Published in final edited form as:

Anim Cogn. 2012 July ; 15(4): 597–607. doi:10.1007/s10071-012-0488-8.

Domestication has not affected the understanding of means-end connections in dogs

Friederike Range,

Messerli Research Institute, University of Veterinary Medicine Vienna, Medical University of Vienna, University of Vienna, Veterinärplatz 1, 1210 Vienna, Austria; Wolf Science Center, Dörfler 48, 2115 Ernstbrunn, Austria

Helene Möslinger, and

Wolf Science Center, Dörfler 48, 2115 Ernstbrunn, Austria

Zs Virányi

Messerli Research Institute, University of Veterinary Medicine Vienna, Medical University of Vienna, University of Vienna, Veterinärplatz 1, 1210 Vienna, Austria; Wolf Science Center, Dörfler 48, 2115 Ernstbrunn, Austria

Abstract

Recent studies have revealed that dogs often perform well in cognitive tasks in the social domain, but rather poorly in the physical domain. This dichotomy has led to the hypothesis that the domestication process might have enhanced the social cognitive skills of dogs (Hare et al. in *Science* 298:1634–1636, 2002; Miklósi et al. in *Curr Biol* 13:763–766, 2003) but at the same time had a detrimental effect on their physical cognition (Frank in *Z Tierpsychol* 5:389–399, 1980). Despite the recent interest in dog cognition and especially the effects of domestication, the latter hypothesis has hardly been tested and we lack detailed knowledge of the physical understanding of wolves in comparison with dogs. Here, we set out to examine whether adult wolves and dogs rely on means-end connections using the string-pulling task, to test the prediction that wolves would perform better than dogs in such a task of physical cognition. We found that at the group level, dogs were more prone to commit the proximity error, while the wolves showed a stronger side bias. Neither wolves nor dogs showed an instantaneous understanding of means-end connection, but made different mistakes. Thus, the performance of the wolves and dogs in this string-pulling task did not confirm that domestication has affected the physical cognition of dogs.

Keywords

Domestication; Means-end connections; Physical cognition; Dogs; Wolves; *Canis familiaris*; *Canis lupus*

Introduction

Intense research in the last few decades focusing on dogs' social interactions and communicative skills has shown that dogs (*Canis familiaris*) are extremely good at reading human social and communicative behavior (Gácsi et al. 2004; Miklósi and Soproni 2006; Szeteci et al. 2003; Virnsányi et al. 2004, 2006). In contrast to the social-communicative domain, however, dogs usually perform rather poorly in the asocial, physical domain. For instance, dogs solve the invisible displacement task in object permanence studies by using simple local rules, like the adjacency rule (Collier-Baker et al. 2004; Watson et al. 2001; Fiset and Leblanc 2007). Similarly, they show a gravity bias (Osthaus et al. 2003a), fail to use means-end connections (Osthaus et al. 2005; but see Range et al. 2011) or to infer the location of food if only indirect causal information is available (Brauer et al. 2006). Based on their advanced cognitive abilities in the social domain and their poor performance in the physical domain, it has been proposed that the domestication process increased the social cognitive skills of dogs (Hare et al. 2002; Miklósi et al. 2003; Udell et al. 2010; see also Wynne et al. 2008) but had a detrimental effect on their physical cognition (Frank 1980). The rationale for the latter is that, in contrast to wolves (*Canis lupus*), environmental effects on feeding and mating are buffered in dogs (and in other domesticated species) by humans and thus, natural selection on dogs' individual problem solving might have been relaxed (Frank 1980; Hemmer 1990).

Despite the recent interest in dog cognition and especially in the effects of domestication, we lack detailed knowledge of the physical understanding of wolves needed to test this hypothesis. One exception is the studies by Frank and colleagues, who tested four wolf and four malamute pups raised similarly in different problem solving and manipulation tasks at the age of 6 and 10 weeks (Frank and Frank 1982, 1985; Frank et al. 1989). They found that the wolves were more successful than the dogs in these tasks. However, the sample size was small, the wolves (and dogs) were from the same litter, and the animals were tested at a relatively young age, opening the possibility that individual differences, genetic, or developmental effects could have influenced the outcome of these studies. Here, we set out to test whether adult wolves and dogs recognize means-end relationships by testing the prediction that wolves would perform better than dogs in a string-pulling task.

Animals often encounter problems that force them to progress through a series of mediating actions in order to reach a certain goal at the end. When performing a goal-directed behavior, understanding this progressive series of steps as the means to an end can be of great advantage for any individual (Schmidt and Cook 2006). Accordingly, means-end understanding is considered to be a key step in cognitive development (Bratman 1981; Osthaus et al. 2005). Usually, relying on means-end connections is studied by offering a choice of two options, one of which is in physical connection to an out-of-reach object of desire. One widely used task is the "string-pulling task," originally introduced by Köhler (1927), where a string within reach of the subject is attached to a piece of food out of reach of the subject. The question is whether the subject realizes that the goal object can be obtained by pulling on the string. The assumption is that subjects who do not understand that pulling the string is the "means" for achieving the desired "ends" of bringing the toy or food

within reach, will only be able to succeed through repeated exposure and associative learning (Thorndike 1898).

