Aus dem Department für Interdisziplinäre Lebenswissenschaften

der Veterinärmedizinischen Universität Wien

Messerli Forschungsinstitut

(Leiter: Univ.-Prof. Dr.rer.nat. Ludwig Huber)

Probing means-end understanding in dogs and wolves: comparing their learning in a magic vs. natural On/Off task

Masterarbeit

Veterinärmedizinische Universität Wien

vorgelegt von

Diana Petri

Wien, im August 2023

Supervisors: Priv.-Doz. Zsófia Virányi, PhD

Second Examiner: Priv.-Doz. S. Marshall-Pescini, PhD

Acknowledgements

I thank Wolfgang Berger for constructing the apparatus. Special thanks to Remco Folkertsma for his statistical support. I also thank the animal trainers at Wolf Science Center, Ernstbrunn and the dog owners for participating in this study. And finally thank you to Priv.-Doz. Zsófia Virányi PhD for all her help and support.

Abstract

Compared to social cognition, the physical understanding of dogs and wolves have been relatively sparsely studied. Nevertheless, the available studies show poor performance of dogs in such cognitive studies, and it has been suggested that domestication, by human care relaxing natural selection on such skills of dogs, may have had a detrimental effect on dogs' problem solving and physical cognition. In line with this, a former study had found that wolves learnt to solve a string-pulling task faster than dogs and had suggested that dogs may have a limited means-end understanding. Means to an end understanding is in general the understanding of more complex systems as when an object can have an impact on the movement of another one. However, as long as the animals need to learn to solve the task, it remains unclear whether the subjects understood the causal structure of the task or whether they solved it based on perceptual cues such as contact between reward and support. To address this problem, we compared whether the animals learn to solve the natural version of the support task faster than a counterintuitive, "magic" version. For this aim, we tested 2 groups of dogs and 2 groups of wolves. In the ON group (natural), animals could get the reward when pulling the board with the reward on it. In contrast, in the OFF group (magic) the animals could get the reward by pulling the OFF board (with the reward on the side of the board) and the reward did not move if the animals pulled the ON board We found no evidence that either the dogs or the wolves would have differentiated between the two conditions. The wolves chose randomly throughout their 20 sessions, whereas two dogs learnt to discriminate between the 2 boards with a similar speed in both groups. These results provide no evidence that either dogs or wolves would have a means-end understanding, but suggest faster associative learning in dogs than in wolves, which is in contrast with former studies. Future studies with bigger sample sizes and with more observations will have to confirm these results.

Zusammenfassung

Im Gegensatz zur sozialen Kognition wurden die kognitiven Fähigkeiten von Hunden und Wölfen in Bezug auf das physikalische Verständnis kaum untersucht. Hier kommt hinzu, dass beispielsweise Hunde in den verfügbaren Studien schlecht abschneiden (Müller et al. 2014). Die Domestizierungshypothese behauptet, das Menschen die soziale und physikalische Kognition von Hunden durch die Domestizierung und Ontogenese beeinflusst haben (Lampe et al. 2017).

Übereinstimmend damit hat eine frühere Studie herausgefunden das Wölfe bei einer string-pulling Aufgabe besser abschneiden als Hunde und impliziert, dass Hunde ein limitiertes Verständins von Mittel zum Zweck besitzen. Beim Mittel zum Zweck Verständnis geht es im Allgemeinen um das Verstehen von komplexen Systemen wie beispielsweise, wenn die Bewegung eines Objektes ein anderes mit bewegt. Dadurch, dass die Tiere lernen müssen dieses Problem zu lösen, bleibt es unklar, ob die Probanden das Problem basierend auf auf dem Verständnis der kausalen Zusammenhänge gelöst haben oder ob sie sichtbaren Hinweisen, wie dem Kontakt zwischen Belohnung und Hilfsmittel, gefolgt sind. Um dieses Problem zu adressieren wurde die magische Gruppe eingeführt, in welcher die Belohnung neben dem Brett war und sich auf magische Weise mit dem Brett bewegte.

Zur Untersuchung, ob Tiere mit einer natürlichen Situation (Belohnung auf dem Brett) oder eine magischen (Belohnung neben dem Brett) schneller lernen, wurden zwei Gruppen mit Wölfen und zwei Gruppen mit Hunden gebildet. In der natürlichen Situation gab es ein Brett mit der Belohnung auf dem Brett, welche der Hund sich durch rausziehen verdienen konnte. Das andere Brett war in dieser Gruppe mit einer Belohnung neben dem Brett, welche auch nicht durch heraus ziehen des Brettes verdient werden konnte versehen. In der magischen Gruppe war es genau umgedreht, d.h. das Tier konnte die Belohnung durch das Ziehen des Brettes mit der Belohnung nebendran erhalten und nicht durch das Ziehen des Brettes auf dem die Belohnung scheinbar lag. Das Ergebnis zeigte keine Hinweise darauf, dass Wölfe oder Hunde ein Mittel zum Zweck Verständnis haben. Allerdings haben wir Hinweise darauf gefunden, Hunde (in beiden Gruppen) schneller lernen, was im Widerspruch mit vorangegangen Studien

steht, die bei Wölfen eine bessere Lernfähigkeit aufwiesen. Künftige Studien mit einer größeren Kohorte und mehr Beobachtungen werden notwendig sein um diese Ergebnisse zu bestätigen.

