Animal Behavior and Cognition
ABC 2024, 11(3): 236-251
DOI: https://doi.org/10.26451/abc.11.03.01.2024

Horses Look Longer at Skilled Humans than Unskilled Ones Despite Similar Outcomes in an Object Manipulation Task

Takuto Sugimoto* **and Satoshi Hirata**

Wildlife Research Center, Kyoto University, Japan

*Corresponding author (Email: sugimoto.takuto.88e@st.kyoto-u.ac.jp)

Citation – Sugimoto, T., & Hirata, S. (2024). Horses look longer at skilled humans than unskilled ones despite similar outcomes in an object manipulation task. *Animal Behavior and Cognition, 11*(3), 236-251. https://doi.org/10.26451/abc.11.03.01.2024

Abstract – Evaluating others based on skillfulness and prosociality is an important ability in choosing a cooperative partner. Such abilities have been studied in nonhuman species such as primates and dogs; however, few studies have examined whether horses could evaluate humans based on their skillfulness. Here, we investigated whether horses (*Equus caballus*) could evaluate humans based on their skillfulness in a situation related to their own interests (food). We showed horses a pair of experimenters, one that could open the lid of a container (skilled actor) and another that could not open the lid of the container (unskilled actor). The unskilled actor 'accidentally' dropped the container during the demonstration phase, causing the lid to open, allowing him to remove the object and thus achieve the same outcome as the skilled actor. The experimental trials involved two conditions, one in which the container had food inside, and one in which it was empty. This was because we thought the horses would have an expectation that the skilled actor would give them food after opening the container, so we predicted that horses would look longer at the skilled actor than the unskilled actor only in the food condition. We measured the gaze duration of horses toward each experimenter. Horses looked longer at the skilled actor overall and looked at the two actors longer in the food condition than in the non-food condition, and there was no interaction between role and condition. Post-hoc tests showed that the horses looked at the skilled actor longer than the unskilled actor in both conditions. They also looked at the skilled actor for longer in the food condition compared with the no-food condition. These results suggest that the horses were able to visually discriminate between humans based on their skillfulness.

Keywords – Horses, Social evaluation, Third-party evaluation, Social cognition, Horse-human relationships, Skillfulness

---

Humans form large-scale cooperative societies, in which they evaluate others in their daily lives and decide to whom they will behave altruistically (Wu et al., 2016). The ability to evaluate others is therefore an important aspect of maintaining these societies. It has been shown that from infancy, when seeking help to solve a problem, humans identify and seek help from the person who is the most competent solver (Warneken et al., 2006). Furthermore, young children can evaluate the prosociality of others, not only through direct experience based on direct interactions with others, but also through indirect experience by observing interactions between third parties (Herrmann et al., 2013; Vaish et al., 2010).

According to the stereotype content model (SCM), there are two fundamental dimensions of social evaluation in humans: warmth and competence (Fiske et al., 2002). In the SCM, the warmth dimension includes characteristics such as friendliness, helpfulness, and trustworthiness, while the competence dimension includes characteristics such as intelligence, skill, and creativity. Warmth reflects others' intentions and competence reflects others' ability to act on their intentions. When we meet other people, we immediately need to know their intentions toward us for survival: "Are they friend or foe?" After inferring their intentions, we need to know whether others can act on their intentions: "Are they capable of it?" Therefore, from an evolutionary perspective, warmth is primary and judged faster than competence because whether others' intentions are good or bad is more important to survival than whether they have the ability to act on their intentions (Fiske, 2015, 2018; Fiske et al., 2006).

Moreover, Kervyn et al. (2010) suggested that warmth and competence judgments tend to be negatively correlated in humans when comparing two individuals, and called this the compensation effect. In other words, if one person is judged more positively on one dimension, the other person being compared tends to be judged higher on the other dimension. For example, if one person is judged to have higher warmth, he/she tends to be judged to have lower competence, while another person tends to be judged to have lower warmth and higher competence. The authors proposed a hypothesis as to why warmth and competence would generally not be seen as being positively correlated: humans may prefer a balanced evaluation of two people or groups to justify the existing social structure. In other words, humans may have a motivation to perceive a system as fair and balanced by thinking that two people or groups have both strengths and weaknesses to ensure that both have an equal amount of positive characteristics. In sum, warmth and competence judgments are two basic dimensions and are deeply intertwined in human social evaluation.

To clarify the evolutionary origin of the ability to evaluate others, researchers have been studying the ability to evaluate others in animals, mainly nonhuman primates that are evolutionarily related to humans. Previous studies have shown that some species of nonhuman primates, like humans, also have the ability to evaluate conspecifics based on their skillfulness (Melis et al., 2006; Ottoni et al., 2005; Stammbach, 1988). For example, chimpanzees (*Pan troglodytes*) have been shown to choose a more efficient partner when given a choice between two partners in a task in which they must simultaneously pull a string in cooperation with another individual to obtain food (Melis et al., 2006).

Furthermore, primates have been shown to have the ability to evaluate not only conspecifics, but also humans, based on their prosociality (chimpanzees: Herrmann et al., 2013; Russell et al., 2008; Subiaul et al., 2008; orangutans, *Pongo pygmaeus*: Hermann et al., 2013; Russell et al., 2008; common marmosets, *Callithrix jacchus*: Kawai et al., 2014). Social evaluation studies in nonhuman animals mainly use human actors to rigorously control for many factors, so it is important to examine whether nonhuman animals can evaluate humans in a way that is comparable with the results of other social evaluation studies in nonhuman animals. For example, Herrmann et al. (2013) exposed orangutans to the following direct experiences with two human actors. One actor gave the orangutan food (friendly actor), while the other actor (mean actor) interfered with the food exchange between the orangutan and the friendly actor. When the orangutans were allowed to choose between the food offered by the two actors, they accepted food from the friendly actor more often. In the same study, chimpanzees and orangutans were also shown the following indirect interactions between three human actors. The friendly actor gave food to a food-seeking actor (E1), while the mean actor interfered with the food exchange between the friendly actor and E1. Subsequently, when orangutans were allowed to choose between the food offered by the friendly and mean actors, they chose to receive food from the friendly actor more often (Herrmann et al., 2013). Thus, it has been shown that apes are able to evaluate humans to some extent based on their prosociality. It should be noted that because these experiments used food exchanges, there is a possibility that the subjects learned the association between the actor and the food and simply chose the actor who was likely to give them food.

