Does a high social status confer greater levels of trust from groupmates? An experimental study of leadership in domestic horses

Mathilde Valenchon, Jean-Louis Deneubourg, Anna Nesterova, Odile Petit

To cite this version:
Mathilde Valenchon, Jean-Louis Deneubourg, Anna Nesterova, Odile Petit. Does a high social status confer greater levels of trust from groupmates? An experimental study of leadership in domestic horses. Behavioural Processes, 2022, 201, pp.1-8. 10.1016/j.beproc.2022.104708 . hal-03818611

HAL Id: hal-03818611
https://hal.science/hal-03818611
Submitted on 26 Oct 2022

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

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Does a high social status confer greater levels of trust from groupmates?
An experimental study of leadership in domestic horses

Mathilde Valenchon\textsuperscript{a,b,*}, Jean-Louis Deneubourg\textsuperscript{c}, Anna P. Nesterova\textsuperscript{d} & Odile Petit\textsuperscript{a,*}

\textsuperscript{a} Physiologie de la Reproduction et des Comportements, UMR INRAE, CNRS, Université de Tours, IFCE, 37380 Nouzilly, France
\textsuperscript{b} Bristol Veterinary School, Langford House, Bristol, BS40 5DU, United-Kingdom
\textsuperscript{c} Center for Nonlinear Phenomena and Complex Systems (Cenoli) - Université libre de Bruxelles, 1050 Brussels, Belgium
\textsuperscript{d} FaunaStats, 16 avenue de l'Europe, 67300 Schiltigheim, France

Contact
*Corresponding authors:
Odile Petit, email address: odile.petit@cnrs.fr
Mathilde Valenchon, email address: mathilde.valenchon@yahoo.fr
Postal address: UMR 7247 Physiologie de la Reproduction et des Comportements INRAe - CNRS - Université de Tours – IFCE, 23 rue du Loess, 67037 Strasbourg
Tel: +33 (0)3 88 10 74 57
Abstract

In collective movements, specific individuals may emerge as leaders. In this study on the domestic horse (*Equus ferus caballus*), we conducted experiments to establish if an individual is successfully followed due to its social status (including hierarchical rank and centrality). We first informed one horse about a hidden food location and recorded by how many it was followed when going back to this location. In this context, all horses lead their groupmates successfully. In a second step, we tested whether group members would trust some leaders more than others by removing the food before the informed individual led the group back to the food location. In addition, two control initiators with intermediate social status for which the food was not removed were tested. The results, confirmed by simulations, demonstrated that the proportions of followers for the unreliable initiator with highest social status are greater than the ones of the unreliable initiator with lowest social status. Our results suggest an existing relationship between having a high social status and a leadership role. Indeed, the status of a leader sometimes prevail at the detriment of the accuracy of the information, because an elevated social status apparently confers a high level of trust.
1. Introduction

One of the most captivating social displays in the biological world is undoubtedly collective movements and concerns many species: swarms of invertebrates, shoals of fish, flocks of birds, herds of ungulates or troops of primates. This polyphyletic phenomenon reflects its highly adaptive function (Boinski and Garber 2000). However, how such coordination for moving together is achieved raises several issues. Among them, the possibility that some individuals exert a disproportionate influence on group decisions, and thus considered as leaders, is still in debate (Strandburg-Peshkin, Papageorgiou et al. 2018) including in humans (von Rueden and van Vugt 2015), and raises questions about its sustainability within a social group. Indeed, confidently following a leader will speed up the collective decision but it implies that the information is concentrated in this sole individual (Couzin, Krause et al. 2005, Sumpter and Pratt 2009, Kao and Couzin 2014, Gavrilets, Auerbach et al. 2016). In case of wrong information, the whole group will suffer the consequences of bad choices (Conradt and Roper 2005).

Even if all group members can initiate collective movements in nearly all species, some of these initiators are joined more successfully than others. In this context, we wonder why some individuals are elevated to leadership roles by group members, resulting in a more efficient decision-making process. These interindividual differences in leadership success can be accounted for by several factors such as age, sex, temperament, energetic state, dominance status, or centrality (see Petit and Bon 2010 for a review). Whatever the species, there is not a unique and consistent factor explaining the success of an initiator. However, it appears that several parameters underlying social status may play a preponderant role. Indeed, on one hand, we know that dominant individuals can play a key role in collective decision-making (Radford 2004, Sueur and Petit 2008, Jacobs, Watanabe et al. 2011). On the other hand, it seems that the network of affiliative relationships amply explains joining latencies with an optimization of the process when initiators have high centrality scores (Sueur and Petit 2008, King, Sueur et al. 2011, Sueur, MacIntosh et al. 2013, Briard, Dorn et al. 2015). All these factors refer to the social position of an individual among its group mates. It is thus likely that a favourable social position may lead to a higher social influence on group decisions. This social influence could reflect the trust that other group mates might place in this conspecific.

