Cognition of dogs: developing a Guesser-Knower task.

"MSc thesis research report"

Baker, A. (2014). Canine brains: test your dog's intelligence.

Name: M. Beurden, van
Chairgroup: Behavioural Ecology Group
Date: 10-04-2019
BHE80336

Cognition of dogs: developing a Guesser-Knower task.

M. van Beurden
April 2019

MSc Thesis
Behavioural ecology Group
Research report

WAGENINGEN UNIVERSITY

Abstract

To what degree dogs are truly altruistic is subject to debate as their assistance of humans could be explained by operant conditioning and constitute learned responses to obtain rewards. Altruism requires a Theory of Mind (ToM), meaning an understanding of the mental states and needs in others. The present study aimed to gain knowledge on ToM and altruism in dogs and contribute to the scientific debate on altruism in dogs. Using a Guesser-Knower task, by means of multiple protocols, it was investigated if dogs (N=34) understood the attentional state of a human informer and his/her knowledge about the location of a hidden food reward. This existing test on ToM was adapted for use within a short time frame. Altruism was assessed by testing dogs (N=40) on signalling behaviours (i.e. informing) following their owners' request (or not, being the control) for help in finding his/her coat, of which the location was known to the dog. In the altruism test, the owners had to guess the correct hiding box (out of 3) based on their dog's signalling behaviour, which they managed in control trials 16 out of 45 times (binomial one-sided $P>0.1$) and in test trials 15 out of 45 times ($P>0.1$). In the trials in which owners guessed the correct hiding box and were sure about the hiding place, dogs showed increased signalling behaviours by looking at the target ($P=0.024$) and approaching the target ($P=0.003$ for Linear Mixed Model 2-way interaction effect of correct choice and no doubt). Dogs showed more tail wagging, looking at the target and referencing at the target when owners requested help than when they did not ($P<0.001$ for the fixed effect treatment). In the training part of the ToM test, 34 dogs chose 500 times the correct bowl (two-way choices) out of 595 trials (binomial one-sided $P<0.001$). The choice for the correct bowl was above chance for all protocols and clearly dogs followed the pointing gestures by the experimenter to find the hidden foods. In the experimental phases twenty-four dogs chose the correct bowl (out of four) 98 times out of 190 trials (binomial one-sided $P<0.001$). Finally, dogs that made more than 50% correct choices were selected and analysed. Nine dogs together chose the correct bowl (out of four) 66 times out of 93 trials (binomial one-sided $P<0.001$). In response to their owners' requests for help in finding a missing coat the dogs showed increased levels of informing behaviour, suggesting an altruistic response. Only some dogs seemed to grasp the attentional state and knowledge of human informants in the Guesser/Knower task. Deviations from the protocol in previous studies, such as the removal of a screen, may have caused some dogs to base their choices on associative learning instead of using the attentional state of the informant. Given that dogs showed signs of altruism it seems unlikely they lack a rudimentary theory of mind. A rudimentary theory of mind implies that ToM is not an all or nothing concept but exist of various degrees where dogs could have a more moderate degree of ToM compared to humans. More research is needed to better understand the degree of ToM in dogs and for example how these are a reflection of different individual life-histories.

Table of contents

1. Introduction

Cognition is a concept defining mental capacities of animals which mostly includes the ability to remember, communicate, and categorize objects or persons (Paiget & Inhelder, 1969). In the last couple of decades, research has been performed in an attempt to discover the cognitive abilities of domestic dogs. Across cognitive studies with dogs variable outcomes have been reported, often without clear explanations. What is clear is that dogs are unique in social cognitive abilities, possibly due to their relationship with humans. Pet dogs respond to human body language such as pointing gestures (Dorey, Udell & Wynne, 2010; Wynne et al., 2016), social cues in finding food (Hare & Tomasello, 2002, 2005; Kaminski, 2009) and they react to human attentional states (Maginnity & Grace, 2014). Studies that focus on demonstrating higher cognitive abilities such as Theory of Mind (ToM) and altruism are typically somewhat inconclusive. Flaws in experimental design such as not taking into account the live history of the subjects (Horowitz,2011), or designs where the goal is not clear (Bräuer et al.,2013), may compromise the interpretation of the findings. The notion that one study can irrefutable demonstrate ToM or altruism in dogs seems unrealistic and rather a large number of studies, with each having its specific imperfections, should point towards the most likely view. The present study contributes to this by conducting behaviour tests with privately owned dogs on ToM and altruism.

Altruism is seen as helping others, where there is a benefit to the receiver and a cost to the actor. In this, the actor has to be motivated to help and should understand the goal of the receiver (Warneken & Tomasello, 2009b). There are different kinds of altruistic behaviour such as instrumental helping, showing behaviour and 'regarding-others' behaviour. Instrumental helping is 'acting for another individual to achieve its behavioural goal' (Warneken and Tomasello, 2006, 2009a, 2009b). The pro-social behaviour of dogs by helping humans has been investigated earlier (Bräuer et al., 2013). A human indicated to want to enter a room, which the dog could open by pushing a button. Different owner to dog communication protocols were used, such as natural communication, pointing at the door, pushing the door, or reaching for the key, and it was observed whether the dog opened the door. When natural communicative behaviour was used to express intention, dogs were motivated to help (Bräuer et al., 2013). Showing behaviour refers to a dog pointing out an object or food that is hidden by means of communicative signals and actions (Miklósi et al., 2000). Dogs did not perform showing/helping for finding a relevant object (notepad) for the human when signalling a dog toy was the alternative (Piotti & Kaminski, 2016). Though the human expressed it was looking for the object the dog did not perform showing behaviour towards to location of the notepad but mainly acted out of self-interest by indicating the location of the toy. Next, only one object was hidden, either a relevant (notepad) object or distractor (stapler) object and the experimenter voiced (or not) to look for the object. Now indeed showing behaviour was demonstrated and particularly when the relevant object was hidden and verbal request to find the object was used (Piotti & Kaminski, 2016). Moreover, dogs tend to be altruistic to other familiar conspecifics by sharing food (Quervel-Chaumette et al., 2015). A bar-pulling task was used to identify 'other-regarding' behaviour. In this test the dog would have a choice to pull a rope that would deliver a tray with food to the next enclosure or pull a rope to deliver an empty tray to the next enclosure. Dogs would hardly deliver an empty tray (n=16, range of 0.6 to 1.5 times) compared to a food tray (mean of 17.2 times). Empty trays were delivered more to stranger dogs then familiar dogs, suggesting an altruistic response to familiars (Quervel-Chaumette et al., 2015). However, dogs were not pro-social towards humans and did not discriminate between a familiar human and a stranger in this bar-pulling test. Important in this study was that the humans were not allowed to communicate with the dog in any way (Quervel-Chaumette et al., 2016). Preliminary studies show that altruism in dogs is difficult to demonstrate, and outcomes depend on whether or not communication is used and what tasks/objects are used to test for altruistic behaviours.

Theory of mind (ToM) implies that an individual believes that ‘’mental states play a causal role in generating behavior and infers the presence of mental states in others by observing their appearance and behavior under various circumstances’’ (Heyes, 1998, p.102). Difficulties in finding evidence for ToM in animals, especially non-primate studies such as dogs, is in part due to the complex distinction between ‘mind reading’ and other cognitive processes like behaviour-reading and associative learning. Perspective-taking tasks are designed to see whether animals behave according to information they get from observing other individuals and situations. Findings from perspective-taking tasks suggest that dogs do read minds and their behaviour is not the result of merely associative learning or behaviour reading (Maginnity & Grace, 2014). Perspective taking tasks are appropriate for pet dogs to do in relation to their owners because of their strong bond with humans. Dogs respond to signals and cues from an informant’s attentional state, which shows for example in a visual perspective-taking task known as Guesser-Knower task (Maginnity & Grace, 2014). A Guesser/Knower task, developed by Povinelli et al., 1990, is a food-hiding task which consist of two or three informants. The three informants place themselves behind four cups, which are blocked from the view of the owner and the dog by the screen. In view of the dog, one informant (baiter) baits one of the cups, with the latter being out of view and behind the screen. The ‘Knower’ informant is present and can see the baiting while the ‘Guesser’ informant cannot, because of looking elsewhere or being out of the room. In the read-out phase, the ‘Guesser’ points to an empty cup and the ‘Knower’ points to the baited cup. A dog’s choice for the baited cup indicates it understands the attentional states and knowledge of the informants. Dogs showed a significant tendency to follow the pointing gestures of the ‘Knower’ over that of the ‘Guesser’ when the latter had been absent, had its eyes closed or when the ‘Guesser’ had looked at the ceiling at the time of baiting (Maginnity & Grace, 2014; Catala et al., 2017). Dogs did not discriminate between informants when both were ‘Guessers’. Variation in choice of informant was not explained by the dogs’ sex or age, but surprisingly obedience trained dogs showed less preference for the ‘Knower’ compared to untrained dogs (Maginnity & Grace, 2014). The experimental set-up rules out explanations such as guidance by olfactory cues, bias for choice of informant and ‘Clever Hans’ effects. Associative learning was unlikely to explain the findings since no changes in preference for the ‘Knower’ was found over trials. The same Guesser/Knower task has been adapted to test whether the dogs’ assessment of human knowledge could go further than just direct behaviour reading (Catala et al., 2017). Both informants behaved in the same way and looked in the same direction to the ground (to their right at an angle of 45 degrees). The difference was that the ‘Knower’ could see the baiting while looking at the ground and the ‘Guesser’ could not. Dogs chose the ‘Knower’ significant more times than the Guesser, suggesting that subtle cues (the line of sight) are enough for dogs to discriminate between informants based on geometrical gaze following (Catala et al., 2017). These two studies suggest that dogs can follow gazing direction and are aware of the attentional state of humans. The robustness of the findings should show in replications of the findings in other comparable experiments and here I aimed to establish this.

In this study I explore different possibilities for testing a dog’s theory of mind, including the replication of a Guesser/Knower paradigm. The first aim is to test for ToM in dogs by means of assessing if dogs understand the attentional state of a human informer. It is expected that dogs can identify a baited food bowl by heeding signals from a knowledgeable informant whilst disregarding signals from an ignorant informant. The second aim is to determine if dogs act altruistically towards their owners by informing them about the location of a missing coat. Do dogs behave altruistically by using signalling behaviours and thus help their owners find a missing object when he/she asks for help? Expected is that dogs will use signalling behaviours to indicate the location of the missing coat. The findings add to the existing framework on ToM in dogs and help to better understand dogs’ abilities in understanding our mental states, including desires and intentions.

2. Materials & Methods

The study consisted out of two parts. During an altruism test the coat of the owner was hidden in one of the three boxes, out of sight of the owner but with the dog observing the hiding process. When the owner returned and asked the dog explicitly were the coat was it was recorded whether dogs performed signalling behaviours. The owner could use info-sharing behaviour to guess which box contained the coat. The second part of the study was about the design, execution and analyses of a ToM test following the Guesser/Knower (G/K) design. The ToM test consisted of four training phases during which the dog was accustomed to the set-up of the experiment and two experimental (read-out) phases (one with two informants and one with a third person involved).

2.1 Altruism test

Forty subjects (domestic dogs with their owners) were chosen from the database (van Herwijnen et al., 2018). Dogs lived with their owner and were brought to Carus facility for testing. To assess whether dogs show signalling behaviours while there is no benefit for the dog itself a box test set-up was constructed. It was tested if dogs signalled the location of a missing coat of which the hiding place (one of three boxes) was known to them but not to the owner. The test was performed in a room with four chairs, three boxes and a screen in the corner of the room (appendix 1, figure 3). Two of the chairs were numbered so the owner knew on which chair to sit and hang their coat on. Each dog had a control trial and a test trial. During the test trial the dog and owner entered the room and got two minutes to explore and adjust to the room. After the two minutes the dog was secured on a leash that was attached on the wall and the owner sat on one of the numbered chairs and hung their coat on the chair. After one minute the owner was instructed to sit behind the screen and put the headphones on which played classical music as to mask sounds that could betray the hiding place. The experimenter then entered the room and made eye contact with the dog (called its name). When eye contact was made the experimenter hid the coat in one of three boxes and left the room. The owner was then instructed to stand behind the chair on which they sat earlier, make eye contact with the dog and ask: "Where is my coat?", whilst raising arms and looking around in a searching way. After five seconds of silence the owner would repeat this. The test leader then asked the owner if they had an idea of where the coat was and regardless of a yes or not, the owner then made a choice on which box contained the coat. During the control trial the procedures were as described, expect that that the owner did not ask the dog where the coat was. The owner made eye contact with the dog and looked at it in a neutral way. The protocol that was used by Jager et al. (2017) is placed in appendix 1.