A few studies have investigated the ability in nonhuman primates to evaluate humans using scenes that do not involve food exchange (tufted capuchin monkeys, *Cebus apella*: Anderson et al., 2013; bonobos, *Pan paniscus*: Krupenye & Hare, 2018). For example, Krupenye and Hare (2018) used the out-of-reach paradigm. Bonobos watched the following interactions between three humans: when an experimenter (E1) dropped a stuffed animal, one of the two actors (helper) picked it up and returned it to E1, but the other person (hinderer) interrupted the exchange between E1 and the helper and took the stuffed animal away. After watching the interactions, the bonobos were allowed to choose between the food offered by the helper and the hinderer. As a result, they chose to receive food from the hinderer more often. Because antisocial behavior is often exhibited to maintain dominance, this preference for the hinderer may actually be a preference for the more dominant experimenter. To test this dominance hypothesis, the authors performed a further experiment. Bonobos watched the following dominance interaction in a video: a dominant agent repeatedly displaced a subordinate agent as a result of competition for a central location. After watching the interactions in the video, the bonobos were allowed to choose between paper cutouts of the agents in four trials. The bonobos chose the dominant agent more often. However, eight of the 24 bonobos chose the dominant agent in almost all trials, so the preference for the dominant agents may have been affected by those eight subjects. Consequently, the authors suggest that the bonobos' preference for the hinderer may be partially based on dominance, but that this is not the sole factor influencing their preference. Therefore, even in no-food conditions, the authors demonstrated that apes are able to evaluate humans to some extent based on their prosociality.

In recent years, the ability to evaluate humans has also been studied in domestic dogs (*Canis familiaris*). Since dogs were domesticated, they have developed close relationships with humans as partners and are thought to have acquired advanced social cognitive abilities in their interactions with humans during domestication (Hare & Tomasello, 2005). Recent studies have also shown that dogs can evaluate humans based on their prosociality, not only in situations involving food exchange (Freidin et al., 2013; Kundey et al., 2011; Nitzschner et al., 2012, 2014; Marshall-Pescini et al., 2011), but also in no-food conditions (Chijiiwa et al., 2015). Chijiiwa et al. (2015) showed dogs a helping situation in which one actor (helper) helped the dog's owner (a neutral person) to open a container to remove an object inside, and the other (non-helper) refused to help the owner by turning away. After watching the interactions, the dogs could choose to receive food from either one of the actors or the owner. The researchers found that the dogs avoided receiving food from the non-helper, but showed no preference for the helper. Thus, it has been shown that dogs can evaluate humans based on their prosociality. However, as Jim et al. (2020) point out, it should be noted that the dogs might have remembered where the owner had removed the object and simply chose the location (i.e., the owner), because the positions of the actor and the owner were not swapped before the choice phase. Therefore, whether dogs can evaluate humans based on their prosociality in no-food conditions is still debatable.

Dogs can evaluate humans not only based on their prosociality, but also on their skillfulness (Chijiiwa et al., 2022; Horn et al., 2012). For example, when dogs were simultaneously shown two people attempting to open containers with food inside, and one was successful while the other was not, female dogs looked longer at and chose the person who could open the container (Chijiiwa et al., 2022). This suggests that not only species genetically related to humans (nonhuman primates such as chimpanzees), but also those that are not (domesticated dogs) can evaluate humans based on their skillfulness.

The ability to evaluate humans has also been studied in horses (*Equus caballus*). Horses live in groups, form complex social hierarchies, and are highly social animals (Ringhofer et al., 2017a). Horses were domesticated approximately 5,500 years ago (Outram et al., 2009) and have had a close relationship with humans since then. Several studies have shown that horses have an excellent ability to communicate with humans (see Jardat & Lansade, 2022 for a review). For example, horses can follow human pointing and find hidden food (Maros et al., 2008; Proops et al., 2010). They can also change their behavior toward a novel object according to the emotions that humans show toward the novel object (Schrimpf et al., 2020). In addition, they are sensitive to the attentional state of humans and have been shown to request food from a human actor who knew where food was hidden rather than from one who did not (Trösch et al., 2019b). They can also match pictures of humans expressing emotions with corresponding vocalizations (Trösch et al., 2019a). In other words, horses can recognize human emotions cross-modally through visual and vocal cues.

Horses can also evaluate humans based on their prosociality through direct experience with two human actors. After being trained by a whip-bearing or food-bearing human actor, horses attempted to touch the food-bearing human actor more (Sankey et al., 2010). In addition, Trösch et al. (2020) investigated whether horses can also evaluate humans based on their prosociality after the following indirect experiences of observing interactions between another horse and a human actor. After watching videos of positive (grooming) and negative (trying to spray) interactions between a human actor and another horse, the horses were allowed to choose between a positive and a negative actor. As a result, the horses touched the negative actor more. The authors suggest that this result could be interpreted as an appeasement behavior; in another study, a third horse was observed to interact in affiliation with an aggressor horse in post-conflict periods (Cozzi et al., 2010). In other words, this result may be due to the dominance preference as shown with bonobos (Krupenye & Hare, 2018). Thus, it remains open to debate whether horses can evaluate humans based on their prosociality through indirect experience.

However, to our knowledge, there is only one study investigating whether horses can evaluate humans based on their skillfulness. In a previous study, we (Sugimoto & Hirata, 2022) investigated whether horses could discriminate between successful and unsuccessful human solvers in a situation related to the horses' own interests (food). The horses were shown that one experimenter could open a container with an object inside, while the other could not. The two experimenters then stood side by side in front of the horses and tried to open a container with/without food (food condition/no-food condition), and the horses' gaze duration was measured. The results showed that the horses looked longer at the person who could open the container in the food condition, and this proportion was greater than in the no-food condition. This suggests that horses can evaluate humans based on their skillfulness. However, in that study the end state in the demonstration phase differed between conditions, so it is possible that the horses simply preferred the person who removed the object.

Here, we investigated whether horses can evaluate humans based on their skillfulness in an unsolvable situation that potentially involved the horses' own interests (food). We used essentially the same methodology as our previous study (Sugimoto & Hirata, 2022), but made modifications to ensure that the final state of the two actors in the demonstration phase was the same. In the present study, we showed horses demonstrations by two actors (demonstration phase): one of the experimenters was able to open the container and remove an object inside (skilled actor), while the other was unable to open the container but 'accidentally' dropped it (i.e., he deliberately dropped it, but in a way that would appear unintentional to the onlooking horse), causing the lid to open and the object inside to be removed (unskilled actor). In each demonstration, the containers were accessible only to the experimenters. A test phase then took place, in which the two actors were trying to open containers with/without food at the same time (food condition/no-food condition). We measured the gaze duration of the horses toward each experimenter during the test phase. This is because looking behavior has been widely used in other studies in tasks involving discriminating between two people (Ringhofer & Yamamoto, 2017b; Trösch et al., 2019b; Lampe & Andre, 2012; Proops & McComb, 2012). If horses can evaluate humans based on their skillfulness in an unsolvable situation related to their own interests, we predicted that in the food condition, horses would look longer at the skilled actor than at the unskilled actor. This was because we thought the horses would have an expectation that the skilled actor would give them food after opening the container. In the no-food condition, the experimenters' activities were not related to the horse's own interests; therefore, from the horses' perspective it would not be a problem-solving situation, so there would be no need to consider the human's skillfulness, and the horses would look at each experimenter randomly. Therefore, we predicted that the horses would look at the skilled actor longer in the food condition than in the no-food condition. In other words, we predicted that the horses would look longer at the skilled actor versus the unskilled actor only in the condition involving food.