Keywords

Means-end understanding, support problem, physical cognition, dog, wolf, domestication

Table of Contents

1. Introduction

1.1 Physical cognition in dogs

All canines are faced with social and physical challenges in their environment, ranging from finding food to cooperation among group members. To find food wolves must hunt and track their prey. They must find back to their den after hunting. Dogs, if kept in human families, do not have to find food, but to cooperate with their group member(s), one or more humans in a family. This cooperation includes working with humans in several settings like search and rescue.

Compared to social cognition, the cognitive abilities of dogs in the physical domain have been relatively sparsely studied. Additionally, the available studies show poor performance of dogs in physical understanding (Müller et al. 2014). It is possible that humans have influenced social and physical cognition of dogs throughout domestication and ontogeny (Lampe et al. 2017). To explore these effects Lampe compared captive wolves and dogs living in packs under the same conditions with pet dogs living in human families. The animals were faced with a series of object choice tasks in which their response to communicative, behavioural, and causal cues were tested. The results show that all three groups performed similar in the communicative and behavioural conditions, but wolves outperformed dogs in their ability to follow causal cues. An example for a causal cue, the experimenter was hiding under the table and a container produced noise while shaken vs. a container that made no sound. Another causal cue was an inclined shape vs. a flat shape (Lampe et al. 2017).

One evolutionary theory that explains the poor performance of dogs is the information processing hypothesis (Frank 1980). This hypothesis suggests that selection pressures advancing causal understanding and problem solving in wild animals have relaxed on domesticated dogs due to a buffering effect e.g., care of humans. Therefore, Frank (1980) has suggested that domestication had a detrimental effect on dog's physical cognitive abilities. Explaining this fact that, in contrast to wolves, dogs' biological feeding and mating are inhibited by humans and thus natural selection is less present (Frank 1980; Hemmer 1990).

In addition to evolutionary adaptation, dogs living in close contact to humans may also learn to rely on human help instead of solving problems independently. This suggests that dogs that live independently could potentially be better problem solvers. Even if so, as free-ranging dogs scavenge for food in waste distributed around human settlements while wolves search for and hunt prey actively. Therefore, the feeding ecology hypothesis proposes that dogs might have evolved reduced causal insight, persistence, and exploration (Lampe et al. 2017).

1.2 Means-end understanding as a fundamental component of physical cognition

Understanding of means-end relationships is best shown by the demonstration of insightful behaviour to solve a novel task on the first trial (Osthaus et al. 2005). Presumably an understanding of means-end relationships, particularly as it pertains to object manipulation, is a key mental prerequisite to advanced cognitive abilities. As a result, evaluation of means-end understanding across species is an important area for comparative cognition research (Bensky et al. 2013). Means to an end understanding is in general the understanding of more complex systems as when an object can have an impact on the movement of another one. To test this understanding in animals string pulling is one of the oldest and most common tests. String pulling in animals was first observed in ravens during a cold Swedish winter. After a fisherman left the hole he had cut into a frozen lake with the baited fishing line inside, a raven flew down to examine the situation. Then it took the fishing line in its peak, and started to pull it upward by stepping on the loop, and repeated this several times until reaching the end of the line. The raven was then rewarded with either a fish or the bait (Jacobs 2018, Larsson 1958).

String pulling as a formal task was then originally introduced by Köhler 1927: he gave the subject a string within reach that was attached to a piece of food out of reach. The string pulling test has been conducted with around 170 species in over 210 studies (Jacobs 2017; Jacobs and Osvath 2015). Split in at least 68 mammalian species and 90 avian species (Lamarre & Wilson, 2021). It provides countless testing conditions varying the number and patterns of strings (Jacobs 2018). Asian Elephants (Irie-Sugimoto et al. 2008), great apes (Herrmann et al. 2008) and keas (Auersperg et al. 2009) have solved the string pulling task and/or the support problem.

The “support problem” was originally introduced by Piaget 1952: here a goal object (toy or food) is out of reach of the subject but is resting on a support, for example a blanket, that is within the subject’s reach. Dogs, in comparison to other species, perform rather poorly in tasks that require understanding causal connections or physical characteristic of objects. Osthaus and colleagues (2004) suggested that this may reflect not only cognitive but also their physical capabilities: “(…) unlike primates or even some birds, dogs are not well equipped anatomically for manipulating objects, and object manipulating is of no obvious ecological relevance to them, and on these grounds string pulling might be a task at which they would not do well.” (Osthaus et al. 2004, p. 38) Therefore the support problem, using boards, may be more suitable to test dogs’ means-end understanding because they can pull the boards more easily than strings by using their paw. Also because the board is lying on the floor in horizontal position and is not in a vertical position.

A variety of species have been tested in the classic setup of the support task: two choices are presented, a reward is placed on a support, next to a second support where the reward is placed besides. For example, Asian elephants were tested in four variations of the support problem if they have a means to an end understanding. First condition, the training condition, two identical trays serving as the support, one with bait on the tray and the other one without. Second condition the bait was placed on one tray while the other bait was placed besides the other tray. The third condition was a transfer test that can be considered as a variation of the “on / off” task to test whether the elephants are able to make distinctions between relevant and irrelevant features of the problem. Therefore, the task was identical to the second condition except with two additional irrelevant features like colour and size of the trays. And in the fourth condition, both trays contained food but one of the trays had a small gap between the subjects and the food which prevented the elephants form reaching the reward. The results showed that one elephant’s performance was above chance level in all conditions, after several sessions. Based on this finding, the authors suggested that this animal was able to understand the relationship between the tray as “means” and the bait as the “end” (Irie-Sugimoto et al. 2007).