This study aims to explore the social influence of individuals in explaining successful leaders and thus focuses on followership which is generally neglected in the literature (Van Vugt, Hogan et al. 2008). We first predict that potential followers would be more inclined to follow an initiator proposing an activity if this initiator possess a high social status (including hierarchical rank and centrality). If this is the case, it might suggest that followers consider those individuals as more trustable leaders, making them more motivated to follow.

In naturally occurring complex social systems, correlations between multiple factors result in feedback effects that make it difficult to distinguish between causes and consequences. It is therefore necessary to conduct experiments to assess the determining factors of decision-making. Thus, we conducted experiments simulating natural situations to establish if an individual is successfully followed due to its social status. By controlling experimentally the time and place of the group departure, we are able to narrow our focus on the influence of a leader’s identity itself by discarding the influence of the pre-departure period (Bourjade, Thierry et al. 2009, Sueur, Deneubourg et al. 2011, Briard, Deneubourg et al. 2017, Briard,
Deneubourg et al. 2021). In the first step of this experiment, we informed one individual about a hidden food location containing highly preferred food and recorded how much and how rapidly it was followed in comparison to other group members. As a result, the whole group had access to a great amount of highly preferred food. We tested individuals with different hierarchical status and centrality scores and compared their performances in terms of proportion of the group members that followed the initiator. In the second step of the study, we tested whether group members would trust some leaders more than others with an extinction procedure. To investigate this extent of trust, we experimentally tested its perseverance for an initiator for which the reliability is challenged. In this second step, we informed a single individual about a food location as usual, but food was removed before the informed individual led the group back to this location. We then measured how this new unreliability impacted the followers’ decision throughout several testing sessions. If group members no longer trust the informed individuals whatever their identity, we should observe a decrease of the following rates over the course of the experiment. Conversely, if congeners continue to follow specific individuals, those with high social status for example, it would provide the best demonstration that a form of trust from the followers that does not solely depend on an assessment of a leader’s reliability as a keystone factor to determine group decision-making.

We conducted this work in the domestic horse (Equus ferus caballus) which is an ideal model for this experimental study as they are easily trainable. The domestic horse forms cohesive social groups of both sexes (Cameron, Setsaas et al. 2009, Linklater and Cameron 2009), and shares physiological attributes, cognitive and social characteristics with many social primates that have been extensively studied when exploring the influence of social status on decision-making processes.

2. Methods

2.1. Experimental horses

Three groups of 6 females were studied (two in 2014: groups A and B, and one in 2015: group C). Subjects were Haflinger, Dartmoor and Shetland ponies aging 7.2 ± 0.6 years (mean ± SE), housed at the “Elevage du Haut-barr” (Saverne, France) and used for reproduction purposes (brood mares and their daughters) (see Table 1 for a summary of individual characteristics). All horses were familiarised with each other for several years within the farm and studied groups were formed and stabilized at least 6 months before the beginning of the experiment. Four of the 12 horses studied in 2014 were part of the 2015 group C (ARM, OTO, SER, and TAK). All horses lived in outdoor pasture all year long with periodical fodder supplementation and water available ad libitum. Each morning when an experiment took place, all horses were equipped with halters that were removed after the last test. The three groups (A, B, C) were tested for the individual initiation tests (Experiment 1) and only the group C was finally tested for the extinction tests (Experiment 2).

2.2. Experiment 1: individual initiations tests

We provoked movement initiations with a protocol based of food hiding in order to assess individual levels of success (Gérard, Valenchon et al. 2020). The three groups were tested. The experiment consisted of 73 valid tests.
At each test, we first gathered the whole group into a temporary pen (approximately 5 m x 5 m) at the departure area. Then, an experimenter concealed the food reward (19 carrots) out of sight at 30 m from the departure area. The tested horse (only one for each test) was then brought to this hidden food location and was allowed to eat only one carrot so it became motivated to come back. Then, it was brought back to the group and as soon as it entered the departure zone, the whole group was released so the informed horse could initiate a movement to the food location and being eventually followed by others. The 19 carrots were divided in a pile of 10 carrots at the centre plus 9 carrots randomly distributed at a distance of 2-3m around the centre so all horses can have access to food even when the whole group is present (i.e. non-monopolisable resource).