Methods

Ethics Statement

The research protocol was approved by the Animal Experimentation Committee of the Wildlife Research Center, Kyoto University, Japan (WRC-2023-008A).

Subjects

Thirty-one horses (21 males and 10 females) participated in this study. Fourteen horses were kept at the Equestrian Club of Kyoto University, Japan, and the rest were kept at the Equestrian Club of Kyoto Sangyo University, Japan. All males were geldings (i.e., they had been castrated). The horses' ages ranged from 3 years to 22 years, with a mean age of 12.81 years (SD = 5.53) (see Table 1). These horses were first trained for racing, and then for equestrian competitions. At the time of the present study, they were engaged in daily training with the equestrian club members for equestrian competitions. The horses were also able to observe the daily actions of the club members' care-taking activities, including opening bags containing food.

Table 1

Breed, Sex, and Age of Subject Horses

Breed	Sex[a]	Age (years)
Thoroughbred	M	22
Thoroughbred	M	21
Thoroughbred	M	19
Thoroughbred	M	18
Thoroughbred	M	17
Thoroughbred	M	16
Thoroughbred	M	16
Thoroughbred	M	15
Thoroughbred	M	14
Thoroughbred	M	13
Thoroughbred	M	11
Thoroughbred	M	7
Thoroughbred	M	7
Thoroughbred	M	7
Thoroughbred	M	6
Thoroughbred	M	6
Thoroughbred	M	6
Thoroughbred	F	19
Thoroughbred	F	16
Thoroughbred	F	13
Thoroughbred[b]	F	9
Thoroughbred[b]	F	9
Thoroughbred[b]	F	8
Thoroughbred[b]	F	7
Thoroughbred[b]	F	7
Thoroughbred[b]	F	3
Nihon Sport Horse	M	19
Nihon Sport Horse	M	11
Nihon Sport Horse	F	20
PZHK	M	20
KWPN	M	15

Note. [a]M, stallion; F, mare; KWPN, Koninklijk Warmbloed Paardenstamboek Nederland/Royal Warmblood Studbook of the Netherlands registered Dutch Warmblood; PZHK, Polish Horse Breeders Association. [b]We excluded these horses from our analyses because of experimental errors.

Apparatus

We used four transparent cylindrical containers (10 cm diameter, 15 cm height) with a lid and two objects (6 cm diameter rolls of vinyl tape). The experiments took place in the hoof-washing area of the Kyoto University Equestrian Club and in the corridor of the stable of the Kyoto Sangyo University Equestrian Club. Two male experimenters, who were strangers to the horses, participated in the experiment. The space between the two experimenters was approximately 80 cm and approximately 1 m away from the horses. An assistant, a member of the equestrian club, also participated in the experiment to handle the horses and was unaware of the purpose of the experiment (Figure 1). We recorded horses' behavior using a video camera (Sony HDR-CX470) mounted on a tripod and positioned between the two actors.

Figure 1

Experimental Set-up During the Test Phase

Procedure

An assistant attached a halter to the horses to guide and hold them, led them to the test site, and had them stand in a stationary position approximately 1 m away from the experimenters. During this time, the assistant could interact with the horses and touch them to calm them down. After the horse was settled, a trial was initiated at the cue of the assistant. The trial consisted of a demonstration phase (once) and a test phase (once), in that order. The test phase had two conditions: one with food in the container and one without food.

In the demonstration phase, the assistant turned the horse's face in the direction of the experimenters and said "Hai" (Japanese for "yes") when the horse looked toward the experimenters. Two experimenters then stood in front of the horse in turn and tried to take the lid off a container with a roll of vinyl tape inside for approximately 5 s (i.e., one experimenter stood in front of each horse at a time and performed their demonstration). One experimenter was able to open the lid of the container and removed the vinyl tape (skilled actor). The other experimenter was unable to open the lid of the container, but 'accidentally' dropped the container, causing the lid to open, therefore achieving the same end result, removal of the vinyl tape from the container (unskilled actor).

In the test phase, the two experimenters stood side by side, facing downward, in front of the horse. The assistant turned the horse's face until it was facing the mid-point between the experimenters and said "Hai" (Japanese for "yes") when the horse looked toward the experimenters. The two experimenters then performed in either the food condition or the no-food condition. In the food condition, the two experimenters simultaneously attempted to open the lid of a container containing food (a carrot) for approximately 10 s without actually opening the lid. In the no-food condition, the two experimenters performed the same action but there was no food in the container.

Each horse received two trials (the food and no-food condition), with at least one week between each trial. In both phases, the two experimenters did not look at the horses during the trial and maintained a neutral facial expression to avoid unintentional cues. During the trials, the assistant also turned his/her face down so as not to influence the horses. The role of each experimenter was counterbalanced within each condition across individual horses and the order of conditions was counterbalanced between horses. The position of the experimenter was fixed, with Experimenter A always on the right side of the horse and Experimenter B always on the left side of the horse.

Data Analysis

We excluded data from six horses in our analyses because of procedural errors (i.e., the assistant did not turn his/her face down and he/she made eye contact with the horses). To determine whether the horses discriminated between the two experimenters based on their skillfulness, we analyzed the proportion of the horses' gaze duration toward each of the experimenters during the test phase. As the duration of the experimenters' performance in the test phase was not exactly the same for each horse, we calculated the proportion of gaze duration to performance time rather than comparing absolute values of gaze duration. Ringhofer and Yamamoto (2017) defined gaze behavior as follows: "*The direction of looking was defined when a horse faced its nostrils directed (within 45°) to a specific target (e.g. caretaker or food buckets) more than 1 s with its ears focused forward* " (p. 400). The 45° criterion was defined based on Lampe and Andre (2012) as until either the left or right eye of the horse was out of sight. When this criterion was met, horses were judged not to be looking at both experimenters. Following the above definition, we calculated the proportion of gaze duration by dividing the time spent looking at either experimenter by the performance time of the test phase. Video analysis was performed using BORIS software (v.8.20.4; Friard & Gamba, 2016). All videos were analyzed by an experimenter. Gaze was recorded from when the assistant gave the cue to when the actors finished performing. To assess the reliability of the video analysis, a third person blind to the purpose of the experiment analyzed 20% ($n = 10$) of all videos. We chose five horses from each condition ($n = 10$ horses), and the third person analyzed and coded the videos. We then examined the correlation between the coding of the first coder and the blind coder for skilled actor/performance time and unskilled actor/performance time using Spearman's rank correlation tests. Reliability between the two coders was sufficiently high (skilled actor/performance time, Spearman's rank correlation coefficient: $r = 0.99$, $p < .001$; unskilled actor/performance time, Spearman's rank correlation coefficient: $r = 1.00$, $p < .001$).