So far, non-human primates (e.g., chimpanzees, *Pan troglodytes*, Köhler 1927, Povinelli 2000) and birds (*Corvus corax*, Heinrich 1995, Heinrich and Bugnyar 2005; spectacled parrotlets, *Forpus conspicillatus*, Krashennikova and Wanker 2010, African gray parrot, *Psittacus erithacus*, Pepperberg 2004, blue-fronted amazons, *Amazona aestiva*, hyacinth macaws, *Anodorhynchus hyacinthinus*, Lear's macaws, *Anodorhynchus leari*, Schuck-Paim et al. 2009, keas, *Nestor notabilis*, Werdenich and Huber 2006) have solved at least some versions of the string-pulling task. While dogs and wolves can learn to retrieve out-of-reach food by means of an attached string (Miklósi et al. 2003; Osthaus et al. 2005), dogs have failed to show understanding of means-end relationships when tested in a 2-choice paradigm involving strings with varying angle orientations as well as crossed strings (Osthaus et al. 2005). The main reason for the failure of the dogs in these tasks was a proximity error, i.e., the dogs pawed or mouthed at the location closest in line to the treat. This failure of dogs may reflect an inability to recognize means-end connections. Alternatively, inherited predispositions to go for food directly may overshadow the recognition of means-end connections, and in combination with the inability to inhibit this response, could lead to the proximity bias of dogs (Lea et al. 2006). This interpretation has also been suggested for dogs' performance in other test paradigms (e.g., detour task) (Pongrácz et al. 2001; Scott and Fuller 1965). Dogs may have inherited this predisposition (to focus on proximity in regard to potential food) from their ancestor or it might be an effect of domestication as outlined above.

In the current study, we set out to test these hypotheses by comparing the performance of well-socialized wolves and kennel kept sledge dogs in a series of string-pulling experiments. In each of the experiments, the subjects needed to retrieve a food treat that was visible behind a physical barrier (fence) and was connected to a string within their reach. Subjects were confronted with either one or two strings, of which only one was attached to the food treat, presented in different layouts with increasing difficulty (parallel, angle, and crossed). Since several studies have shown that primates and dogs have difficulties solving the crossed condition (Tomasello and Call 1997; Povinelli 2000; Osthaus et al. 2005), we did not counterbalance the sequence of experiments in order to avoid the risk that failure with an insoluble task would demotivate the animals and prevent them from paying close attention to later conditions. Based on the published results by Osthaus et al. (2005), we predicted that while dogs would still be able to solve the easy tasks, they would fail with the more complicated versions. Wolves, on the other hand, were expected to solve even the more difficult versions of the string-pulling tasks, if the failure of dogs is at least partly due to a detrimental effect of domestication on their physical cognition. However, if canines in general are prone to commit the proximity error, we would expect no difference in the performance between wolves and dogs.

Methods

Subjects

All wolves ($n = 9$) that participated in this study originated from North America and were born in captivity. Six wolves (3 males and 3 females) were housed at Wolf Park (<http://www.wolfpark.org>), USA, and three wolves (2 males and 1 female) at the Wolf Science Center (<http://www.wolf-science.at>), Austria. All of the wolves were hand-raised in peer groups after being separated from their mothers in the first 10 days after birth. They were bottle-fed and later hand-fed by humans and had continuous access to humans for the first 3–5 months of their life. From this age on, there were no humans continuously present in the enclosures, but the wolves participated in training and educational programs (Wolf Park) and cognitive and behavioral experiments (Wolf Science Center) several times per week and hence had intensive social contact with humans. The wolves were kept in packs in outside enclosures up to 2,000–8,000 m² (Wolf Park) and 3,000 m² (Wolf Science Center). The enclosures were equipped with trees, bushes, logs, and shelters. Water for drinking was permanently available. The wolves received a diet of meat, fruits, milk products, and dry food throughout the study period.

The sledge dogs ($N = 10$), Siberian Huskies, were housed at the Mountain Wolf Farm, Austria. All dogs were raised by their mother until the age of 8 weeks. At the age of 8 weeks, all dogs but one was kept in packs of three to five animals in enclosures of up to 40–90 m². The single dog was kept in the house until 3 months of age before being transferred to the enclosures. From that age on, all dogs received training sessions twice a week. At the age of 8 months, they started to pull a sledge. Due to their training and feeding routine, the dogs received social contact with humans on a daily basis. All enclosures were equipped with shelters. The dogs were fed with high-energy dry food every day. Water was freely available.

The dogs had never participated in any behavioral or cognitive task. The wolves had some experience with participation in cognitive tasks, but not with string-pulling paradigms or any other instrumental means-end task. All animals were regularly walked on a leash and were separated from their pack mates for short periods of time. Both practices are part of the usual animal care procedures at Wolf Park, the Wolf Science Center, as well as at the Mountain Wolf Farm. Participation in all testing sessions was voluntary. For the sex and age distribution as well as relatedness among all the participating animals, please see Table 1.

Materials

We used a dark brown wooden tray (62 × 72) and beige colored ropes (60 cm long) for the experiments. As a reward, we used yellow colored pieces of cheese (1 cm thick, 2 cm wide and 6 cm long). Although the contrast between the board and the ropes and rewards was very good, all ropes were of the same color in order to avoid the animals choosing based on color association instead of connectivity. The rope(s) (60 cm long) were laid out on the outside of the enclosures in front of the gate at a distance of 20 cm. A 5 cm gap between the ground and the gate of the enclosures allowed for pushing the board underneath the gate, so that the wolves and the dogs could reach the end of the rope(s) (see Fig. 1 for the setup).

Experimental design

Due to the limited sample size of wolves, we used a within subject design, with every animal being tested in four experiments, in increasing order of their expected difficulty. To minimize learning effects across the different experiments, we limited the number of trials for Experiment 1–6 trials, for Experiment 2 and 5–20 trials, and for Experiment 3 and 4–24 trials. Fewer trials were conducted if an animal reached significance at the individual level (Binominal test: $P < 0.05$) before reaching the maximum number of trials, to avoid the risk that they would fixate on a particular strategy if allowed to carry on. Even if an animal did not solve a particular experiment within the maximum number of trials, it proceeded to the next experiment since it could have been possible that either (1) the animals learned over trials to solve the string-pulling task or (2) that against our expectations, a later experiment was easier for the animals to solve than an earlier experiment. The last, 5th experiment repeated the second experiment to test whether the performance of the animals had changed during multiple testing.