One study by Frank and Frank (1985) found that wolves performed more successful than dogs in a means end task. The study weaknesses are a small sample size and all animals were tested in a variation of puzzle boxes at young age. Later Range et al. (2012) showed that neither wolves nor dogs showed instantaneous understanding of means-end connection in a string-pulling task. Experiment 1 a single rope with food attached to the end was laid diagonally. The expectation was that dogs would paw the ground closest to the reward before trying to pull the rope with the reward on it. All animals, wolves as well as dogs, solved the task but often committed this so-called proximity error. The proximity error, pawing or mouthing at a location closest to the reward, is thought to be an inherited predisposition to go for food directly. In canine it seems to overshadow the recognition of means-end connection and in combination with the inability to inhibit this response could lead to proximity bias of dogs (Lea et al. 2006). Importantly, Range et al. (2011) showed that dogs can solve a means-end task even if proximity of the unsupported reward was a cofounding factor. For this aim, they presented the animals with two parallel ropes one connected to a reward and the other one with a food reward but not connected to the rope. The aim was to test whether the wolves and dogs can solve a 2 choice string-pulling task taking their means-end relation into account, if proximity of the food reward cannot influence the solution. When analysing the number of successful pulls, they found that wolves improved their performance over 20 trials while dogs' performance got worse. These results show that few dogs learned to solve the task and also the wolves needed to learn, instead of showing an instantaneous understanding of means-end connections. If these results indicate an understanding of means-end connection needs further testing. As long as the animals need to learn to solve the task, it remains unclear whether the subjects understand the causal structure of the task or whether they solve it based on perceptual cues such as contact between reward and support or repeatedly observing moving which support leads to the movement of the reward. To determine if the subjects relied on perceptual cues to solve the problem it is necessary that successful subjects are subsequently presented with modified versions of the classic setup where causally relevant aspects of the setup have been changed. Unfortunately, the task can also be solved accidently by the animal. Some animals solve the problem through trial and error learning when given many repetitions of the same condition, but not when the conditions are intermixed (Jacobs & Osvath 2015).

In sum, based on the poor performance of the dogs and wolves in the on/off task (Müller, Riemer, Virányi, Huber, & Range, 2014, Range et al., 2012), it has been suggested that, while some dogs and wolves can learn to solve the support problem, they appear to do so by associating perceptual cues instead of understanding the causal underpinnings of the task.

Moreover, it remains unclear whether the improved performance of the wolves, reported in some studies, is due to their general associative learning skills superior to those of dogs or to some recognition of means-end connections. That is, despite the recent interest in dog cognition and especially in the effects of domestication, we lack detailed knowledge of the physical understanding of wolves and dogs, their similarities and differences.

To address this problem, in this study we developed a "magic version" of the support task where the animals needed to choose the off board to get the reward within their reach. To investigate whether the animals have some understanding of means-end connections, we wanted to test whether they learn to solve the natural problem faster than the magic one. For this aim, within each species, we tested 2 groups of animals. In the ON group (natural condition where the reward on top of the board (ON board) was accessible), the animals could get the reward when pulling the board with the reward on it. The alternative, the second board, was the OFF board where the reward was placed next to the board and did not move if the animal tried to pull this board. In contrast, in the OFF group (magic condition) the animals could get the reward by pulling the OFF board (with the reward on the side of the board) and the reward did not move if the animals pulled the ON board (where the reward was placed on top of the board) (Figure 1).