Three experimenters were present to build and rapidly remove the temporary pen made of 6 plastic posts and 20mm-large plastic tape. Three cameras were arranged in order to cover the departure area, the route taken by the group and the arrival area.

The number of followers per initiation was recorded (from direct observations and videos). Any horse seen moving away (at least 20 steps) from the departure in the same direction as the initiator was identified as a follower (Briard, Dorn et al. 2015). Only the trials when the informed horse was the first to move to go directly from the departure zone to the food reward were analysed. Three trials a day were conducted and we had at least two hours break between trials.

2.3. Experiment 2: extinction tests

Extinction tests aimed to assess the perseverance of a group to follow an initiator despite this initiator not leading them to a food reward. This experiment was conducted on group C only. The protocol was the same as for individual initiations tests except the reward (with the exception of two carrots) was removed by an experimenter (discreetly) after the initiator was informed and before it came back followed by the group. The two carrots were left so the initiator had still access to a food reward as it arrived first (to keep it motivated to go back to the reward location), but the rest of the group was not rewarded. Two “unreliable initiators” (i.e. no reward anymore) and two control “reliable initiators” (i.e. intact reward) were tested randomly. The two misinformed initiators were: the horse with the highest social status (HS) and the one with the lowest social status (LS). The two informed initiators were two horses with intermediate social status (see paragraph (d) for social status determination).

During the first phase, we completed five rewarded initiation tests for each initiator, meaning the whole reward remained present for the group. Then, in the second phase, we conducted seven other trials per initiator. These seven trials were extinction trials for the two unreliable initiators (food reward was removed for the followers) and rewarded trials for the two reliable initiators (food reward remained present for all). For each initiator, we compared the amount of followership between the six first trials (five control trials + the 1st trial of the second phase when the group discovered the absence of food reward for misinformed initiators) and the next six trials (6 rewarded trials for reliable initiators or 6 extinction trials for unreliable initiators).

2.4. Assessing social status


Each group was observed and filmed continuously for 6 h daily, distributed between 8 a.m. and 7 p.m., during the four weeks preceding the beginning of experiment 1, for a total of 50 hours. Agonistic interactions were recorded continuously. For each interaction, we recorded the identity of the emitter of each approach, threat or aggression, and both the identity and the response of the receiver (offensive behaviour, avoidance or no reaction). Unidirectional interactions - when only one horse was approaching or emitting offensive behaviour and the other clearly avoided them - were used to build a dominance hierarchy. We carried out an analysis of dominance hierarchy using SOCPROG 2.4 (Whitehead 2009) to obtain David’s scores for each rank and then checked for hierarchy linearity (de Vries 1995, de Vries 1998, de Vries, Stevens et al. 2006).

Spatial proximities between horses were recorded every 15 min using instantaneous scan sampling (Altmann 1974), 200 scans total). Two individuals were considered to be close when they were <1 m apart. Based on these scans, we built a proximity matrix based on the number of scans where two individuals were observed close to each other, for each possible dyad. We carried out a social network analysis to obtain a Centrality score for each horse (Eigen vector index (Jacobs and Petit 2011)).

Dominance and centrality scores have been attributed to each individual (Table 1). For the group C used for the experiment 2, the dominance hierarchy was significantly linear (de Vries test for linearity, h’=0.89, P=0.04) and dominance scores and centrality scores were positively correlated (Spearman correlation test, r=1, P=0.003). Therefore, the horse (SER) with the highest David’s score and Centrality index was chosen to be the unreliable initiator with the highest social status (HS), and the horse (TAK) with the lowest David’s score and Centrality index was chosen to be the unreliable initiator with the lowest social status (LS). Horses ARM and RAT were chosen as control reliable initiators 1 and 2 due to their intermediate status.

