

Department of Interdisciplinary Life Sciences

University of Veterinary Medicine Vienna

Messerli Research Institute

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

**Examining the test-retest reliability of the Guesser-Knower task with
dogs**

Diploma Thesis

University of Veterinary Medicine Vienna

Submitted by

Hannah Veitl

Vienna, March 2026

Supervisor: Univ.-Prof. Dr.rer.nat. Ludwig Huber
Messerli Research Institute
Department of Interdisciplinary Life Sciences
University of Veterinary Medicine Vienna
Lucrezia Lonardo, PhD.
Messerli Research Institute
Department of Interdisciplinary Life Sciences
University of Veterinary Medicine Vienna

Reviewer: Univ.-Prof. Jean-Loup Rault, PhD.
Animal Welfare Science
Department of Animal Nutrition and Animal Welfare Sciences
University of Veterinary Medicine Vienna

Declaration of Authenticity

I hereby declare that I have written this thesis independently and that I have not used any other sources and aids other than those listed. All text passages taken from other sources have been cited.

I have carried out the decisive work myself and have listed all contributors with their contributions to the work.

This thesis has not been submitted or published elsewhere.



Vienna, 11.03.2026

Hannah Veitl

ZUSAMMENFASSUNG

Die kognitiven Fähigkeiten von Tieren und besonders Tieren, die in enger Gemeinschaft mit uns Menschen leben, werden viel erforscht. Ein wichtiger Aspekt der Kognitionsforschung ist die "Theory of Mind". Diese beschreibt die Fähigkeit eines Individuums, sich selbst und anderen Individuen geistige Zustände zuzusprechen (1). Haushunde (*Canis familiaris*) sind besonders interessant für diese Art der Forschung, da sie besonders gut darin sind, uns Menschen zu lesen (2,3). Trotzdem gibt es noch keine eindeutige Übereinstimmung, wie weit Tiere andere Individuen verstehen und wie gut geeignet aktuelle Tests sind, um dies zu überprüfen. Es gibt kaum Forschung über die psychometrischen Eigenschaften von nonverbalen Tests, die verwendet werden, um Rückschlüsse über kognitive Fähigkeiten von Tieren zu ziehen, und keine, die Hunde betrifft.

Catala et al. (4) führten eine Studie mit 16 Hunden durch, indem sie mit der Guesser-Knower-Aufgabe die Fähigkeit von Hunden testeten, die Perspektiven anderer Individuen einzunehmen, besonders die Fähigkeit, zwischen einem wissenden und unwissenden Experimentator zu unterscheiden, die jeweils richtig und falsch den Ort eines versteckten Futterstücks anzeigen (5). In dieser Studie wurde eine signifikante Präferenz für den wissenden Experimentator („Knower“), unter verschiedenen Bedingungen gefunden, was auf die Fähigkeit hindeutet, Perspektiven anderer einzunehmen und ihren Blick geometrisch zu verfolgen. Wir replizierten diese Studie mit 93 Hunden und erweiterten das Experiment um eine Bedingung, um diese Fähigkeiten weiter eingrenzen zu können. Wir konnten die Ergebnisse von Catala et al. (4) teilweise replizieren, aber wir konnten auch zeigen, dass Hunde keine Präferenz für den Knower zeigen unter der Bedingung, dass sich beide Informanten identisch verhalten, indem sie in die gleiche Richtung schauen, aber aufgrund ihrer Position einer in die Richtung des Versteckprozesses schaut und der andere weg von diesem Vorgang schaut, während das Futter von einem dritten Experimentator versteckt wird. Dies würde gegen die Fähigkeit von Hunden sprechen, den Blick eines anderen Individuums geometrisch zu verfolgen. Außerdem testeten wir die Performance von Hunden über mehrere Sitzungen mit einigen Wochen zwischen den einzelnen Terminen. Wir fanden eine sehr variable Performance der Hunde in den verschiedenen Sitzungen, was für eine niedrige Test-Retest Reliabilität sprechen würde.

ABSTRACT

The cognitive abilities of animals and even more so animals who live in close proximity to us humans are subjects of many studies. An important aspect of social cognition is Theory of Mind which describes the ability of individuals to attribute mental states to other individuals and themselves (1). Domestic dogs (*Canis familiaris*) are especially interesting for this kind of research as they show exceptional abilities at reading human cues (2,3). However, there is still no general consensus about how well animals understand other individuals and how well current tasks are suited for testing for this ability. There has been very little research about the psychometric properties of nonverbal tests used to draw inferences about animal cognition and none with dogs.

Catala et al. (4) performed an experiment with 16 domestic dogs using the Guesser-Knower task in order to investigate the ability of dogs to take another individual's perspective, and in particular the ability to distinguish between a knowledgeable and ignorant experimenter who indicate correctly and incorrectly, respectively, where a piece of food is hidden (5). In their study they found a significant preference for the knowledgeable individual ("Knower") across several conditions, pointing to the possibility that dogs are able to take another individual's perspective and follow their gaze geometrically. We not only replicated their study with 93 dogs but also introduced a new condition to further investigate dogs' ability to take another individuals perspective. We could replicate some results by Catala et al. (4) but could also show that the preference for the Knower does not hold up in a larger sample size in a condition in which both informants behave identically, by looking towards the same direction, but because of their position in the room one looks in the direction of the food hiding and the other away from it while the food is hidden by a third experimenter. Our findings speak against the possibility of dogs being able to follow gaze geometrically. Furthermore we tested dogs' performances across several sessions with several weeks between appointments. We found dogs' performances to be highly variable across sessions, indicating a low test-retest reliability.

List of abbreviations

GP	Guesser Present
GA	Guesser Absent
GBT	Guesser Back Turned
GLA	Guesser Looking Away

List of contents

1.	INTRODUCTION	7
1.1.	THEORY OF MIND	7
1.2.	PERSPECTIVE-TAKING ABILITY	9
1.3.	GUESSER-KNOWER PARADIGM	10
1.4.	TEST-RETEST RELIABILITY	13
2.	ANIMALS, MATERIALS AND METHODS	14
2.1.	ANIMALS	14
2.2.	MATERIALS AND METHODS	15
2.2.1.	Set up	15
2.2.2.	Familiarisation phase	17
2.2.3.	Pre-training phase	18
2.2.4.	Testing procedure.....	19
2.3.	STATISTICS	22
3.	RESULTS	23
3.1.	CORRELATION BETWEEN SESSIONS.....	23
3.2.	COMPARISON WITH STUDY OF CATALA ET AL. (4).....	28
4.	DISCUSSION	30
4.1.	TEST-RETEST RELIABILITY BETWEEN SESSIONS	30
4.2.	COMPARISON WITH THE STUDY OF CATALA ET AL. (4).....	31
4.3.	CONCLUSION	33
5.	BIBLIOGRAPHY	34
6.	LIST OF TABLES	36
7.	LIST OF FIGURES	37
8.	ACKNOWLEDGEMENTS	38

1. Introduction

The extent of animal cognition has always been a fascinating topic for humans, even more so when we think about animals who live in close proximity to us. An example for this would be the dog (*Canis familiaris*) who has been domesticated over thousands of years ago and has a remarkable ability to follow human cues. This is not surprising considering that they have been fulfilling numerous different functions as working and social companions. Their roles range to this day from working companions such as hunting, herding or medical assistance dogs to social partners, such as therapy or family dogs (6). Their close relationship with humans makes it especially interesting and also relatively easy for cognition research as dogs are more accessible and used to human contact in comparison to other animals. Due to the domestication process dogs have changed in their social behaviour and their interaction with humans (7). Studies support the possibility that adult dogs are even better at following human cues than cues from their conspecifics (2). Research on dogs' cognitive abilities is valuable when it comes to our understanding of human influence on animals' cognition through domestication, especially when comparing dogs' abilities to those of wolves (7). Especially in foraging situations, dogs are very skilful at following human behavioural cues such as pointing or gazing. However, the question, whether dogs can actually understand perspectives of other individuals and use this knowledge remains unanswered. Catala et al. conducted a study where they attempted to answer this question using the guesser-knower-paradigm (4). We wanted to replicate this study with a larger sample size and in a longitudinal study, in order to assess the test-retest reliability of this task.

1.1. Theory of Mind

Premack and Woodruff (1) have defined theory of mind as the ability of an individual to attribute mental states to himself and others. As one can already conclude from this description, theory of mind is not simple to measure as it is not a directly observable skill. Mental states that humans can infer would be for example intention, pretence, knowledge or ignorance, of which the latter two are especially relevant for our experiment (1).

There has been research on non-human primates, corvids and dogs to learn more about whether animals possess at least some of the abilities described under the umbrella of Theory of Mind.