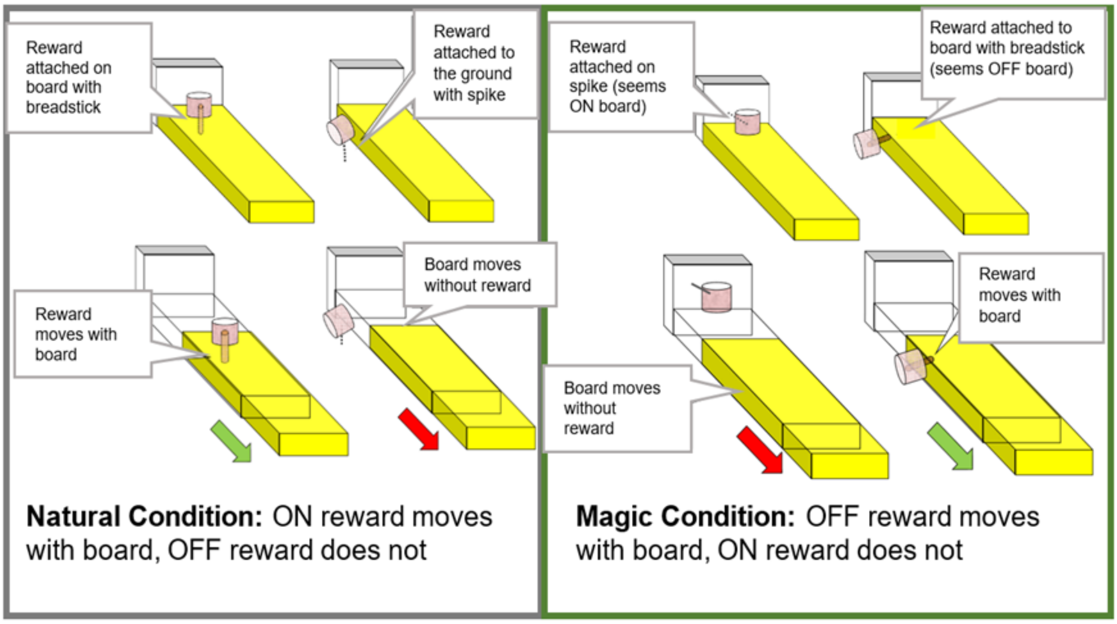

Figure 1 - ON / OFF - Natural/Magic Condition

Our hypothesis was that if solving the task was based on free associations (i.e., picking up and learning salient features of the task), the animals would learn in the causally correct, natural version and in the causally incorrect, magic version similarly fast. In this case, subjects would only be able to succeed via repeated exposure and their performance would improve with a similar speed in both the natural and magic groups. However, if the animals had a certain level of means-end understanding, they would learn faster in the causally correct, natural version than in the counterintuitive, magic version.

Regarding our species comparison, we expected that, if domestication had negatively impacted dogs' means-end understanding, wolves would differentiate better and/or sooner between the natural and magic versions of the task than dogs. In contrast, if wolves are "only" faster learners than dogs, they would outperform dogs (sooner) in both the natural and magic conditions.

2. Methods

2.1 Subjects

Nine wolves (3 females and 6 males, age average 7.6 years and range from 4 to 11 years old) and 8 dogs (2 females and 6 males, age average 9.625 years and range from 8 to 13 years old) participated in the study that was conducted at the Wolf Science Center, in Ernstbrunn, Austria (Table 1). All the wolves were hand-raised in peer groups after being separated from their mothers. They were bottle fed and later hand-fed by humans for the first months of their life. Beginning with 4 months of age, no human was continuously present in the enclosures, but the wolves participated in training and cognitive and behavioural experiments at the Wolf Science Center. Therefore, an intensive social contact to humans was guaranteed. With one exception, all the dogs (all mongrels, with one exception) were also raised at the Wolf Science Center. Three of them were still living in packs in enclosures of the Wolf Science Center at the time of testing, whereas the other 5 dogs had been living with WSC trainers for the last years. One of these dogs (an Airdale terrier) was adopted from a breeder and had always lived with her owner. These dogs accompanied their owners during their work at the WSC daily and were, thus, familiar with the testing environment and the testing routines applied with all subjects.

Table 1- List of Subjects

Name	Species	Sex	Age (years)	Living Condition	Group
Zazu	Dog	Male	8	Private	Natural
Asali	Dog	Male	12	Private	Natural
Kilio	Dog	Male	13	Private	Natural
Freya	Dog	Female	8	Private (raised as pet)	Natural
Imara	Dog	Female	8	Pack	Magic
Enzi	Dog	Male	8	Pack	Magic
Hakima	Dog	Male	12	Private	Magic
Hiari	Dog	Male	8	Pack	Magic
Wamblee	Wolf	Male	8	Pack	Natural
Kenai	Wolf	Male	10	Pack	Natural
Chitto	Wolf	Male	8	Pack	Natural
Yukon	Wolf	Female	11	Pack	Natural
Tekoa	Wolf	Male	4	Pack	Natural
Tala	Wolf	Female	8	Pack	Magic
Geronimo	Wolf	Male	11	Pack	Magic
Taima	Wolf	Female	4	Pack	Magic
Amarok.	Wolf	Male	8	Pack	Magic

2.2 Setup and Apparatus

The wolves were tested outdoors, in a separation compartment of their living enclosure. Therefore, they were separated from the boards with a wire mesh fence whose holes were big enough to allow for pulling the boards through and for reaching the food (Figure 2 and 3). The dogs were tested inside, in a testing room of the Wolf Science Center. Here the apparatus was enclosed in a small wire mesh compartment that had a plexiglass front panel with 2 holes where the dogs could pull the boards out and take the food reward. The Handler was responsible for treating the dog. At a distance of 1.5 meter in front of the apparatus was the starting position for the dog. After the dog made a choice, the dog had to return to H at the starting position. The Experimenter placed behind the opaque partition was not visible for the dogs while preparing the boards with the rewards. Not only the front view also the sides were made in transparent.

Figure 2 - dog pulling at test apparatus

Figure 3 - wolf pulling at test apparatus

The apparatus consisted of two metal boards mounted on a light-grey platform. An opaque partition was mounted 50 cm behind the front barrier (plexiglass for dogs and wire mesh for wolfs) to prevent the animals from seeing the experimenter (E) and the baiting process. The platform consists of two boards with three wooden lines in the front and one in the back what makes pulling the board easier for the animal. Additionally, the boards have drilled holes on the top and on the side. These holes help to fix the rewards. The size of the holes on the board and the back wall was exactly the size of a pretzel stick respectively a screw. To obtain the reward the animal had to pull out the whole length of the board. This is how the baited boards looked: The experimental reward consisted of two components: Extrawurst sausage and pretzel sticks. The pretzel sticks have been broken down in 3cm long pieces. The Extrawurst was cut with a cookie cutter into blocks (2,7cm x 2,7cm x 1,5cm) before the test sessions. These cubes were solid enough to keep their shape. One side of the pretzel is stuck into the sausage and plugged with the other half into a drilled hole in the board so sausage can not fall off accidentally. With this solution, the food cubes can be fixed also to the side of the boards (in the “off” condition) and can be easily removed by the animals who can eat the food cubes as well as the bread sticks inside.