2.1.1 Analysis

Dogs (n=40) were video recorded and analysed with an ethogram (appendix 2) in Noldus Observer. The time spent in a certain zone (7 zones) was recorded (%). Stress signals such as whining etc. were recorded in rate per minute (rpm). Signalling signals were recorded in rpm per time zone. Chi-square tests were used with a two-factorial design to identify whether observed values deviated from expected values. Standardized residuals were used to identify significant deviations from expected counts, residuals > |2| were considered to be significant. Binomial tests were used to determine whether factors such as choosing the correct box happened above chance (33%). A mixed linear model (REML) in GenStat was used to test the relationships between dog behaviour scores and the owner raising the thumb (0,1), being correct in localizing the hiding box (0,1) and trial (control or test). Thus, the statistical model included three main fixed effects with two-way interactions. The model used was: $Y_{abcd} = \mu + Trial_a + Owner\_thumb\_up_b * Owner\_correct_c + ID_d + e_{abcd}$. In the model the random factor ID (dog) accounted for the 2 repetitions for each dog. For significant interactions the contrasts were considered significant when the difference between predicted means was higher than two times the standard error of differences (p<0.05 when Δ>SED*2).

2.2 Theory of Mind test

During the Theory of mind test (ToM) it was investigated whether dogs identified the baited bowl (one out of four) correctly in a Guesser/Knower task (G/K) based on making use of perceptual knowledge of a human informant. Each of two informants pointed towards a specific bowl of the 4 available after which dogs could select one of the four bowls. Only one experimenter witnessed the baiting of the bowl and dogs follow the pointing gestures of this person and ignore those of the uninformed experimenter. Six protocols were used derived from the procedure described by Maginnity & Grace, 2014. The protocols were adjusted to find one that resulted in the highest success. Two pre-liminary protocols were excluded from the study. In this paragraph the focus is on protocol six, because the most dogs participated in this protocol. Protocols three, four and five are explained in detail in appendix 3. Protocol six consisted of two parts with six phases. Training consisted of four phases and dogs continued to the next phase after reaching a certain (phase specific) number of correct choices. The experimental part consisted of two phases.

2.2.1 Subjects

Fifty owner-dog dyads were chosen from the database of the survey of van Herwijnen et al (2018) or through the social network of the experimenters. All the test subjects lived with their owners and were brought to Carus facility for testing. The dogs were naïve in Guesser-Knower testing, but some dogs had earlier experience with behaviour testing. All the dogs were tested individually with their owner present, owners were instructed on how to act during testing. The performed tests were not considered to be animal experiments by the Animal Care and Use Committee of Wageningen University August 2017, there were no invasive treatments in the life of the owners and / or their dogs.

2.2.2 Apparatus

All the tests were done in the dog behaviour test room (7 x 7 m) at the research facility Carus at Wageningen University. Dogs had been in and about the room for about half an hour prior to tests. The room was video recorded during the tests with four Axis M10 network camera from each corner. The setup of the test room during training was based on Maginnity & Grace, 2014. Materials used in the test room were a yoga mat, a green dog leash, eight plastic dog bowls, two hand grabbers, a poster on the wall and a removable screen (figure 1A). The yoga mat indicated the dog release point (DRP). A leash was secured to the wall, so the dogs could not reach the informants but could reach the bowls. Two meters from the DRP five pieces of tape were placed on the ground in a halve circle with a range of sixty cm away from each other. The pieces of tape indicated the positions of the bowls. Position five (figure 1A, green bowl) was used only in the first phase were only one bowl was used. The plastic bowls (diameter of 25 cm) were secured to the ground with grey duct tape. Other plastic bowls (diameter of 25 cm) were placed on top of the bowls that were secured on the ground, so each position contained two bowl with one a false bottom. All the lower bowls contained a commercial dog treat (Caniland soft happen, Straus mit kartoffel), so all the bowls smelled like food. The treats were kept in the pockets of the informants. Two informants placed themselves behind the bowls. A removable pin board screen (2mx1.20m) was placed one meter behind the informants, the screen was placed so the dogs could see the informants from everything above the knees. A scheme was hung on the pin board screen for the informants to see which bowl was baited and which position they should occupy in each trial and phase. The informants used hand grabbers to point at the baited bowl. A poster was hung on the wall, for the owners to look at during baiting. During phase six (experiment) a third informant: the baiter was introduced which placed itself between the two informants (figure 1B).

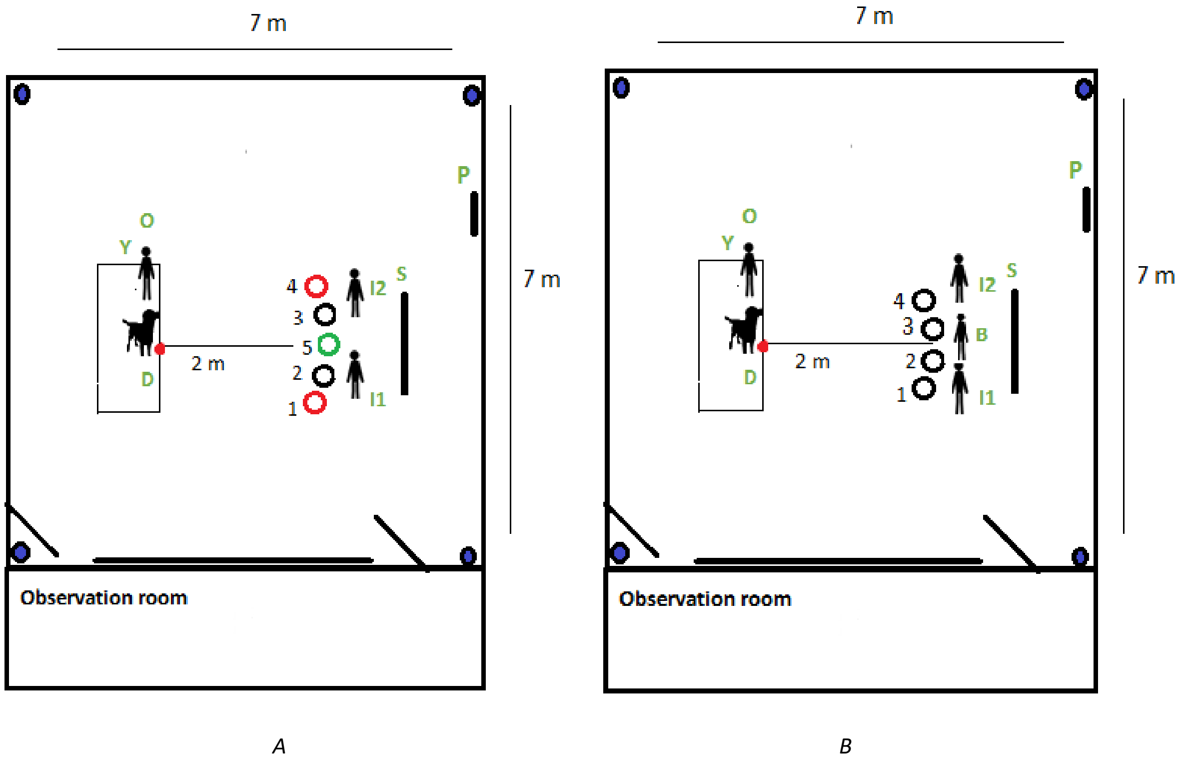

*Figure 1: Set up of room at Carus during training (**A**) and the experimental Guesser/Knower test (**B**) of the ToM test in protocol 6. The owner (O) was instructed to stand beside the dog (D) and let the dog sit on the dog release point (red dot) on the yoga mat (Y). The owner was informed to look at the poster (P) on the wall. A removable screen (S) was placed 3 meters from the dog's release point. The numbers indicated the positions of the bowls. **A:** Two informants (I1 and I2) placed themselves behind the bowls. During phase one only the green bowl on position 5 was used, in phase two the two red bowls on position 1 and 4 were used, in phase three and four the red and black bowls on position 1,2,3 and 4 were used. **B:** The three informants: the baiter (B) and the two informants (I1 & I2) placed themselves behind the bowls. In the experiment four bowls on position 1,2,3 and 4 were used. The baiter handled the baiting of the treat.*

2.2.3 Training

The goal of the training was to accustom the dog to the setup of the experiment (sounds, informants, grabbers) and learn dogs to select the baited bowl by following the pointing gesture of the grabber. The informants and roles differed per trial to even the exposure to both informants and to prevent bias for one informant and side. Positions and roles were changed after at most two consecutive trials. Informants always kept the grabber closed in their right hand and their arms along their sides. The informant who did not bait the bowl, looked at the wall behind the owner and said nothing during the time the reward was placed by the second informant. The informant who baited the bowl, interacted with the dog during baiting and kept his/her eyes on the bowl during pointing. Following baiting, the baiter informant would count to three after which both informants pointed towards a different bowl simultaneously.

The basic set up for performing the G/K task was the same for all phases (table 1). Dog and owner entered the test room with the test leader and the owner and dog were directed to position themselves on the mat. The owner leashed the dog, stood on the yoga mat and commanded the dog to sit on the mat on the DRP beside them. Owners were instructed to keep the dog on a short leash, so the dog would not walk to the bowl prematurely. Owners were asked not to use hand-gestures but commands such as 'go', or 'search' were allowed. The two informants placed themselves behind the bowl in front of the pinboard screen. Before the start of the first trial the informant that was the baiter instructed the owner to look at the poster during all trials so no 'Clever Hans' could occur. The other informant (not the baiter) looked at the wall, facing away from the dog and bowls, and did nothing. The baiter then got the attention of the dog (called its name) and showed it the food treat and, put it in the bowl. After baiting the baiter instructed the owner to look away from the poster. After a two second delay the baiter pointed at the baited bowl and instructed the owner to release the dog and encourage it to go to one of the bowls. The baiter kept her/his gaze on the bowl and the informant that not baited the bowl kept facing the wall. The choice of the dog was defined as: nose in the bowl. When the dog made the correct response, the dog ate the treat in the baited bowl and was praised by the baiter. A 'no response' was recorded when dogs did not respond within one minute after which the dog was called back without reward or vocal praise. Following a response, the baiter then instructed the owner to call the dog back to the starting position. In the last trial, or last two trials, of the phase (depended on phase) both informants pointed at the baited bowl to let the dog experience two informants pointing at the bowl. When the dog did not make the correct choice or any (i.e. 'correct response'), phase one was repeated. Before the start of phase two, where a second bowl was introduced, owners were instructed to retrieve the dog shortly after they made a choice, so when an incorrect choice was made dogs were not able to eat the treat from the baited bowl. When the dog made an incorrect choice, phase two was repeated. In phase three a third and fourth bowl were added. In case of incorrect choices phase three was repeated. In the last phase of training, phase four, false baiting was introduced. False baiting consists of touching all the bowls at least once while pretending to bait. Only one bowl was actually baited. The baiter would keep two pieces of food in his/her hand, when baiting the bowls, the food was dropped in the bowl and picked up again. The informant would leave one piece of food in the assigned baited bowl and continue false baiting with the second piece. By using a second piece of food dogs would always hear the food being dropped even when the bowl was not baited. The number of trials in phase four was depended on the number of correct choices. When in the first set of four trials less than three trials were correct another set of 4 trials was performed. When again less than three choices were correct another set was executed. When after three sets of four trials the dog made less than three correct choices/no choice per set the test was stopped. When three or more correct choices were made in phase four a break of 5 minutes was inserted, before the experimental phases started.

Table 1: Protocol of each phase used during the training part of the ToM test in protocol 6 .The goal of the phase, the number of bowls used, the position of the bowls (figure 1A), the number of trails executed and the number of trials that needed to be correct to move on to the next phase and the number of trials were two informants pointed at the same bowls are described. The consequences for not having the right number of correct choices is displayed.