Until now there is no conclusive evidence that would support the claim that either of those species have an explicit, human-like theory of mind. When testing for theory of mind it is important not to accidentally test simpler abilities like reading behaviour which makes accurately testing for theory of mind difficult (8). There have been many previous studies which attempted to further deepen our understanding of dogs' sensitivity to human attention and perspective taking.

Call et al. conducted a study in 2003 to test their sensitivity to humans' attentional state, focusing on the role of eyes and body orientation. Being sensitive to others attentional state is an important part of the umbrella term theory of mind. In their study, the experimenter entered the room with the dog, placed the food on the floor and then forbade the dog to take the food with a verbal command. For the first experiment, in different conditions, the experimenter would either leave the room or stay and give the dog a release command or not say anything. For the second experiment the experimenter would again forbid the dog to take the food and then either watched the dog, played on a handheld computer, closed the eyes or turned the back to the dog. Dogs retrieved less food or approached it indirectly when the experimenter was watching them than when they were inattentive or not in the room. This study also showed that dogs used the eyes of humans as an indicator of humans' attentional state (9).

Similar to the study above, Schwab and Huber tested the sensitivity for attentional cues in dogs but instead of using an unfamiliar experimenter in an unfamiliar environment, they conducted the experiments in the owner's living room, and the owner took the role of the experimenter. The owner told the dog to lie down and then either watched the dog, read a book, watched TV, turned their back to the dog or left the room entirely. In general, the results replicated those of Call et al. with dogs taking the food faster when the owner were either distracted or left the room. However, in their study the dogs took the food faster when the owners read a book or watch TV, while in the study of Call et al. the dogs took the food quicker when the eyes of the experimenter were closed than when they were distracted. This indicate that dogs' behaviour differs when they know the experimenter and therefore can estimate the severity of possible consequences when disobeying the command and in turn rises the question whether these sensitivity to attention might be simply be learned by dogs through cohabitation (10).

So, dogs' sensitivity to attentional cues is extraordinarily developed, however this alone does not indicate the possession of a humanlike theory of mind as it can be subscribed to a simpler cognitive process like behaviour reading. Between a theory of behaviour (animals connect certain behaviours with certain outcomes through experience) and actually having a theory of mind there is a grey area which could be explained by degrees of theory of mind or a rudimentary theory of mind (11).

1.2. Perspective-taking ability

An important aspect of theory of mind is perspective-taking, the ability to take the point of view of another individual.

Bräuer et al. (12) conducted a study where they investigated if dogs are capable of level-I perspective taking by experimenting with barriers. In their first experiment the experimenter placed a piece of food and then prohibited the dog to eat it. Depending on the condition, there was barrier placed either perpendicular and left to the experimenter (who thus was able to see the food) or parallel and in front of the food and blocking the experimenter's visual access to the food. Dogs took the forbidden food significantly more often when the experimenter could not see the food-taking. They did a similar experiment where the barrier either had a small window, so the dogs' approach to the food was not visible but the food-taking itself was visible, or the barrier was smaller, and the approach could be seen but not the food-taking. While there was no significant difference between the window- and the small-barrier-condition, the dogs took significantly more often the food when the barrier was big and windowless than when either of those conditions were given. These findings are compatible with dogs being sensitive to the visual access of another individual and being capable of level-I perspective taking (12). However, it is also possible that dogs simply based their choice whether to take or leave the food on certain environmental cues and the use of an egocentric strategy: e.g., one barrier allowed the dog to see the experimenter, from the dog's point of view, and was avoided more often than the other. Such an egocentric strategy does not require the ability to think about the experimenter's perspective, but still results in passing the task (i.e., behaving as if one was capable of understanding the experimenter's visual access). The same inability to distinguish between behaviour-reading and mind-reading applies to most subsequent studies on dogs'

perspective-taking. For example, Kaminski et al. (13) conducted a study where a human asked the dog to bring them a toy. There were two toys visible to the dog in the room and between the human and the toys there were two barriers, only one of which was transparent. Therefore, the human had only visual access to one of the toys (but, again, also the dog had visual access to the human only in proximity of the transparent barrier). The human asked the dogs to bring a toy, but did not specify which toy the dog should bring. The dogs showed overall a preference for the toy that was visible to the human, even more so when the toys were different (13). This further supports the finding of Bräuer et al. (12) that dogs seem to be able of level-I perspective taking, but once more these tests did not control for behaviour-reading.

1.3. Guesser-Knower paradigm

In our experiment we used the Guesser-Knower paradigm, which was first used by Povinelli et al. with four chimpanzees (*Pan troglodytes*) (14). This paradigm is used to examine the ability of animals to infer that seeing leads to knowing. For the procedure there were three experimenters: one experimenter who operated the hood that covered the food trays and the cups which covered the food, one knower and one guesser. The knower hid the food and therefore knew where the food was located. The guesser did not watch the food hiding process and assumed different states of inattentiveness. The food was placed visibly on trays between the subjects and the experimenters. First the guesser left the room, then the cups were placed over the food trays, the knower hid the food under one of the cups while performing false baiting movements. The subjects could see the food hiding process but not under which cup the food had been hidden. Then the guesser came back into the room and both guesser and knower pointed to different cups, the knower always pointing to the baited cup. Then the third experimenter removed the hood, and the subjects could pull the tray closer to them, lift the cup and obtain the food if they chose the correct cup. When choosing incorrectly, the knower let the subject know by showing them the food under the baited cup, which the subject did not receive in this case. While all subjects showed a preference for the Knower, their performance only improved a bit over the six weeks of this phase. They repeated this experiment and additionally let the knower and the operator wear little blue hats as discriminative cues in hopes to improve the performance of the subjects. Surprisingly this additional marker did not have a

significant influence on the performance. For the third phase, the guesser stayed in the room during the food hiding process, performed for the first time by the third experimenter, but before the hider placed the food under a cup, the guesser put a paper bag over his head and therefore could not see where the food was being hidden. Three of the four chimpanzees showed a significant preference for the knower in all three phases. This seems to support the hypothesis, that they are able to make an inference about the knowledge of another individual or lack thereof based on different visual experiences, but that they need hundreds of trials to learn such discrimination and transfer it to novel conditions (14).

In 2014 Maginnity and Grace (15) performed similar experiments with dogs using the guesser knower task. In their experiment the dog with its owner was seated on one side of the room and the experimenters on the other side, looking at the dog. Between them was a removable barrier and behind the barrier there were four cups. The barrier allowed the dog to see the upper bodies of the experimenters but not the cups. Their “Absent condition” was most similar to the study of Povinelli et al., the knower assumed the same role as in their study and the guesser left the room for the hiding process, then re-entered the room and both experimenters pointed at a cup, the knower pointing to the baited cup. In this condition, dogs showed a significant preference for the knower ($M = .73$; 95% CI .66–.79). For a subsequent condition that was performed once the dogs had already been tested in the previous one, a third experimenter handled the food and was seated between the guesser and the knower. The knower covered their cheeks with their hands, and the guesser covered their eyes during the hiding of food. Even when the knower was not handling the food, dogs showed again a significant preference for the knower. For another subsequent condition (tested with the same subjects), the food baiting was again performed by a third experimenter while the knower watched them attentively and followed the baiting with their gaze. The guesser looked up at the ceiling and again, the dogs showed a significant preference for the knower. To test for unintentional cues or expectancy effect, they also had one condition where both the guesser and knower either did or did not watch the baiting process and one where neither of the experimenters pointed to a cup. Pooled over attentive (both informants watched the baiting) and inattentive (neither informant watched the baiting) trials, overall Knower preference was at 50%, not above chance level. Overall, dogs showed a significant preference for the knower and performed better than nonhuman primates as well (14–

16). The guesser-knower task seems to be showing very well how good dogs are at reading behavioural cues to human attention.

Catala et al. (4) also used the guesser-knower task to further investigate perspective taking in dogs. The general procedure was the same as in the study of Maginnity and Grace, but additionally to the Guesser Absent and Guesser Present condition, they introduced a new condition in their studies, the Guesser Looking Away condition, to at least eliminate behavioural differences (though not environmental ones) between experimenters. The setup was similar to Maginnity and Grace, with the dog and their owner on one side of the room and the experimenters and cups for the food on the other side behind a barrier. Again, the barrier allowed the dog to see the upper bodies of the experimenters but not their hands and the cups. For the Guesser Looking Away condition, the food handling was performed by a third experimenter who knelt between the two other experimenters. Both the guesser and the knower looked down and to the same side so that their lines of sight were parallel to each other. This way, only the knower had visual access to the baiting as the line of sight of the guesser went past the side of the barrier. Dogs showed in the Guesser Absent and also the Guesser Looking Away condition a significant preference for the knower. This is compatible with dogs discriminating between a knowledgeable and an ignorant experimenter based on their looking behaviour (4). In our study we aimed at (i) replicating these results with larger sample sizes, (ii) examining the test-retest reliability by conducting a longitudinal study and (iii) introducing a new condition in order to test construct validity (the construct validity aspect is left out of this thesis as it is presented in another thesis: Dygryn, *in prep.*). There has been research which suggests that there are some aspects which have a significant influence on dogs' performance in the guesser-knower task, such as the familiarity with the objects that are being used to obstruct the guesser's view (17). If minor methodological choice impact dogs' performance considerably, it might be possible that this task is not a stable test for perspective taking in dogs. If we want to draw solid conclusions from the guesser-knower task, we have to examine if it provides stable results across a longer time frame. This is why we conducted a longitudinal study with a large sample size so we can further examine the test-retest reliability of the guesser-knower task.

