Supporting Information

Marshall-Pescini et al. 10.1073/pnas.1709027114

SI Materials and Methods

Subjects. All available dog and wolf subjects were tested in all possible pairs. For the wolf–dog comparison, a subsample of wolf dyads was used to ensure that the wolf and dog dyads had comparable experiences: all completely naïve pairs of both species were included, and dyads in which one partner had attempted the task before with another partner but without success (Table S2). Because in some cases wolf (but not dog) dyads were composed of an individual who had been previously successful in another dyad, and this could have affected their success rates, these wolf dyads were not included in the dog–wolf comparison (see Table S1 for all wolf dyads tested in all conditions).

Apparatus and Test Location. The apparatus was a 1.5-m × 75-cm food delivery tray with a rope passing through loops in the tray. The apparatus was placed on one side of the fence, and the ropes were placed in such a way that the ends lay on the ground in the testing enclosure. The rope was approximately 520-cm long, with 120 cm dangling from each end of the apparatus. The apparatus was such that if only one rope was pulled, the tray could not move, but rather the rope slid out of the loop system and the nonpulled end of the rope became unavailable. Adjacent to each rope, 20-cm apart from each other, two food-delivery areas were set, each containing one dead chick and one chunk of raw meat. To move the tray toward them and hence successfully obtain the food, both subjects needed to pull the two ends of the rope at the same time.

Testing took place in the two main testing enclosures at the Wolf Science Center, Vienna. The starting location of the animals was on the opposite side of the testing enclosure, both animals placed within the same cubicle (see Fig. S1, showing the start location of each wolf, the location of the apparatus, and the location of the experimenter behind the barrier). Before running the test, the animals were allowed to explore the empty test enclosure for 5 min. During testing the experimenter was behind a screen, out of sight of the animals. Before the trial started the experimenter called the animals' attention and showed them the food being placed on the apparatus.

Procedure of Experimental Conditions in Detail.

Spontaneous condition. The food tray was placed outside the animals' reach, with the ropes going through the bars of the enclosure into the testing area. The members of the dyad were then simultaneously released into the enclosure. The animals' spontaneous behavior toward the apparatus and toward each other was recorded. Trials lasted 2 min, or ended either once the task was completed successfully and the animals had finished the food, or when the rope was pulled solely on one side, making the other end unavailable and hence making the tray impossible to move. At the end of each trial the animals were called back to the start position, while the experimenter set up the task again for the following trial. Following test trials, if an individual never pulled the rope during the trial, motivation string-pulling trials were presented (see details below). All dyads were given an average of two sessions per week for a total of six to eight sessions.

Sessions 1 and 2 consisted of six test trials in which the dyad was allowed access to the apparatus. In sessions 3–6, the number of test trials was no longer fixed but varied from between two and six trials, depending on the animal's behavior. If neither of the animals pulled the rope for two consecutive trials, the session was stopped to avoid demotivating the animals. All animals received motivation string-pulling trials following these sessions (see details below). If in session 6 animals were presented with fewer than four trials, two additional sessions were presented. In sessions 7 and 8 the animals again received a fixed number of six trials, independent of their performance, followed by motivation string-pulling trials if animals did not pull the ropes for two consecutive trials. Overall, the mean number of trials across animals was 36, with a range between 32 and 48 trials (Table S2).

Motivation string-pulling trials varied depending on session. In all cases, each animal was allowed in the enclosure without its partner. After sessions 1 and 2, a piece of meat was placed on the ground on the opposite side of the fence, out of direct reach of the animal. The meat was attached to a rope, which could be pulled by the animal to retrieve it. Following test trials in sessions 3 and 4 the food was placed on a wooden box outside the test enclosure, with the rope dangling within reach of the animal. These changes to the presentation were done to encourage the use of the mouth rather than the paw during string pulling. From sessions 5–8 in motivation string-pulling trials, the meat was placed on the test apparatus, alternating the position of the food between the two food slots, with the meat directly attached to the rope. These changes to the presentation were carried out to further encourage animals to view the apparatus as a potentially reinforcing object, even if they had had no success during testing. The number of motivation string-pulling trials depended on the animals performance in that as many trials were given as necessary for the animal to successfully retrieve the meat in three consecutive trials, with no prompting or encouragement.