Phase	**Goal**	**Number of bowls**	**Position(s) bowl**	**Number of trials**	**Number of correct trials needed**	**Number of trials where 2 inf. point**	**When not correct?**
1	Accustom to set up and experience a reward	1	5	3	2	1	Repeat phase 1
2	Let experience a choice	2	1 and 4	3	2	1	Repeat phase 2
3	Let experience a choice between more bowls	4	1,2,3,4	4	3	2	Repeat phase 3
4	Introduce false baiting. Base choice on pointing of the grabber instead on informant. Accustom to experiment set up.	4	1,2,3,4	4/8/12	3	2/4/8	Extra set of trials (4) when again less than 3 correct another set of trials (4). After 3 sets less than 3 correct: end test

2.2.4 Read-out phase

After a break a couple of refresh trials were done before the read-out phase ('experiment') started. Depended on the number of correct choices, 2,4 or 8 trials of phase four were repeated. When after two sets of four trials less than three correct choices per set were made the test was stopped. The experiment consisted of a 'Guesser-Knower' (G/K) test with two phases (table 2). Phase five existed of a simpler version of the G/K test from Maginnity & Grace, 2014. During this phase only two informants were present. The informant that baited the bowl was the Knower, the other informant was the Guesser and he/she was not present during baiting. After the refresh trials phase five started, the informants took their places behind the bowls. The informant's role alternated each trial and were never the same for more than three consecutive trials. When the dog payed attention to the informants the Guesser stepped behind a pinboard screen, out of the dog's sight. The Knower then attracted the attention of the dog by calling its name and (falsely) baited the bowls. After a 'Yes' by the Knower the Guesser returned from behind the screen and retook its previous position. Meanwhile the Knower looked at the wall behind the owner. After a two second delay the Knower counted to three and at three the informants simultaneously pointed at a different bowl. The Knower pointed at the correct bowl (baited bowl) while the Guesser pointed at an incorrect bowl (non-baited bowl). The pointing gestures of the informants never crossed to prevent unclear gesturing. The owner was then instructed by the Knower to release the dog, the dog then had to make a choice. When the dog made the correct choice, it ate the treat and the Knower gave a vocal praise. When an incorrect choice was made the owner retrieved the dog without a reward and a vocal praise and the next trial commenced. When less than four correct choices were made the test was stopped. When four or more correct choices were made phase six started.

The G/K test in phase six was comparable to experiment two in Maginnity & Grace, 2014. A third informant was introduced and this third experimenter took on the role of the baiter and positioned itself between the two informants (figure 1B). The baiter in this test communicated with the dog, handled the food and baited the bowl. When the dog focussed on the informants the Guesser walked behind the pinboard screen. The baiter got the attention of the dog and hid the treat by means of false baiting. During the baiting of the bowl the Knower watched the baiting with an attentive gaze. When the treat was hidden the baiter positioned itself back in the middle and indicated the Guesser to come back by saying "Yes". The Guesser would step back from the pin board screen and positioned himself/herself on its previous position. After a two second delay the baiter would count to three and at three the Knower and Guesser simultaneously pointed at a different bowl following procedures as described.

Table 2: Protocol of phase five and six used during the experimental part of the ToM test in protocol 6. The goal of each phase, number of trials performed, number of correct trials needed to continue to the next phase and who baited the bowl are described. The consequences for not having the right number of correct choices are displayed.

Phase	Goal	Number of trials	Who baited the bowl	Number of correct trials needed	When not correct?
5	To see whether dogs follow the Knower and understand the Knower hid the treat.	6	Knower	4	End of test
6	To see whether dogs follow the Knower and understand the Knower saw the baiting by a third informant.	6	Baiter	4	End of test

2.2.5 Analysis

The dogs' (N=34) choices were recorded together with the positions of the informants and bowls which led to 785 trials in total (phase 1: 90, phase 2: 89, phase 3: 192, phase 4: 224, phase 5: 144, phase 6: 46). Noted were the correct bowl, chosen bowl and last success bowl/touched bowl and the choice of the dog was recorded as one of four discrete categories: Knower/baiter, Guesser, Other bowl and No choice. Chi-square and binomial tests were used to detect Knower preference and to see whether the correct number of choices for the bowl was higher than chance (phase 1: 100%, phase 2 and 3: 50%, phase 4,5 and 6: 25%). Chi-square standardized residuals > $|2|$ were assumed to identify counts that deviated significantly from expected values. A mixed linear model (REML) was used to analyse the probability of dogs choosing correctly or not (i.e. expressed as a binomial response variate, where 1 was choosing the correct bowl and 0 was choosing the incorrect bowl) across different protocols and phases. Thus, the model consisted of two fixed factors (protocol 3 to 6 and phase 1 to 6) and a two-way interaction. The model used was. $Y_{abc} = \mu + Phase_a + Protocol_b + VoterID_c + e_{abc}$. In the model the random factor VoterID (dog) accounted for between 1-30 repetitions per dog. Contrasts were considered to be significant if differences between predicted means were over two times the standard error of differences to be significant ($p < 0.05$ when $\Delta > SED*2$).

3. Results

3.1 Altruism

Dogs were tested for informing their owner about the location of a missing coat that they had observed being hidden in one of three boxes. During test trials, but not during controls, owners asked their dog for help in finding the coat. In total 40 dogs completed the control and the test trial, yielding 90 records. The percentage of male and female dogs was 56% and 44%, respectively. Fifteen out of twenty-two male dogs were intact, and five out of eighteen females were intact. The age ranged from younger than six months to older ten twelve years with an average of six years. Characteristics of the dogs that participated in the altruism test can be found in appendix 4. Behavioural analyses focussed on stress behaviour, signalling behaviour and the associations of these behaviours with the correct choices of owners regarding the hiding place (box) and differences in test trials and control trials.

3.1.1 Data set formation

Behaviours that were rare, occurring in less than 9 out of 90 records, and which could not be combined into meaningful new behavioural parameters, were yawning, paw lifting, shaking and reference to the owner multiple. The rare behaviours whining and yelping were summed up to form a new column. New parameters were made which summarized signalling behaviours in the target zone (zone where coat was hidden), and non-target zone (all the other zones). The same was done for percentage of time spent in zones.

3.1.2 Chance on correct box

The owner guessed the correct box 31 times out of the 90 trials. The proportion of owners that chose the correct box is not larger than chance (33%) over all trials (binomial on-sided $p>0.1$). Choosing the correct box in test trials (15 out of 45 choices, $p>0.1$) and control trials (16 out of 45 choices, $p>0.1$) separate were in both cases not above chance. Clearly, owners were not able to correctly guess the location of the coat, even during test trials when dogs were asked for help.

Combinations of raising the thumb (no thumb up, thumb up) and choosing the correct box (incorrect, correct) was near significant across all trials (χ^2= 4.4, df=1, p=0.04). Counts (residuals) were 51 (0.55 no thumb up, incorrect), 21 (-1.1, no thumb up, correct), 8 (-1.1, thumb up, incorrect) and 10 (1.5, thumb up, correct). There was a tendency towards owners guessing the box correctly when they indicated to be sure about their guess. No significant differences were found in control trials only (χ^2= 1.1, df=1 p=0.27), counts were 26 (no thumb up, incorrect), 11 (no thumb up, correct), 4 (thumb up, incorrect) and 4 (thumb up, correct). Like described for all trial, there seemed to be a trend within the dataset on test trials only (χ^2= 3.4, df=1 p=0.07, with as residual of 1.3 for the combination of thumb up and the owner choosing the correct box. Counts were 25 (no thumb up, incorrect), 10 (no thumb up, correct), 4 (thumb up, incorrect) and 6 (thumb up, correct). Thus, there was a trend towards owners guessing the correct box when they were in no doubt, particularly during the test trials.

3.1.3 Signalling-, stress behaviours and correct choices

Mixed linear models on 90 records (40 dogs) were used to test for relationships between stress or signalling behaviours and the choice of the owner (‘owner choice’: incorrect, correct) during test trials or controls (‘trial’). The statistical models with two fixed effects included the two-way interaction (table 3). There was no effect of the interaction trial and owner choice on any of the stress behaviours. The main effect trial affected tail wagging (p<0.001). During the test trials (5.7±0.72 rate per minute) more tail wagging was performed than during the control trials (3.5±0.72).

Table 3: Dogs were tested for informing their owner about the location of a missing coat that they had observed being hidden. During test trials, owners asked their dog for help in finding the coat and stress-related behaviours were recorded (column 1). Ninety records on 40 dogs were analysed with linear mixed models and the p-values and predicted means (± standard error) per fixed effect level are presented. Fixed effects included Trial, with control (C) and test (T). Owner choice consisted of the two conditions incorrect choice (0) and correct choice (1). The statistical model with two fixed effects included two-way interactions. The behaviours were analysed as rate per minute (rpm).

Stress behaviour (rpm)	**Fixed effects**			
	µ±se	Trial	Owner Choice	Trial.Owner Choice
Bark	1.4±0.81	P=0.61	P=0.51	P=0.18
Whine or Yelp	4.0±1.35	P=0.84	P=0.99	P=0.29
Tongue flick	2.2±0.46	P=0.77	P=0.52	P=0.09
Look away	0.7±0.18	P=0.37	P=0.53	P=0.72
Tail wag	4.8±0.69	**P<0.001**	P=0.21	P=0.93
Panting	5.5±0.86	P=0.37	P=0.58	P=0.28

Mixed linear models on 90 records (40 dogs) were furthermore ran to test for relationships between signalling behaviours and the choice of the owner ('owner choice': incorrect, correct) and whether the owner raised their thumb ('owner thumb': no thumb up, thumb up), during test and control trials ('trial'). Thus, the statistical models included with three fixed effects including two-way interactions (table 4).

Table 4: Dogs were tested for informing their owner about the location of a missing coat that they had observed being hidden During test trials, owners asked their dog for help in finding the coat. Ninety records on 40 dogs were analysed with linear mixed models and the p-values and predicted means (± standard error) per fixed effect level are presented. Fixed effects included Trial, with control (C) and test (T). Owner choice consisted of the two conditions incorrect choice (False, F) and correct choice (True, T). Owner thumb up consisted of the two condition no thumb up (↓) and thumb up (↑). The statistical model with three fixed effects included two-way interactions. In the interaction effect the first character indicates whether the owner chose the correct box or not and the second character indicates whether the owner raised their thumb or not. Predicted means are expressed in rate per minute, for time in target area (% of observation time). Predicted means for interaction effects differ significantly when these do not share any superscript letter. The behaviours were analysed as rate per minute (rpm).

Signalling behaviour (rpm)	**Fixed effect** Trial (C, T)	Owner choice (F, T)	Owner thumb (↓,↑)	Owner choice.Owner thumb up
Look at target – p-value	0.01	0.026	<0.001	0.024
Control / False / ↓ / False ↓	0.8±0.3	0.8±0.3	0.5±0.2	0.4±0.2 [a]
Test / True / ↑ / False ↑	1.8±0.3	1.8±0.3	2.2±0.4	1.2±0.5 [ab]
/ True ↓				0.5±0.4 [a]
/ True ↑				3.2±0.6 [b]
Approach target	0.26	<0.001	0.04	0.003
	0.6±0.1	0.1±0.2	0.4±0.1	0.2±0.1 [a]
	0.7±0.2	1.2±0.2	0.9±0.2	0.0±0.0 [a]
				0.6±0.2 [a]
				1.9±0.3 [b]
Reference	0.02	0.11	0.45	0.067
	0.1±0.1	0.1±0.1	0.2±0.1	0.0±0.1
	0.5±0.1	0.4±0.1	0.3±0.1	0.1±0.2
				0.3±0.1
				0.6±0.2
Time in target area	0.84	0.16	0.42	0.9
	6.9±2.9	4.7±3.0	4.9±2.1	4.8±3.9
	6.3±3.1	8.6±3.1	8.4±3.8	9.2±4.1
				4.6±3.9
				8.1±4.5

The interaction between choosing the correct box and owner raising the thumb was significant for the signalling behaviours: Looking at the target (p=0.024) and approaching the target (p=0.003). For looking at the target, the combination of owner thumb up and owner guessing correctly (True, ↑, 3.2±0.6 rate per minute, table 4) differed significantly from the combinations no thumb up (both correct, 0.5±0.4, and incorrect guesses, 0.4±0.2). For approaching the target, the interaction of owner raising the thumb and owner guessing correctly (True, ↑, 1.9±0.3) was significant different from all the other combinations (≤0.6, table 4). When more signalling behaviours were performed by the dogs (look at and reference at the target) more owners guessed the correct box and were confident about having it right.