1.4. Test-retest Reliability

The test-retest reliability measures how stable results are across multiple replications of the same testing procedure. In our experiment we not only wanted to see whether we could replicate the results of Catala et al. (4) with a larger sample size, but also if there were differences between different sessions. For this, we tested dogs in three identical sessions with about three months between sessions.

While the study of Catala et al. (4) originally replicated the study of Maginnity and Grace (15) and could further support their results that dogs preferred the Knower in a Guesser-Knower task, the sample size of both studies was rather small with 16 dogs each. It has also not been tested so far if the same dogs perform with the same accuracy over a longer period of time or if they improve. For chimpanzees, there appeared to be a learning curve in the first phase of the study by Povinelli et al. (14) which stabilized after three to four weeks of repeated. However, this might not be necessarily connected to the understanding of humans' perspective but rather to their ability to follow human pointing gestures, at which dogs perform generally better and without any prior training.

As there has not been any research on the test-retest reliability of perspective-taking tasks in dogs, it is important to do further research to provide more insight on the integrity of nonverbal tests. As the test-retest reliability describes the level of congruency between several measurements of the same individual under the same conditions, it provides an insight if differences in test results stem from actual differences in performance rather than instability of the measuring instrument or in this case, measuring task (18). Without sufficient test-retest reliability the results might be strongly variable which limits the validity of the whole task and its results (19).

2. Animals, Materials and Methods

2.1. Animals

We tested pet dogs (*Canis familiaris*) of different ages, breeds and sexes. To be eligible to participate, dogs needed to have been adopted as pets before their 13th week of life (approx. the end of dogs' critical period for socialisation). Overall we tested 93 dogs for the first session, 87 dogs for the second session and 84 dogs for the third session. We tested 52 female and 41 male dogs with an average age of 69.35 months. In the first session we had to exclude 15 dogs, in the second session six dogs who participated in session one did not come back due to different personal or health reasons of the owner. Of those six dogs, five also did not participate in the third session. For the third session we had to exclude three more dogs and one dog who participated in the first two sessions did not participate in the third session.

Table 1: List of excluded dogs

<i>Name</i>	<i>Test date</i>	<i>Session</i>	<i>Phase excluded</i>	<i>Trial excluded</i>	<i>Reason</i>
Nala12	04/02/2025	1	test	Trial 11	4x no choice
Skye	04/02/2025	1	training	Step 5	4 mistakes during pre-training, not going straight to the cup
Abby6	07/02/2025	1	training	Step 3	3 consecutives wrong choices during pre-training
Nala15	17/02/2025	1	training	Step 2	3 consecutive mistakes, seem too afraid
Blue6	17/02/2025	1	training	Step 1	3 mistakes in a row
Ross	18/02/2025	1	training	Step 1	3 no choice in a row, too old, suspicion of blindness
Akela3	25/02/2025	1	test	Trial 8	Too long to come back, 1 hour to do the pre training and only 8 trials
Pixi	06/03/2025	1	training	Step 5	4 total mistakes during pre-training, dog not following the pointing
Nuu	07/03/2025	1	training	Step 5	wet food everywhere, no following the pointing, 4 mistakes in total
Onix	07/03/2025	1	training	Step 2	4 mistakes in total, not interested

Puck3	10/03/2025	1	training	Step 4	4 mistakes in total
Hunter2	14/03/2025	1	test	Trial 24	video lost
Garou	09/03/2024	1	training	Step 5	4 mistakes in total
Paige	25/03/2025	1	training	Step 5	4 mistakes in total, cannot see well because of her eyebrows
Sherlock	04/04/2025	1	test	Trial 12	Too long to come back to the owner, loss of motivation
Sixtus	08/10/2025	3	training	Step 5	4 mistakes in total
Ragnar	20/10/2025	3	training	Step 5	4 mistakes in total
Mickey3	29/10/2025	3	training	Step 5	4 mistakes in total
<i>Did not come to the second session</i>				<i>reason</i>	
				Cookie12	change of behaviour of the dog
				Bonda	doesn't want to come back
				Boomer2	health issue from the owner
				Hilde	never answered
				Paul3	injury left eye
				Emma10	not available
<i>Did not come to the third session</i>				<i>reason</i>	
				Cookie12	change of behaviour of the dog
				Bonda	doesn't want to come back
				Boomer2	health issue from the owner
				Nika2	health issue from the owner
				Paul3	doesn't want to come back
				Hilde	was planned but could not come in due to personal reasons

2.2. Materials and Methods

2.2.1. Set up

The testing took place at the Clever Dog Lab at the University of Veterinary Medicine Vienna. The experimental room (Figure 1) measured 605 x 333 cm, with cameras placed on the wall above the entrance behind the dog, on the ceiling in the middle of the room and on the wall

behind the experimenters on the opposite side of the room. The blindfolded owner was sitting on a chair behind the dog. The dog's starting position was a 38 x 38 cm square between the legs of the owner. In front of the dog was a barrier, which could be lowered to the floor between trials. The barrier was a chipboard barrier (220 x 56 cm) with L-feet on left and right side which we blocked with wooden blocks when we needed the barrier upright. For lowering the barrier, the blocks were removed by the two informants and the barrier carefully lowered. Between the upright barrier and the dog was a distance of 120 cm and behind the barrier the experimenters were kneeling on cushions. Between the experimenters and the barrier, four opaque cups were placed at equidistant positions to the dogs (each 140 cm away from the dog). Between each cup there was a distance of 45 cm. The cups were numbered from one to four, starting on the left from the experimenters' point of view and in front of the cups were lines marked at a distance of 20 cm to the cups. These lines were used for scoring as the crossing of a line with the dog's nose was marked as the chosen cup. For the experimenters there were three cushions placed equidistant to each other behind the barrier and the cups. The two informants, alternating between the roles of guesser and knower within a testing session for each dog, were kneeling on the outer two cushions (fixed positions, counterbalanced across dogs). The cushion in the middle was used by the Hider in the "Guesser looking away" condition to perform the food hiding process. In the other conditions, the Hider was sitting on a chair behind the two informants, facing the dog and noting the dog's choices.

A bowl of water on a towel for the dog was always in the room next to the entrance and available for the dog at any point during the experiment. The informants had each a treat bag clipped on their belts on the backside, which contained Frankfurter sausage, cheese or treats provided by the owner (depending on each dog's dietary requirements). The cups were cylindrical cups with a diameter of 8.5 cm and a height of 10.5 cm with removable lids. The cups were rubbed with the treats of choice before the experiment so all of them smelled like food, which prevented dogs from smelling the treat location. In the cups, paper towels were placed to avoid a distinct sound of the treat being placed inside the cup. The cameras that we used were Panasonic HC-V777 High-Definition Video Cameras with full HD resolution, with a frame rate of 50 frames per second.