We analyzed whether the proportion of gaze duration differed between experimental conditions and roles using a generalized linear mixed model (GLMM) with a binomial distribution. Experimental condition (food/no-food) and role (skilled actor/unskilled actor) were included as fixed effects with an interaction between the two, the proportion of gaze duration as a response variable, and individual horse as a random effect. The model formula was:

glmer (cbind (proportion_looking_time, 100-proportion_looking_time) ~ condition * role
+ (1|horse), family = binomial, data = data).

Likelihood ratio tests (type II tests) were performed to test whether fixed effects were significant. Multiple comparisons using the Bonferroni method were performed as post-hoc tests to investigate whether the proportion of gaze duration toward the skilled actor differed between conditions and whether the proportion of gaze duration toward each experimenter differed in each condition. All statistical analyses were conducted using R version 4.1.2 (R Core Team, 2021). We used the cor.test function from the stats package (R Core Team, 2021) for the Spearman's rank correlation tests, the glmer function from the lme4 package (Bates et al., 2015) for modeling, the anova function from the base package (R Core Team, 2021) for likelihood ratio tests, the emmeans function from the emmeans package (Lenth et al., 2023) for multiple comparisons, and the ggemmeans function from the ggeffects package (Lüdecke, 2018) for calculating the predicted values from the model.

Results

The gaze duration of each horse toward each experimenter and the performance time of the test phase are shown in Table 2, and the proportions of each horse's gaze duration toward the skilled actor and unskilled actor in each condition are shown in Figure 2. Likelihood ratio tests showed that the main effects of condition and role were significant, but the interaction was not significant (condition, $\chi^2(1) = 15.14$, $p < .001$, 95% CI [−0.41, −0.14]; role, $\chi^2(1) = 92.93$, $p < .001$, 95% CI [−0.72, −0.44]; interaction, $\chi^2(1) = 2.82$, $p = .093$, 95% CI [−0.03, 0.38]; Table 3).

Figure 2

Proportion of Gaze Duration of Each Horse Toward the Skilled Actor and the Unskilled Actor in Each Condition

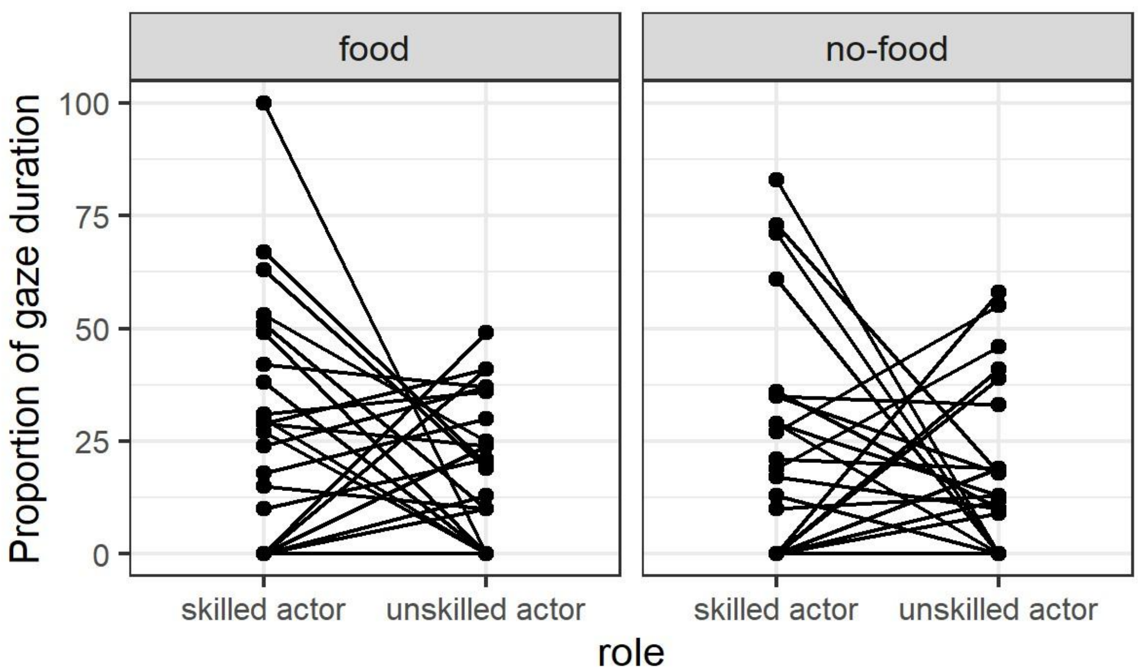

Table 2

Gaze Duration of Each Horse Toward the Skilled Actor and the Unskilled Actor and Total Performance Time (Seconds)

Breed	Food condition			No-food condition		
	Skilled actor	Unskilled actor	Total	Skilled actor	Unskilled actor	Total
Thoroughbred	2.8	4.0	9.8	3.8	1.0	10.5
Thoroughbred	1.8	3.0	10.0	0.0	0.0	10.0
Thoroughbred	4.0	0.0	10.5	1.0	1.3	10.3
Thoroughbred	6.7	2.0	10.0	2.7	5.5	10.0
Thoroughbred	1.5	1.0	9.8	0.0	0.0	10.0
Thoroughbred	0.0	4.8	9.8	2.3	2.0	10.5
Thoroughbred	10.0	0.0	10.0	3.5	3.3	10.0
Thoroughbred	3.0	0.0	11.0	0.0	1.3	10.5
Thoroughbred	0.0	2.5	10.5	3.5	1.8	10.0
Thoroughbred	0.0	0.0	10.2	7.3	0.0	10.3
Thoroughbred	3.0	0.0	10.0	1.8	1.0	10.3
Thoroughbred	0.0	4.3	10.5	1.3	0.0	10.0
Thoroughbred	0.0	0.0	10.8	0.0	0.0	10.3
Thoroughbred	2.5	3.8	10.3	0.0	1.0	10.8
Thoroughbred	0.0	1.0	10.3	2.8	0.0	9.8
Thoroughbred	1.0	2.2	10.5	0.0	2.0	10.3
Thoroughbred	6.5	2.0	10.3	0.0	4.0	10.3
Thoroughbred	4.3	3.8	10.3	8.5	0.0	10.3
Thoroughbred	5.3	2.5	10.0	6.3	0.0	10.3
Thoroughbred	5.3	1.0	10.3	0.0	0.0	10.8
Nihon Sport Horse	3.3	3.8	10.5	7.3	1.8	10.0
Nihon Sport Horse	0.0	1.3	10.3	2.0	5.0	10.8
Nihon Sport Horse	3.0	2.5	10.5	0.0	4.3	10.5
PZHK	0.0	0.0	10.5	3.0	1.3	10.3
KWPN	5.0	0.0	10.3	0.0	6.0	10.3

Note. KWPN, Koninklijk Warmbloed Paardenstamboek Nederland/Royal Warmblood Studbook of the Netherlands registered Dutch Warmblood; PZHK, Polish Horse Breeders Association.