Furthermore, looking at target (p=0.011) and reference at target (p=0.021) were expressed significantly different between the trials. Looking at the target happened more in the test trial (1.79±0.32) compared to the control trial (0.83±0.3). Referencing at the target also happened more in the test trial (0.45±0.11) compared to the control trial (0.05±0.12). No differences (p>0.1) between trials, or between correct/incorrect choices, were found in percentage of time spend in the target area, or in the nontarget area. More signalling behaviour (looking at and referencing at the target) was performed when the owner communicated with the dog.

3.2 Theory of mind test

Dogs were tested for theory of mind in a Guesser/Knower test, in which dogs had to approach the baited bowl (out of four) as pointed out by a knowing informant, thus ignoring the pointing behaviour of the person without relevant knowledge. Four different protocols where used. Dogs were trained in different phases to reach the final read-out phase and each protocol consisted of six phases. Fifty dogs and their owners participated in the test, 16 dogs were excluded from statistical analyses. Two of them had no complete data set due to computer problems, two other dogs were too aroused to continue and twelve dogs were tested with preliminary protocols. Sixty-five percent of the participating dogs were female and the average age was six years. Four dogs participated in protocol 3, four in protocol 4, five in protocol 5 and twenty-one dogs participated in protocol 6. Characteristics of the dogs can be found in appendix 6.

3.2.1 Training phases

Training consisted out of the four first phases in each protocol. The goal of these four phases were to accustom the dog to the set-up of the experiment and learn to follow pointing gestures made with the pointing tool (i.e. the “grabber”). In phase 1, one bowl was used, in phase 2, two bowls were used, in phase 3, four bowls were used and in phase 4 false baiting was introduced. Thirty-four dogs participated in the training and ten of these dogs failed to continue to the final (read-out) experiment. Thirty-four dogs chose 500 times the correct bowl out of 595 trials during training (binomial one-sided p<0.001). Similar binomial tests on outcomes per phase yielded similar results (p<0.001). A chi-square test found differences between the observed and expected values of correct bowl choices in the different phases for all protocols (χ^2= 210.2, df=3 p<0.001) (table 5). Counts (residuals) for the incorrect bowl and correct one for the different phases were 3 (-3.9) and 87 (2.2) for phase 1, 5 (-3.5) and 84 (1.9) for phase 2, 22 (-3.5) and 170 (1.9) for phase 3 and 65 (1.6) and 159 (-0.9) for phase 4. During the first 3 phases the ratio of choosing the incorrect and correct bowl was about 1:17, which changed to 1:3 in phase 4. In phase 4, dogs made more errors than the earlier phases. Detailed information on the choices can be found in table 5.

Table 5: Dogs were tested for theory of mind in a Guesser/Knower test, during training dogs had to approach the baited bowl (out of 1,2 or 4 bowls) as pointed out by an informant. 595 records of 34 dogs have been analysed with a chi-square test that looked at differences between correct (C) and incorrect (I) choices per protocol (3,4,5,6) and phase (1,2,3,4). Per protocol the number of observed choices for the correct bowl and incorrect bowl per phase are shown with their totals. The "p" indicates the chance effect for making the correct choice. Per phase the number of expected choices for a correct bowl and incorrect bowl are shown with their totals.

Protocol	Phase				Total
Observed values	1 p=1	2 p=0.5	3 p=0.25	4 p=0.25	
3 C	10	7	32	18	67
I	0	3	4	0	7
4 C	8	8	29	22	67
I	0	0	7	2	9
5 C	6	8	30	23	67
I	2	0	2	1	5
6 C	63	61	79	96	299
I	1	2	9	62	74
Total C	87	84	170	159	500
I	3	5	22	65	95
Expected values C	69	68	146	171	454
I	21	21	46	53	141

Mixed linear models on 595 records (34 dogs) were used to test for relationships between choosing the correct bowl (incorrect, correct), expressed as a binomial response variate, and the different protocols (3,4,5,6) and phases (1,2,3,4). Thus, the logistic statistical models ran with two fixed effects including the two-way interaction. The chance of the correct bowl being chosen was significantly affected by the two-way interaction ($p<0.001$) (table 6). No effect of protocol was found in phase one and phase three. In phase two, the probability of choosing correctly was higher in protocol 6 (0.97±0.04) than in protocol 3 (0.73±0.11). In phase four of protocol six, the probability of choosing correctly was only (0.61±0.03) compared to ≥0.92 for the remainder of protocols (table 6). Also, the main effect phase ($p<0.001$) was significant but protocol was not ($p=0.11$). Across phases performances dropped from 0.94 (phase 1), 0.93 (phase 2), 0.88 (phase 3) to 0.87 (phase 4).

Table 6: Dogs were tested for theory of mind in a Guesser/Knower test, during training dogs had to approach the baited bowl (out of 1,2 or 4 bowls) as pointed out by an informant. In total 595 records on 34 dogs were analysed with a mixed linear model. The statistical model with two fixed effects included a two-way interaction. Main effect phase (1,2,3,4) was significant in different protocols (3,4,5,6) in whether there are differences in choosing the correct choice bowl (correct, incorrect), predicted means (± standard error) per fixed effect level are presented. Predicted means for interaction effects of phases within the given protocol differ significantly when these do not share any superscript letter.

Protocol	Phase			
	1	2	3	4
3	1.03 ± 0.11^{a}	0.73 ± 0.11^{a}	0.88 ± 0.06^{a}	0.99 ± 0.08^{b}
4	1.0 ± 0.12^{a}	1.0 ± 0.12^{ab}	0.81 ± 0.06^{a}	0.92 ± 0.07^{b}
5	0.75 ± 0.12^{a}	1.01 ± 0.12^{ab}	0.94 ± 0.06^{a}	0.96 ± 0.07^{b}
6	0.99 ± 0.04^{a}	0.97 ± 0.04^{b}	0.9 ± 0.03^{a}	0.61 ± 0.03^{a}

The REML procedure was used again on 595 trials (34 dogs) to test the probability of dogs choosing correctly for effects of trials (1-31) and protocols (3,4,5,6), including the two-way interactions ($p<0.001$). Figure 2 shows the generated predicted means for fraction of chance of choosing the correct bowl per protocol for a certain number of trials. The straight lines in the graph reflect that trial was fitted linearly in the model as a co-variate. Protocol 6 stood out in that less correct choices were made when more trials were performed (dogs reached a higher phase). In protocols 5 and 4, more trials resulted in dogs performing more correct choices. In protocols 3 and 4 the probability of making correct choices remained unchanged. Protocol 3 and 4 seem to work best for the dogs where protocol 6 was too difficult.

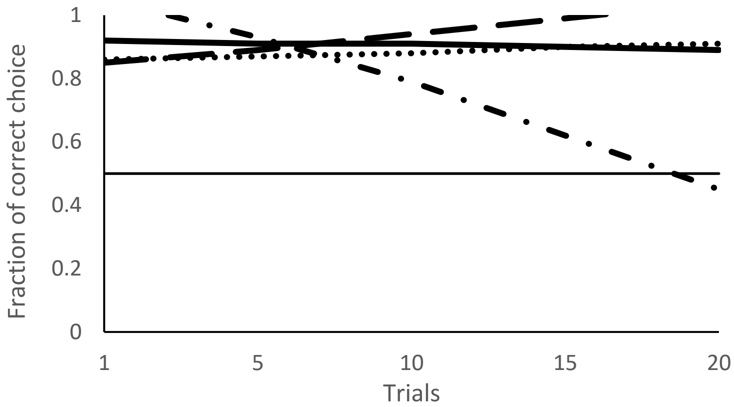

Figure 2: Dogs were tested for theory of mind in a Guesser/Knower test, during training dogs had to approach the baited bowl (out of 1,2 or 4 bowls) as pointed out by an informant. In total 595 records on 34 dogs were analysed for predicted probabilities for making correct choices in relation to number of trials in training, where low trial numbers represent the early training phases. The thin black uninterrupted line presents the random chance of choosing the correct bowl (50%). Protocol 3 is the solid line, protocol 4 is the dotted line, protocol 5 is the dashed line and protocol 6 is the dashed dotted line.

3.2.2 Experimental phase

After the training phases, two experimental (read-out) phases were conducted. In phase five two informants took part in the test. The Knower hid the treat whilst the Guesser stood out of view and behind a screen when the Knower baited the bowl. In phase six a third informant, the baiter, was introduced. The baiter would handle the baiting while the Knower looked at the baiting and the Guesser stood behind a screen. Twenty-four dogs participated in experimental phase five and eight of these in phase six. Only four of the eight dogs completed phase six with 50% or more correct choices, one dog from protocol five (50 % correct choices) and three dogs from protocol six (all more than 50% correct choices). Across phases 5 and 6, 24 dogs chose 98 times the correct bowl out of 190 trials (binomial one-sided p<0.001 and for details see table 7).

Table 7: Dogs were tested for theory of mind in a Guesser/Knower test, in which dogs had to approach the baited bowl (out of 4) as pointed out by a knowing informant, thus ignoring the pointing behaviour of the person without relevant knowledge. Hundred-ninety records of 24 dogs were analysed with a chi-square test that looked at differences between correct (C) and incorrect (I) choices per protocol (3,4,5,6) and phase (5,6). Per combination of protocol and phase the number of observed choices for the correct bowl and incorrect bowl are shown. In each phase the chance of success was 25% (out of four bowls).

Protocol Observed values		Phase 5	6	Total
3	C	9		9
	I	9		9
4	C	9		9
	I	13		13
5	C	5	3	8
	I	11	1	12
6	C	52	20	72
	I	36	22	58
Total	C	75	23	98
	I	69	23	92
Expected values	C	110	35	145
	I	34	11	45

The chance of choosing the correct bowl were not larger then chance (25%) for each protocol (3,4,5,6) (binomial one-sided p>0.1), but for protocol 6 in phase 5 (52 correct choices out of 88, one-sided binomial p<0.001). The ratios of correct choices to incorrect ones did not differ between protocols (χ^2= 5.7, df=3, p=0.13) or phases (χ^2= 0.06, df=1, p=0.81). No difference in relative number of incorrect or correct choices were found for the different protocols or phases. The REML procedure was used on 190 trials (24 dogs) to test if the probability of dogs choosing correctly depended on effects of phases (5,6) and protocols (3,4,5,6). The two-way interaction was not significant (p=0.55). Nor was the probability on choosing a correct bowl different between phases (p=0.2) or protocols (p=0.66). Predicted mean probabilities for (±standard error) choosing correctly were 0.45±0.07 in phase 5 and 0.37± 0.08 in phase 6. A similar REML analysis on 190 trials (24 dogs) did not detect effects of trials (1-20, p=0.19), protocols (3,4,5,6, p=0.57) or two-way interactions (p=0.35).

The final evaluations zoomed in on the performances of dogs that made many correct choices. These were dogs that made more than 50% correct choices in phases 5 and 6, leading to nine subjects (table 8). The average age of this group was six years, four were male and five were female. These 9 dogs chose the correct bowl 66 times out of 93 trials (binomial one-sided p<0.001). In phase 5, nine dogs chose 47 times the correct bowl out of 59 trials (binomial one-sided p<0.001). In phase 6, six dogs chose 19 times the correct bowl out of 34 trials (binomial one-sided p<0.001). Apparently, this select group of dogs solved the problem of identifying the baited bowl and chose this above chance levels in both phases.

Table 8: Dogs were tested for theory of mind in a Guesser/Knower test, in which dogs had to approach the baited bowl (out of 4) as pointed out by a knowing informant, thus ignoring the pointing behaviour of the person without relevant knowledge. Dogs were selected that had more than 50% correct choices in phase 5 and 6 (n=9) with ninety-three records. The name of the dog, sex, age, breed, in which protocol they participated, the number and percentage of correct (C) and incorrect (I) choices is shown.