Experimental procedure

The experiments were conducted outside either in a familiar small holding enclosure, physically and visually separated from the main enclosure, or within the home enclosure with only the subject present at the time of the experiment. The wolves and dogs showed no distress at being separated from the other pack members. During the entire experiment, the subject was allowed to move freely around in the enclosure. A member of staff remained at the back of the enclosure facing away from the experimental setup and retrieved the ropes after each trial.

Habituation phase—Prior to testing, each subject was given the opportunity to play with the ropes to help decrease their motivation to play during the experiments. The habituation was terminated when the subject ceased to have physical contact with the ropes for 2 min. To habituate the wolves and dogs to the moving of the tray, after placing food (without the ropes) on it, the tray was moved toward the subject. Once they had taken the food, the tray was removed. Experimental trials were conducted only after the animals did not display a startle response to the movement of the tray, which took only a few trials for all animals tested.

Experimental trials—At the beginning of each trial, the rope(s) were laid out on the wooden tray according to the condition. The animal was prevented from watching this procedure by an opaque barrier placed in front of the gate. Once the rope(s) were in place (2 marks at the end at the tray equidistance from the centerline indicated where the food should be placed), the barrier was removed and the experimenter called the attention of the subject toward the arrangement by calling its name. During this presentation, the experimenters' right hand was placed on a handle in the middle of the tray and the other hand was kept behind her back. When the subject had looked at the arrangement on the tray for 3 consecutive seconds through the wire mesh of the gate, the experimenter moved the tray forward, so that the end of the rope(s) could be reached by the subject. The experimenter never looked directly at the subject; the experimenter monitored the animal's response out of the corner of her eye, while looking straight down at the ground. The tray and the second

rope were removed as soon as the subject grabbed one of the strings and started pulling it. If the subject chose the correct string, it was allowed to eat the food.

All experiments were based on the voluntary participation of the animals; thus, if an animal left the testing area, we stopped the experimental session and continued on another day. Each subject was tested for one session per day with a maximum of 32 trials (range 1–32). All trials were videotaped with a camera placed on a tripod behind the experimenter.

Analyses

Videos were analyzed using Solomon Coder beta 090626© Andras Peter. In order to assess the motivation of the animals to participate in each task and to get the food in all experiments, we coded the latency of starting to pull one of the ropes, defined as the time when the experimenter started to move the board forward within reach of the animals until the subject moved one of the ropes either with the mouth or paw toward its direction. This first pull on one of the ropes was considered the choice of the animal and was either correct if it was the baited rope or incorrect if it was the non-baited rope. In Experiment 1, we also coded if the animal was pawing close to the food defined as the subject pawing while its head was positioned above the baited side of the board.

Test videos were coded by one of the authors. A second coder, who was blind to the experimental questions, coded 20 % of the trials. The two observers agreed on the choice of the animals in 385 of 389 trials (99 %) that were coded across all conditions. Spearman rank correlations revealed high inter-observer reliability in the latency to pull the first rope in all conditions ($r_s = 0.86$; $N = 388$; $P < 0.001$).

In each experiment, we first analyzed the complete dataset (see below for details). If we found a significant interaction of species with any of the other predictors, we split the dataset and analyzed the two species separately in order to explore the nature of the interaction. If we found a main effect of condition but no interaction, we pooled the data of wolves and dogs and conducted post hoc tests to determine which conditions differed.

In order to determine whether species or learning over trials influenced the performance of the animals in Experiments 2 and 5, we fitted the success as the binomial response term in a generalized linear model. The species (wolves vs. dogs) and trial number (1–20) were fitted as fixed effects with animal identity as a random factor. Furthermore, to analyze whether species learning over the trials or experimental condition influenced the performance of the dogs and wolves in Experiments 3 and 4, we fitted the number of correct trials as the binomial response term in a generalized linear model with the total number of trials as the denominator. The species (wolves vs. dogs), learning (first six trials vs. last six trials), and condition (overlap vs. control) were fitted as fixed factors with animal identity as a random factor.

According to the Kolmogorov–Smirnov test, the behavioral data were not normally distributed. Consequently, we used nonparametric statistics to compare the performance of dogs and wolves calculated as the percentage of successful trials to chance level or to certain other conditions. All tests were two-tailed, and alpha was set at 0.05. When we analyzed

subsets of data, probabilities were corrected by a sequential Bonferroni procedure (Hochberg 1988). Statistical analyses were performed in R 2.10.1 (R Development Core 2009) and Instat 3.

Experiment 1: One diagonal rope

In Experiment 1, wolves and dogs were tested for their ability to use a single rope to pull in food spontaneously. Performance was measured by the animals' latency to pull the string into the enclosure, so that they could reach the food reward and eat it. A single rope with food attached to the end was laid out diagonally (Fig. 2a). Based on Osthaus's et al. (2005) results, we expected the dogs to paw the ground at the closest position in respect to the food reward, before actually trying to pull the food close with the rope. Our first question was whether wolves are less prone than dogs to this proximity error in this simple situation. Each animal received 6 trials in a row; the direction (left-to-right and right-to-left trials) was pseudo-randomized, so that no more than 2 trials in a row were conducted in the same direction.

Results

Overall, all animals solved the task easily. Before doing so, however, they often committed the proximity error. Over the six trials, 4 of the 9 wolves pawed at the location closest to the food at some point, although never in the very first trial. Of those 4 animals, three were the younger, one-year-old animals tested at the Wolf Science Center. In comparison, 8 of the 10 dogs pawed at the location closest to the food reward—three in the very first trial. The difference found between the number of animals committing the proximity error was not significant (Fisher's exact test: $P = 0.169$). We also found no significant difference between wolves and dogs when comparing the number of mistakes they made (Mann–Whitney U test: $U = 55.5$, $N_w = 9$, $N_d = 10$, $P = 0.400$).

We found no difference in the latency to pull in the food reward in the first trial, or its mean over all trials, between wolves and dogs (first trial: Mann–Whitney U test: $U = 38.5$, $N_w = 9$, $N_d = 10$, $P = 0.623$; mean of all trials: Mann–Whitney U test: $U = 44.5$, $N_w = 9$, $N_d = 10$, $P = 0.999$), indicating that both species had similar levels of motivation.