Table 3

Parameter Estimates and 95% Confidence Intervals

Term	Estimate	Lower CI	Upper CI
(Intercept)	−1.17	−1.58	−0.77
Condition (no-food)	−0.28	−0.41	−0.14
Role (unskilled actor)	−0.58	−0.72	−0.44
Interaction	0.17	−0.03	0.38
SD (random effect)	0.94	0.70	1.34

The post-hoc tests showed that the difference between the proportion of gaze duration toward the skilled actor in the food and no-food condition was significant ($Z = 4.03$, $p = .0001$, 95% CI [0.14, 0.41]; Table 4). The predicted value for the proportion of gaze duration toward the skilled actor was 24% (95% CI [17%, 31%]) in the food condition and 19% (95% CI [14%, 26%]) in the no-food condition (Figure 3). In other words, the horses looked at the skilled actor longer in the food condition than in the no-food condition. Conversely, the difference between the proportion of gaze duration toward the unskilled actor in the food and no-food condition was not significant ($Z = 1.32$, $p = .186$, 95% CI [−0.05, 0.26]; Table 4). The predicted value for proportion of gaze duration toward the unskilled actor was 15% (95% CI [11%, 20%]) in the food condition and 14% (95% CI [10%, 19%]) in the no-food condition (Figure 3). In other words, there was no significant difference between the proportion of gaze duration toward the unskilled actor in the food and no-food conditions.

The post-hoc tests also showed that the difference between the proportion of gaze duration toward each experimenter was significant in both conditions (food condition, $Z = 8.08$, $p < .001$, 95% CI [0.44, 0.72]; no-food condition, $Z = 5.42$, $p < .001$, 95% CI [0.26, 0.55]; Table 4). In the food condition, the predicted value for the proportion of gaze duration toward the skilled actor was 24% (95% CI [17%, 31%]), and toward the unskilled actor was 15% (95% CI [11%, 20%]) (Figure 3). In the no-food condition, the predicted value for the proportion of gaze duration toward the skilled actor was 19% (95% CI [14%, 26%]), and toward the unskilled actor was 14% (95% CI [10%, 19%]) (Figure 3). In other words, the horses looked longer at the skilled actor in both conditions.

Figure 3

Predicted Values for the Proportion of Gaze Duration Toward the Skilled Actor and the Unskilled Actor in Each Condition

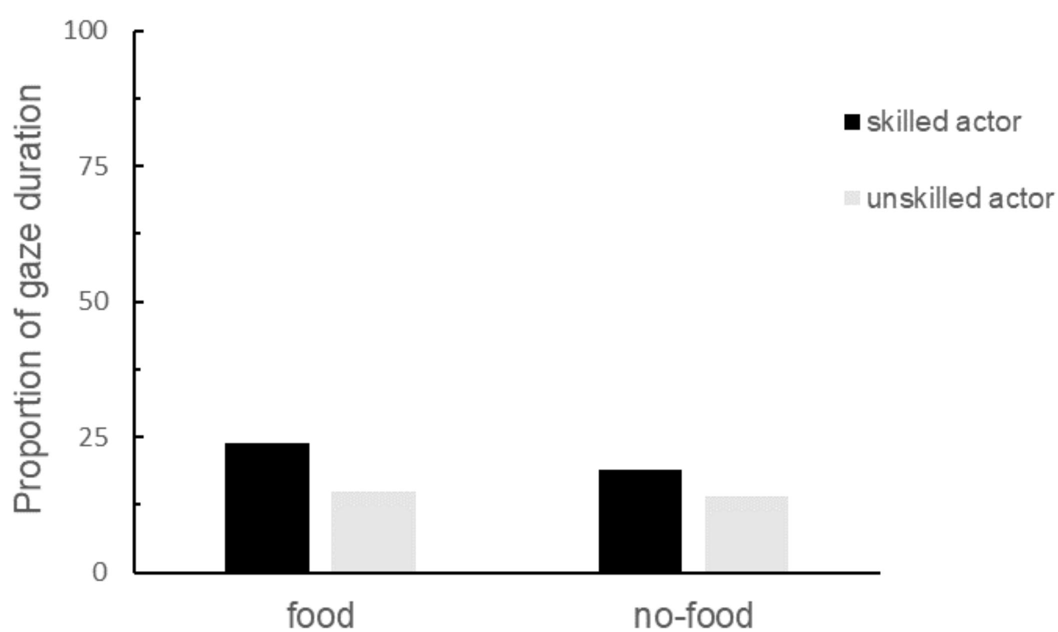

Table 4

Estimates and 95% Confidence Intervals in Multiple Comparisons Using the Bonferroni Method

Contrast	Estimate	Lower CI	Upper CI
Food – No-food (role = skilled actor)	0.28	0.14	0.41
Food – No-food (role = unskilled actor)	0.10	−0.05	0.26
Skilled actor – Unskilled actor (condition = food)	0.58	0.44	0.72
Skilled actor – Unskilled actor (condition = no-food)	0.41	0.26	0.55

Discussion

The aim of this study was to investigate whether horses can evaluate humans based on their skillfulness in a situation related to their own interests (food). If horses could evaluate humans based on their skillfulness, we predicted that they would look longer at the skilled actor than the unskilled actor in the food condition, and longer at the skilled actor in the food condition than in the no-food condition. We found that the horses looked longer at the skilled actor than the unskilled actor overall and looked at the two actors longer in the food condition than the non-food condition, and there was no interaction between role and condition. The post-hoc test results and lack of interaction also revealed that the horses looked preferentially at the skilled actor in the food condition, and contrary to our expectations, they also looked preferentially at the skilled actor in the no-food condition (food condition: skilled actor 24% [17%, 31%], unskilled actor 15% [11%, 20%]; no-food condition: skilled actor 19% [14%, 26%], unskilled actor 14% [10%, 19%]). Furthermore, the proportion of gaze duration toward the skilled actor was greater in the food condition than in the no-food condition (24% [17%, 31%] in the food condition, 19% [14%, 26%] in the no-food condition). These results partially support our hypotheses and show that horses can visually discriminate between humans based on their skillfulness in an object manipulation situation. Thus, these results suggest that horses may be sensitive to cues that would indicate more skilled humans.