Do	Sex	Age	Breed	Protocol	Phase 5		Phase 6	
					C:I	%C:I	C:I	%C:I
Max	M	6	Labrador	6	5:1	84:16	5:1	84:16
Mug	F	11	Australian shepherd	6	5:1	84:16	3:3	50:50
Roefus	M	4.5	Long haired dachshund	6	6:0	100:0	2:0	33:67
Kenna	F	2.5	Golden retriever	6	8:1	89:11	5:1	84:16
Caio	M	2.5	Mix husky/malamute	6	8:3	72:28	2:4	33:67
Tara	F	5	Manchester Terrier	3	4:2	67:33		
Pretty	F	9	Border collie	4	3:2	60 :40		
Dorit	F	12	Labrador	4	5:1	84:16		
Jedi	M	2	Novan Scotia Duck Tolling	5	3:1	75:25	2:2	50:50

Probabilities for choosing the correct bowl were analysed with REML on effects of phase and protocol as described, but now on 93 trials of the best performing 9 dogs. The two-way interaction (p=0.23) nor the main effect of protocol (p=0.72) was significant, but phase was (p=0.02). The predicted mean for choosing the correct bowl in phase five was 0.74±0.09 and for phase six 0.57±0.08. The selected dogs made more correct choices in phase five compared to phase six. The REML with the fixed effects trial (1-12) (co-variate) and phase (5 and 6) did not reveal a significant interaction effect (p=0.18) or main effect trial (p=0.4) or protocol (p=0.94).

4. Discussion

Dogs demonstrate remarkable skills in reading human signals and aiding their owners, but so far it remains unclear to what degree dogs grasp the feelings, intentions or beliefs of humans and act altruistically. Training and associative learning in general could explain many of the prosocial behaviours that dogs show towards humans, representing a selfish motivation to obtain rewards like food or praise. To gain further knowledge on a dog's capacity to understand the mental states of humans and have a Theory of Mind (ToM), I tested dogs on ToM with a Guesser-Knower task and on altruism with an informing test. Dogs tended to inform their owners about the location of a missing coat when owners asked for help, but owners may not always have recognized the signalling behaviours. Only a small portion of the present study dogs seemed to understand when a human informant was knowledgeable about the location of a food reward on te basis of this human having (not) observed the act of hiding. Clearly, dogs help-out humans, like assistance dogs supporting their owners in performing daily activities and therapy dogs interacting with clients. The question is if this, partly, mirrors true altruism in the sense of helping with intent or merely results from operant conditioning and from a selfish motivation to earn rewards. For example, dogs do tend to show pro-social behaviours towards conspecific's in a bar-pulling test (Quervel-Chaumette, et al 2015), but those pro-social behaviours did not extend towards humans (Quervel-Chaumette et al., 2016). Though this may have been due to the fact that the humans were instructed not to communicate with the dog in anyway. This absence of contact may have been aversive to the dogs, leading to the stop of social behaviour towards the human (Quervel-Chaumette et al., 2016). The present findings support that it is not unusual for dogs to act altruistically towards their owner by informing them about a missing object, but only few demonstrated a clear understanding of mental states in humans in the Guesser-Knower task. As a species, dogs are likely to have a ToM, but there could be considerable variation between individuals.

Communication and making the goal clear are important factors to consider when it comes to the design of an experimental set up for altruism and ToM. When humans do not communicate clearly with the dog, the latter may not understand what it has to do and fail to show pro-social altruistic behaviours (Quervel-Chaumetter et al 2015; 2015; Bräuer et al., 2013; Johnston, Huang & Santon, 2018). Dogs were found more likely to use showing behaviour when the human communicated with the dog, compared to when they did not, and when the dog knew that the object that was used was relevant for the human (Piotti &Kaminski, 2016). Dogs are sensitive to human cues and attentional states (Udell & Wynne, 2008) and when humans use vocal or body cues to stimulate the dog this raises the activity level of the dog (Range et al., 2009; Bräuer et al., 2013; Udell, 2015). For instance, dogs would open the door for a human informant to obtain a desired object when the human pointed at the button to open the door or talked to the dog in a natural way and explicitly communicated the goal to the dog. The drawback is that by means of pointing or communicating in this way it may be that dogs were instrumentally directed to the button instead of figuring out what the human wants (Bräuer et al., 2013).

In the present study, dogs also had to indicate a relevant object of the owner (his/her coat) in situations where the owner did (test trial) or did not (control) communicate. More tail-wagging occurred in the test trials, i.e. with communication, which may have been due to arousal or stress caused by the asking for help of the owner. Arousal could represent an anticipatory response and in general represents the intensity of emotional experience (Lang & Davis, 2006). Stress may have played a role as the question of 'Where is my coat?' may have been unclear to dogs. This interpretation is unlikely, though, due to the fact there were no differences in other stress behaviours between the control and test trials. Clear communication facilitates altruistic behaviours in dogs and more showing behaviour was performed when the informant communicated with a high-pitched voice with the dog (Piotti & Kaminski, 2016). A distractor object (stapler) and a relevant object (notepad) were hidden and before the start of the test the informant used the notepad (wrote on in with a pen) so the dog would know its importance to the owner. When one of the objects was hidden, the informant used either vocalisation or not to ask help from the dog to find the object. Vocal trials led to more and longer gazes towards the location were the target was hidden. Also, when the important object was hidden the dog performed more showing behaviour. This indicates that it is important for the dog to know that the object matters to the owner to understand the task and help out (Piotti & Kaminski, 2016). Ostensive cues (high pitched voice, eye contact) may increase the attention of the dog which can help the dog to understand the needs of the human (Hill, 1965). The present study dogs showed more looking at the target and referencing at the target during test trials than during control trials. Seemingly, the owner asking the dog where the coat was while keeping eye contact was sufficient for dogs to understand the owner's need and request.

Important to keep in mind is that it is uncertain why dogs are motivated to help. Like, do they really want to help their owner or do they just want to leave by retrieving the owner's coat. Since there was no difference in stress behaviours between the test and control trial the latter seems unlikely though. Whine or yelp (a mean of 4.0 times per minute) and panting (5.5) did occur and such behaviours have been linked to some minor degree of stress earlier (Beerda, 1999). Emotional states have components of valence (positive, negative) and arousal, and behaviours like panting and whining may represent arousal that could be both positive and negative. Dogs seem to be motivated to help even when they continuously received no praise (Bräuer et al., 2013; Kaminski et al., 2011), and it is unlikely that dogs experienced negative affect during the altruism test. The behavioural differences between test trials and controls support that dogs acted altruistically towards their owner. It could be that owners did not always understand the signals the dog showed. An early study on human-dog relationship showed that dogs behave socially depended in a problem-solving task when they have a companionship relation with their owner. In this study the relationship between the owner and dog was determined by means of a questionnaire (companionship or working relationship). Dogs were tested in a simple problem-solving task where they had to make food available by pulling a dish from under a fence. Dogs with a companion relationship showed a decrease in performance, suggesting that the strong dependency of a dog may prevent dogs to complete a task successfully (Topál, Miklósi & Csányi, 1997). Indicating that dogs in the current study may have not been able to solve the problem because of their relationship with their owner. Furthermore, dogs do not perform signalling behaviour in everyday life, meaning owners may have never seen the signals the dog showed and may have not caught up on the signals even when the dogs did perform the altruistic signals.

Not all dogs in the current study seemed not to have a Theory of mind (ToM), which is in contrast with the results of Maginnity & Grace, 2014 and Catala et al., 2017, were dogs chose the Knower significantly more than the Guesser. This when the guesser left the room, closed his eyes, look at the ceiling or used geometrical gaze following. However, when looking at individual dogs they do seem to have some kind of ToM. These contrasting outcomes lead to Horowitz's (2011) notion of a rudimentary theory of mind. Different abilities could explain the dogs' behaviour in the Knower Guesser task. Also, a failure (choosing the Guesser) does not have to be seen as a negative outcome (read 'that dogs do not have a theory of mind'), because of imperfections in the experimental design. In the current study, we removed the screen in protocol 6. This resulted in less correct choices and may have been crucial in the design of a ToM test. In protocol 6, the object of interest was visible for the dog, so it could use the (preferred) eye-object line instead of following the grabber (Udell, Dorey & Wynne, 2011). This associative learning may have led to the dogs to choose the last touched bowl because they could see the baiting. In the other protocols (where dogs made less incorrect choices) the dogs could not see the baiting because of the screen and the rule would have been impossible to follow. Differences in outcomes of ToM test may reflect that there are degrees of Theory of mind (Gómez, 1996; Whiten, 1997). The variation of success across trials also in this study may demonstrate an intermediate ability, it is not a case of all or nothing. There can be many levels of skill at using or understanding attentional states. From noticing attentional states in others to linking attentional states of other to information (here about the correct bowl) (Horowitz,2011).

However, as did we, Cooper et al., 2003 found no significant choice for the Knower. Fifteen dogs were tested with six trials, in each trial the Guesser would leave the room. Fourteen out of fifteen dogs chose the Knower in the first trial. In consecutive trials dogs did, however, not choose the Knower significantly above chance (>50%) and the number of correct choices decreased over trials. Cooper et al., 2003 proposed that the dog may fail the test because the dog may try to apply information from previous learned trials. In the first trial all dogs succeed whereas in the consecutive trials dogs got confused. It could be that this initial information leads to conditioned responses. Dog may use pro-active interference, this means that the dog's memory over who was the previous Knower and what was the previous success bowl interferes with the roles and bowl in the current trial. This leads to the dog choosing the last correct bowl or choosing the informant who was the Knower in the previous trial (Roberts & Macpherson, 2011). The preferences can be explained as associative learning instead of problem solving. In the current study dogs could also have been confused by the difference of previous and current success bowl or informant which could have led them to choose an incorrect bowl. Furthermore, dogs are assumed to invent their own rules in these perspective-taking tasks. Such as 'follow the person that hides the treat and is present during baiting (Cooper et al., 2003). However, in phase six were the third informant the baiter is introduced this rule does not apply anymore. The baiter hides the treat and does not point at a bowl, the dog could not simply follow the 'rule' anymore. It may then be confused on who to follow and chose the incorrect bowl.

Differences in experimental design between the current study and Maginnity & Grace,2014; Catala et al., 2017 may have led to difference in outcomes. In Maginnity & Grace, 2014 they performed a number of experiments to see which factors may have influenced the choices of the dogs. In the first experiment of Maginnity & Grace, 2014 they only used two informants and the experiment consisted of two conditions. One where the Guesser was absent during the baiting by the Knower and another where the Guesser was present during the baiting by the Knower. In the condition were the Guesser was absent dogs chose the Knower significantly more times, but in the conditions were the Guesser was present the choice for the Knower was also higher. Thus, dogs based their choice on the food handling cues and not directly on attentional states. In the additional experiments they used three informants in all phases so the dog could not base its choice on the food handling cues. In the current study dogs may have used both food handling and attentional states of informants to guess the location of the food. Three informants should have been used during all trials to factor out the food handling cue. This food handling cue and the visibility of the bowls during the trials of the experiment may have confused and influenced the dog's choice, especially in phase six were the set-up changed from two to three informants. Furthermore, in Maginnity & Grace, 2014 and Catala et al., 2017 they used more trials in each phase before the experimental phase started. In the false baiting phase dogs would have to had at least six consecutive trials, whereas in the current study dogs only had to have three out of four correct choices. It could be that the number of trials were not enough for the dog to really understand the concept of the test. Another difference was that the informants in the current study stood behind the bowls while the informants in other Guesser/Knower tasks sat on their knees. The dog may not have looked at the faces of the informant and thus missed (in phase six) that the Knower informant looked at the baiting. Also, in the analysis of the data in Maginnity & Grace, 2014 and Catala et al., 2017 they omitted the choices for another bowl leaving only the choices for a Guesser bowl or a Knower bowl. In the current study these choices were not omitted.