Following the Spontaneous condition, dyads that had successfully obtained the food on at least four of six trials in the last two sessions conducted were presented with the Two-tray condition.

In packs where no dyad was successful, a training regimen (*SI Materials and Methods*) involving positive reinforcement was set up to allow animals to acquire the behavior of pulling both ends of the rope simultaneously when alone before testing them in dyads again (36).

Individual string-pulling training. The overall aim of the training was to facilitate animals' understanding of the mechanics of the task: that is, that pulling on a single rope would not allow them to obtain food, but pulling on both simultaneously would. Hence, to achieve this aim we adapted the procedures previously adopted by Seed et al. (33) with rooks and Ostojic and Clayton (37) with dogs.

In the first training phase: A 90-cm-long rope, folded in two so that the two ends were dangling from the end of the apparatus (the length of the dangling rope was 30 cm), was used. The ends of the rope were kept close together by cable ties placed at 10-cm intervals along the whole length of the rope. The rope was fixed in the middle of the apparatus so that the food tray moved forward (allowing subjects to obtain two pieces of meat) whenever the animals simultaneously pulled both ends of the rope. If animals reached criteria (see below) on this first step, the first cable tie (closest to the subject) was removed, so that the two ends dangled a few centimeters apart. In successive progressions we removed all of the cable ties, so that gradually the ends of the rope were further apart (the removal of the cables consisted of four training steps).

In the second training phase a small mesh (1-cm × 1-cm holes) was attached to the fence. This allowed us to have precise control of the distance between rope ends presented to the animals. In the first step of this second training phase, rope ends were first put through two holes next to each other (distance 0 cm). When the animals reached criteria (see below), they were presented with the rope ends 1-cm apart. We progressively increased the distance between rope ends until dogs successfully pulled the tray forward when presented with rope ends 6-cm apart.

In all cases (both for training phase 1 and training phase 2), the criterion to move on to the next step of training was that animals successfully pulled the platform in a minimum of five consecutive trials. However, if animals performed three consecutive unsuccessful trials, rather than continue with the current step, animals went back to the previous training step. Furthermore, after reaching criterion on the final step of each training phase, subjects had to successfully solve the step again on the next training session without going back to the previous phase before moving onto the next phase of the test (37).

Training was considered complete when animals successfully pulled the tray alone, in four of six trials with the rope ends 6-cm apart in two consecutive sessions. Having completed the training, individuals were paired with a partner who had also completed the training, and the dyads were tested in the Posttraining condition.

Two more steps were presented before retesting animals in pairs. For the first step, with the ropes 6-cm apart, the animals were released from the shifting system from the same location used during testing (i.e., on the opposite side of the enclosure in relation to where the apparatus was placed). Animals had to reach criteria (four successful trials of six) in this step to move onto the final step. The second step was identical to the previous step. However, in this case, as in test trials, the experimenter was behind a screen and not visible to the animals. This final step was carried out to ensure animals would still perform the string-pulling task independently (with no human social contact), after a long training period where the trainers were sitting next to the apparatus. Animals had to be successful again in four of six trials.

Posttraining condition. The Posttraining condition was identical to the Spontaneous condition, in which each dyad was presented with a single apparatus and both animals were simultaneously released into the test enclosure. A total of six sessions consisting of six trials per session were conducted. The same pass criterion of four of six successful trials in three consecutive sessions was set. If this criterion was reached the dyad was presented with the Two-tray condition.

Two-tray condition. Dyads that were successful in the Spontaneous or Posttraining condition were presented with two identical apparatuses placed 10-m apart from each other, in the same enclosure. The animals were released simultaneously from the start position as described above. A total of six sessions consisting of six trials each were carried out per dyad and again, trials lasted 2 min, or ended either once the task was completed successfully and the animals had finished the food or when the rope was pulled solely on one side, making the other end unavailable and hence making the tray impossible to move. A dyad was considered successful in a trial if they were able to obtain food from both trays. Regardless of performance in this condition, individuals were then presented with the Delay condition.

Delay condition. In this condition, a single tray with two food sources was presented; however, instead of simultaneously opening the enclosure door for the two animals, partners were placed in adjacent enclosures and the subject was released 10 s before the partner. A total of six sessions each consisting of six trials was presented.