Experiments 2–5

To further investigate whether the wolves and dogs used the rope in Experiment 1 simply because this was the only accessible object without recognizing its connectedness to the food, the further experiments employed two ropes, one with food attached to the end and one without. Simply pawing at any accessible rope would provide a random success rate, whereas if the animals understood the connection via the rope, they should pull in only the baited rope.

Experiment 2: Two parallel ropes, one connected to food, two food rewards

The aim of Experiment 2 was to test whether wolves and dogs can solve a 2-choice string-pulling task taking the means-end relation into account, if proximity of the food reward cannot influence the solution. Accordingly, we used two ropes laid out perpendicular to the

gate (20 cm distance between the ropes), each with a reward at the end. However, one of the two ropes was shorter and not connected to the food reward (10 cm gap) (Fig. 2b). The position of the connected rope (left/right) was pseudo-randomized with the restriction that it could not be presented more than 3 times in a row on the same side. We conducted 6–20 trials depending on whether the animal performed significantly above chance in fewer than 20 trials based on a binominal test. However, after the maximum of 20 trials, whether successful or not, the subjects proceeded to the next experiment.

Results—All animals participated in this experiment and, except one wolf, received all 20 trials. One wolf successfully solved the first 6 trials (Binomial test, $P = 0.03$) and was not tested any further in this experiment. No other wolves and dogs were successful at the individual level. When analyzing the number of successful pulls, we found a significant effect of trial number and a significant interaction between trial number and species, suggesting that while wolves and dogs did not differ significantly from each other in solving the overall task, the wolves as a group improved their performance over the 20 trials, whereas the dogs' performance worsened (Table 2). Accordingly, further analyses revealed that while the performance of the wolves (including the animal that solved the task in 6 trials) did not differ significantly from chance performance in the first 10 trials (Wilcoxon signed-rank test: $T^+ = 17.5$, $N = 9$, $P = 0.578$), it increased significantly above chance in the last 10 trials (Wilcoxon signed-rank test: $T^+ = 28$, $N = 8$, $P < 0.016$) (Fig. 3). The performance of the dogs also did not differ from chance in the first 10 trials (Wilcoxon signed-rank test: $T^+ = 19$, $N = 10$, $P = 0.469$); however, in the last 10 trials, their performance decreased significantly below chance (Wilcoxon signed-rank test: $T^+ = 4$, $N = 10$, $P = 0.014$) (Fig. 3), suggesting that the dogs developed a preference for the short, unconnected string.

Finally, we found no significant difference in the latency to start pulling a string between the wolves and the dogs in the first trial (Mann–Whitney U test: $U = 29$, $N_w = 9$, $N_d = 10$, $P = 0.204$) nor over all trials (Mann–Whitney U test: $U = 33$, $N_w = 9$, $N_d = 10$, $P = 0.347$), indicating no difference in motivation between the two species.

Experiment 3: Two diagonal ropes, one baited

Two ropes, one with food attached and one without, were laid in parallel at an acute angle to the fence with a distance of 20 cm to each other. This resulted in a setup where the baited rope could “overlap” with the accessible end of the non-baited rope. If both ropes were laid out with a tilt to the right, and the food attached to the left one of the ropes, then this was called the “overlap” condition, as the end of the non-baited rope was in line with the food (Fig. 2c). When the food was attached to the right rope, no overlap between the baited rope and the accessible end of the non-baited rope occurred; therefore, this was referred to as the “control” condition (Fig. 2d). If canids were able to use the means-end properties of the ropes, they would always pull the rope attached to the food. However, if the animals simply pulled the rope closest to the food, they would choose the wrong rope in the overlap condition while performing well in the control condition. Each animal was tested 12 times in each condition. Within each condition, the ropes were tilted 6 times to the right and 6 times to the left. The conditions as well as the side to which the rope were tilted were pseudo-

randomized with the restrictions that the same condition as well as side of tilting could not be presented more than 3 times in a row and that the exact same placement could not be presented more than 2 times in a row.

Results—The analyses of the successful pulls in Experiment 3 revealed that only the fixed factor “condition: overlap/control” had a significant influence on the performance of the animals (Table 3). We did not find a difference in performance between wolves and dogs nor any effect of learning across trials nor any interaction. After pooling the data accordingly, further analyses showed that while the performance of the wolves and dogs together did not differ significantly from chance performance in the overlap condition (Wilcoxon signed-rank test: $T^+ = 56$, $N = 19$, $P = 0.352$), they were significantly above chance in the control condition (Wilcoxon signed-rank test: $T^+ = 171$, $N = 19$, $P = 0.0001$), suggesting that their choice was strongly (but not exclusively) influenced by the proximity of the food to the ends of the two ropes. However, one wolf and one dog seemed to learn how to solve the task over the 24 trials performing significantly above chance with 10 correct choices in the last 12 trials (Binomial: $P = 0.04$).

Finally, we found no significant difference in the latency to pull in a string between the wolves and the dogs in the first trial (Mann–Whitney U test: $U = 34.5$, $N_w = 9$, $N_d = 10$, $P = 0.412$), indicating that motivation to participate in this experiment was similar in both species. However, overall, there was a significant difference in the latencies of wolves and dogs to start pulling a string (Mann–Whitney U test: $U = 10.5$, $N_w = 9$, $N_d = 10$, $P = 0.006$; adjusted α -value = 0.025), with the dogs being slower.

Experiment 4: Two crossed ropes

In this experiment, we tested the animals on an even more difficult version of the former two conditions. In the first condition, the two ropes were crossed with a distance of 40 cm between the ropes at the ends (Fig. 2f). Here, choosing by proximity would not lead to success. In the control condition (Fig. 2e), the setup of the ropes looked very similar to the crossed condition, but here the ropes did not cross in the middle but instead each was bent in a V, so that choosing by proximity again would lead to success (the distance between the two ropes at the ends was again 40 cm). Each wolf and dog was tested 12 times in each condition, half of the trials provided food on the right and the other half on the left side. The conditions as well as the side that was baited were pseudo-randomized with the restrictions that the same conditions as well as side could not be presented more than 3 times in a row and that the exact same placement could not be presented more than 2 times in a row.