In a design of a ToM test it is important that differences in life-history are taken in account, as they can influence the behaviour of the dog (Horowitz, 2011). In a perspective taking task of Udell, Dorey & Wynne, 2011 a dog had to beg for food. Different designs were used on how the informant presented itself to the dog, one informant would be the attentive human the 'seer', while in the other informant would be the inattentive human the 'blind' informant of which visual attention was blocked. In experiment one the seer informant would stand with their face to the dog while the blind informant turned their back on the dog. Dogs chose significant more the seer informant over the blind informant to beg for food. However, in other experiment, were the blind informant covered their face with a book, had a bucket over their head or had a camera in front of their face, dogs did not learn to discriminate between the attentive and inattentive informant. Horowitz, 2011 proposed that the performance of the dog changed while the obstructing item changed. The knowledge of the properties of the object and the object itself is depended on the amount of exposure to the object. For instance, for dogs that have an owner that reads books it is clear that the book means that the owner is inattentive, whereas they may not attribute the same properties to a bucket on the head of an informant. In this case, outcomes reflected the understanding of a dog about different objects instead of testing whether the dogs understand the inattentive state of an informant. Different life-histories of dogs may influence how dogs perceive the situation and how they perform in tests on mental states (Horowitz, 2009). In a study with dogs in rough-and-tumble play set-up they looked at the 'guilty-look' of dogs. Dog's that attended obedience trainings showed more guilty-looking behaviours compared to dogs that did not attend obedience training (Horowitz, 2009). It may be that this submission behaviour is learned during the obedience training just as the exposure to a book has been learned at home where the owner reads a book. Also, in Maginnity & Grace, 2014 dogs attended professionally obedience training had higher preferences for the Guesser compared to dogs that were not professionally trained. In the current study it could be that difference in life-history also influenced the dog's choice in the Guesser-Knower task. Dog's may have made choices based on knowledge of trainings or learned behaviours at home, which could have led to the dogs not performing correctly in the task. These choices may have led to the choice for the Guesser but does not directly have to indicate that they do not have a theory of mind.

5. Conclusions and recommendations

Dogs showed signalling/helping behaviour in the altruism box test. More signalling behaviours were shown when the owner communicated with the dog and asked for his/her coat. However, the owner did not indicate the coat more often in the test trial compared to the control trials, which may indicate that owners may not read the dogs' signalling behaviours. Motivations for the dog to show these altruistic behaviours remain speculative. Motivational test can be done between tests to check whether dogs are still motivated to help or handle out of self-interest. For the dog to understand a task on mental states and abilities the task must be made as clear as possible, for example that the object is relevant for the owner. The object could be handled before the test is performed or the objects could be an object that is used at the owner's home. More research needs to be done on whether owners understand signalling behaviours in their dogs. The difficulties in studies on mental states in dogs lie more in the task (making it as clear as possible, make object relevant for owner, owner understanding signalling behaviours) then in the willingness of the dog to help.

Some dogs chose correctly in a Guesser/Knower task based on the attentional state of a human informant were other dogs seem to fail. In the first three phases of training dogs chose the correct bowls. However, in phase four (false baiting) less correct choices were made due to the removal of the screen, after which dogs tended to follow food handling cues which led to more incorrect choices. During the experimental phases more incorrect bowls and less correct bowls were chosen, indicating that dogs may invent their own rules or get confused by proactive interference. Individual dogs show however more correct choices in the experimental phase which may indicate that there are degrees of theory of mind. In further studies on Theory of mind it needs to be taken in consideration to use three informants and a screen to prevent choices being made based on food handling cues. Furthermore, more research needs to be done on the different levels of Theory of Mind and on how life histories may influence the choice of a dog in a Guesser/Knower task.

References

Baker, A. (2014). Canine brains: test your dog's intelligence.

Bräuer, J., Schönefeld, K., Call, J.(2013). When do dogs help humans? *Applied Animal Behaviour Science,* 148-1(2), 138-149.

Beerda, B. (1999). Stress and well-being in dogs.

Catala, A., Mang, B., Wallis, L., Huber, L. (2017). Dogs demonstrate perspective taking based on geometrical gaze following in a Guesser–Knower task. *Animal Cognition* 20-4, 581-589.

Cooper, J.J., Ashton, C., Bishop, S., West, R., Mills, D.S., Young, R.J. (2003). Clever hounds: social cognition in the domestic dog (Canis familiaris). *Applied Animal Behaviour Science* 81,229-244.

Dorey, N.R., Udell, M.A.R., & Wynne, C.D.L. (2010). When do domestic dogs, Canis familiaris, start to understand human pointing? The role of ontogeny in the development of interspecies communication. *Animal Behaviour* 79, 37–41.

Gómez, J. C. (1996). Nonhuman primate theories of (nonhuman primate) minds: Some issues concerning the origins of mind- reading. In P. Carruthers & P. K. Smith (Eds.), Theories of mind. *Cambridge: Cambridge University Press* 330-343.

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in *Dogs Science*, 298(5598), 1634-1636.

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

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *The Behavioral and Brain Sciences*, 21, 101–114. doi:10.1017/ S0140525X98000703.

Herwijnen van, I.R., van der Borg, J. A., Naguib, M., & Beerda, B. (2018). The existence of parenting styles in the owner-dog relationship. *PloS one*, 13(2).

Hill, A.B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, 58(5),295.

Horowitz, A. (2009). Attention to attention in domestic dog (Canis familiaris) dyadic play. *Animal Cognition*, 12, 107–118.

Horowitz, A. (2011). Theory of mind in dogs? Examining method and concept. *Learning and Behavior* 39(4), 314-317.

Jager, Smit and van Woensel. (2017). The influence of dog-owner parenting styles on the attachment, owner-directed info-sharing altruism and dogs' quality of life.

Johnston, A.M., Huang, Y., & Santon, L.R., (2018). Dogs do not demonstrate a human-like bias to defer to communicative cues. *Learning and Behavior, 449-461.*

Kaminski, J. (2009). Dogs (Canis familiaris) are adapted to receive human communication. *Neurobiology of 'Umwelt' Springer,*103-107.

Kaminski, J., Neumann, M., Bräuer, J., Call, J., Tomasello, M. (2011). Dogs (Canis familiaris) communicate with humans to request but not to inform. *Animal Behaviour*, 82 (4), 651–658.

Lang, P.J., Davis, M. (2006). Emotion, motivation, and the brain: Reflex foundations in animal and human research. *Progress in Brain Research*, 158-chapter 1.

Maginnity, M.E., Grace, R.C. (2014) Visual perspective taking by dogs (Canis familiaris) in a Guesser–Knower task: evidence for a canine theory of mind? *Animal Cognition,* 17(6),1375–1392.

Miklósi, Á., Polgárdi, R., Topál, J., Csányi, V. (2000) Intentional behaviour in dog-human communication: an experimental analysis of 'showing' behaviour in the dog. *Animal Cognition*, 1; 3(3):159–66.

Piotti, P., Kaminski, J. (2016). Do dogs provide information helpfully? *PLoS ONE*, 11-8,1-19.

Povinelli, D.J., Nelson, K.E., Boysen, S.T. (1990) Inferences about guessing and knowing by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology,* 104,203–210.

Paiget, J., Inhelder, B. (1969). The Psychology of the Child. *New York: Basic Books*.

Quervel-Chaumette, M., Dale, R., Marshall-Pescini, S., Range, F. (2015). Familiarity affects other-regarding preferences in pet dogs. *Scientific Reports*,5,1-7.

Quervel-Chaumette, M., Mainix, G., Range, F., Marshall-Pescini, S. (2016). Dogs do not show pro-social preferences towards humans. *Frontiers in Psychology*, 7,1-9.

Range, F., Heucke, S. L., Gruber, C., Konz, A., Huber, L., and Virányi, Z. (2009). The effect of ostensive cues on dogs' performance in a manipulative social learning task. *Applied Animal Behaviour Science*, 120, 170–178.

Roberts, W.A., Macpherson, K. (2011). Theory of mind in dogs: Is the perspective-taking task a good test? *Learning and Behavior,*39(4),303-305.

Topál, J., Miklósi, A., Csányi, V. (1997). Dog-Human Relationship Affects Problem Solving Behavior in the Dog. *Anthrozoös*, 10(4), 214-224.

Udell, M. A. R., & Wynne, C. D. L. (2008). A review of domestic dogs' (Canis familiaris) human-like behaviors: Or why behavior analysts should stop worrying and love their dogs. *Journal of the Experimental Analysis of Behavior*, 89, 247–261.

Udell, M. A. R., Dorey, N.R., Wynne, C. D. L. (2011). Can your dog read your mind? Understanding the causes of canine perspective taking. *Learning and Behavior,* 39(4),289-302.

Udell, M. A. R. (2015). When dogs look back: inhibition of independent problem-solving behaviour in domestic dogs (Canis lupus familiaris) compared with wolves (Canis lupus). *Biology Letters,* 11, 20150489. doi: 10.1098/rsbl.2015. 0489.

Warneken, F., Tomasello, M., (2006). Altruistic helping in human infants and young chimpanzees. *Science 311*, 1301–1303.

Warneken, F., Tomasello, M., (2009a). Varieties of altruism in children and chimpanzees. *Trends Cognitive Science*, 13, 397–402.

Warneken, F., Tomasello, M., (2009b). The roots of human altruism. *British Journal of Psychology*, 100, 455–471.

Whiten, A. (1997). The Machiavellian mindreader. In A. Whiten & R. W. Byrne (Eds.), Machiavellian intelligence II: Extensions and evaluations. Cambridge*: Cambridge University Press* 144-173.

Wynne, C.D. (2016). What is special about dog cognition? *Current Directions in Psychology Science*, 25(5),345–350. doi:10.1177/0963721416657540.

Appendix 1: Protocol of the Altruism test by Jager et al. (2017).

Altruism test
The altruism test, tests whether the dog displays info-sharing behaviour towards its owner.
The test is conducted in an observation room in Carus, the dog-facility of Wageningen University. There are two numbered chairs present (#1 and #2), and there are two not numbered chairs present. Three numbered cardboard moving boxes (small white boxes of 50L, bought at Action), are placed evenly across the tape on the floor. The numbered chairs are placed around the middle box, and the two unnumbered chairs are placed on the sides.
The dog is on the long leash of 4m25 meter tethered to the wall opposite the chairs and boxes and is able to reach the boxes, but not able to touch them. In the corner of the room is a screen behind which the owner stands while the stranger is in the room. The dog can go to the owner, and thus is aware that the owner remains present during the test. The time will be monitored using a stopwatch. The position of the screen (left: #1 and right: #2) the chair used by the owner (left: #1 and right: #2) and the box used for the owner's coat (left: #1, middle: #2 and right: #3) will be randomly picked before the start of the test. This was done using the Research randomizer (https://www.randomizer.org/) separately for the box, the chair and the location of the screen. In each box, **a neutral coat is placed beforehand**. The owner is instructed beforehand how the "search action" works. An image of this will be hung inside the testing room and placed on the chair when the test is active.

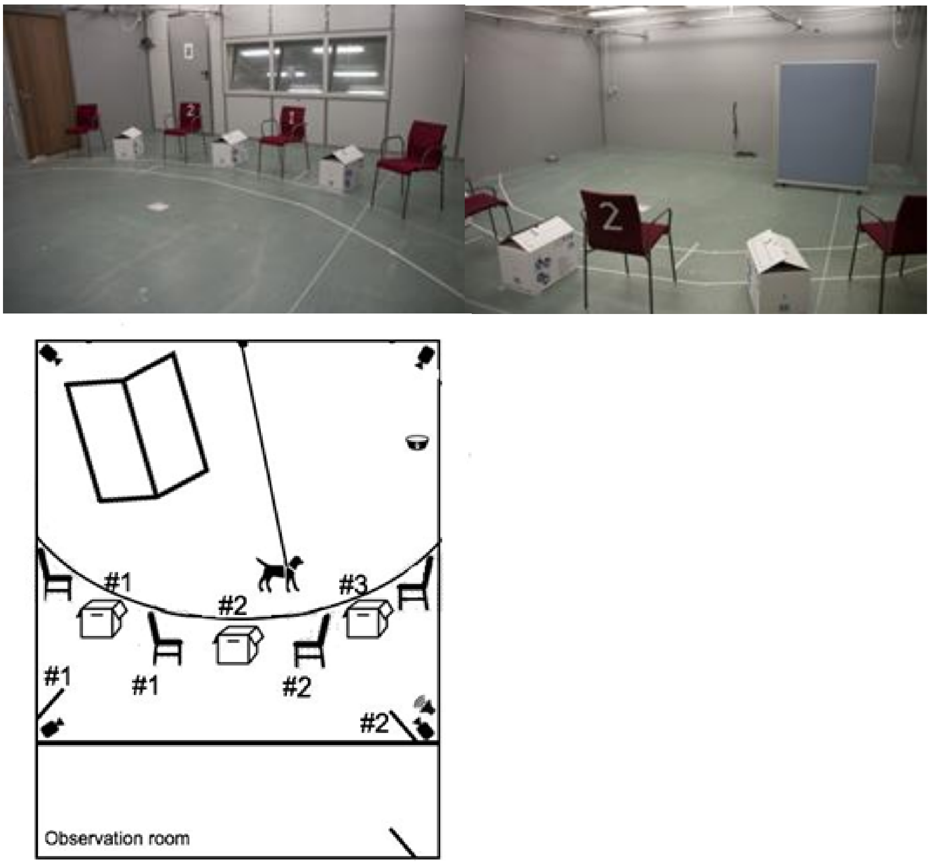

Figure 3. Experimental setup Altruism test (Jager, Smit and van Woensel 2017); four chairs are present, of which two are numbered. Also, three numbered cardboard boxes, a water bowl and a screen are present. The dog is fixed to the wall with a long leash and can reach its owner (also when the owner is behind the screen) but can't reach the stranger. The screen is placed against the left or right wall, so the dog cannot move around the screen.
Materials used:

- The owner's coat
- Two chairs
- Three identical boxes
- Long leash attached to the wall
- Observer XT 10.5
- Ethogram
- The song: Evergreen from the album Viaje Clásico Tranquilo or mixed classical music from youtube.