Behavioral Analyses.

Assessment of dyad's relationship. Daily observations of all pack members were carried out at the Wolf Science Center by long-term students, following a training period and interrater reliability assessment with senior staff. Ten-minute focal animal sampling, focused on the social interactions with other pack members, was carried out for each member of a pack during different times of day (Table S3). The ethogram adopted comprises behaviors which have been found to be good indicators of dominance in a number of studies in both wolves and dogs (53–55).

The Pocket Observer program (3.2 Software) was used for data collection, then imported into the Observer XT 10.5 program (both from Noldus Information Technology) for further analyses.

Based on such observations, we calculated: (*i*) an "affiliation score" for each dyad: that is, the bidirectional frequency of affiliative behaviors exchanged by individuals A and B, divided by observation time (hours) for subjects A + B; and (*ii*) the rank distance between members of the dyad: that is, subtracting the David's score of individual A from that of individual B. The David's score for each individual was calculated based on the frequency of dominant and submissive behaviors displayed toward other pack members. The David's score is considered the most accurate measure of an individuals' dominance status within the pack, since it takes into account the relative strength of all pack members, thereby also allowing an assessment of relative strength across groups with different numbers of members within (56).

No significant difference emerged between the wolves' and dogs' affiliation score ($\chi^2 = 0.3$, $P = 0.58$) and rank distance ($\chi^2 = 1.04$, $P = 0.31$) and no correlation emerged between rank distance and affiliation score for wolves ($R = 0.27$, $P = 0.395$).

String-pulling success. In all conditions, a trial was considered to end in a successful cooperative interaction if the dyad succeeded in obtaining the food by both individuals simultaneously pulling on the rope ends so that the tray moved forward sufficiently for them to reach the food. On a number of occasions (fewer than five trials overall), an individual obtained a piece of food by reaching their paws or snout through the fence and moving the tray forward enough to obtain the food; this was not considered as a success and the trial in such cases was repeated.

Behavior coding of session 1 of the Spontaneous and Posttraining conditions. The first session of both the Spontaneous and Posttraining conditions were observed on video and a detailed coding of the animals' behaviors was carried out (see Table S4 for ethogram). For analyses purposes, dominant and aggressive behaviors were summed in a single category, as were submissive and withdrawing behaviors. Interobserver reliability was carried out on 20% of the data, and was found to be high for all behaviors (all Spearman's $\rho > 0.93$). Due to video malfunction, the Spontaneous and Posttraining sessions for one dyad (Wamblee–Yukon) could not be coded in detail; therefore, this dyad is not included in these analyses.

Statistical Analyses.

Wolf–dog comparison. To compare the success rate of wolves and dogs in the task, we ran a GLMM with proportion of successful trials over all trials (since this could vary between dyads), including trials from both the Spontaneous and the Posttraining conditions, since we had a comparable number of wolf and dog dyads in both conditions (Table S2). Species and condition (and their interaction) were included as explanatory factors and dyad as the random factor (since some dyads were tested both in the Spontaneous and Posttraining conditions).

We further evaluated whether wolves and dogs showed a different pattern of behaviors during session 1 (i.e., when all animals were comparable in their experience with the apparatus). Hence, considering only session 1 for both Spontaneous and Posttraining conditions, we ran a number of GLMM models with the frequency (pulling the rope, nonfunctional behaviors, gaze alternation from partner to apparatus) and durations (time spent in proximity of the apparatus) of specific behaviors (normalized by trial duration) as the dependent factor, the species and condition (and their interaction) as the explanatory factor and the dyad as random factor. We additionally ran a GLMM model looking at whether species and condition (and interaction) affected the (*i*) latency it took both animals to be in proximity of the apparatus and (*ii*) the likelihood they would display dominant/aggressive interactions in a trial (binomial distribution). Finally, we assessed whether species had an effect over all trials in the likelihood that both individuals would simultaneously manipulate the apparatus (display nonfunctional behaviors) in a trial (binomial distribution). For this, we ran a GLM with the number of trials a dyad was shown to be simultaneously active on

the apparatus corrected for the total number of trials that dyad was tested in (since this could vary) as the dependent variable and species as the explanatory factor.