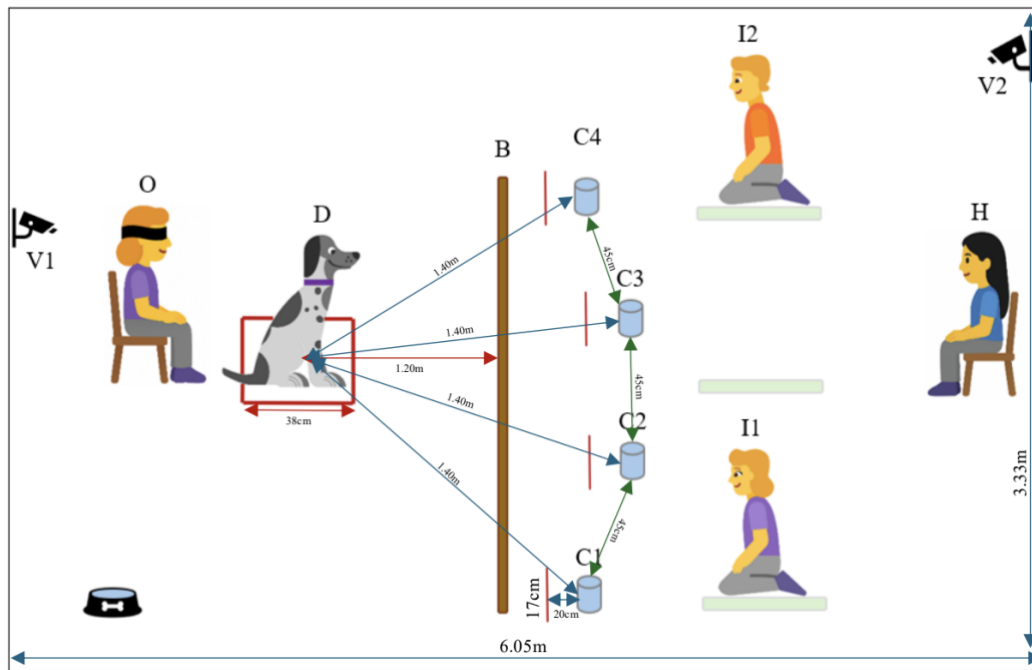


Figure 1: Set up. Blindfolded Owner (O) sitting behind dog (D), facing the barrier (B) with the two informants (I1, I2) behind it. Hider (H) sits behind experimenters facing the Owner and dog. Cups (C1-C4) placed equidistant from dog's position.

2.2.2. Familiarisation phase

When the owner and dog first entered the testing room, the owner was instructed to remove the leash, so the dog was able to explore the room freely before we started testing. Then we started a short familiarisation with the barrier which was laying on the floor initially. The Hider and the experimenter on the left (from the point of view of the Hider) performed the familiarisation by raising and then lowering the barrier little by little. The Hider provided the dog with treats for any approach to the barrier. The duration of this process depended on the dogs and was stopped when the dog was not bothered by the barrier anymore.

The owner was then instructed to call the dog back and place it centrally between their legs in the square which marks the staring position. The owners held the dogs by their harness or collar and put on their blindfold. They were instructed to release the dog when the Hider said “Okay” and if the dog was not moving after ten seconds, they should encourage it with a vocal command, such as “Go” or “Search” but not to push the dog in any direction. After every trial, we asked the owner to call their dog back, for this they were allowed to remove the blindfold.

2.2.3. Pre-training phase

The video recording was started as soon as the owner and the dog entered the testing room. The setup for the pre-training was as described above, the only difference was that three of the cups were not yet placed on the floor but on the windowsill behind the Hider. There were five steps for the pre-training, for the first step, only one cup was on the floor, for the next step one cup was added and so on until the fourth step. The informant on the left started with the hiding process. They made eye contact with the dog, showed the dog the treat in their hand and made sure that the dog saw the treat. If the dog was not paying attention, the informant made a clucking sound. Then the informant placed the treat in the predetermined cup, made eye contact with the dog again and pointed to the cup in which the food was hidden. After two seconds, the Hider instructed the owner to release the dog. After the dog chose a cup, the owner was instructed to call it back. As already mentioned above, choice was marked when the dog crossed the line in front of the cups with their nose. The informants took turns with hiding the food (and consequently pointing) but during the pre-training only one informant per trial hid the food and pointed at the baited cup, so the dog did not have to choose between the two informants during the pre-training. In step five, the barrier was raised during the hiding process and sham baiting all four cups was added. For sham baiting, the informant opened the lid of every cup and tapped the paper towels with the treat to simulate the sound of placing the treat into the cup. The informants always started the sham baiting from left to right from their point of view. In order for the pre training to be considered successful, the dog had to make two consecutive correct choices in step one to four and six consecutive correct choices in step five. If the dog chose a cup that was pointed at by the informant, it would be marked as a correct choice. If the dog chose a cup that was not pointed at or did not choose a cup at all after 60 seconds, the trial was marked as either wrong or no choice, respectively and would be repeated. If the dog made four mistakes overall or three mistakes in a row, the pre-training was stopped and the dog excluded from the study. If the pre-training was successful, we took a short break of five minutes, and the owner could take a short walk outside with their dog. After owner and dog left the room, the video recording was stopped and saved in a folder for all pre-training video recordings. The sequence of positions (one to four) of the pointed at cup was predetermined and identical for all dogs during the pre-training but counterbalanced so that overall each cup had the same

probability of being baited, unless a dog made mistakes (in which case the same trial as before was repeated).

2.2.4. Testing procedure

Before starting with the test phase, the owner was informed that for this phase, the Hider will count down “Three, two, one, point” but that the owner could ignore this and should again release the dog after the Hider says “Okay”.

The dogs were tested in three identical sessions with approximately three months between sessions. For every session, each dog was presented with six trials of four conditions, so in total 24 trials per session and 72 overall. Overall, 17 counterbalancing tables were created to ensure counterbalancing of: identity of the experimenters playing each role (guesser, knower, hider); position of the baited cup; positions of the pointed at cups; order of presentation of conditions, with each condition being equally likely to appear at every trial number (1 to 24) and with the constraint that each factor could not have the same value for more than two consecutive trials, for similarity with Catala et al. (2017) and Maginnity & Grace (2014). For each dog the counterbalancing table was the same for all three sessions.



Figure 2: a) Guesser Present; b) Guesser Absent; c) Guesser Back Turned; d) Guesser Looking Away; e) Pointing at respective cups

Dogs were tested in four conditions. The “Guesser Present” (GP) condition (Figure 2a) was the control condition in which both informants witnessed the food hiding and pointed both to a baited cup. This served to investigate if the dogs had a preference for one of the informants. The two informants raised the barrier, the Knower made eye contact with the dog and showed the treat and proceeded to sham bait each cup, starting on the left. During this process the Knower left the food in two cups as indicated by the counterbalancing table. In this condition, the Guesser followed the baiting process with their gaze. Both informants then lowered the barrier, sat back and looked down to the floor. The Hider then counted down: at “Two” both informants looked up and simultaneously looked at the dog, at “One” both started to move towards their respective cup and at “Point” both informants touched the lid of their cup with their index fingers (Figure 2e). The informants always pointed with the arm closer to their respective cup in order to avoid contralateral pointing. If the informants’ arms had to cross for a trial, the arm of the left informant was always above the arm of the right informant. After two seconds, the Hider instructed the owner to release the dog. As for the pre-training, a choice was made if the dog crossed the red line in front of one of the cups. In this condition, a correct choice was marked if the dog chose on of the cups at which the informants pointed and received the treat. If the dog crossed the line of a cup to which no informant pointed, it was marked as “Irrelevant choice” and if the dog did not choose any cup, it was marked as “No choice”. For both situations, the trial was repeated for a maximum of three times. If one of the experimenters made a mistake, such as pointing to the wrong cup, the trial was also repeated. After this, the Hider instructed the owner to call the dog back and place it again at the starting position and the next trial began.

The procedure for the “Guesser Absent” (GA) condition (Figure 2b) was almost the same as for the “Guesser Present” condition. The difference was that only one cup is baited by the Knower, and the Guesser left the room before the hiding process starts. The Guesser could watch the hiding process on video over the computer outside of the test room and returned to the room when the hiding process was finished. During the leaving and returning of the Guesser, the Knower kept eye contact with the dog to compensate for the movement of the Guesser. The informants then lowered the barrier again and pointed to their cup with the countdown of the hider. In this and all following conditions, a correct choice was marked only if the dog chose the cup at which the Knower pointed and only then received the treat. If the dog chose the cup

pointed at by the Guesser, it was marked as “Incorrect choice” and the dog did not receive a treat. The dog was then called back by the owner and the next trial began. The criteria for “Irrelevant choice” and “No choice” stayed the same for all conditions.

In the “Guesser Back Turned” (GBT) condition (Figure 2c), the procedure was the same as for the “Guesser Absent” condition. However, during the hiding process, the Guesser did not leave the room but turned around so that their back was visible to the dog. After the hiding process, the Guesser turned around, both informants lowered the barrier, and the remaining procedure was the same as in the “Guesser absent” condition. Again, during the turning of the Guesser, the Knower kept eye contact with the dog to compensate for the movement. This condition is novel to the previous research on dog cognition in a Guesser-Knower task.

The procedure for the “Guesser Looking Away” (GLA) condition (Figure 2d) was different in several ways from the first three conditions. The barrier was again raised by the two informants, then both of them moved their cushions closer to the wall. Then they both looked at two marked spots, their gazes were in a parallel line in the same direction. The point at which the Knower look at, was marked on the barrier and allowed visual access to the food baiting – however, the Knower did not follow the baiting with their gaze. The Guesser looked at a spot at the wall, which made it impossible to see the hiding process. The Hider then knelt down on the cushion between the two informants, took a look left and right to check if both informants looked in the same direction. The Hider then made eye contact with the dog and showed it the treat. They then started sham baiting all cups from the left to right and placed the treat in the predetermined cup during this process. The Hider then got up and sat down in their chair again, the informants lowered the barrier and pointed at their respective cups on the countdown. The scoring system is the same as for all other conditions (except the “Guesser present” condition).