When the test is being done the following episodes take place:

Episode 0: Exploration

00:00 - 02:00 The owner and dog enter through door #1. After 2 min of off leash exploration and the owner visiting the space behind the screen, the owner attaches the dog on the leash and will stand behind the chair. When the dog is attentive (if not call its name), the coat is hung on the chair and the owner immediately walks behind the screen and puts on headphone for 30 sec. At the end of these 2 minutes, the test leader instructs the owner via the microphone to attach the dog to the long leash, and to sit down on a chair adjusted randomly after hanging their coat on the back of that chair and lay the leash behind it. (“Nu mag u de hond aan de lange groene lijn aan de muur bevestigen, vervolgens uw eigen riem van de halsband klippen, uw jas ophangen over de rugleuning van stoel nummer X, uw riem achter de stoel neerleggen en gaan zitten op stoel nummer X”)

Episode 1: Owner on chair

02:00 - 03:00 The owner is a nonparticipant sitting on the assigned chair while the dog explores. The owner only interacts with the dog (i.e. by talking to it) if it specially asks for attention. The dog is free to explore the room. After 1 minute the test leader instructs the owner to attach the dog on the long leash and the owner to sit behind the screen. (“Nu mag u achter het scherm in de hoek gaan zitten, zometeen komt (*naam stranger*) binnen en dan mag u gewoon achter het scherm blijven zitten”).

Episode 2: Owner behind screen, stranger enters

03:00 - 05:00 The owner is behind the screen and is non-participant. 03:30 The music will start. At 04:00 the stranger enters the room through door #2, waits until she/he has made eye contact with the dog and calls the dog by its name, and then takes the owner’s coat of the chair and puts it in a randomly assigned box. After this the stranger leaves through door #2.

Episode 3: Owner searches coat 1,

05:00 - 05:30 The owner calls the dog to them. Then the owner returns to stand behind chair, (1) looks at the dog and (2) asks “Waar is mijn jas?” and (3) raises arms, owner looks around searching for 5 seconds.

1. Oogcontact met je hond
2. ‘Waar is mijn jas?’
3. Vragende houding (plaatje)
4. **5** sec rondkijken
5. Oogcontact met je hond
6. ‘Waar is mijn jas?’
7. Vragende houding (plaatje)
8. **5** sec rondkijken
9. Klaar.

Figure 4. Instruction figure for the owner. (Jager, Smit and Van Woensel 2017); During the test the leaflet is placed on the chair of the owner as a reminder of the sequences.

<u>Episode 4:</u> Owner searches coat 2,
05:30 - 06:00 Repeat sequence 1,2,3 and again look around searching for 5 seconds.
<u>Episode 5:</u> What does the owner think
06:00 - 06:30 The test leader will ask the owner where they think the coat is based on the signals the dog is giving them. ("Denkt u te weten in welke doos uw jas zit? Wilt u dan uw duim opsteken? In welke doos denkt u dat uw jas zit als u kijkt naar de signalen van uw hond? Wilt u deze doos aanwijzen? ") Then the test leader tells the owner where their coat is. ("Dank u wel, dit is het einde van de test. U mag zonder iets te zeggen uw jas uit doos nummer X halen. U mag (*naam hond*) weer aan uw eigen lijn aanlijnen en de ruimte verlaten via deur nummer 1.")
For the control of the test episode 1 and 2 will be repeated. After that the owner calls the dog to them and the owner stands behind the chair and places their hands on the back of the chair: looks at the dog, without a questioning look, looks around for 5 seconds in a neutral way, looks at the dog again and ends with looking around, modelling the previous procedure.

Time (min)	**Who**	**Line**
00:00	Test Leader	(while standing outside door #1 with the owner) U mag zometeen de ruimte betreden met uw hond en uw hond aflijnen. U mag samen met de hond de ruimte verkennen, zorg hierbij dat u ook achter het scherm loopt. Als U de vraag "Waar is mijn jas?" moet stellen, voer dan de handelingen van dit plaatje uit (show picture). (test leader closes both doors)
02:00	Test Leader	Nu mag u (*naam hond*) aan de lange groene lijn aan de muur bevestigen(...) uw jas ophangen over de rugleuning van stoel nummer X (...) uw riem achter de stoel neerleggen (...) en gaan zitten op stoel nummer X
03:00	Test Leader	Nu mag u achter het scherm in de hoek gaan zitten. Zometeen komt (*name stranger)* binnen, u mag achter het scherm blijven zitten.
03:30	Test Leader	Nu mag u uw handen over uw oren plaatsen. Zometeen wordt de muziek gestart, u mag dan uw handen over uw oren houden.
03:45	Test Leader	music starts: Evergreen - Viaje Clásico Tranquilo or classical music from youtube
04:00	Stranger	(Enters room through door #2) (Get attention (call dogs name) & make eye contact with the dog) (Get coat from chair X and put it in box X) (Leave through door #2)
4:45	Test Leader	Music stops

05:00	Test Leader	Test: U mag uw hond bij u roepen, u mag achter stoel nummer X gaan staan (...), oogcontact maken met uw hond en 1 keer duidelijk aan uw hond vragen *Waar is mijn jas? (see picture on chair)* Control: U mag uw hond bij u roepen, u mag achter stoel nummer X gaan staan (...), en zo natuurlijk mogelijk rondkijken.
05:30	Test Leader	Test: U mag nu weer even een neutrale houding aannemen. *(Na een paar seconden)* U mag nog 1 keer duidelijk vragen: *Waar is mijn jas?* Control: n.a.
6:00	Test Leader	Denkt u te weten in welke doos uw jas zit? Wilt u dan uw duim opsteken? Als u het niet weet zou u dan uw schouders willen ophalen? JA -> In welke doos denkt u dat uw jas zit als u kijkt naar de signalen van uw hond? Wilt u naar deze doos wijzen? NEE-> Wilt u toch een doos aanwijzen in welke u gokt dat uw jas zit?
06:30	Test Leader	Dankuwel, dit is het einde van de test. U mag zonder iets te zeggen (*naam hond*) weer aan uw eigen lijn aanlijnen, uw jas uit box X halen en de ruimte verlaten via deur nummer 1.

Appendix 2: **Ethogram used in the Altruism test**

Ethogram for the Altruism test (Obtained from BSc and MSc research, 2017). Dogs (n=40) were tested for informing their owner about the location of the owner's coat, which had been hidden in view of the dog and in absence of the owner. Only during test trials did the owners ask for the help and looked for the missing coat. Stress and signalling behaviours were scored.

Behaviour		Description
Position in room*	Zone in which the dog has its front legs	Equal sized areas in the front of the chairs and boxes, up to 1m distance from the object (%)
Stress signals	Bark	Head and lips forward, mouth opening, and shutting repeatedly (rpm)
	Yelp	Loud (relative to whine) high pitched short vocalisation (rpm)
	Whine	Soft, high pitched, whistling vocalisation in short repeated bursts (rpm)
	Tongue flick	Showing brief licking movements with tongue directed towards nose and head oriented towards recipient, without physical contact (rpm)
	Yawn	Mouth open wide for a period of a few seconds whilst exhaling (rpm)
	Paw lift	Front limb is raised and lowered, often in quick succession (rpm)
	Shake	Rotation of the body starting at the head and moving caudally (rpm)
	Look away from owner	Turning only the head away from the recipient, while staying on the same spot (rpm)
	Tail wag	Non-accelerated, regular sideward movement of the tail, about in one plane (rpm)
	Pant	Mouth wide open with tongue protruding, often moving in and out of the mouth (rpm)
Signalling behaviour	Approach	In normal pace walking (not accelerated) towards the recipient (zone 1-7 and owner) up to a distance of 1m or less (rpm)
	Look at	Look to a certain stimulus (zone 1-7 and owner) (rpm)
	Reference once	Gaze from owner to stimulus (zone 1-7), or vice versa (rpm)
	Reference multiple	Gaze between owner and stimulus (zone 1-7). In the order of owner-stimulus-owner, or stimulus-owner-stimulus (rpm)

* duration expressed as % of the total observation time.

Appendix 3: Protocol 3-5 of the Theory of mind test

Protocol 3

Apparatus:

A removable curtain (220 x 55 cm) hanging from the ceiling by rope was placed two meters from the DRP. The height of the curtain was made so that when the curtain is placed the dogs could not see the baiting behind the curtain. The curtain could be moved out of sight when not needed. Owners were instructed to look at the poster during baiting for the phases without curtain to prevent unintentional signalling (figure 1A).

Training:

The basic set-up for phase 1,2 and 3 were the same (table 1) and were the same as in protocol 6. However, no second informant pointed in the last two trials at the baited bowl compared to protocol 6. In each trial only one informant pointed and in the first 2 phases less trails were performed were in phase three more trials were performed. The treat was only given after the dog made the correct choice so during false baiting only one treat was used compared to two treats in protocol 6. In phase one and two the same positions were used as in protocol 6. Phase three consisted out of two parts, first four training trials were performed and when the dog made three or more correct choices the dog continued to an exam, in this exam four or more correct choices needed to be made to continue to phase four. When in the training part of phase three the dog did not make three or more correct choices an extra training set of four trials was performed before the dog started the exam trials. After phase three a break of 3-5 minutes was inserted. After the break phase four started, where false baiting and the curtain were introduced. In the first two trials of phase four only two bowls were used on position one and four. When these trials were correct two bowls were added. The informant that pretended to hide the treat grabbed the curtain and put it in front of the bowls. The informant then retook its position, got the attention of the dog (call its name) and hid the treat in one of the bowls using false baiting out of sight of the dog. The informant that baited the bowl then removed the curtain so the bowls would be visible for the dog, retook its position and pointed at the baited bowl after a two second delay. The other informant would do nothing and looked at the wall behind the owner. When the dog made the correct choice, the treat was put in the correct bowl so the dog could at the treat and a vocal praise was given. When the dog did not make the correct choice, the owner retrieved the dog on the DRP and no vocal praise or reward were given. The informants took in their new positions and a new trial began.

Table 1: *Protocol of each phase used during the training part of the ToM test in protocol 3. The goal of the phase, the number of bowls used, the position of the bowls, the number of trails executed, the number of trials that needed to be correct to move on to the next phase and the consequences for not having the right number of correct choices are displayed.*

Phase	Goal	Number of bowls	Position(s) bowl	Number of trials	Number of correct trials needed	When not correct?
1	Accustom to set up and experience a reward	1	5	2	2	Repeat phase 1
2	Let experience a choice	2	1 and 4	2	2	Back to phase 1
3	Let experience a choice between more bowls	4	1,2,3,4	9 (4 training, 5 exam)	3 (training) 4 (exam)	Extra set of training trials
4	Introduce false baiting. The curtain is introduced. First 2 bowl then 4	2,4	1 and 4 1,2,3,4	6	4	Stop the test

Experiment:
In protocol three there was only one experimental phase: phase five, the Guesser/Knower task with two informants. In this phase two informants performed the test. One informant was the Knower it handled the treat and baited the bowl, whereas the second informant was the Guesser and left the room during baiting. After phase four the owner was instructed on the experiment. The experiment consisted of six trials. The Knower would grab the curtain and put it in front of the bowl. When the dog had its attention on the informants or the informant got the attention of the dog (called its name) the Guesser would leave the room and walked in the observation room. The Knower would then get the attention of the dog (called its name) and hid the treat with false baiting, out of sight of the dog. The Guesser waited in the observation room and returned to its position when it saw that the Knower retook its position. The Knower would then take away the curtain, retook its position counted to three and at three pointed at the baited bowl. When the Guesser retook its position, it looked at the wall behind the owner until the Knower counted to three and at three pointed at a non-baited bowl. When the dog made the correct decision, the treat was put in the bowl and a vocal praise was given. When the dog did not make the correct choice, the owner retrieved the dog on the DRP and no vocal praise or reward were given. The informants took in their new positions and a new trial began. After six trials the test was stopped.