The video recording documented one session (10 trials) per dog in a take. The camera was put on a tripod so you could see the dog in starting position and the testing apparatus (side view, fixed angle). Every video included a short introduction of the session with date, name of the animal and number of session by the experimenter.

Figure 4 - 3cm pretzel sticks – used to fixate the Extrawurst blocks

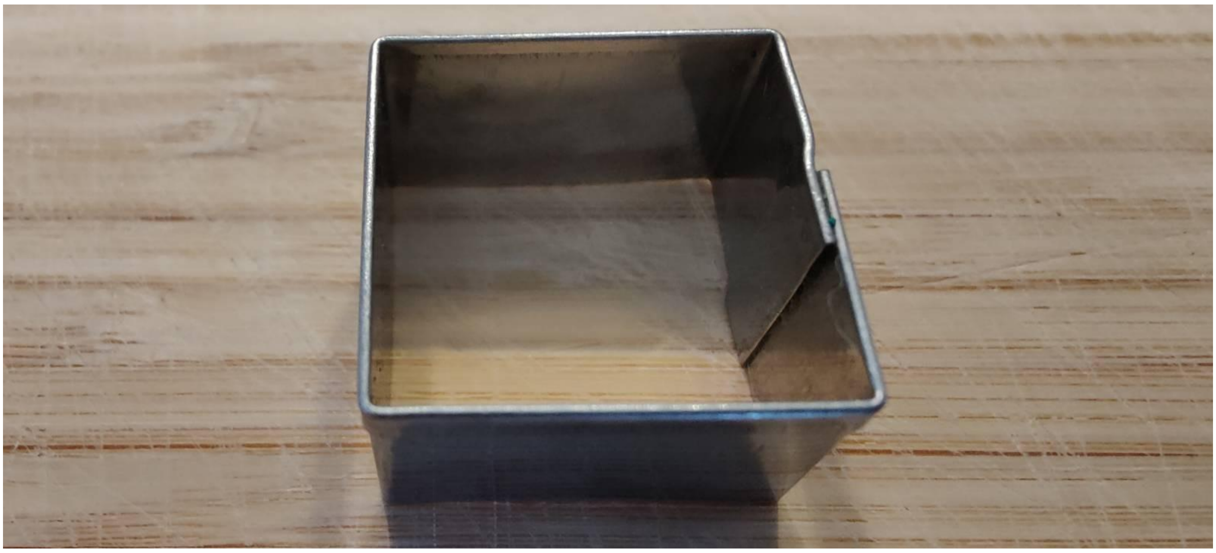

Figure 5 - cookie cutter (2,7cm x 2,7cm x 1,5cm) – used to cut blocks of Extrawurst

Figure 6 - Experimental Setup Apparatus

Figure 7 - Experiment Overview

Figure 8 - Dog in starting position

2.3 Procedure

The animals were separated from their pack and tested individually. Each session began with starting the video recording. Then the Handler (H) called the subject to the start position. For this, H called the dog to herself and held the dog on their collar as long as the Experimenter E prepared the apparatus. When the apparatus was ready E pushed the boards out and the Handler released the dog to make his choice. After the dog's choice the Handler called him back and rewarded him. For the wolves, H (standing outside the wolf's pen) threw a piece of reward to the end of the pen opposite to the side where the apparatus was positioned. He also covered this part of the fence by pulling an opaque sliding door in front of it. When the apparatus was ready, H threw another piece of treat to the opposite side of the pen and opened the sliding door only when the wolf was 1.5m away from the apparatus. In this way, also the wolves were positioned at this distance to the apparatus at the beginning of each trial, similarly to the dogs. After the subject had made a choice, H called the wolf back to her or closed the sliding door again, thereby separating the wolf from the apparatus.

2.4 Training phase

The Training phase for the dogs started with a short familiarization of the room and the apparatus on the first day. During the training phase the second partition was placed directly behind the Plexiglas barrier to prevent the dogs form seeing the reward on the board to avoid possibly learning effect preferring this board and to block access to the second board. Then the handlers trained the animals to pull out one board using a shaping procedure, where a single baited board will be pushed out (pseudo randomly on the left or on the right side) until the front part of the board becomes accessible to the dog. The food reward will initially be placed on the board just behind the front partition. The distance between the reward and the partition will be gradually increased in subsequent shaping trials. Dogs that successfully retrieved the reward placed at the furthest distance of the board four consecutive times (twice on each side) proceeded to the testing phase.

2.5 Testing phase

For dogs in the testing phase, every session started with a short checking of the room. The test phase lasted several days and was scheduled depending on the animal's availability. But usually, every day or with a break of 2 to 3 days. Each animal received 20 sessions (10 trials each), in total 200 trials. There were short breaks between sessions, and the animals received no more than 2 or 3 sessions per day. At the beginning of each trial, the reward was placed on the boards according to the experimental condition (natural or magic). Each test trial will start by the Experimenter informing the Handler the dog is ready to start, because E cannot see behind the partition. Then E will push out the apparatus so that its front with the first wooden mark becomes accessible to the dog and the Handler released the animal to make a choice. Choice was defined as moving one of the boards for the first time. After this, the animals were allowed to further manipulate the chosen board till they reached the reward, upon which the handler called them away from the apparatus. If they left the apparatus or tried to switch to the other board before accessing the reward (usually after having experienced that the reward did not move with the board manipulated), the handler called them back to the start position and E pulled the apparatus back, thereby making the boards inaccessible for the animals. The correct side was counterbalanced within subjects, and varied pseudo randomly so that the same side never was correct more than twice in a row.