Factors affecting wolf dyad cooperative success. To evaluate performance across conditions for all wolf dyads (Table S1), we ran a GLMM with rate of success as the dependent variable, condition as the explanatory factor, and dyad identity as the random factor.

To assess potential learning effects within each condition (Spontaneous/Posttraining conditions, Two-tray and Delay), we ran a GLMM with rate of success as the dependent variable, session (including only the first six sessions of the Spontaneous condition, since these were carried out by all dyads) as the explanatory variable, and dyad as the random factor.

To evaluate the potential effect of the animal's social relationship on the dyad's performance, we considered data from the Two-tray condition for two reasons: (*i*) this condition better controls for the different levels of experience of the animals since they must all have had a measure of success in the Spontaneous condition before being tested in the Two-tray condition; and (*ii*) this condition requires a greater level of coordination between partners, since they need to synchronize their action both in space and time to successfully obtain food from both apparatuses. The latter level of coordination is more likely to be sensitive to the type of relationship between individuals. Accordingly, we ran a GLMM with the rate of success (i.e., number of trials in which animals solved both apparatuses) as the dependent measure, the dyad's affiliation score and rank distance as the explanatory variables, and the dyad identity as a random factor.

Finally, we analyzed the dyad's performance in the Delay condition, assessing whether the success rate of a dyad in the Spontaneous and Two-tray conditions affected the subject's success in the Delay condition. We ran a GLM with the individuals' rate of success in the Delay condition as the dependent variable, and the dyads' percentage of successful trials in the Spontaneous and Two-tray conditions as explanatory variables.

All analyses were carried out in R v3.2.2 (57) package "lme4." Model assumptions were checked and corrected for if not met. A model reduction procedure based on *P* values was adopted, starting from dropping interactions between factors if these were found not to be significant.

Fig. S1. Test set-up in the Spontaneous/Posttraining conditions (*A*) and the two-tray (*B*) condition. Figure courtesy of Akshay Rao and Corinna Kratz (Wolf Science Center, Ernstbrunn, Austria).

Table S1. All wolf dyads (sex combination and pack membership) tested in all conditions

Pack	Dyad	Sex	Spontaneous total_trials (success)	Posttraining (P)/retesting (R) total_trials (success)	Two-tray total_trials (success)	Delay total_trials (success)
1	Kaspar_Shima	M–F	45 (25)	NA	42 (11)	Kaspar 36 (23) –Shima 36 (3)
1	Kaspar_Tala	M–F	36 (34)	NA	48 (24)	Tala 36 (22)
1	Aragorn_Chitto	M–M	35 (27)	NA	36 (29)	Chittto 36 (20)
1	Aragorn_Shima	M–F	36 (23)	NA	36 (3)	Aragorn* 36 (8)
1	Kaspar_Aragorn	M–M	36 (31)	NA	36 (4)	Aragorn* 36 (26)
1	Tala_Shima	F–F	36 (10)	NA	36 (14)	NA
1	Kaspar_Chitto	M–M	36 (7)	R-36 (36)	36 (36)	NA
1	Chitto_Shima	M–F	37 (3)	R-36 (32)	36 (36)	NA
1	Chitto_Tala	M–F	34 (3)	R-36 (36)	36 (36)	NA
1	Aragorn_Tala	M–F	33 (2)	R- 36 (36)	36 (31)	NA
2	Amarok_Kenai	M–M	32 (1)	P-36(0)	NA	NA
2	Geronimo_Amarok	M–M	37 (0)	NA	NA	NA
2	Geronimo_Kenai	M–M	36 (0)	NA	NA	NA
3	Geronimo_Yukon	M–F	NA	P-36 (28)	36 (31)	Geronimo 36 (33) –Yukon 36 (27)
3	Wamblee_Yukon	M–F	40 (2)	P-36 (5)	NA	NA
4	Nanuk_Una	M–F	48 (3)	P-36 (33)	36 (28)	Nanuk 36 (34) –Una 36 (32)

F, female; M, male. Retesting refers to dyads that failed the first time they were tested together; each individual was tested with another partner, and then retested with the initial partner. Individuals were tested in the Delay condition only once, with the first partner with whom they were successful. The exception to this was Aragorn*, who was first tested with Shima and performed rather poorly. Upon inspection of the movies, it became obvious that his poor performance was attributed to Shima; he was tested again in the Delay condition with Kaspar.