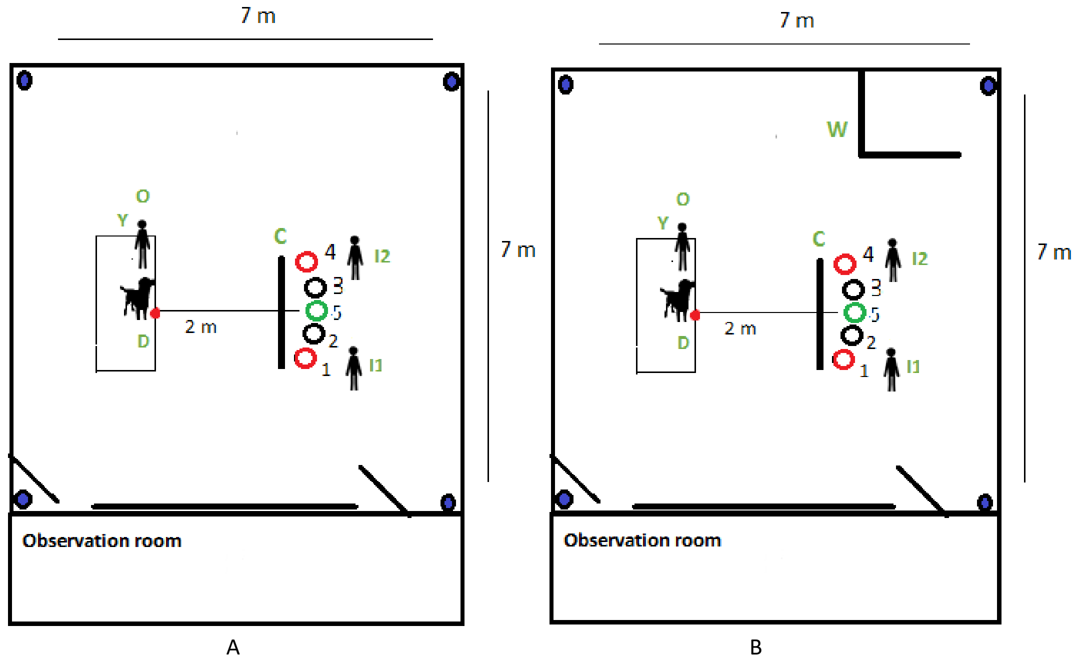

Figure: *1: Set up the room at Carus during protocol 3 (**A**) and protocol 4 and 5 (**B**) of the ToM test. The owner (O) was instructed to stand beside the dog (D) and let the dog sit on the dog release point (red dot) on the yoga mat (Y). The owner was informed to look at the poster (P) on the wall during trials were no curtain was present. A removable curtain (C) was placed before the bowls. The numbers indicated the positions of the bowls. **A:** Two informants (I1 and I2) placed themselves behind the bowls. During phase one only the green bowl on position 5 was used, in phase two the two red bowls on position 1 and 4 were used, in phase three and four the red and black bowls on position 1,2,3 and 4 were used. The curtain was put in front of the bowls and the Guesser left the room during the G/K test. **B:** The set-up of the screen, bowls and informants was the same as in protocol 3. Only a wooden screen was placed in the right corner were the Guesser would place itself during the G/K test.*

Protocol 4 & 5

Apparatus:
A removable curtain (220 x 55 cm) hanging from the ceiling by rope was placed two meters from the DRP. The height of the curtain was made so that when the curtain is placed the dogs could not see the baiting behind the curtain. The curtain could be moved out of sight when not needed. Furthermore, a wooden screen was placed in the right corner of the room. The horizontal side of the screen was placed against the wall. The vertical side of the screen was placed 75 centimetres from the wall so the

informant could place itself easily behind the screen. Owners were instructed to look at the poster during baiting for the phases without curtain to prevent unintentional signalling (figure 1B).

Training:
Protocol 4: The basic set-up for phase 1,2 and 3 were the same (table 2) and were the same as in protocol 6. However, no second informant pointed in the last two trials at the baited bowl compared to protocol 6. The treat was put in the bowl during baiting so was the same as in protocol 6. The first two phases are the same as in protocol 3. Only one bowl is used in phase one and two bowls are used in phase two and in each phase two trials were performed. In phase three six trials were performed were the dog had to have four or more correct choices to continue to phase four, when the dog did not make four or more correct choices another set of six trials was performed. Then a break of 3-5 minutes was inserted. In phase four false baiting and the curtain was introduced, six trials were performed with first two bowls and then with 4 bowls, four or more correct choices needed to be made to continue to the experiment. The protocol for phase four was the same as in protocol three.
Protocol 5: The first two phases are the same as in protocol four. The only difference in protocol five in the training phase is phase three, in phase three eight trials were performed instead of six trials compared to protocol 4.

Table 2: Protocol of each phase used during the training part of the ToM test in protocol 4 and 5 . The goal of the phase, the number of bowls used, the position of the bowls, the number of trails executed, the number of trials that needed to be correct to move on to the next phase and the consequences for not having the right number of correct choices are displayed.

Phase	Goal	Number of bowls	Position(s) bowl	Number of trials	Number of correct trials needed	When not correct?
1	Accustom to set up and experience a reward	1	5	2	2	Repeat phase 1
2	Let experience a choice	2	1 and 4	2	2	Back to phase 1
3	Let experience a choice between more bowls	4	1,2,3,4	6	4	Extra set of trials
4	Introduce false baiting. The curtain is introduced. First 2 bowl then 4	2,4	1 and 4 1,2,3,4	6	4	Stop the test

Experiment:
Protocol 4: In protocol four there was also only one experimental phase: phase five, the Guesser/Knower task with two informants. The protocol was the same as in protocol three. The only difference was that the Guesser walked to the wooden screen set-up in de corner and hid her/his so the dog would know that the Guesser was not present.

Protocol 5: In protocol five the experimental phase consisted of two phases, phase five and phase six. The set-up and execution of phase five was the same as in protocol four only five instead of six trials were used. When the dogs made three or more correct choices the dogs continued to phase six. In phase six the third informant, the baiter, was introduced. The baiter placed itself between the two informants, handled the curtain, treats and the dog. The baiter would grab the curtain and put it in front of the bowls When the informants got the attention of the dog, the Guesser walked to the wooden screen and hid. The baiter then got the attention of the dog and hid the treat with false baiting out of sight of the dog. The Knower looked attentively at the baiting of the treat by the baiter. When the treat was hidden the baiter would give an indication to the Guesser that it was finished: Yes. The Guesser retook its position and the baiter removed the screen and stayed next to the screen. The baiter would then count to three and the Knower and Guesser pointed at three at the assigned bowls.

Appendix 4: **Characteristics dogs Altruism test**

In the altruism box test 40 dogs were selected from the database of van Herwijnen et al.,2018. Each dog would get its own VoterID so it was clear which dog had performed which trial and the number of trials performed. Characteristics such as sex, age and the breed of the dog are described below.

VoterID	Sex	Age	Breed
29515	F	8	Beauceron
29621	F	8	Rottweiler
29629	M	5	Rottweiler
29643	M	4	Saint-Bernard Long hair
29647	F	12	Welsh Corgi Cardigan
29687	F	7	Unknown
29691	M	2	Podengo Português
29698	F	7	Golden retriever
29721	M	5	Flatcoated retriever
29725	F	6	Belgian shepherd
29726	F	7	Labrador
29727	M	9	Standard dashund
29731	F	2	Mix pointer/Mastin
29745	M	9	Rottweiler
29746	F	3	Swiss white shepherd
29754	F	7	Old german shepherd
29769	F	5	Hovawart
29780	M	3	Swiss shepherd
29793	M	9	Grand Basset Griffon Vendéen
29805	M	8	Shepherd x Labrador
29814	F	2	English cocker spaniel
29815	M	4	Golden retriever
29827	F	8	Mix
29876	M	13	Erdelyi Kopo
29885	M	6	Briard
29911	F	5	Welsh Springer Spaniel
29914	M	13	Frisian pointer
29918	F	7	Mix
29926	F	10	Labrador retriever
30942	M	4	Sheltie
30964	M	3	Border Collie
30971	F	5	English cocker spaniel
31916	M	1	Vizsla
31925	M	1	Weimaraner
31959	M	4	Pure breed
31962	M	9	Australian shepherd
31973	M	12	Beagle
32038	M	2	Dutch Partridge Dog
32049	M	2	Staffordshire bull terrier
32086	F	5	Scottisch shepherd

Appendix 5: **Interaction differences of owner raising thumb (no thumb up, thumb up) and owner guessing correctly (incorrect, correct) and signalling behaviour.**

*Table 1. Dogs were tested for informing their owner about the location of a missing coat that they had observed being hidden. During test trials owners asked their dog for help in finding the coat. Ninety records of 40 dogs have been analysed with mixed linear models. The statistical model with 3 fixed effects included a 2-way interaction. This interaction the choice of the owner for the location of the coat (incorrect (False), correct (True)) and owner thumb (no thumb up, (↓), thumb up (↑)) and was significant different in approach the target and look at target. The difference between predicted means (Δ), standard error of differences times two (SED*2) and whether the interaction is significant (Δ>SED*2) is presented per signalling behaviour.*

	Approach target		
Interaction	Δ	SED *2	Δ>SED*2?
False,↓*False,↑	-0.42	0.4484	no
False,↓*True,↓	0.1824	0.6482	no
False,↓*True,↑	-1.6657	0.6316	yes
False,↑*True,↓	0.6024	0.7128	no
False,↑*True,↑	-1.2457	0.701	yes
True,↓* True,↑	-1.8481	0.8382	yes

	Look at target		
Interaction	Δ	SED *2	Δ>SED*2?
False,↓*False,↑	-0.745	0.782	no
False,↓*True,↓	-0.12	1.1	no
False,↓*True,↑	-2.765	1.16	yes
False,↑*True,↓	0.63	1.222	no
False,↑*True,↑	-2.020	1.29	yes
True,↓* True,↑	1.15	1.48	no

Appendix 6: **Characteristics dogs Theory of mind test**

In the Theory of Mins test 34 dogs were selected from the database of van Herwijnen et al.,2018. Each dog would get its own VoterID so it was clear which dog had performed which trial and the number of trials performed. Characteristics such as sex, age, the breed of the dog and in which protocol they participated, are described below.

VoterID	Sex	Age	Breed	Protocol
33156	F	5	Manchester Terrier	3
33051	F	2	Small Münsterlander	3
32193	M	6	Welsh Corgi	3
34874	M	9	Shiba inu	3
54602	F	6	Mix Saint-Bernard golden retriever	4
33325	F	9	Border Collie	4
54606	F	3	Labrador	4
32504	F	12	Labrador	4
32446	M	6	Old-German Shepherd	5
34904	M	2	Novan Scotia Duck Tolling Retriever	5
54800	F	9	Labrador	5
33258	F	4	Labrador retriever	5
54666	F	4	Mix	5
39185	M	5.5	Labrador	6
54604	F	1	Belgian Shepherd	6
54607	M	8 months	Mix	6
32244	F	11	Australian Shepherd	6
54802	F	1.5	Mix	6
32800	M	4.5	Longhaired Dachshund	6
54799	M	1	Spinone Italiano	6
32248	M	6.5	Welsh springer spaniel	6
33672	M	5.5	Labrador retriever	6
29846	M	8	Boomer	6
33385	M	10	Jack russell Terrier	6
32638	F	8	Basset Fauve the Bretagne	6
29569	M	8	Belgian Shepherd	6
35143	F	5	Pug	6
54807	F	2.5	Golden retriever	6
32440	F	12	Labrador	6
10000	M	8	Labrador	6
20000	F	2	Labrador	6
33663	F	2.5	Labrador retriever	6
32495	M	2.5	Mix Husky malamute	6
54806	F	11	Labrador retriever	6