Previous studies examining the ability of horses to evaluate humans have focused on their prosociality (Sankey et al., 2010; Trösch et al., 2020). These studies show that horses evaluate humans based on their prosociality. There is only one study on the ability of horses to evaluate humans based on their skillfulness (Sugimoto & Hirata, 2022). In that study, we showed that horses visually discriminate between humans based on their skillfulness. However, in that study the end state in the demonstration phase differed between conditions, so it is possible that the horses simply preferred the person who removed the object from the container. In the present study, we made modifications to ensure that the final state of the two actors in the demonstration phase was the same. The present findings suggest that horses visually discriminated between humans based on their skillfulness, which are in line with our previous study (Sugimoto & Hirata, 2022).

The reason why the horses looked at the skilled actor longer in the food condition than the no-food condition may be that they were simply more interested in the former than the latter because food was present. In the no-food condition, in which the actors tried to open empty containers, the horses may not have shown much interest because there was no prospect of any reward for them. Moreover, we only measured the horses' looking behavior and did not investigate whether the horses could form a behavioral preference based on the information they obtained about humans' skillfulness. Thus, we cannot conclude from this study alone that horses can evaluate humans based on their skillfulness. To investigate the ability of horses to evaluate humans, future studies should investigate a behavioral preference, such as which person the horses choose to receive food from after providing them with information about the skillfulness of two people.

The overall trend of horses looking at the skilled actor longer than the unskilled actor in the food condition might have been driven by a small number of individuals. The following is a breakdown of the data presented in Figure 2, showing the actual differences in percentage values by using percentage points (pt). One horse showed a 100% point difference in the expected direction; five additional horses showed differences of 40% point or more in the expected direction. Two horses showed differences of 40% point or more in the unexpected direction. Additionally, while seven horses showed differences greater than 0%pt and less than 40% point in the expected direction, eight horses showed differences greater than 0% point and less than 40% point in the unexpected direction. Three horses did not look at the skilled actor or the unskilled actor at all. Therefore, it is unlikely that the significant effects in the food condition were driven by a small number of specific individuals.

In this study, the difference between the proportion of gaze duration toward the skilled actor (24% [17%, 31%]) and toward the unskilled actor (15% [11%, 20%]) in the food condition was 9%, and would actually be between 0% points and 20% points. To our knowledge, there are two studies that measured looking behavior of horses and provided the proportion of gaze duration. One is a study by Trösch et al. (2019b), which showed that horses can evaluate humans based on their attentional state. In their study, the difference in the proportion of gaze duration was 10% points (experimenter that knew where the food was hidden: 17%; experimenter that did not know where the food was hidden: 7%). The other is a study by Wathan et al. (2016), which showed that horses can discriminate between facial expressions of conspecifics based on photographs. In their study, the differences in the proportion of gaze duration were 4% points (positive facial expression: 7%; negative facial expression: 3%) and 5% points (relaxed facial expression: 7%; negative facial expression: 2%). Although we only have these two studies as benchmarks, given the threshold they established, we consider the difference between the proportion of gaze duration in this study (9% points) to be big enough. However, the difference between them in this study would actually be between 0% points and 20% points. Therefore, careful judgment should be used in interpreting this difference, as the actual difference may not be that great. In the future, we hope that more studies measuring horses' looking behavior will be conducted to better conceptualize the magnitude of the difference in gaze duration.

We found the unexpected result that the horses looked longer at the skilled actor in the no-food condition. There are three possible reasons for this finding. The first is that the extra activities of the unskilled actor, which were the 'accidental' dropping of the container and crouching down to pick it up during the demonstration phase, may have caused the horses to avoid the unskilled actor. To equate the behavior of the two actors in the demonstration phase, in future experiments, the skilled actor should also drop the container unintentionally and crouch down to pick it up after removing the object from the container during the demonstration phase. The second is that the horses may just have looked longer at a failure. Because longer gaze duration is often associated with surprise, the horses may have been surprised that the skilled actor could not open the container, and thus may have looked at the skilled actor longer in both conditions. The third is that the horses were able to evaluate the experimenters' skillfulness, even in situations not related to their own interests, and preferred the skilled actor in the no-food condition. Future studies should use situations unrelated to the horses' interests and investigate whether horses can evaluate humans based on their skillfulness. For example, it might be possible to investigate which experimenter horses would prefer after seeing a skilled actor or an unskilled actor helping a person with an unsolvable task. Such studies could investigate whether horses can evaluate humans based on their skillfulness in situations that are not related to their own interests.

We did not confirm if the horses could discriminate between intentional and unintentional actions, thus, additional caution is needed to interpret the results. However, before conducting this experiment, we investigated whether the dropping of the container by the unskilled actor is thought to be unintentional by showing this performance to ten people. Nine of them considered this dropping to be unintentional. Moreover, Trösch et al. (2020) suggested that horses perceive human goals or intentions from their behavior. Therefore, we assumed that the horses perceived this dropping as unintentional. However, there is no way to show that the horses completely perceived this dropping as unintentional. In the future, more studies should be conducted to investigate whether horses can recognize others' goals or intentions.

An additional limitation of the experimental design of our study is that we are not able to determine whether the animals developed preferences based on observable behavior alone without the need to attribute underlying characteristics or intentions because there must always be a visible behavior that signals underlying states such as competence or prosociality, and hence it is not possible to know whether the subjects rely only on the observable behaviors or additionally infer the unobservable state. A possible way to overcome this challenge would be to conduct various tests using different behaviors to see if the subjects consistently prefer another individual with a certain characteristic across various contexts (i.e., the generalization test). Future research is needed in this regard.

Using the same experimenters for the same horses in both conditions may also have influenced the horses' looking behavior in the second condition via the memory of the experimenter's behavior in the first condition. We swapped their roles and conducted the two conditions one week apart to reduce the likelihood that the information about the experimenters in the first condition would influence the horses' looking behavior in the second condition. However, we cannot completely rule out the possibility that the memory of the first condition influenced the result of the second condition. Thus, future experiments should be performed with different experimenters from the first condition.

This study also suggests that horses can visually discriminate between humans based on their skillfulness through indirect experience of observing human performance. Therefore, future studies should test whether horses can evaluate humans based on their skillfulness after directly interacting with a skilled actor or an unskilled actor. For example, using the method used by Piotti et al. (2017) with dogs, future studies could confront a horse with a device containing food that it cannot remove on its own (an unsolvable task) and give the horse the direct experience of being helped by a skilled actor or an unskilled actor. When the horses are then faced with the unsolvable task again, future studies could test which experimenter the horses look at and touch for longer. Although this method was not used in this study because of the complexity of the apparatus and the habituation procedures, it may be possible in the future to use this method to clarify whether horses can evaluate humans based on their skillfulness through direct experience with a skilled actor or an unskilled actor. Fiske et al. (2006) also claim that the traits of competence and warmth are important in evaluating others. In this study, we investigated horses' preference for humans based solely on their skillfulness. Therefore, future studies should investigate whether horses can evaluate humans based on their prosociality.