2.5. Simulation study

A general probability to follow has been calculated from the first experiment and was the same for all initiators (0.9). This probability was considered to be degraded by the failures (i.e. the unrewarded trials) and we made the assumption that this decrease occurred as soon as the first failure occurred. The probability that an individual will not follow will then be related to the number of failures by the following relationship:

\[ P_S = P_0 + \frac{P_m}{1 + \left(\frac{E}{S}\right)^n} \]  

Where \( P_S \) is the probability to follow an initiator. This probability decreases with the number of failures. \( P_0 \) the basic probability to follow an initiator, \( P_m \) the maximal probability to follow a reliable initiator, \( E \) the number of failures (number of trials with no food) and \( S \) the threshold for tolerating a failure (depending on the social status in our assumption), \( n \) the sensitivity to the number of failures. The extreme is an all-or-none response of the potential followers.
\[ P_S = P_0 \quad \text{if } E \geq S. \quad (2,a) \]
\[ P_S = P_0 + P_m \quad \text{if } E < S \quad (2,b) \]

P₀ and S depend on the identity or the status of the initiator: a highly potentially reliable initiator shows a high value of P₀ and S. This formula allowed us to calculate the probability to follow each unreliable initiator for each trial as well as the probability to follow the reliable initiators. Finally, we obtained a threshold for tolerating failure for the two unreliable initiators.

Eventually, we conducted simulations on the number of followers according to the respective threshold we obtained for both unreliable initiators to confirm our experimental results.

### 2.6. Statistical analyses

Social status analyses (Davids score, linearity tests and Eigenvector centrality scores) have been performed using SOCPROG 2.4. Statistical analyses were performed using R software (R Core Team 2019). The level of statistical significance was set at P<0.05, and the level of tendency at P<0.10. In the text, we report median and interquartile ranges in the form “M=(1st interquartile–3rd interquartile)”. For experiment 1, inter-individual comparisons have not been statistically investigated due to a lack of inter-individual variability in the number of followers. For experiment 2, we used generalised linear model (GLM) with binomial distribution to test how the testing phase, individual’s identity, and trial number influenced the proportion of the group members that followed the informed individual. We used car package (Fox and Sanford 2019) for this analysis. We started with the full model that included the testing phase, individual identity, interaction between these two terms, and trial number as predictor variables.

To perform model selection, we relied on the second-order Akaike’s Information Criterion (AICc) and Relative Variable Importance (RVI) as was suggested by Burnham & Anderson (2002) and Symonds & Moussalli (2011). As best-fitting models, we considered the model with the lowest AICc and any model within two AICc units. We then used model.avg function of the MuMIn package (Barton 2019) to calculate RVI for our predictor variables. Our best-fitting model contained variables with an RVI of at least 0.7 and it had the lowest AICc score. It included testing phase and individual identity as predictor variables. However, the interaction term between these two variables and trial number were not retained. To compare performance of different individuals we performed post-hoc analysis on the individual identity variable in the GLM. We used emmeans function of the emmeans package (Lenth 2019) to perform multiple pairwise comparisons with a Tukey correction.

### 2.7. Ethics statement

The experimental protocols followed EU Directive 2010/63/EU guidelines for animal experiments and were approved by the Ethical Committee (CREMEAS) under agreement number AL/01/10/07/11.

### 3. Results

#### 3.1. Experiment 1
A total of 73 trials have been analysed (i.e. when the informed individual was the first to go
directly to the food location) and resulted in 3 or more trials per initiator for a total of 15
initiators (Table 2). We found no variation in the number of followers according the initiator’s
identity since a large majority of the trials was followed by the whole group (Group A: 96.15%, Group B: 100%, Group C: 78.57 %). Three initiators failed to validate the criteria to
be retained in the analysis (DLI in Group A, CHE in Group B and OTO in Group C).

3.2. Experiment 2
During the first six trials (5 trials of phase 1- rewarded - plus the 6th trial when followers
discovered the food was missing for the first time, see protocol details above), 4 or 5
followers always followed each one of the four initiators. During the 6 following trials (phase
2), the number of followers ranged from 5 to 1 for the reliable initiators (rewarded trials).
Concerning the unreliable initiators, this number ranged from 4 to 1 for the initiator with
highest social status (extinction trials), and from 2 to 1 for the initiator with the lowest social
status (extinction trials, Fig. 1). The best fitting model was explained by horse’s identity and
testing condition (Supplementary Table XXX). There was a significant effect of Phase 2: the
proportion of group members that followed decreased during the unrewarded (second) phase
of the test (GLM, N =4 8, β = -3.075, SE=0.525, Z = -5.86, P < 0.001). The proportion of
group members that followed also depended on the identity of the initiator (Table 3).
Specifically, the ability of the two reliable initiators (ARM, Reliable 1 and RAT, Reliable 2)
to recruit the followers was 7 and 10 times respectively greater than the ability of the
unreliable initiator with lowest social status (TAK, Unreliable 2) (Ration value
Reliable1/Unreliable2=7, z = 3.6, P<0.0001; Ration value Reliable2/Unreliable2 = 10.1, z = 4.00, P <
0.0001, Table 2). On the contrary, the differences between the unreliable initiator with
highest social status (SER, Unreliable 1) and the other individuals were not significant
(Ration value Reliable1/Unreliable1 = 2.9, z = 2.00, P = 0.19; Ration value Reliable2/Unreliable1 = 4.1, z =
2.500, P = 0.06; Ration value Unreliable1/Unreliable2 = 2.4, z = 1.800, P = 0.25, Table 3). Finally,
no difference was found between two reliable initiators (Ration value Reliable1/Reliable2=0.7, z = -
0.6, P = 0.93).