Results—The analyses of the successful pulls in Experiment 4 revealed no effect of learning in either species. The fixed factor “condition: crossed/control”, however, had a significant influence on the performance of the animals (Table 3). Moreover, we found a significant interaction between condition and species, suggesting that wolves and dogs differed in their performance in one of the two conditions. Further analyses showed that while wolves performed at chance level in both the crossed and the control condition (Wilcoxon signed-rank tests: $T^+ = 5$, $N = 9$ (5 ties), $P > 0.999$; and $T^+ = 13$, $N = 9$ (4 ties), $P = 0.188$, respectively), dogs performed significantly below chance in the crossed condition

(Wilcoxon signed-rank test: $T^+ = 3$, $N = 10$ (2 ties), $P = 0.039$), but significantly above chance in the control condition (Wilcoxon signed-rank test: $T^+ = 41.5$, $N = 10$ (1 tie), $P = 0.019$). The dog results can be explained by a strong proximity error with the dogs always pulling the string closest to the food. In contrast, 5 of the 9 wolves fell into a very strong side preference (Binomial tests, all $P < 0.007$). One wolf learned over trials how to solve the task with 11 correct choices in the last 12 trials (Binomial, $P = 0.006$). None of the dogs approached significant individual performance in this task.

We also found a significant difference in the latency to start pulling in a string between the wolves and the dogs in the first trial (Mann–Whitney U test: $U = 14.5$, $N_w = 9$, $N_d = 10$, $P = 0.014$; adjusted α -value = 0.025) with the dogs being slower than the wolves. This difference remained significant when comparing the mean latencies of all wolves over all 24 trials (Mann–Whitney U test: $U = 15$, $N_w = 9$, $N_d = 10$, $P = 0.013$; adjusted α -value = 0.025), with the dogs still being slower.

Experiment 5

To test whether the performance of the animals had improved during the previous experiments, we retested the animals in the initial discrimination task (Experiment 2, Fig. 2b).

Results—Overall, when analyzing the number of successful pulls, we found no significant difference between wolves and dogs, no effect of trial number, or any interaction (Table 4; Fig. 3). Pooling the wolf and dog data, we found that their performance did not differ from chance (Wilcoxon signed-rank test: $T^+ = 18.0$, $N = 11$ (8 ties), $P = 0.206$).

Finally, we found no significant difference in the latency to pull in a string between the wolves and the dogs in the first trial (Mann–Whitney U test: $U = 32.5$, $N_w = 9$, $N_d = 10$, $P = 0.323$). However, overall, there was a significant difference in the latencies of wolves and dogs to start pulling a string (Mann–Whitney U test: $U = 7$, $N_w = 9$, $N_d = 10$, $P = 0.001$; adjusted α -value = 0.025), with the dogs being slower.

Discussion

Our results demonstrate that while the wolves and dogs that we tested differed in regard to their biases (proximity versus side preference) when trying to solve a string-pulling task, neither the dogs nor the wolves seemed to have an insightful understanding of means-end relationships. However, a few animals of both species learned to solve some specific tasks, and in Experiment 2, the wolves even learned at the group level how to solve the task relatively quickly.

In Experiment 2, two rewards were laid out at the same distance to the subject and thus the proximity of the reward could not influence the decision made by the animals. At the group level, the wolves did not show any instantaneous understanding of the means-end relationships, but were found to improve within the 20 test trials and to perform above chance level in the last 10 trials. Whether this improvement was based on using a “rule-of-thumb” (e.g., pulling the longer rope) or indicates some basic understanding of the means-

end connection needs further exploration. In contrast to the wolves, the dogs developed a preference for the shorter string that was not connected to the food reward. Osthaus et al. (2005) also found that the dogs, instead of learning how to solve a task, decreased in their performance over 20 trials in the crossed condition, choosing the unrewarded rope more often over the course of the experiment. It is unclear why the dogs insisted on a solution that was not reinforced by food rewards. Moreover, the choice cannot be explained by alternative rewards: the dogs were not praised for making a wrong choice and they also did not play with the shorter string once obtained, but dropped it immediately. The behavior of the dogs is also surprising because in an “on/off” task, dogs have been found to rely on means-end relationships when proximity did not interfere with the correct solution (Range et al. 2011).

In the subsequent experiments, we found that the wolves, although they committed the proximity error in Experiment 3, quickly changed to a side preference in Experiment 4. The results of the dogs in the 1st, 3rd, and 4th experiment, on the other hand, underscore the previous results by Osthaus et al. (2005) that dogs are prone to commit the proximity error in string-pulling tasks. One might argue that in our study, applying a series of experiments, the animals' performance was partly due to learning over the experimental conditions. However, Osthaus et al. (2005), who tested naïve dogs in the same conditions of Experiment 3 and in the crossed condition of Experiment 4, found the same results as we did, suggesting that the experience that our dogs could have collected in the previous experiments did not influence their performance in the later tasks. Further, the predisposition of dogs to approach a reward in the direct, shortest way is apparent also in another task—the detour task—where dogs show severe difficulties in walking away from a visible reward in order to detour a fence or transparent barrier and thus reach the reward on the other side (Chapuis et al. 1983; Pongrácz et al. 2001, 2003; Scott and Fuller 1965). Studies by Frank and Frank (1987, 1982) on four 6-week-old wolf puppies replicating methods used by Scott and Fuller (1965) suggested that wolves were much more successful at the task than their domestic counterparts. However, as discussed in the introduction, sample size was small in their studies, and more importantly, the barrier was opaque as opposed to transparent. The development of different biases of our wolves and dogs (side preference vs. proximity) is also reflected in the latencies of the current study: since the dogs' choice was based on proximity to the food, they still had to examine the setup of the string to choose the one closest to the reward, which likely led to increased latencies in comparison with the wolves, which mainly chose based on a side preference, which does not require examination of the setup of the strings.