Overall we performed 2224 valid trials in the first session, 2087 valid trials in the second session and 2015 in the third session. Time between sessions were 89.07 days on average between session one and two with a standard deviation of 11.35 days. There were 139.59 days between session two and three with a standard deviation of 22.77 days and 227.87 days between first and last session with a standard deviation of 19.94 days.

2.3. Statistics

We fitted one model including all three test conditions (excluding the Guesser Present condition) and one model for the data of each of the four conditions separately. All models included only the random intercept of subject ID, to account for repeated observations of the same subjects. Wald tests of the intercepts and 95% confidence intervals were used to draw inferences about performance being different from chance level. To assess test-retest reliability, we analysed whether the performance of dogs across test conditions (so excluding the Present condition) correlated between each pair of sessions, including only those dogs that had data in both sessions. We aggregated the data (proportion of trials in which the dogs chose the Knower) for each dog, condition and session and calculated pairwise-complete Pearson's correlations for each condition and combination of sessions.

3. Results

3.1. Correlation between sessions

To see if test-retest reliability was high, we investigated if preferences for the Knower were stable across sessions. The average Knower preference across all sessions was for the “Guesser Present” (GP) condition 56% (n=1583), for the “Guesser Absent” (GA) condition 55% (n=1582), for the “Guesser Back Turned” (GBT) condition 59% (n=1582) and for the “Guesser Looking Away” (GLA) condition 49% (n=1579).

In session 1 the Knower preference for the GP condition was 56% (n=557), for the GA condition 48% (n=556) for the GBT condition 61% (n=556) and for the GLA condition 50% (n=555). For the first trial in session 1, Knower preference was at 58% (n=93) for the GP condition, 52% (n=93) for the GA condition, 60% (n=93) for the GBT condition and 48% (n=93) for the GLA condition.

In session 2 the Knower preference for the GP condition was 59% (n=522), for the GA condition 57% (n=522) for the GBT condition 58% (n=522) and for the GLA condition 50% (n=521). For the first trial in session 2, Knower preference was at 69% (n=87) for the GP condition, 56% (n=87) for the GA condition, 67% (n=87) for the GBT condition and 44% (n=87) for the GLA condition.

In session 3 the Knower preference for the GP condition was 54% (n=504), for the GA condition 61% (n=504) for the GBT condition 56% (n=504) and for the GLA condition 48% (n=503). For the first trial in session 3, Knower preference was at 50% (n=84) for the GP condition, 64% (n=84) for the GA condition, 57% (n=84) for the GBT condition and 44% (n=84) for the GLA condition.

In the graphics below the consistency of Knower choices is shown. Every point represents one or more dogs with the colour indicating how many dogs overlap at a certain value. The diagonal line indicates where dogs would fall if they performed identically in the two sessions shown on the x- and y-axis. While some individuals showed strong preferences for Knower in both sessions, most dogs performed around chance level with a high inter-individual variability. For the comparison of the first and second session, the p-value was 0.001 which indicates a

significant correlation between dogs performances in session one and two. The r-correlation-value of 0.35 indicates a combined increase of variables.

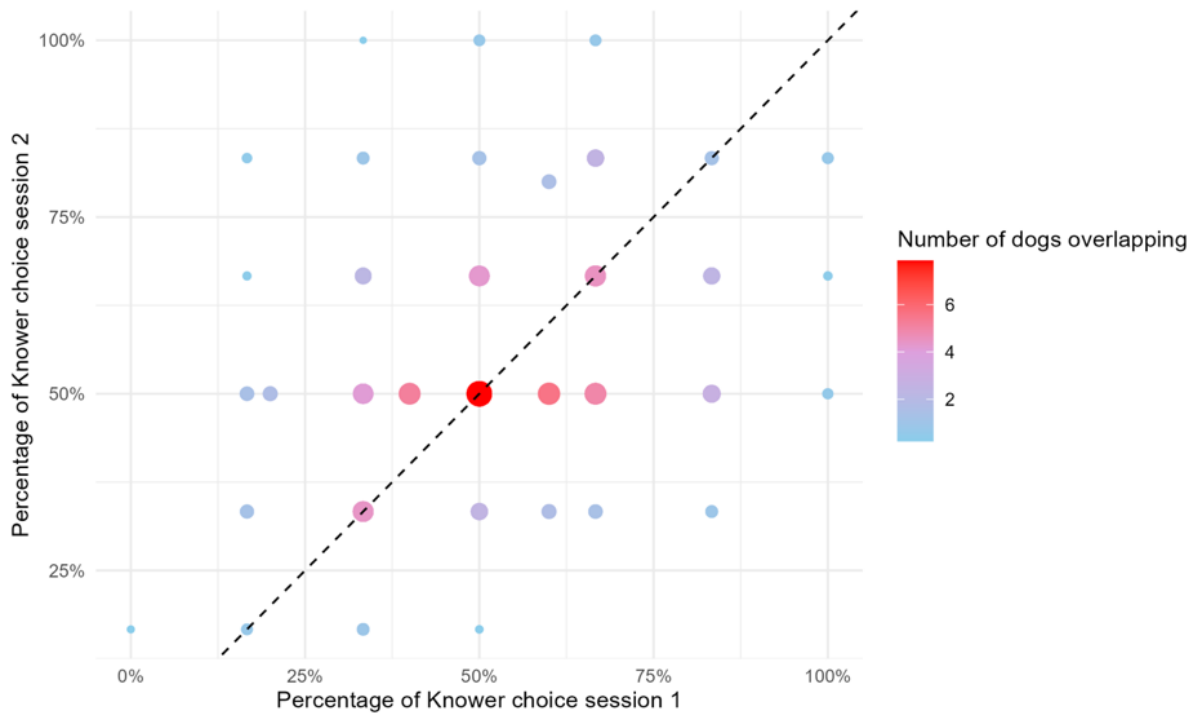


Figure 3: Density plot showing dogs' Knower preference between session 1 and 2 in each test condition (excluding Present condition). Within each session, each dog contributes to three dots (one per condition). The dots' size and colour represent the number of overlapping dogs with a given performance. The diagonal dashed line shows identical performance between sessions.

Although there was a significant correlation between the performances of session one and two, this was not true for the other sessions. For the comparison between the first and third session, the p-value was 0.662 and the r-correlation-value 0.05 which indicates no significant correlation between dogs' performances in session one and three.

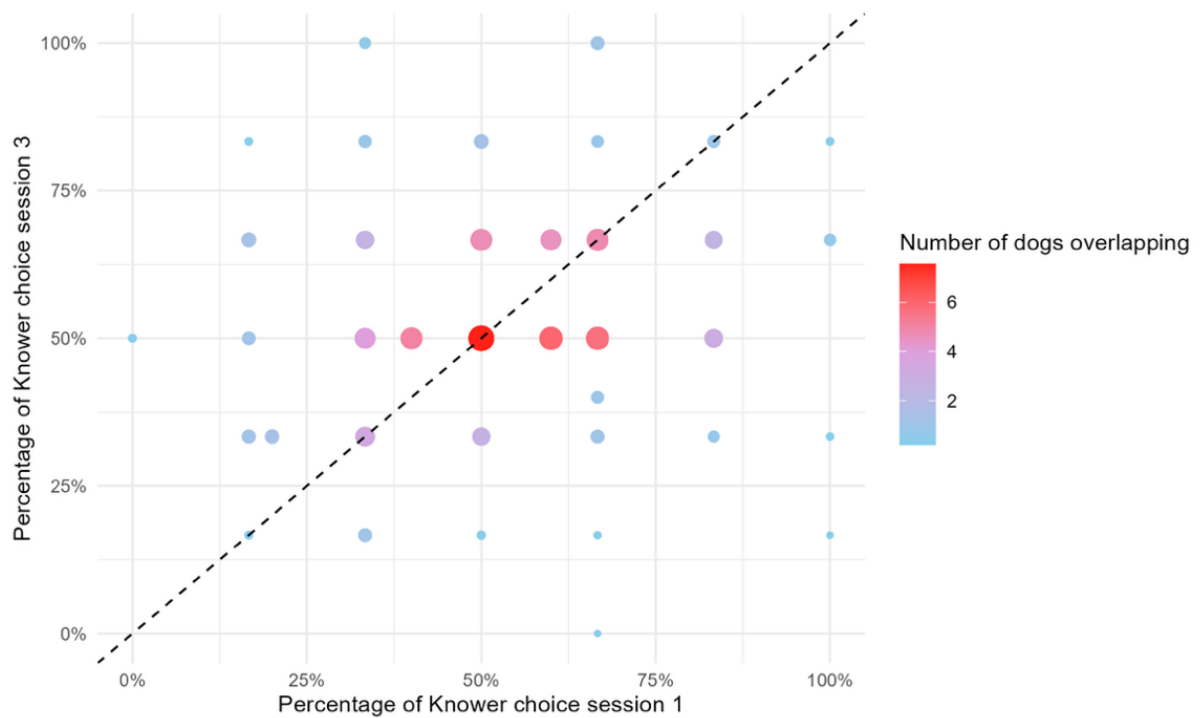


Figure 4: Density plot showing dogs' Knower preference between session 1 and 3 in each test condition (excluding Present condition). Within each session, each dog contributes to three dots (one per condition). The dots' size and colour represent the number of overlapping dogs with a given performance. The diagonal dashed line shows identical performance between sessions.