2.6 Behavioural coding

The animals were video recorded during the tests and their choices (correct or incorrect) were coded for all trials.

2.7 Statistical analysis

We fitted a logistic Generalized Linear Mixed Model (GLMMs; (Baayen, 2008)) fitted via maximum likelihood using the statistical program R (version 4.2.3; (R Core Team, 2022)) using the function ‘glmer’ of the package ‘lme4’ (Bates et al., 2015), with the optimizer “bobyqa” (Powell, 2009) with 100.000 iterations.

As a response we used correct choice (no/yes). As test predictors, we included species (as a factor with levels dog and wolf); group (as a factor with levels on and off); session number (as a covariate ranging from 1 to 20); and all their interactions up to the third order. Additional control predictor was accessible side (a factor with levels left and right). Before being included in the model, session number was z-transformed to ease model convergence and achieve easier interpretable model coefficients (Schielzeth, 2010).

To account for repeated observations of the same individual as well as to avoid pseudo-replication, we included the random intercept effects of subject. Additionally, to model day specific variation in motivation/mood, we included a factor combining subject and date of session subject.date nested within subject. To avoid overconfident models and to keep the Type I error rate at the nominal level of 0.05 (Barr et al., 2013; Schielzeth & Forstmeier, 2009), we included all possible identifiable random slopes of accessible site and session number in subject and subject.date. We checked for possible unidentifiable correlations between random slopes and intercepts (with absolute correlation parameters estimated as 1) (Matuschek et al., 2017), which were not present.

After fitting the model, we confirmed that the none of the model assumptions were violated and assessed model stability. We verified absence of collinearity by calculating the Variance Inflation Factor (VIF) using the R package “car” version 3.0-12 (Fox & Weisberg, 2019), which revealed that collinearity was not an issue (max VIF = 1). We visually inspected an confirmed that the best linear unbiased predictors (BLUPs) per level of the random effects were approximately normally distributed (Baayen, 2008). To estimate the stability of the model (Nieuwenhuis et al., 2012), we excluded the levels of random effects one at a time and compared the resulting estimates with those obtained from the model based on all data. This revealed that the model was only of moderate stability, likely caused by the low level of individuals tested (see results table).

To avoid an increased type 1 error risk due to multiple testing (Forstmeier & Schielzeth, 2011), we first tested the overall effect of the test predictors. Therefore, we compared the full with all terms included to a null model lacking the test predictors but only including accessible side and similar in the random effects structure.

We calculated confidence intervals for the model estimates by applying the function 'bootMer' of the package 'lme4', using 1,000 parametric bootstraps.

3. Results

The full model provided a significantly better fit to the data than the null model ($\chi2$= 21.92, df = 7, $p < .004$). We therefore inspected the effects of each test predictor. We did so by reducing model complexity and dropping non-significant interactions, from higher order to lower order terms, from the model one at a time and compare the simpler with the more complex model utilizing likelihood ratio tests (Dobson & Barnett, 2018). This revealed a non-significant effect of the three-way-interaction ($\chi2 = 0.310$, df = 1, p = 0.6). Reducing the model to only significant interactions and main effects resulted in a final reduced model, which showed a significant interaction effect between session number and species ($\chi2 = 11.965$, df = 1, $p < 0.001$). Specifically, while the probability to be correct was similar over session number for wolves, dogs' probability to be correct showed an increase with session number.