The results from this study may provide new insight into the evolutionary origins of the ability to evaluate others. Previous studies have shown that dogs, which are phylogenetically distantly related species but share their environment with humans through domestication, can evaluate humans not only on their prosociality but also on their skillfulness (Chijiiwa et al., 2015, 2022). In the current study, our findings with horses may suggest the importance of well-developed social cognitive skills acquired through domestication in the evolution of social evaluation. Cognitive skills needed by domestic animals to evaluate others, especially humans, may include the ability to recognize individuals, to learn and remember human behavior by interacting with humans and/or observing human behavior, to discriminate between human characteristics such as human vocalizations and facial expressions, and to integrate human characteristics (cross-modal recognition). Moreover, they need to combine this information and recognize human goals or intentions underlying their behavior, and would therefore evaluate humans based on their past interactions and/or their observations. In fact, previous studies have shown that horses have these abilities (see Jardat & Lansade, 2022 for a review). Future studies need to investigate the ability to evaluate humans or conspecifics in other domestic animals, such as cows and pigs to see if this idea is correct.

Another possibility is that living in groups with complex social hierarchies may be important in the evolution of social evaluation. Previous studies of cats, which are domesticated animals like dogs and horses, have provided no evidence that cats evaluate humans based on their prosociality or skillfulness (Chijiiwa et al., 2022; Leete et al., 2020). Dogs are descended from wolves (*Canis lupus*) (Vila et al., 1997), a highly social species (Mech & Boitani, 2003), and feral dogs live in pack-like social groups with long-term social bonds (Cafazzo et al., 2010). Horses also live in groups with complex social hierarchies (Ringhofer et al., 2017). Conversely, cats are basically solitary predators, although free-ranging cats can flexibly adjust their sociality, for example, if the environment is rich in food, they can live in groups (Vitale, 2022). Therefore, differences in social group structures may affect the ability to evaluate others. However, there are few studies investigating the ability to evaluate others in other highly sociable species (Asian elephant, *Elephas maximus*: Jim et al., 2021; bottlenose dolphins, *Tursiops* spp.: Johnson et al., 2018;

cleaner fish, *Labroides dimidiatus*: Bshary & Grutter, 2006). We hope that more studies will be conducted in species with complex social structures.

Our finding that horses visually discriminated between humans based on their skillfulness in an object manipulation situation could enhance training methodologies and improve human-horse communication. If future research shows that horses evaluate skilled humans more positively than unskilled humans, we can improve the effectiveness and development of training methods by making sure that humans do not behave ambiguously or inconsistently in front of horses, and therefore contribute to a positive evaluation by the horses. More studies on social cognition of horses as well as on the ability of horses to evaluate others would contribute to better human-horse relationships.

Acknowledgements

We wish to thank Monamie Ringhofer and Hitomi Chijiiwa for their advice on the experimental design. We also thank the equestrian clubs of Kyoto and Kyoto Sangyo universities for allowing us to use their facilities and work with their animals. We thank Clio Reid, PhD, from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript. This work was supported by JSPS KAKENHI (19H00629).

Author Contributions: Takuto Sugimoto: conceptualization, methodology, formal analysis, investigation, data curation, writing – original draft, visualization, project administration, funding acquisition. **Satoshi Hirata:** investigation, writing – review and editing, supervision, project administration.

Funding: This work was supported by JSPS KAKENHI, grant number 19H00629.

Conflict of Interest: The authors declare no conflicts of interest.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, T.S, upon reasonable request.

References

Anderson, J. R., Kuroshima, H., Takimoto, A., & Fujita, K. (2013). Third-party social evaluation of humans by monkeys. *Nature Communications, 4*, 1561.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Bshary, R., & Grutter, A. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature, 441*, 975–978.

Cafazzo, S., Valsecchi, P., Bonanni, R., & Natoli, E. (2010). Dominance in relation to age, sex, and competitive contexts in a group of free-ranging domestic dogs. *Behavioral Ecology*, *21*(3), 443–455.

Chijiiwa, H., Horisaki, E., Hori, Y., Anderson, R. J., Fujita, K., & Kuroshima, H. (2022). Female dogs evaluate levels of competence in humans. *Behavioural Processes, 203*.

Chijiiwa, H., Kuroshima, H., Hori, Y., Anderson, J. R., & Fujita, K. (2015). Dogs avoid people who behave negatively to their owner: Third-party affective evaluation. *Animal Behaviour*, *106*, 123–127.

Cozzi A, Sighieri C, Gazzano A, Nicol, C. J., & Baragli, P. (2010). Post-conflict friendly reunion in a permanent group of horses (Equus caballus). *Behavioural Process*, *85*, 185–190.

Fiske, S. T. (2015). Intergroup biases: A focus on stereotype content. *Current Opinion in Behavioral Sciences, 3*, 45–50.

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science, 27*(2), 67–73.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2006). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Science, 11*(2), 77–83.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902.

Freidin, E., Putrino, N., D'Orazio, M., & Bentosela, M. (2013). Dogs' eavesdropping from people's reactions in third party interactions. *PLoS ONE, 8*(11).

Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution,7*, 1325–1330.

Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences, 9*, 439–444.

Herrmann, E., Keupp, S., Hare, B., Vaish, A., & Tomasello, M. (2013). Direct and indirect reputation formation in nonhuman great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus*) and human children (*Homo sapiens*). *Journal of Comparative Psychology, 127*, 63–75.

Horn, L., Virányi, Z., Miklósi, A., Huber, L., & Range, F. (2012). Domestic dogs (*Canis familiaris*) flexibly adjust their human-directed behavior to the actions of their human partners in a problem situation. *Animal Cognition, 15*, 57–71.

Jardat, P., & Lansade, L. (2022). Cognition and the human–animal relationship: A review of the sociocognitive skills of domestic mammals toward humans. *Animal Cognition, 25*, 369–384.

Jim, H. L., Marshall-Pescini, S., & Range, F. (2020). Do dogs eavesdrop on human interactions in a helping situation? *PLoS ONE, 15*(8).

Johnson, C. M., Sullivan, J., Jensen, J., Buck, C., Trexel, J., & St. Leger, J. (2018). Prosocial predictions by bottlenose dolphins (*Tursiops* spp.) based on motion patterns in visual stimuli. *Psychological Science, 29*(9), 1405–1413.