Table S2. Wolf and dog dyads (sex combination and pack membership) compared in their success rates when tested in the Spontaneous and Posttraining conditions, including number of trials completed by each dyad, number of successful trials, and number of failures due to one individual pulling out the rope (the default failure being a time-out)

Pack	Dyad	Sex	Species	Condition	Total trials	No. success trials	No. rope-out fails	Condition	Total trials	No. success trials	No. rope-out fails
1	Nuru-Layla*	M–F	Dog	Spontaneous	40	0	2	NA			
1	**Zuri-Layla***	F–F	Dog	Spontaneous	36	0	0	NA			
1	**Nuru_Zuri**	M–F	Dog	Spontaneous	32	0	1	Posttraining	36	0	0
2	Asali_Bora	M–F	Dog	Spontaneous	38	0	3	Posttraining	36	0	22
3	Maisha_Binti	M–F	Dog	Spontaneous	32	0	1	Posttraining	36	0	0
4	Meru_Nia	M–F	Dog	Spontaneous	33	0	1	Posttraining	36	1	6
2	Asali_Banzai	M–F	Dog	NA				Posttraining	36	0	20
2	Bora_Banzai	F–F	Dog	NA				Posttraining	36	0	11
5	Imara_Hiara	F–M	Dog	Spontaneous	36	1	26	NA			
6	Sahibu_Gombo†	M–M	Dog	Spontaneous	9	0	6	NA			
1	Kaspar_Shima	M–F	Wolf	Spontaneous	45	25	14	NA			
1	Chitto_Tala	M–F	Wolf	Spontaneous	34	3	23	NA			
2	Geronimo‡_Amarok	M–M	Wolf	Spontaneous	37	0	26	NA			
2	**Geronimo‡_Kenai**	M–M	Wolf	Spontaneous	36	0	20	NA			
2	**Amarok_Kenai**	M–M	Wolf	Spontaneous	32	1	7	Posttraining	36	0	22
3	Wamblee_Yukon	M–F	Wolf	Spontaneous	40	2	26	Posttraining	36	5	9
3	Geronimo‡_Yukon	M–F	Wolf	NA				Posttraining	36	28	7
4	Nanuk_Una	M–F	Wolf	Spontaneous	48	3	24	Posttraining	36	33	2

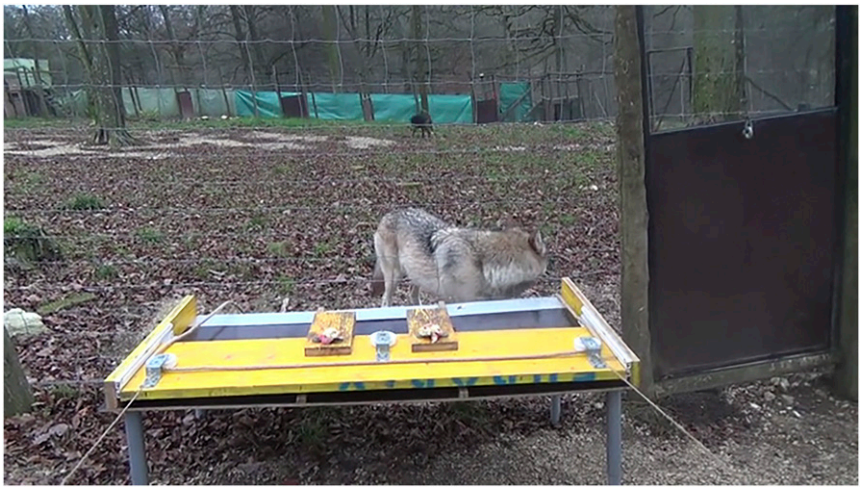
In bold are dyads tested in the Spontaneous condition, where one or both partners had been tested with a previous partner but unsuccessfully.
*Layla, when being trained in the individual training condition, stopped showing any interest in the task and was dropped from the study.
†Testing for this dyad was interrupted after only nine trials because during the last three trials they showed an escalation in the intensity of threatening behaviors toward each other.
‡After being tested in the Spontaneous condition, Geronimo was removed from the pack due to an illness. After recovery it was no longer possible to integrate him with his former pack mates; hence, it was not possible to retest him in the Posttraining condition with Amarok and Kenai. Geronimo was, however, tested with his new pack-mate Yukon.