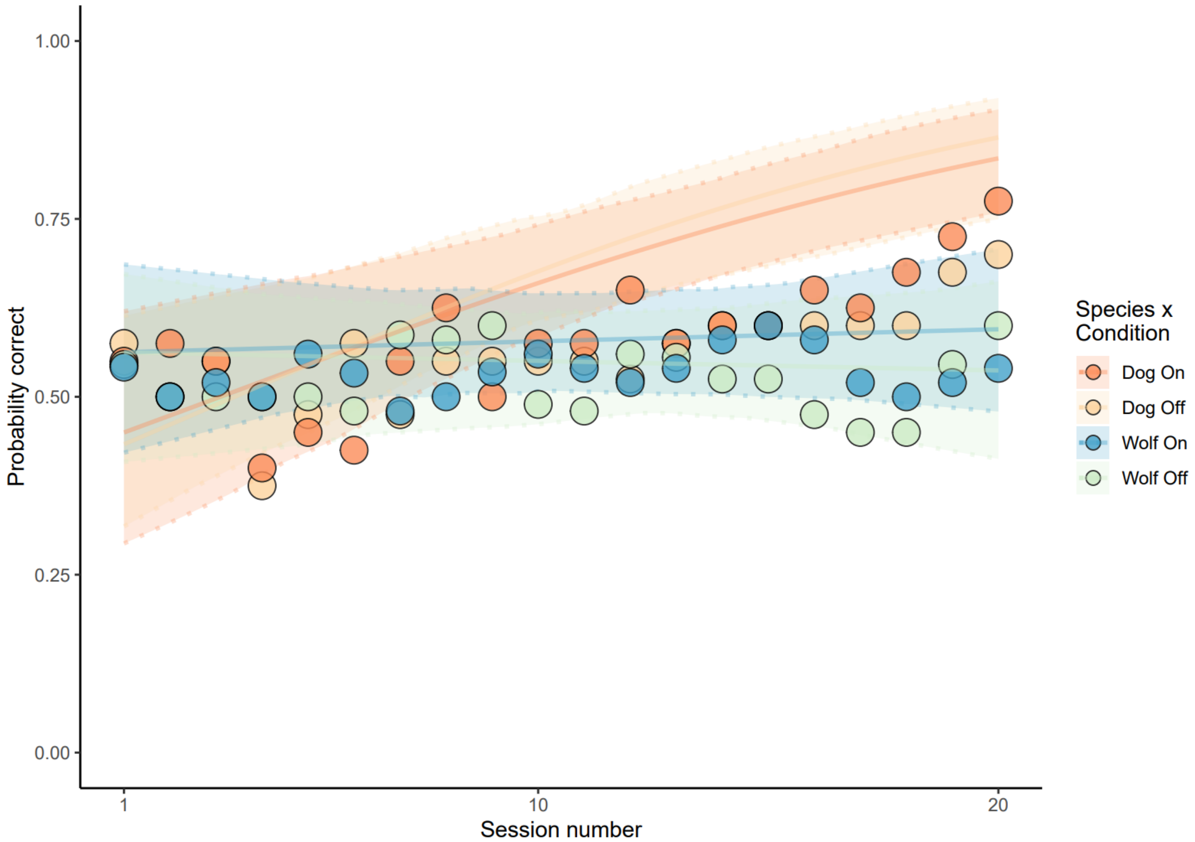

Figure 9 - Graph Probability Correct (Dogs & Wolfs)

Term	Estimate	*SE*	lower cl	upper cl	χ^2	*df*	*P-value*	min	max
(Intercept)	0,927	0,716	-0,527	2,281	NA		(3)	0,636	1,356
specieswolf	-0,471	0,169	-0,808	-0,131			(3)	-0,621	-0,306
z.session.nr	0,581	0,101	0,361	0,801			(3)	0,507	0,678
conditionON	0,052	0,167	-0,315	0,392	0,090	1	0,764	-0,086	0,165
accessible_sideR	-0,448	1,441	-3,180	2,353	0,096	1	0,757	-1,336	0,080
specieswolf:z.session.nr	-0,568	0,131	-0,858	-0,300	**11,965**	**1**	**0,001**	-0,671	-0,465

Figure 10 - Results of the final reduced model

4. Discussion

In this study we used the ON / OFF task to test dogs and wolves, each divided into a natural group and a magic group. The ON group (natural) with the reward on top of the board, and the OFF group (magic) with the reward on the side of the board. If the animals had solved the task by means-end understanding, the natural ON group should have shown faster progress in learning compared to the magic OFF group. We found no evidence however that either the dogs or the wolves would have differentiated between the two conditions. This effect was missing in the two species for different reasons, however. The wolves chose one of the two boards randomly throughout their 20 sessions, and we found no evidence that their performance would have improved with the sessions. In contrast, the dogs' performance increased during sessions and did so similarly in both conditions: they started to perform above chance at least in their last sessions. Pulling the correct board does not always require means-end understanding. A subject pulling the board and the reward is moving leads to the association between the two. An action that might be repeated without means-end understanding. If the successful performance is only based on associative learning, direct perceptual feedback is essential. Our dog results seem to confirm the former suggestion by Müller et al., that dogs can learn to solve the support problem based on perceptual cues. Alternatively, it is in principle possible that the dogs perceived the natural condition more natural than the magic one, but still learnt to pull on the OFF board in the magic group with the same speed as pulling the ON board in the natural group because of their high flexibility in learning. They can adopt new cues when old ones become unreliable. In our case the dogs learn that in the magic condition the reward comes with the board on the side. That's new but the dogs can adopt this new cue and learn to pull the board with the hovering reward. Other studies also indicate that dogs are flexible learners and can adopt quickly to any rule. An interesting preliminary study showed that two dogs successfully operated a robot through a straight course using tug, button, and proximity affordance (Byrne et al. 2019). In a former study on means-end understanding (e.g., Range et al.2012) using the ON/OFF task the dogs solved the task quickly (12 trials) perhaps by learning and owed to their young age, the mean age of the animals was 6 years. This was not the case in this study, the dogs needed at least 100 trials to start to master the task, but taking their age into account.

What could have slowed learning down in this study (as in many other, two-choice, studies) is side bias. A common problem with all patterned string problems is that subjects could develop a side bias. The side bias is a successful strategy half of the time, meaning the subject obtains the reward in 50 percent of the trails (Jacobs & Osvath 2015). Dogs and wolves seem to be prone to side preference in both conditions the natural as well as the magic version.

Interpreting the results, the wolves seem to choose a board more randomly in all sessions than the dogs. Choosing randomly may show disinterest. The wolves chose any board to end the trial, not trying to understand/learn the task. One reason can be a lack of motivation (Jacobs & Osvath 2015). This can be fixed by using a reward that is more desired or by food deprivation before the tests. Another reason for the lack of learning by the wolves can be because they were rewarded, independent of their choice, shortly after their choice in each trial: for calling them to start the test. In contrast, the dogs were not systematically rewarded for returning to the Handler, thus, the bait gained by pulling the right board may have been more important for them.

