1097/01.wnr.0000048023.74602.60)
 - 38 Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., Landis, T. & Schomer, D. L. 1997 Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport* **8**, 2749–2754. (doi:10.1097/00001756-199708180-00021)
 - 39 Rhodes, G. 1985 Lateralized processes in face recognition. *Br. J. Psychol.* **76**, 249–271. (doi:10.1111/j.2044-8295.1985.tb01949.x)
 - 40 Vallortigara, G., Snyder, A., Kaplan, G., Bateson, P., Clayton, N. S. & Rogers, L. J. 2008 Are animals autistic savants? *PLoS Biol.* **6**, 208–214. (doi:10.1371/journal.pbio.0060042)
 - 41 Proops, L. & McComb, K. 2010 Attributing attention: the use of human-given cues by domestic horses (*Equus caballus*). *Anim. Cogn.* **13**, 197–205. (doi:10.1007/s10071-009-0257-5)
 - 42 Duberstein, K. J. & Gilkeson, J. A. 2010 Determination of sex differences in personality and trainability of yearling horses utilizing a handler questionnaire. *Appl. Anim. Behav. Sci.* **128**, 57–63. (doi:10.1016/j.applanim.2010.09.012)
 - 43 Cameron, E. Z., Setsaas, T. H. & Linklater, W. L. 2009 Social bonds between unrelated females increase reproductive success in feral horses. *Proc. Natl Acad. Sci. USA* **106**, 13 850–13 853. (doi:10.1073/pnas.0900639106)
 - 44 Murphy, J., Sutherland, A. & Arkins, S. 2005 Idiosyncratic motor laterality in the horse. *Appl. Anim. Behav. Sci.* **91**, 297–310. (doi:10.1016/j.applanim.2004.11.001)
 - 45 Brooks, D. E., Komaromy, A. M. & Kallberg, M. E. 1999 Comparative retinal ganglion cell and optic nerve morphology. *Vet. Ophthalmol.* **2**, 3–11. (doi:10.1046/j.1463-5224.1999.00047.x)
 - 46 Harman, A. M., Moore, S., Hoskins, R. & Keller, P. 1999 Horse vision and an explanation for the visual behaviour originally explained by the 'ramp retina'. *Equine Vet. J.* **31**, 384–390. (doi:10.1111/j.2042-3306.1999.tb03837.x)