Kawai, N., Yasue, M., Banno, T., & Ichinohe, N. (2014). Marmoset monkeys evaluate third-party reciprocity. *Biology Letters*, *10*(5), 20140058.

Kervyn, N., Yzerbyt, V., & Judd, C. M. (2010). Compensation between warmth and competence: Antecedents and consequences of a negative relation between the two fundamental dimensions of social perception. *European Review of Social Psychology, 21*(1), 155–187.

Krupenye, C., & Hare, B. (2018). Bonobos prefer individuals that hinder others over those that help. *Current Biology, 28*(2), 280–286.

Kundey, S. M. A., De Los Reyes, A., Royer, E., Molina, S., Monnier, B., German, R., & Coshun, A. (2011). Reputation-like inference in domestic dogs (*Canis familiaris*). *Animal Cognition, 14*(2), 291–302.

Lampe, F. J., & Andre, J. (2012). Cross-modal recognition of human individuals in domestic horses (*Equus caballus*). *Animal Cognition, 15*, 623–630.

Leete, J., Vonk, J., Oriani, S., Eaton, T., & Lieb, J. (2020). Do domestic cats (*Felis silvestris catus*) infer reputation in humans after direct and indirect experience? *Human-Animal Interaction Bulletin, 8*, 35–53.

Lenth, V, L., Bolker, B., Buerkner, P., Gine-Vazquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2023). emmeans: Estimated marginal means, aka least-squares means. R package version 1.8.8.

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), 772.

Maros, K., Gácsi, M., & Miklósi, Á. (2008). Comprehension of human pointing gestures in horses (*Equus caballus*). *Animal Cognition, 11*, 457–466.

Marshall-Pescini, S., Passalacqua, C., Ferrario, A., Valsecchi, P., & Prato-Previde, E. (2011). Social eavesdropping in the domestic dog. *Animal Behaviour, 81*(6), 1177–1183.

Mech, L. D., & Boitani, L. (2003). Wolf social ecology. *USGS Northern Prairie Wildlife Research Center, 318*.

Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science, 311*, 1297–1300.

Nitzschner, M., Kaminski, J., Melis, A. P., & Tomasello, M. (2014). Side matters: Potential mechanisms underlying dogs' performance in a social eavesdropping paradigm. *Animal Behaviour, 90*, 263–271.

Nitzschner, M., Melis, A. P., Kaminski, J., & Tomasello, M. (2012). Dogs (*Canis familiaris*) evaluate humans on the basis of direct experiences only. *PLoS ONE, 7*(10).

Ottoni, E.B., de Resende, B.D. & Izar, P. (2005). Watching the best nutcrackers: what capuchin monkeys (*Cebus apella*) know about others' tool-using skills. *Animal Cognition, 8*, 215–219.

Outram, K. A., Stear, A. N., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., Thorpe, N., & Evershed, P. R. (2009). The earliest horse harnessing and milking. *Science, 323*, 1332–1335.

Piotti, P., Spooner, M. R., Jim, L. H., & Kaminski, J. (2017). Who to ask for help? Do dogs form an opinion on humans based on skillfulness? *Applied Animal Behaviour Science, 195*, 93–102.

Proops, L., & McComb, K. (2012). Cross-modal individual recognition in domestic horses (*Equus caballus*) extends to familiar humans. *Proceedings of the Royal Society B, 279*, 3131–3138.

Proops, L., Walton, M., & McComb, K. (2010). The use of human-given cues by domestic horses, *Equus caballus*, during an object choice task. *Animal Behavior, 79,* 1205–1209.

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing, Vienna, Austria.
Ringhofer, M., Inoue, S., Mendonca, R. S., Pereira, C., Matsuzawa, T., Hirata, S., & Yamamoto, S. (2017a). Comparison of the social systems of primates and feral horses: Data from a newly established horse research site on Serra D'Arga, northern Portugal. *Primates, 58*, 479–484.
Ringhofer, M., & Yamamoto, S. (2017b). Domestic horses send signals to humans when they face with an unsolvable task. *Animal Cognition, 20*, 397–405.
Russell, Y. I., Call, J., & Dunbar, R. I. M. (2008). Image scoring in great apes. *Behavioural Processes, 78*(1), 108–111.
Sankey, C., Richard-Yris, M.-A., Henry, S., Fureix, C., Nassur, F., & Hausberger, M. (2010). Reinforcement as a mediator of the perception of humans by horses (*Equus caballus*). *Animal Cognition, 13*(5), 753–764.
Schrimpf, A., Single, M. S., & Nawroth, C. (2020) Social referencing in the domestic horse. *Animals 10*(1), 164.
Stammbach, E. (1988). Group responses to specially skilled individuals in a *Macaca fascicularis* group. *Behaviour*, *107*, 241–266.
Subiaul, F., Vonk, J., Okamoto-Barth, S., & Barth, J. (2008). Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers. *Animal Cognition, 11*(4), 611–623.
Sugimoto, T., & Hirata, S. (2022). Horses' preferential looking to humans based on problem-solving ability. *The Japanese Journal of Animal Psychology, 72*(1), 17–25
Trösch, M., Pellon, S., Cuzol, F., Parias, C., Nowak, R., Calandreau, L., & Lansade, L. (2020). Horses feel emotions when they watch positive and negative horse-human interactions in a video and transpose what they saw to real life. *Animal Cognition, 23*, 643–653.
Trösch, M., Cuzol, F., Parias, C., Calandreau, L., Nowak, R., & Lansade, L. (2019a). Horses categorize human emotions cross-modally based on facial expression and non-verbal vocalizations. *Animals, 9*(11), 862.
Trösch, M., Ringhofer, M., Yamamoto, S., Lemarchand, J., Parias, C., Lormant, F., & Lansade, L. (2019b). Horses prefer to solicit a person who previously observed a food-hiding process to access this food: A possible indication of attentional state attribution. *Behavioural Processes, 166*, 103906.
Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Development, 81*, 1661–1669.
Vila, C., Savolainen, P., Maldonado, J., Amorim, R. I., Rice, E. J., Honeycutt, L. R., Crandall, A. K., Lundeberg, J., & Wayne, K. R. (1997). Multiple and ancient origins of the domestic dogs. *Science, 276*, 1687–1689.
Vitale K, R. (2022). The Social Lives of Free-Ranging Cats. *Animals*, *12*(1),126.
Warneken, F., Chen, F., Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development, 77,* 640–643.
Wathan, J., Proops, L., Grounds, K., & McComb, K. (2016). Horses discriminate between facial expressions of conspecifics. *Scientific Reports, 6*, 38322.
Wu, J., Balliet, D., Van Lange, P. A. M. (2016). Reputation, gossip, and human cooperation. *Social and Personality Compass, 10*, 350–364.