3.3. Simulation study
The probability to follow the unreliable initiator with highest social status (SER) was less
affected by the discovery that food was no longer present than for the unreliable initiator with
lowest social status (TAK) (Fig. 2).
In our case, after a first failure, for both unreliable initiators (regardless their social status), the
mean number of followers decreases but then remains constant from trial to trial (linear
regression between the number of followers as a function of the successive trials, null
hypothesis the slope is equal to zero, P > 0.05). However, the mean number of followers per
trial is greater for the initiator with highest social status than for the other unreliable initiator
(respectively SER: 2.66 and TAK: 1.17). These results suggest that equation 2b is a good
approximation of the individual response to an initiator, the two thresholds (number of
failures) for decreasing the probability to the initiator (S, equation 2) are equal to 1 and
confirm our assumption that an individual with a high social status continues to be trusted
even if it provides false information. On the contrary, an individual with a low social status is
poorly followed immediately after the discovery of the food absence. The probability that each potential follower ($P_0$ in equation 1) follows the initiator with a low social status is lower than the $P_0$ for the initiator with a high social status (respectively $P_0 = 1.17 / 5 = 0.23$ for TAK and $P_0 = 2.66 / 5 = 0.5$ for SER). This result also means that the success of the reliable initiators was not influenced by the failures experienced by group members in trials with unreliable initiators.

We performed 10000 simulations of 6 successive trials in which we applied the $P_0$ of each unreliable initiator. $P_0$ being the probability that each individual follows the initiator. We obtained that the proportions of followers for the unreliable initiator with highest social status (SER) and the unreliable initiator with lowest social status (TAK) are statistically different ($P < 0.05$) (Fig. 2).

4. Discussion

In the first experiment of this study, we provided optimal conditions of departure by giving the same value to all initiators in terms of food access for the whole group. As a consequence, we did not observe any difference between the different initiators in the amount of followership. Such positive influence of possessing right information on followers has already been found in pigeons (Watts, Nagy et al. 2016) and horses (Andrieu, Henry et al. 2016). This absence of inter-individual variability contrasts with what is observed in spontaneous collective movements, where initiators were found to be of unequal value for followers which join them less quickly or not at all (Lea, Gunst et al. 2003, Sueur and Petit 2008). It suggests that phenomena that are non-specific to an initiator identity are also at play in the followers’ decisions and emphasises the importance of the pre-departure period. Indeed, we know that each group member can notify its motivation to move, negotiate or vote during this period and thus contributes to the collective decision at this early step or that group dispersion and activity are decisive for moving collectively (Bourjade, Thierry et al. 2009, Ramseyer, Boissy et al. 2009, Sueur, Deneubourg et al. 2010, Sueur, Deneubourg et al. 2011, Briard, Deneubourg et al. 2021).

Since in optimal experimental conditions each horse appears to be able to lead successfully, in the second experiment, we placed potential followers in a more challenging environment which mimicked natural cases where an initiator could lead the group to a bad location (in terms of food expectation). These conditions simulated the crucial impact of a leader on collective choices. With this experiment, we aimed to reveal the strength of the trust the followers put into a leader, by asking directly the followers. In such a context, followers successfully displayed flexibility in their willingness to follow a leader according to its reliability. Indeed, an immediate drop in the followership was observed within the unreliable initiator with the lowest social status, while followers kept following both reliable initiators. It demonstrates that horses are able to assess the reliability of an initiator. This could be mediated by associative learning mechanisms, especially since it involves the presence or absence of food reward. Such influence of recent experience on former choice on subsequent decision has been found in human and non-human primates (Steelandt, Dufour et al. 2012, Pele, Broihanne et al. 2014). Being able to discriminate social information based on the reliability of their emitting source has also been shown in other social contexts, such as for
alarm calls for instance (vervet monkeys: Cheney and Seyfarth 1988, yellow-bellied marmots:
Blumstein, Verneyre et al. 2004, carrion crows: Wascher, Hillemann et al. 2015). Such ability
is undeniably crucial to optimise the benefit in the use of social information.