A wider comparative framework should help to detect the evolutionary origins of the capability of using means-end connections and the ecological factors facilitating the development of this skill. However, direct comparisons with data from other studies are unfortunately impossible due to methodological differences, such as the number of trials animals received, using strings of the same or of different colors, and most importantly the lack of intermixing test and control conditions. Nevertheless, based on published data, several primate and bird species outperform wolves and dogs in the overlap condition (Finch 1941; Harlow and Settlage 1934; Werdenich and Huber 2006), which may, however, be explained by testing the wolves and dogs with intermixed overlap and control conditions in contrast to the other studies, which utilized both conditions successively. Intermixing the

conditions is likely to negatively influence learning speed since the perceptual features change from trial to trial and thus prevent animals from relying on simple rules (e.g., do not pull the string closest to the reward). On the other hand, the crossed condition seems to be similarly difficult for most animal species. At the group level, even primates cannot immediately solve that problem and may require an extended period of learning even when the crossed and control conditions are not intermixed (Tomasello and Call 1997; Povinelli 2000). Interestingly, however, both in primates and in ravens, there are great interindividual differences with a few individuals who can solve the problem instantly or after a short learning period (Finch 1941; Heinrich 2000). Similarly, both in our dog and wolf sample, a few animals seemed to learn how to solve even the more difficult tasks. In Experiment 2 (two strings and two rewards), one wolf chose correctly in the first 6 trials suggesting that he understood means-end relationships in this setup spontaneously. However, this wolf failed in the repetition of the same condition at the end of this study, which might be interpreted either as a sign that it did not understand the means-end connection in Experiment 2 but used some “rule-of-thumb”, or that it got confused by the more difficult intermediate problems that it could not solve. In Experiment 3 (two diagonal strings), one wolf and one dog improved over trials and significantly performed above chance in the last 12 trials. In Experiment 4, one wolf performed significantly above chance in the last 12 trials. These individual data are interesting since they show that at least some animals can learn relatively quickly how to solve such tasks. Obviously, these superficial comparisons across species based on methodologically different studies cannot provide a good basis for evolutionary reasoning. Further studies are needed to reveal what rules the successful wolves and dogs used to solve the string-pulling tasks and how much of the variation we see between the performances of different species is driven by different methodologies.

Species comparisons need to be controlled for the life-long experiences and training of the subjects being compared (see for example Marshall-Pescini et al. 2008; Osthaus et al. 2003b; Range et al. 2009; Topál et al. 1997). In this study, although we tried to match the keeping conditions of the wolves and the dogs as much as possible, their life-long experiences and training history were not fully identical (e.g. hand-raising and some clicker training in wolves, versus mother-raising and complete lack of clicker training in dogs) nor were their enclosures equipped similarly (wolves: containing bushes, trees and a lake, dogs: normal kennels without trees or bushes). Accordingly, it cannot be ruled out entirely that the differing performance of dogs and wolves was influenced by their somewhat dissimilar previous experiences rather than domestication. Whenever possible, animals that are raised, kept and trained in an identical way should be compared with identical methods to answer questions concerning the genetically based effects of domestication. Finally, some of the animals were related to each other, which might also have influenced the results.

Keeping these limitations in mind, overall, this study suggests that while the wolves and dogs tested differ in regard to the biases they develop (proximity vs. side preference), neither species has an insightful understanding of means-end relationships. Although the wolves were able to learn how to solve Experiment 2 in contrast to the dogs, their improvement might be due to superior general associative learning skills rather than some recognition of means-end connections. Further studies are needed to clarify whether and

how domestication has influenced the recognition of means-end connections and causal understanding in other experimental paradigms in wolves and dogs.

Acknowledgments

We thank Wolf Park and the Mountain Wolf Farm for being able to test their animals, Pat Goodman on helpful comments on the manuscript, Corsin Müller for helping with the statistical analyses, Katharina Kramer for reliability coding, Kurt Kotrschal for helping to establish the Wolf Science Center, 5 reviewers and Stephen Lea for their helpful comments to improve the manuscript, and many students and volunteers for their devotion and assistance with raising the animals of the Wolf Science Center. The project was financially supported by the Hochschuljubiläumsstiftung der Stadt Wien (H-2076/2008) and Austrian Science Fund (FWF) project P21244-B17. We further thank many private sponsors and Royal Canin for financial support and the Gamepark Ernstbrunn for hosting the Wolf Science Center.