For the comparison between session two and three, the p-value was 0.929 which again indicates no significant correlation between dogs' performances in those two sessions. The r-correlation-value was -0.01.

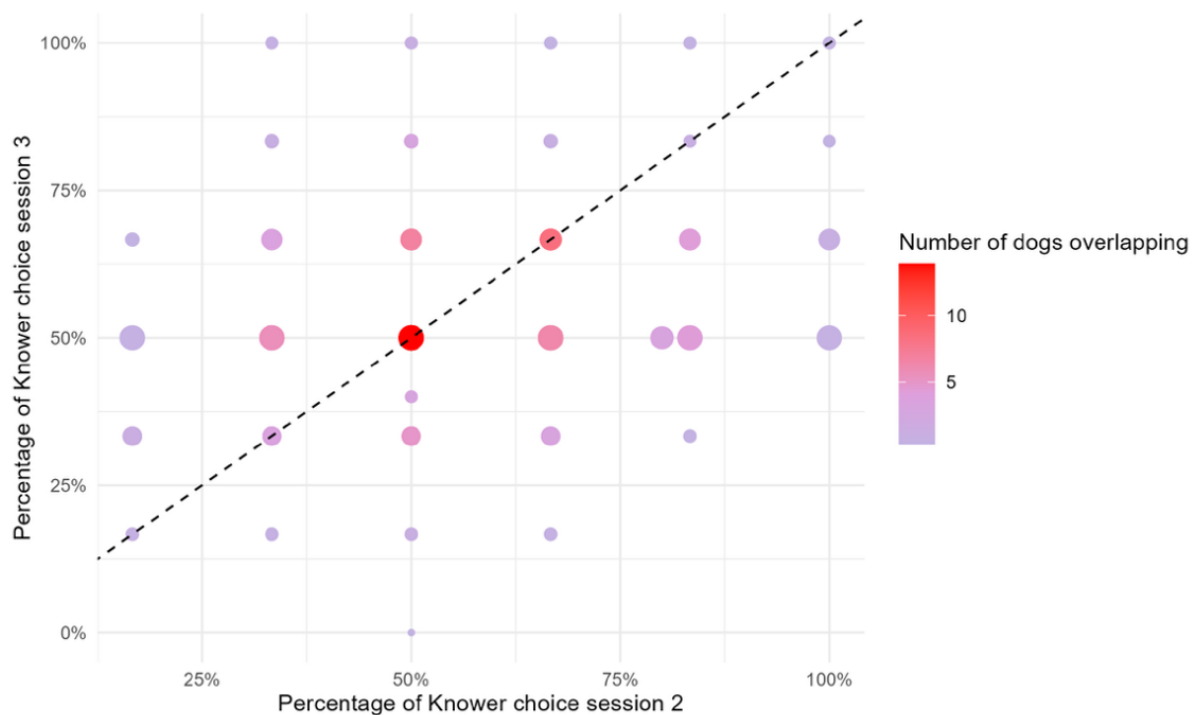


Figure 5: Density plot showing dogs' Knower preference between session 2 and 3 in each test condition (excluding Present condition). Within each session, each dog contributes to three dots (one per condition). The dots' size and colour represent the number of overlapping dogs with a given performance. The diagonal dashed line shows identical performance between sessions.

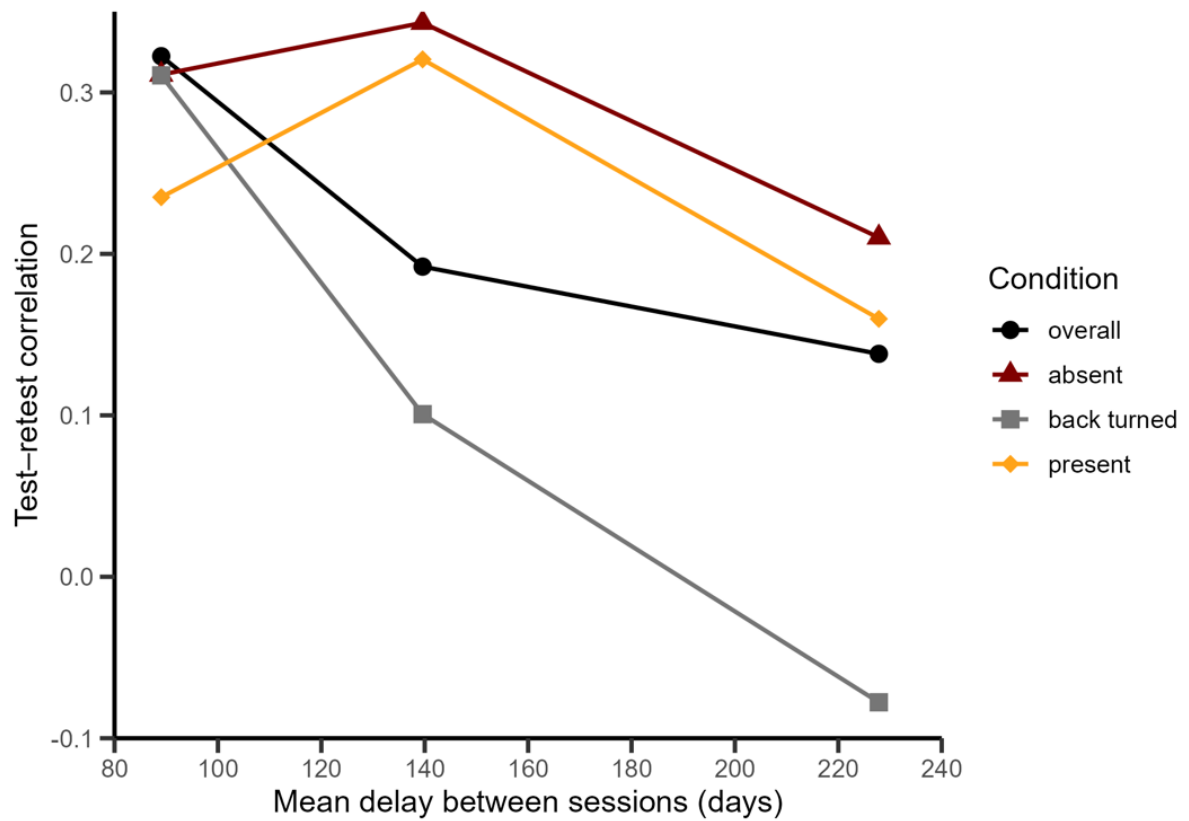


Figure 6: Test-retest correlations across time delays between sessions (in days). Points represent Pearson's correlations coefficients between session pairs (1-2, 2-3, 1-3). Lines connect correlations within each condition (except for the Guesser Looking Away condition that was excluded because dogs' performance did not deviate significantly from chance). Overall correlations (Guesser Absent, Guesser Present, Guesser Back Turned conditions) are shown in black.