Moreover, our results suggest an even more subtle phenomenon: not only the followers have
been able to adapt their choice to follow -or not- based on the reliability of the initiator, but
the identity of the unreliable initiator played a role in how fast this initiator lost their trust.
Indeed, whereas the number of followers dropped drastically and immediately for the
unreliable initiator with the lowest social status, this was not the case for the unreliable
initiator with the highest social status where the decrease was delayed and more gradual. Our
results, obtained on a limited number of individuals, were confirmed by our simulations
which demonstrated that the proportions of followers for the unreliable initiator with highest
social status differed from the one of the unreliable initiator with lowest social status. These
calculated probabilities may represent a threshold for trusting a group mate, and this threshold
may depend on the leader’s social status since our two unreliable initiators occupied the two
most extreme social positions within their group (i.e. the most dominant and central horse vs.
the most subordinate and peripheral one). The fact that the extinction pattern is delayed and
more gradual with an unreliable initiator with the highest social status suggest that even if
followers were able to detect its unreliability, as shown before, there was something left to
maintain their motivation to follow and it strongly suggests an existing relationship between
having a high social status and a leadership role. It is often assumed that animals with a high
social status, especially high-ranked individuals (i.e. in dominance hierarchy), are more
successful as a leader because they are usually older and more experienced, and should
therefore possess accurate information about their environment (McComb, Shannon et al.
2011). In many species, age and hierarchical ranks are indeed correlated (e.g. in horses
(Sigurjónsdóttir, Machteld et al. 2003, Briard, Dorn et al. 2015)), and leaders are also often
high-ranking and/or older group members (Sueur and Petit 2008, Sueur and Petit 2008). As a
consequence, from observational studies, it is impossible to disentangle the intrinsic abilities
that contribute to lead groups successfully. On the contrary, our experimental results suggest
that groupmates’ decision to follow depends on the social status of the initiator, and not age
since neither centrality or David’s scores followed the age order. Moreover, actual signs of
unreliability did not prevent the group to keep following the high-status initiator, suggesting
that a high level of leadership, once acquired, does not depend directly on how well-informed
the leader is, or, at least, that a highly influential leader will benefit from a margin of errors
that a low-status initiator would not benefit from. Altogether, it really emphasises that the
influence of a leader’s social status cannot be solely explained by the actual and/or current
reliability of the information it possesses. We now need to explore whether highly influential
individuals are also leaders in other domains like conflict management (Petit and Thierry
1994, Widdig, Streich et al. 2006), group protection and defence (Meunier, Molina-Vila et al.
2012, Strandburg-Peshkin, Clutton-Brock et al. 2020) or observational learning (Frith and
Frith 2012).

What component of social status explain highest trust in followers and what is their causal
link remain open questions. It is important to notice that in our study, we considered both the
affiliative and agonistic dimensions of the social status, that were correlated in our studied group, but it is not necessarily the case in horses (Cameron, Setsaas et al. 2009). There is a general tendency to oversimplify the social position of a leader to its dominance status can lead to miss the point. Bonanni et al. (2010) indeed demonstrated in free-ranging dogs that leadership is influenced by both affiliative and agonistic relationships. Affiliative relationships appear to be especially important for trust, as shown by chimpanzees that trust their friends more than their non-friends in a modified version of the trust game (Berg, Dickhaut et al. 1995, Tooby and Cosmides 1996, Engelmann and Herrman 2016). In the same way, Fruteau et al.’s study suggests that social centrality is more determinant than dominance’s rank to gain the trust of other group members (Fruteau, Voelkl et al. 2009). Indeed, in their experiment, the authors found than subordinate vervet monkeys were more groomed after providing food access to group mates, whereas no change in troop hierarchical order was reported. These results also suggest that a feedback loop between social status and trust (and thus tolerance to error) should exist even though it is difficult to determine its direction. In order to understand this causal relationship between social status and trust, we will need to manipulate again the level of social credibility of leaders by increasing the one of lower social status individuals.

It is important to note that, in our study, we only conducted six tests for each initiator and we focused on a temporary loss of reliability in a specific food context. This protocol was indeed designed to prevent any habituation or counterstrategies from group members since they might have decided to search for food reward on their own. Such process could be seen during the last trials when followership started to be disrupted even with control initiators. However, even with those limitations, we believe such extinction paradigm should be tested further, in more animals and more groups. Studying the contrast in extinction patterns is a good way to start quantifying this part of trust that animals put into a conspecific and that cannot be solely explained by trial and error mechanisms.