It is important to note, however, that we can form only limited conclusions based on this data because it seems that 20 sessions were probably too few: the dogs just started to learn at the end. With more test sessions we may have found group differences, which would tell a different story. And of course, the low sample size makes the evidence weak. More individuals combined with more observations would be a recommendation for further studies.

Not neglectable is the fact that the age of the dogs and wolves were not comparable. Motivation and concentration on the task are different in different stages of development. A former study has shown that canine performance on a two choice discrimination task varied with age and senior canines made significantly more errors than young (Snigdha et al. 2011). With two exceptions we have only old and senior, meaning older than 8 years, animals in this study. This can have influenced the results for both wolves and dogs' performance but cannot explain the species difference though, as we had younger wolves but no younger dogs.

Also, worth to mention are the differences between sexes in cognitive processes. Müller et al. 2011 tested male and female domestic dogs in an object permanence task. Their results indicate that there is a sex-specific performance in a physical cognition task (Müller et al. 2011). Taking above mentioned parameter into account our suggestions for further studies are a larger sample size that would allow for controlling the effects of age and sex on the performance of the animals.

Another limitation of the study is that, taking gravity into account, it is not “natural” that the reward on the side of the board is hovering. Relating to this fact a recommendation for further studies is to change the setup slightly by adding a surface on the side of the board that apparently carries the reward.

5. References

Baayen, R. H. (2008). Analyzing Linguistic Data: A Practical Introduction to Statistics using R. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511801686

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 10.1016/j.jml.2012.11.001. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1 SE-Articles), 1–48. https://doi.org/10.18637/jss.v067.i01

Dobson, A. J., & Barnett, A. G. (2018). An introduction to generalized linear models. Chapman and Hall/CRC.

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. Behavioral Ecology and Sociobiology, 65(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Fox, J., & Weisberg, S. (2019). An {R} Companion to Applied Regression (Third). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Frank H (1980), Evolution of canine information processing under conditions of natural and artificial selection. Z Tierpsychol 5:389-399

Frank H, Frank MG (1982), Comparison of problem-solving performance of 6-week-old wolfs and dogs. Anim Behav 30:95-98

Irie-Sugimoto N, Kobayashi T, Sato T, Hasegawa T (2007), Evidence of means-end behaviour in Asian elephants (Elephas maximus). Springer, Anim Cogn (2008) 11:359-365

Jacobs I (2018), String Pulling. Springer, Encyclopaedia of Animal Cognition and Behaviour, https://doi.org/10.1007/978-3-319-47829-47829-6_1505-1

Jacobs I, Osvath M. (2015), The String-Pulling Paradigm in Comparative Psychology, American Psychological Association, Journal of Comparative Psychology 2015, Vol. 129, No. 2, 89-120

Garcia-Pelegrin E, Schnell AK, Wilkins C, Clayton NS (2020), An unexpected audience. Science, 18 September 2020, VOL 369 ISSUE 6510, p1424-1427

Lampe M, Bräuer J, Ksminski J, Virányi Z (2017), The efects of domestication and ontogeny on cognition in dogs and wolfs. Nature, Scientific Reports, 7: 11690, DOI:10.1038/s41598-017-12055-6

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. Journal of Memory and Language, 94, 305–315. https://doi.org/https://doi.org/10.1016/j.jml.2017.01.001

Müller CA, Riemer S, Virányi Z, Huber L, Range F (2014), Dogs learn to solve the support problem based on perceptual cues. Springer Online, Anim Cogn (2014) 17:1071-1080

Müller CA, Mayer C, Dörrenberg S, Huber L, Range F (2011), Female but male dogs respond to a size constancy violation. Biol. Lett. published online 27 April 2011, Doi: 10.1098/rsbl.2011.0287

Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B. (2012). Influence.ME: Tools for detecting influential data in mixed effects models. R Journal, 4(2), 38–47. https://doi.org/10.32614/rj-2012-011

Osthaus B, Lea SEG, Slater AM (2004), Dogs (Canis lupus familiaris) fail to show understanding of mens-end connections in a string pulling task. Springer Online, Anim Cogn (2005), 8: 37-47

Piaget J (1952), The origins of intelligence in children. International University Press, New York

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. Cambridge NA Report NA2009/06, University of Cambridge, Cambridge, 26.

R Core Team. (2022). R: A Language and Environment for Statistical Computing. https://www.r-project.org/

Range F, Hentrup M, Virányi Z (2011), Dogs are able to solve a means-end task. Springer Online, Anim Cogn, DOI 10.1007/s100071-011-0394-5

Range F, Möslinger H, Virányi Z (2011), Domestication has not affected the understanding of means-end connection in dogs. Springer Online, Anim Cogn, DOI 10.1007/s10071-012-0488-8

Schielzeth H. (2010). Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution, 1(2), 103–113. https://doi.org/https://doi.org/10.1111/j.2041-210X.2010.00012.x

Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. Behavioral Ecology : Official Journal of the International Society for Behavioral Ecology, 20(2), 416–420. https://doi.org/10.1093/beheco/arn145

Snigdha S, Christie LA, De Rivera C, Araujo JA, Milgram NW, Cotman CW (2010), Age and distraction are determinants of performance on a novel visual search task in aged Beagle dogs. AGE (2012) 34:67–73, DOI 10.1007/s11357-011-9219-3

6. Table of Figures