3.2. Comparison with study of Catala et al. (4)

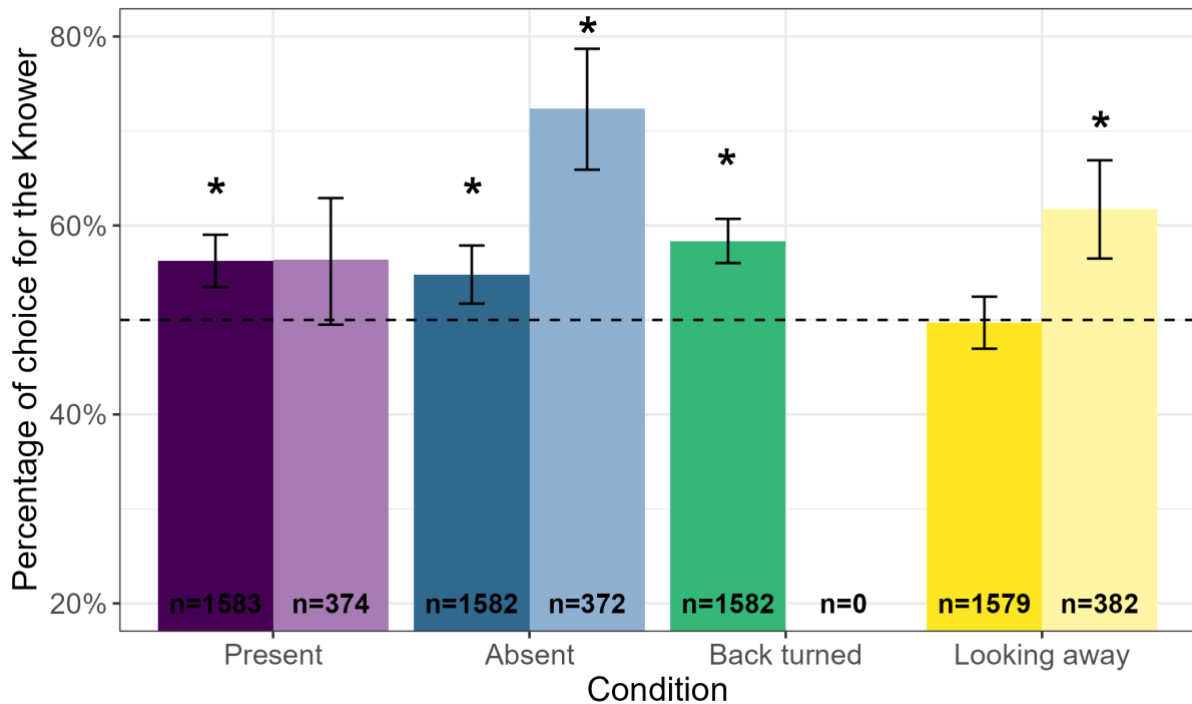


Figure 7: Grouped bar chart comparing the current study (darker colours) with the study of Catala et al. (3) (lighter colours)

Since our study is replicating the study of Catala et al. (4), we also compared our results with this study. For the “Guesser Present” condition we performed 1583 valid trials, 1582 trials for the “Guesser Absent” condition, 1582 valid trials for “Guesser Back Turned” and 1579 trials for the “Guesser Looking Away” condition. In comparison Catala et al. (4) performed 374 valid trials for the “Guesser Present” condition, 372 valid trials for the “Guesser Absent” condition and 382 valid trials for the “Guesser Looking Away”. As the “Guesser Back Turned” condition was newly introduced by us, there is no comparable value in the study by Catala et al. (4). Overall the average Knower preference in all conditions except the control condition was 54 % (95 % CI 53–56 %). In all conditions except “Guesser Looking Away” we found a Knower preference: in the “Guesser Present” condition 56 % (95 % CI 54–59 %, $p < 0.001$), in the “Guesser Absent” condition 55 % (95 % CI 52–58 %, $p = .001$), in the “Guesser Back Turned” condition 58% (95% CI 50–70%, $p < .001$). For the “Guesser Looking Away” condition the Knower preference was at 49 % (95 % CI 47–52 %, $p > .05$). While the study of Catala et al. (4) found similar results in the “Guesser Present” condition although with a smaller sample and therefore bigger confidence intervals, , and for the Guesser Absent (in both studies, dogs

performed above chance level) their results for the “Guesser Looking Away” condition differs from ours. For the “Guesser Looking Away” condition they found a significant Knower preference of 61.7 % while in our study there was no significant preference found (49%).

4. Discussion

4.1. Test-retest reliability between sessions

Interestingly, the only significant correlation between sessions was found between the first and second session. This means that dogs' who performed well in the first session overall had a tendency to performed well in the second session as well. This however, was not true for the third session, which indicates test-retest reliability across sessions was not stable. This also means that dogs did not show a learning effect during the experiment as there was no improvement seen for individual dogs across sessions. Reliability across sessions was low, with a significant correlation emerging only between the first and second sessions when pooling across test conditions. Importantly, the magnitude of the correlations decreased as the temporal distance between sessions increased (Figure 6), with no significant correlations involving the third session. In general, a decline of reliability as the delay between sessions increases is expected (20). However, the fact that in this case the overall correlation coefficients were small, and that their decline over time was pronounced, is inconsistent with the task capturing a durable cognitive ability and instead points to substantial state-dependent variance and sensitivity to contextual or procedural factors. Such methodological factors might also explain the lack of significant correlations involving the third session. Another changing factor over sessions were the experimenters. Due to the length of the study, there were changes in the experimenters since the experiments themselves were performed by exchange and diploma students who were part of the experiment only for a fixed time. While we tried to use the same experimenters for each appointment, this was not always possible. Another important factor was the number of dogs: based on a power analysis, we need at least 85 dogs in all session to be able to find a significant correlation between the sessions. Due to several reasons (see Table 1) we could only test 84 dogs in the third session. This means that there is also the possibility that the lack of significant correlation between session 1 and session 3 and session 2 and session 3, could also stem from a lack of statistical power.

Dogs' unstable performance could also result in general environmental factors, such as noises outside the room, weather and temperature. Due to the usual business at the Clever Dog Lab, we could not guarantee complete silence for the whole appointment, as there were other dogs waiting to be tested or tested and of course there was also often conversational voice or even

construction work. Some dogs showed increased noise sensitivity which might have influenced their focus and therefore their performance. Due to the length of the study, the sessions were held in different seasons and especially during the second session, which was performed in May, June and July, some dogs might have performed worse due to the warm weather. While we had air-conditioning in the testing room, the possibility remains that hot weather influences dogs' behaviour even when being in a cooler environment temporarily.

It would be interesting to replicate the experiment with even more attention to the conditions under which the experiment is held. For future studies it should be examined if the test-retest reliability is higher if experimenters do not change between sessions and maybe even if all experimenters are of the same gender, as some dogs seemed to prefer one gender over the other. It would also be advisable to test dogs in similar environmental conditions, such as weather, season and surrounding noise.

4.2. Comparison with the study of Catala et al. (4)

Our study not only replicated and extended some results of Catala et al. (4) but also provided further insight in dogs' performances over an extended timeline. The main difference between our study and the study of Catala et al. (4) was the sample size, with theirs consisting of 16 dogs and ours of 93 dogs. For the control condition, which was the "Guesser Present" condition, we found very similar results to the original study, supporting their findings. However, we did find a significant preference for the experimenter who performed the baiting, which suggests the influence of food handling cues.

For the "Guesser Absent" condition we could also find a preference for the Knower above chance level, just like in Catala et al. (4); however while in Catala et al. (4) the Knower preference was at 72.3 %, in our study it was only at 55 %. One reason for this slight difference in result might be a change in the set up. In the original study, for the "Guesser Absent" condition the Guesser left the room through a door behind the experimenters and thus in front of the dog. Since we could not access this door for our study as there was an ongoing experiment in the adjacent room to which this door lead, we had to use the door located behind the owner and dog. This often lead to dog focusing more on the experimenter leaving or entering (the

Guesser) than on the Knower. While we tried to keep the dogs' attention on the scene before him through rustling or knocking, the dogs often followed the hiding process attentively but seemed to be distracted when the Guesser entered again and turning around to see what was going on behind them. It would be interesting to perform a separate test to see if the position of the door through which the Guesser is leaving, has any impact on the dogs' performance or if this result is simply due to the larger sample size and represents dogs' performance in this condition more accurately for the general dog population. Another reason for finding lower performance in our study, compared to Catala, might have been the presentation of conditions in a pseudo-randomised instead of a blocked order. We pseudo-randomised the order of presentation of our conditions because the presentation of blocked conditions in Povinelli's seminal study with chimpanzees has been criticised as likely leading to learning effects (21).

The finding which differed the most from the study of Catala et al. (4) was our result in the "Guesser Looking Away" condition. While they found in this condition a significant Knower preference as well, we could not replicate this result. With a Knower preference of 49 %, dogs did not perform above chance level in this condition only. It is possible that due to the larger sample size, the general dog population was better represented and therefore the majority of dogs are not able to follow another individual's gaze geometrically. Another aspect which increased the difficulty of this condition compared to the other conditions in our study was the removal of a food handling bias as the Hider performed the baiting in this condition. This meant that dogs could not rely on food handling cues for their decision which experimenter's pointing gesture to follow. So, our results suggest that the food handling does have a significant influence on dogs' decisions, which speaks against their ability to take another individual's perspective and rather for their observational skills. As our study aimed to be replicating the study of Catala et al. (4) to check construct validity and test-retest reliability, we did not check for food handling bias in all conditions. However, it would be interesting to conduct further studies with a specific focus on the food handling bias. It would be interesting if the Knower preference for the "Guesser Present", "Guesser Absent" and the "Guesser Back Turned" condition changes depending on if the food was hidden by a third experimenter rather than by the Knower. If Knower preference changes compared to this study, the difference in results could be explained by a food handling bias.