References

1. Bratman M. Intention and means-end reasoning. *Philos Rev.* 1981; 90:252–265.
2. Brauer J, Kaminski J, Riedel J, Call J, Tomasello M. Making inferences about the location of hidden food: social dog, causal ape. *J Comp Psychol.* 2006; 120:38–47. [PubMed: 16551163]
3. Chapuis N, Thinus-Blanc C, Poucet B. Dissociation of mechanisms involved in dogs' oriented displacements. *Q J Exp Psychol.* 1983; 35:213–219.
4. Collier-Baker E, Davis JM, Suddendorf T. Do dogs (*Canis familiaris*) understand invisible displacement? *J Comp Psychol.* 2004; 118:421–433. [PubMed: 15584779]
5. Finch G. The solution of patterned string problems by chimpanzees. *J Comp Psychol.* 1941; 32:83–90.
6. Fiset S, Leblanc V. Invisible displacement understanding in domestic dogs (*Canis familiaris*): the role of visual cues in search behavior. *Anim Cogn.* 2007; 10:211–224. [PubMed: 17165041]
7. Frank H. Evolution of canine information processing under conditions of natural and artificial selection. *Z Tierpsychol.* 1980; 5:389–399. [PubMed: 7193386]
8. Frank H, Frank MG. Comparison of problem-solving performance of 6-week-old wolves and dogs. *Anim Behav.* 1982; 30:95–98.
9. Frank H, Frank MG. Comparative manipulation test performance in 10-week-old wolves (*Canis lupus*) and Alaskan Malamutes (*Canis familiaris*) – a Piagetian interpretation. *J Comp Psychol.* 1985; 99:266–274.
10. Frank, H.; Frank, M. The University of Michigan canine information-processing project (1979e1981). In: Frank, H., editor. *Man and Wolf*. Dr W. Junk; Dordrecht: 1987. p. 143–167.
11. Frank H, Frank MG, Hasselbach LM, Littleton DM. Motivation and insight in wolf (*Canis lupus*) and Alaskan Malamute (*Canis familiaris*)—visual-discrimination learning. *Bull Psychon Society.* 1989; 27:455–458.
12. Gácsi M, Miklósi Á, Varga O, Topál J, Csányi V. Are readers of our face readers of our minds? Dogs (*Canis familiaris*) show situation-dependent recognition of human's attention. *Anim Cogn.* 2004; 7:144–153. [PubMed: 14669075]
13. Hare B, Brown M, Williamson C, Tomasello M. The domestication of social cognition in dogs. *Science.* 2002; 298:1634–1636. [PubMed: 12446914]
14. Harlow HF, Settlage PH. Comparative behaviour of primates. VII. Capacity of monkeys to solve patterned string tests. *J Comp Psychol.* 1934; 18:423–435.
15. Heinrich B. An experimental investigation of insight in Common Ravens (*Corvus corax*). *Auk.* 1995; 112:994–1003.
16. Heinrich, B. Testing insight in ravens. In: Heyes, C.; Huber, L., editors. *The evolution of cognition*. MIT Press; Cambridge: 2000. p. 289–305.
17. Heinrich B, Bugnyar T. Testing problem solving in ravens: string-pulling to reach food. *Ethology.* 2005; 111:962–976.
18. Hemmer, H. Domestication. The decline of environmental appreciation. Cambridge University Press; Cambridge: 1990.

19. Köhler, W. The mentality of apes. Vintage Books; New York: 1927.
20. Krasheninnikova A, Wanker R. String-pulling in spectacled parrotlets (*Forpus conspicillatus*). Behaviour. 2010; 147:725–739.
21. Lea SEG, Goto K, Osthaus B, Ryan CME. The logic of the stimulus. Anim Cogn. 2006; 9:247–256. [PubMed: 16909234]
22. Marshall-Pescini S, Valsecchi P, Petak I, Accorsi PA, Previde EP. Does training make you smarter? The effects of training on dogs' performance (*Canis familiaris*) in a problem solving task. Behav Process. 2008; 78:449–454.
23. Miklósi Á, Soproni K. A comparative analysis of animals' understanding of the human pointing gesture. Anim Cogn. 2006; 9:81–93. [PubMed: 16235075]
24. Miklósi Á, Kubinyi E, Topál J, Gácsi M, Virányi ZS, Csányi V. A simple reason for a big difference: wolves do not look back at humans but dogs do. Curr Biol. 2003; 13:763–766. [PubMed: 12725735]
25. Osthaus B, Slater AM, Lea SEG. Can dogs defy gravity? A comparison with the human infant and a non-human primate. Dev Sci. 2003a; 6:489–497.
26. Osthaus, B.; Lea, SEG.; Slater, AM. Training influences problem-solving abilities in dogs (*Canis lupus familiaris*); Proceedings of the annual BSAS conference; 2003b. p. 103
27. Osthaus B, Lea SEG, Slater AM. Dogs (*Canis lupus familiaris*) fail to show understanding of means-end connections in a string-pulling task. Anim Cogn. 2005; 8:37–47. [PubMed: 15338446]
28. Pepperberg IM. "Insightful" string-pulling in Grey parrots (*Psittacus erithacus*) is affected by vocal competence. Anim Cogn. 2004; 7:263–266. [PubMed: 15045620]
29. Pongrácz P, Miklósi Á, Kubinyi E, Gurobi K, Topál J, Csányi V. Social learning in dogs: the effect of a human demonstrator on the performance of dogs in a detour task. Anim Behav. 2001; 62:1109–1117.
30. Pongrácz P, Miklósi Á, Kubinyi E, Topál J, Csányi V. Interaction between individual experience and social learning in dogs. Anim Behav. 2003; 65:595–603.
31. Povinelli, DJ. Folk physics for apes. Oxford University Press; New York: 2000.
32. Range F, Heucke SL, Gruber C, Konz A, Huber L, Virányi ZS. The effect of ostensive cues on dogs' performance in a manipulative social learning task. Appl Anim Behav Sci. 2009; 120:170–178.
33. Range F, Hentrup M, Virányi ZS. Dogs are able to solve a means-end task. Anim Cogn. 2011; 14:575–583. [PubMed: 21445577]
34. Schmidt GF, Cook RG. Mind the gap: means-end discrimination by pigeons. Anim Behav. 2006; 71:599–608.
35. Schuck-Paim C, Borsari A, Ottoni EB. Means to an end: neotropical parrots manage to pull strings to meet their goals. Anim Cogn. 2009; 12:287–301. [PubMed: 18766389]
36. Scott, JP.; Fuller, JL. Genetics and the social behavior of the dog. University of Chicago Press; Chicago: 1965.
37. Szetei V, Miklósi Á, Topál J, Csányi V. When dogs seem to lose their nose: an investigation on the use of visual and olfactory cues in communicative context between dog and owner. Appl Anim Behav Sci. 2003; 83:141–152.
38. Thorndike EL. Animal intelligence: an experimental study of the associative processes in animals. Psychol Rev Monogr Suppl. 1898; 2:1–109.
39. Tomasello, M.; Call, J. Primate cognition. Oxford University Press; Oxford: 1997.
40. Topál J, Miklósi Á, Csányi V. Dog-human relationship affects problem solving behavior in the dog. Anthrozoos. 1997; 10:214–224.
41. Udell MAR, Dorey NR, Wynne CDL. What did domestication do to dogs? A new account of dogs' sensitivity to human actions. Biol Rev. 2010; 85:327–345. [PubMed: 19961472]
42. Virányi, Zs; Topál, J.; Gácsi, M.; Miklósi, Á.; Csányi, V. Dogs respond appropriately to cues of humans' attentional focus. Behav Process. 2004; 66:161–172.
43. Virányi, Zs; Topál, J.; Miklósi, Á.; Csányi, V. A nonverbal test of knowledge attribution: a comparative study on dogs and children. Anim Cogn. 2006; 9:13–26. [PubMed: 15895261]