Finally, the flexibility that followers displayed in our experiments also raises interesting questions for the comprehension of equine social cognition. Indeed, followers appeared to adapt adequately their responses to contexts and initiators identities, suggesting flexibility when deciding which individual to follow and when. Such flexibility evokes the concept of social learning strategies, the “flexible rules that specify or bias when or how individuals should use social information, under various circumstances, to meet functional goals” (Kendal, Boogert et al. 2018). Social learning strategies likely involve both associative and social learning (Kendal, Boogert et al. 2018), and eventually sophisticated social cognition in some cases. The ability of horses to be selective and flexible in what, when, and from which individual they should get and use social information could explain why many authors struggled in demonstrating (observational) social learning in this species, despite the strong assumption they should possess this capacity (classic learning a simple operant: (Lindberg, Kelland et al. 1999), discrimination: (Baer, Potter et al. 1983, Baker and Crawford 1986, Clarke, Nicol et al. 1996), instrumental: (Ahrendt, Christensen et al. 2012), or detour spatial task: (Rørvang, Ahrendt et al. 2015, Burla, Siegwart et al. 2018). Interestingly, the only two studies that possibly showed premises of copying mechanisms suggest that a demonstrator’s influence depends on its social status (Krueger and Heinze 2008, Schuetz, Farmer et al. 2017).
The absence of clear experimental evidence of copying abilities contrasts with the knowledge we have of the horses’ non-social cognitive abilities (reviews: (Hausberger, Stomp et al. 2019)), and the existence of flexible and complex social phenomena that have been studied in more naturalistic conditions, such as in the context of — like in the present study - collective movement (Briard, Dorn et al. 2015, Briard, Deneubourg et al. 2017, Gérard, Valencon et al. 2020) - or third-party interactions and reconciliation (domestic horses (Cozzi, Sighieri et al. 2010), feral horses: (Schneider and Krueger 2012), and Przewalskii horses: (Krueger, Schneider et al. 2015)). Therefore, our study highlights the importance of including the flexibility and selectivity factors when considering social phenomena, and that the study of leadership using experimental paradigms constitutes a great framework to explore this topic.

References
Briard, L., J. L. Deneubourg and O. Petit (2017). "How stallions influence the dynamic of collective movements in two groups of domestic horses, from departure to arrival."
Behavioural Processes 142: 56-63.10.1016/j.beproc.2017.05.014


Gavrilets, S., J. Auerbach and M. van Vugt (2016). "Convergence to consensus in heterogeneous groups and the emergence of informal leadership." *Scientific Reports* 6(1): 29704.10.1038/srep29704


Acknowledgements
We warmly thank Thierry Bourgeois, Nans Burgarella, Rose Delacroix, Océane Liehrmann, Luc Martin, Virginie Nierat, Edouard Paumier, Louise Prévost, Julie Renard, Pierre Uhlrich, and Lucile Villot for helping with the experimental task, Cristian Pasquaretta for his guidance in statistics, and Rachel Annan for proofreading the manuscript. The project was funded by the University of Strasbourg Institute for Advanced Study and the Fondation des Treilles (Young Researcher Prize) that we thank for their support.

Animal welfare note
The experimental protocols followed the EU Directive 2010/63/EU guidelines for animal experiments and were approved by the Ethical Committee (CREMEAS) under agreement number AL/01/10/07/11.
### Table 1. Individual characteristics