4.3. Conclusion

Overall we found that individual differences in dogs' performances between sessions were not stable and while we could not identify a learning effect, we could also only partly find a correlation between performances at all. Only between the first and second session a significant correlation was found. This means that if dogs performed well in the first session, they generally performed well in the second session too but not necessarily in the third session. The magnitude of the correlation coefficients between sessions declined steeply with increasing time delay between sessions.

In comparison to the study of Catala et al. (4) we could only partly replicate their results. While we could support their result of a Knower preference in the "Guesser Present", "Guesser Absent" and "Guesser Back Turned" condition, this was not true for the "Guesser Looking Away" condition. These results strengthen the argument that dogs are sensitive to humans' attentional state and pointing gestures (2,3) but do not support the possibility that dogs are able to follow gazes geometrically. Since we also found a significant preference for the experimenter handling the food in the "Guesser Present" condition, it might be possible that food handling bias plays a bigger role in the decision-making process of dogs than originally anticipated.

Overall, this study has also shown the importance of replicating cognitive studies with animals with larger sample sizes in order to avoid coming to conclusions which might not represent the overall population of animals. While it is of course not always possible to perform studies of this sample size, it is advisable to replicate studies to test their reliability and further strengthen conclusions regarding animal cognition.

5. Bibliography

1. Premack D, Woodruff G. Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences* 4; 515–526; 1978. doi:10.1017/S0140525X00076512
2. Hare B, Tomasello M. Domestic Dogs (*Canis familiaris*) Use Human and Conspecific Social Cues to Locate Hidden Food. *Journal of Comparative Psychology* Vol. 133, No. 2; 173–177; 1999.
3. Gácsi, M, Miklósi, Á, Varga, O, Topál, J, Csányi, V. Are readers of our face readers of our minds? Dogs (*Canis familiaris*) show situation-dependent recognition of human's attention. *Animal Cognition* 7; 144–153; 2003. doi:10.1007/s10071-003-0205-8
4. Catala A, Mang B, Wallis L, Huber L. Dogs demonstrate perspective taking based on geometrical gaze following in a Guesser-Knower task. *Animal Cognition* 20; 581–589; 2017. doi:10.1007/s10071-017-1082-x
5. Krupenye, C. Guesser-Knower-Paradigm. *Encyclopedia of Animal Cognition and Behavior*; 2017. doi:10.1007/978-3-319-47829-6_1605-1
6. Lea SEG, Osthaus B. In what sense are dogs special? Canine cognition in comparative context. *Learn Behav.* 1. Dezember 2018;46(4):335–63. doi:10.3758/s13420-018-0349-7
7. Range F, Marshall-Pescini S. Comparing wolves and dogs: current status and implications for human 'self-domestication'. *Trends Cogn Sci.* 1. April 2022;26(4):337–49. doi:10.1016/j.tics.2022.01.003
8. Whiten A. When does smart behaviour-reading become mind-reading? *Theories of Theories of Mind*; 277–292; 1996.
9. Call J, Bräuer J, Kaminski J, Tomasello M. Domestic Dogs (*Canis familiaris*) Are Sensitive to the Attentional State of Humans. *Journal of Comparative Psychology* Vol. 117, No. 3; 257–263; 2003. doi:10.1037/0735-7036.117.3.257
10. Schwab C, Huber L. Obey or Not Obey? Dogs (*Canis familiaris*) Behave Differently in Response to Attentional States of Their Owners. *Journal of Comparative Psychology* Vol. 120, No; 169–175.; 2006. doi:10.1037/0735-7036.120.3.169
11. Horowitz A. Theory of mind in dogs? Examining method and concept. *Learning Behaviour* 39; 314–317.; 2011. doi:10.3758/s13420-011-0041-7
12. Bräuer J, Call J, Tomasello M. Visual perspective taking in dogs (*Canis familiaris*) in the presence of barriers. *Applied Animal Behaviour Science* 88; 299–317.; 2004. doi:10.1016/j.applanim.2004.03.004
13. Kaminski J, Bräuer J, Call J, Tomasello M. Domestic dogs are sensitive to a human's perspective. *Behaviour* 146; 979–998; 2008. doi:10.1163/156853908X395530

14. Povinelli D, Boysen S, Nelson K. Inferences About Guessing and Knowing by Chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* Vol. 104, No. 3; 203–210.; 1990. doi:10.1037/0735-7036.104.3.203
15. Maginnity G, Grace R. Visual perspective taking by dogs (*Canis familiaris*) in a Guesser-Knower task: evidence for a canine theory of mind? *Animal Cognition* 17(6); 1375–92; 2014. doi:10.1007/s10071-014-0773-9
16. Kuroshima H, Fujita K, Fuyuki A, Masuda T. Understanding of the relationship between seeing and knowing by tufted capuchin monkeys (*Cebus apella*). *Animal Cognition* 5; 41–48.; 2002. doi:10.1007/s10071-001-0123-6
17. Udell M, Dorey N, Wynne C. Can your dog read your mind? Understanding the causes of canine perspective taking. *Learning Behaviour* 39; 289–302.; 2011. doi:10.3758/s13420-011-0034-6
18. Aldridge VK, Dovey TM, Wade A. Assessing Test-Retest Reliability of Psychological Measures. *Eur Psychol.* 1. Oktober 2017;22(4):207–18. doi:10.1027/1016-9040/a000298
19. Leppink J, Pérez-Fuster P. We need more replication research – A case for test-retest reliability. *Perspect Med Educ.* 1. Juni 2017;6(3):158–64. doi:10.1007/s40037-017-0347-z
20. Bohn M, Eckert J, Hanus D, Lugauer B, Holtmann J, Haun DBM. Great ape cognition is structured by stable cognitive abilities and predicted by developmental conditions. *Nat Ecol Evol.* 1. Juni 2023;7(6):927–38. doi:10.1038/s41559-023-02050-8
21. Heyes, C. Social Cognition in Primates. *Animal Learning and Cognition*; 281–305; 1994. doi:10.1016/B978-0-08-057169-0.50016-4

6. List of Tables

Table 1: List of excluded dogs.....	14
-------------------------------------	----

7. List of Figures

Figure 1: Set up. Blindfolded Owner (O) sitting behind dog (D), facing the barrier (B) with the two informants (I1, I2) behind it. Hider (H) sits behind experimenters facing the Owner and dog. Cups (C1-C4) placed equidistant from dog's position.	17
Figure 2: a) Guesser Present; b) Guesser Absent; c) Guesser Back Turned; d) Guesser Looking Away; e) Pointing at respective cups.....	19
Figure 3: Density plot showing dogs' Knower preference between session 1 and 2 in each test condition (excluding Present condition). Within each session, each dog contributes to three dots (one per condition). The dots' size and colour represent the number of overlapping dogs with a given performance. The diagonal dashed line shows identical performance between sessions.	24
Figure 4: Density plot showing dogs' Knower preference between session 1 and 3 in each test condition (excluding Present condition). Within each session, each dog contributes to three dots (one per condition). The dots' size and colour represent the number of overlapping dogs with a given performance. The diagonal dashed line shows identical performance between sessions.	25
Figure 5: Density plot showing dogs' Knower preference between session 2 and 3 in each test condition (excluding Present condition). Within each session, each dog contributes to three dots (one per condition). The dots' size and colour represent the number of overlapping dogs with a given performance. The diagonal dashed line shows identical performance between sessions.	26
Figure 6: Test–retest correlations across time delays between sessions (in days). Points represent Pearson's correlations coefficients between session pairs (1–2, 2–3, 1–3). Lines connect correlations within each condition (except for the Guesser Looking Away condition that was excluded because dogs' performance did not deviate significantly from chance). Overall correlations (Guesser Absent, Guesser Present, Guesser Back Turned conditions) are shown in black.	27
Figure 7: Grouped bar chart comparing the current study (darker colours) with the study of Catala et al. (3) (lighter colours)	28

8. Acknowledgements

First I want to thank my supervisors Lucrezia Lonardo, PhD. and Univ.-Prof. Dr.rer.nat. Ludwig Huber for making this work possible and for guiding me through the process of experimenting and writing this thesis. I also want to thank Univ.-Prof. Jean-Loup Rault, PhD. for taking the time to review my thesis.

A special thanks goes to all the owners and their amazing dogs for participating in this experiment and of course to my colleagues who performed the experiment with me, without whom it would not have been possible to do this.

Furthermore I want to thank my parents for being my biggest supporters in my academic studies and my life and my friends for always being there for me and supporting me.