44. Watson JS, Gergely G, Csányi V, Tópal J, Gácsi M, Sarkozi Z. Distinguishing logic from association in the solution of an invisible displacement task by children (*Homo sapiens*) and dogs (*Canis familiaris*): Using negation of disjunction. *J Comp Psychol.* 2001; 115:219–226. [PubMed: 11594490]
45. Werdenich D, Huber L. A case of quick problem solving in birds: string pulling in keas, *Nestor notabilis*. *Anim Behav.* 2006; 71:855–863.
46. Wynne CDL, Udell MAR, Lord KA. Ontogeny's impact on human–dog communication. *Anim Behav.* 2008; 76:e1–e4.

Fig. 1.
Experimental setup of the **a** wolf (Experiment 4, control condition) and **b** dog experiments (Experiment 3, overlap condition)

Fig. 2. Layout of the ropes in Experiments 1–5. For each experiment, the different conditions are depicted. The side of the reward in each experiment was alternated

Fig. 3.

Box plots showing the performance of the wolves and dogs in the first and last 10 trials of Experiment 2. *Shaded boxes* represent the interquartile range, *bars within shaded boxes* are median values, and *whiskers* indicate the 5th and 95th percentile. *Circles* present outliers. The *dashed line* indicates chance level of performance. *Asterisk* indicates $P > 0.05$

Table 1

Sex, age, housing, and genetic relationships (first letter stands for the mother and second for the father) of the participating wolves and dogs

	Sex	Date of birth	Pack	Relatedness
Wolf				
Marion	F	1998	WP	AA
Kailani	F	2004	WP	BB
Ayla	F	2004	WP	BB
Apollo	M	1995	WP	CC
Renki	M	2004	WP	BB
Wolfgang	M	2005	WP	BD
Shima	F	2008	WSC	DE
Aragorn	M	2008	WSC	DE
Kaspar	M	2008	WSC	E?
Dog				
Jack	M	2001	MWF	FP
Damon	M	2002	MWF	GQ
Firo	M	1999	MWF	HR
Duke	M	2001	MWF	II
Brownie	M	2003	MWF	JS
Otchum	M	2003	MWF	KT
Oakley	M	2006	MWF	LU
Balto	M	2002	MWF	II
Tamani	F	2002	MWF	NV
Dekota	F	2003	MWF	OW

WP Wolf Park, *WSC* Wolf Science Center, *MWF* Mountain Wolf Farm

Table 2

Summary of the results of the generalized Linear models determining whether species or learning over trials influenced the performance of the animals in Experiment 2 and 5

Experiment	Factor	<i>df</i>	<i>t</i> value	<i>P</i>
Experiment 2	Species	1, 17	−0.838	0.414
	Trial number	1, 345	−2.902	<i>0.004</i>
	Species × trial number	1, 345	2.154	<i>0.032</i>
Experiment 5	Species	1, 17	1.739	0.099
	Trial number	1, 345	1.252	0.212
	Species × trial number	1, 345	−0.515	0.607

Italicized values represent significant results

Table 3

Summary of the results of the generalized linear models determining whether species, learning (first six trials vs. last six trials), or experimental condition influenced the performance of the dogs in Experiment 3 and 4

Experiment	Factor	df	t value	P
Experiment 3	Species	1, 17	-1.149	0.266
	“Two diagonal strings”			
	Trials (6 first/6 last)	1, 51	-0.679	0.499
Experiment 4	Condition: overlap/control	1, 51	-3.252	<i>0.002</i>
	Species	1, 17	-1.231	0.235
	“Two crossed strings”			
	Trials (6 first/6 last)	1, 51	-0.175	0.862
	Condition: cross/control	1, 51	-3.611	<i>0.001</i>
	Species × condition	1, 51	2.222	<i>0.031</i>

Italicized values represent significant results

Table 4

Correct choices of all animals in Experiments 2–5

Animal	Species	Exp 2 (20 trials)	Exp 3		Exp 4		Exp 5 (20 trials)
			Overlap (12 trials)	Control (12 trials)	Crossed (12 trials)	Control (12 trials)	
Shima	W	9	2	9	6	8	8
Kaspar	W	13	5	10	6	7	9
Aragorn	W	9	2	9	6	11	10
Renki	W	6 ^a	2	11	3	6	12
Ayla	W	12	0	11	8	7	10 ^b
Marion	W	13	7	9	8	6	13
Apollo	W	9	9	6	6	6	10
Kailani	W	9	2	9	5	5	8
Wolfgang	W	14	7	9	6	6	11
Otchum	D	8	6	9	2	6	4
Firo	D	6	7	11	2	12	10
Brownie	D	10	7	8	5	8	10
Oakly	D	6	5	10	6	9	8
Duke	D	13	6	11	8	5	10
Demon	D	8	7	10	6	4	5
Jack	D	9	5	11	1	11	10
Tamani	D	10	6	8	5	9	9
Cira	D	11	8	10	1	9	9
Dekota	D	9	5	10	2	10	6

^aThis animal was only tested 6 times, since he was correct in each of them^bAlya was tested only 13 times in that condition