Table S3. Relevant subset of the ethogram used at the Wolf Science Center to allow evaluation of partner relationship in terms of affiliation and rank distance

Behavior	Definition
Dominance behaviors	
Stand tall	Subject straightens up to full height, with a rigid posture and tail, may include raised hackles, ears erect and tail perpendicular or above the back.
Paw on	To place one or both forepaws on the other's back.
Ride up	To mount another one from behind or from the side, exhibiting a thrusting motion with the hips.
Head on	The subject approaches another's shoulder/back and puts its head on it.
Muzzle Bite	To grasp the muzzle of another subject either softly or with enough pressure to make the other whimper.
Approach dominant	To approach another subject within one body length for at least 5 s, with the tail perpendicular or above the plane of the back and the ears erect and pointed forward and with a rigid posture.
Submissive behaviors	
Crouch	Lowering the head, sometimes bending the legs, arching the back, lowering the tail between the hind legs, and avoiding eye contact.
Passive submission	To lie on the back showing the stomach and holding the tail between the legs. The ears are held back and close to the head and the subject raises a hind leg for inguinal presentation.
Active submission	The subject has its tail tucked between the hind legs sometimes wagging it while he is in a crouched position (with hindquarters lowered) and may attempt to paw and lick the side of actors'/aggressor's muzzle. The behavior may also include urination.
Withdrawing	The subject withdraws from another moving away slowly in the opposite direction, displaying a submissive posture. It occurs when a subject has been threatened or attacked by another, or a fight has taken place.
Flee	To run away from another with tail tucked between the legs and body ducked. It occurs when a subject has been threatened or attacked by another, or after a fight.
Avoidance	In response to another reducing the distance toward it, the subject moves away displaying a submissive posture. The subject may also look at the individual he is trying to avoid.
Approach submissive	To slowly approach another within one body length remaining within that distance for at least 5 s. The approach is characterized by a ducked posture and tail between the legs. Subject can also be moving in a wavy line and in a hesitant (stop–start) manner.
Affiliative behaviors	
Grooming	To nip, lick or scratch the fur or skin of another.
Lie friendly	To lie on the back, tail-wagging, maybe kicking with the foreleg against another subject sometimes with open mouth.
Stand friendly	The subject stands with tail perpendicular to or below the plane of the back, wagging it, ears pointed forward, while another is approaching it or orienting/looking toward it.
Body contact	Two subjects stay (for at least 10 s) with at least a part of their bodies in contact and in a relaxed position.
Social sniff	To sniff another's body part except its anogenital area.
Body rubbing	To rub one's body against any part of the receiver's.

Table S4. Ethogram used for the behavior coding of session 1 of Spontaneous and Posttraining conditions

Behaviors	Measure	Description
Dominant	Frequency	Stand tall, paw on, ride up, head on, muzzle bite
Aggressive	Frequency	Growl, snarl, snap, lunge
Submissive	Frequency	Crouch, active and passive submission
Withdraw	Frequency	Subject is within two body lengths of the apparatus and withdraws to outside of two body lengths from the apparatus within 2 s of partner's approach or following an dominant/aggressive interaction
Close apparatus	Duration	Both partners simultaneously within one body length of the apparatus
Pull rope	Frequency	Subject pulls the rope (taut, not with slack), so that a movement in the rope is detected (not just holding in the mouth)
Nonfunctional behaviors on tray	Frequency	Subject bites, paws or scratches at the tray or the fence immediately in front of it
Gaze apparatus-partner	Frequency	Changing head orientation from the apparatus to the partner or vice versa



Movie S1. Wolf dyad tested in the Spontaneous condition (session 6, trial 1).

Movie S1

Movie S2. Dog dyad tested in the Posttraining condition (session 1, trial 2).

Movie S2

Movie S3. Wolf dyad tested in the Delay condition (session 6, trial 3).

Movie S3

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)