<table>
<thead>
<tr>
<th>ID</th>
<th>Group</th>
<th>Year of study</th>
<th>Age (yrs)</th>
<th>Breed</th>
<th>David’s score</th>
<th>Centrality index</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAN</td>
<td>A</td>
<td>2014</td>
<td>2</td>
<td>Dartmoor</td>
<td>-5.46</td>
<td>0.52</td>
</tr>
<tr>
<td>DAY</td>
<td>A</td>
<td>2014</td>
<td>1</td>
<td>Shetland</td>
<td>-5.42</td>
<td>0.32</td>
</tr>
<tr>
<td>DLI</td>
<td>A</td>
<td>2014</td>
<td>23</td>
<td>Dartmoor</td>
<td>13</td>
<td>0.48</td>
</tr>
<tr>
<td>HAV</td>
<td>A</td>
<td>2014</td>
<td>19</td>
<td>Dartmoor</td>
<td>5</td>
<td>0.54</td>
</tr>
<tr>
<td>NOR</td>
<td>A</td>
<td>2014</td>
<td>18</td>
<td>Shetland</td>
<td>-1.78</td>
<td>0.17</td>
</tr>
<tr>
<td>TAK*</td>
<td>A</td>
<td>2014</td>
<td>7</td>
<td>Shetland</td>
<td>-5.33</td>
<td>0.28</td>
</tr>
<tr>
<td>ARM*</td>
<td>B</td>
<td>2014</td>
<td>4</td>
<td>Dartmoor</td>
<td>-11.57</td>
<td>0.55</td>
</tr>
<tr>
<td>CAL</td>
<td>B</td>
<td>2014</td>
<td>2</td>
<td>Dartmoor</td>
<td>-8.63</td>
<td>0.28</td>
</tr>
<tr>
<td>CHE</td>
<td>B</td>
<td>2014</td>
<td>2</td>
<td>Dartmoor</td>
<td>6</td>
<td>0.51</td>
</tr>
<tr>
<td>CYB</td>
<td>B</td>
<td>2014</td>
<td>2</td>
<td>Dartmoor</td>
<td>-2.71</td>
<td>0.36</td>
</tr>
<tr>
<td>OTO*</td>
<td>B</td>
<td>2014</td>
<td>12</td>
<td>Dartmoor</td>
<td>2.33</td>
<td>0.26</td>
</tr>
<tr>
<td>SER*</td>
<td>B</td>
<td>2014</td>
<td>8</td>
<td>Haflinger</td>
<td>14.57</td>
<td>0.41</td>
</tr>
<tr>
<td>ARM*</td>
<td>C</td>
<td>2015</td>
<td>5</td>
<td>Dartmoor</td>
<td>0.58</td>
<td>0.23</td>
</tr>
<tr>
<td>BEL</td>
<td>C</td>
<td>2015</td>
<td>4</td>
<td>Shetland</td>
<td>-9.00</td>
<td>0.09</td>
</tr>
<tr>
<td>OTO*</td>
<td>C</td>
<td>2015</td>
<td>13</td>
<td>Dartmoor</td>
<td>4.29</td>
<td>0.67</td>
</tr>
<tr>
<td>RAT</td>
<td>C</td>
<td>2015</td>
<td>10</td>
<td>Shetland</td>
<td>0.52</td>
<td>0.15</td>
</tr>
<tr>
<td>SER*</td>
<td>C</td>
<td>2015</td>
<td>9</td>
<td>Haflinger</td>
<td>15.00</td>
<td>0.68</td>
</tr>
<tr>
<td>TAK*</td>
<td>C</td>
<td>2015</td>
<td>8</td>
<td>Shetland</td>
<td>-15.00</td>
<td>0.07</td>
</tr>
</tbody>
</table>

* studied in both 2014 and 2015

### Table 2. Number of followers per initiation according to initiator’s identity for each group (A, B, C) during experiment 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Initiator</th>
<th>Initiation 1</th>
<th>Initiation 2</th>
<th>Initiation 3</th>
<th>Initiation 4</th>
<th>Initiation 5</th>
<th>Initiation 6</th>
<th>Initiation 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>SER</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OTO</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ARM</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAL</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CYB</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHE</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group B</td>
<td>CAN</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DAY</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DLI</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HAV</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NOR</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAK</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group C</td>
<td>ARM</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>BEL</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAT</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>SER</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAK</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Comparison of individual performances during the testing phase. For each combination of two initiators, the ratio of the proportion of individuals that followed one individual over the proportion of individuals that followed another individual has been tested.

<table>
<thead>
<tr>
<th>Individual A / Individual B</th>
<th>ratio</th>
<th>SE</th>
<th>z.ratio</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliable1 / Reliable2</td>
<td>0.7</td>
<td>0.4</td>
<td>-0.600</td>
<td>0.93</td>
</tr>
<tr>
<td>Reliable1 / Unreliable1</td>
<td>2.9</td>
<td>1.5</td>
<td>2.000</td>
<td>0.19</td>
</tr>
<tr>
<td>Reliable1 / Unreliable2</td>
<td>7.0</td>
<td>3.8</td>
<td>3.600</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Reliable2 / Unreliable1</td>
<td>4.1</td>
<td>2.3</td>
<td>2.500</td>
<td>0.06</td>
</tr>
<tr>
<td>Reliable2 / Unreliable2</td>
<td>10.1</td>
<td>5.8</td>
<td>4.000</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Unreliable1 / Unreliable2</td>
<td>2.4</td>
<td>1.2</td>
<td>1.800</td>
<td>0.25</td>
</tr>
</tbody>
</table>