universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Do dogs coordinate with a conspecific partner in the Assurance Game?"

verfasst von / submitted by

Selina Vanessa Schöndorfer, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021 / Vienna, 2021

Table of Contents

Acknowledgements:

I want to thank my supervisor Friederike Range and my mentor and co-supervisor Mayte Martinez, who always had an open ear and put a lot of effort into supporting me during this thesis. I also want to thank Karin Bayer and the Clever Dog Lab Team for welcoming me at their lab and creating a wonderful work atmosphere. Last but not least, I am very thankful for my dog owners, who trusted me with their dogs, came to the CDL reliably during a worldwide pandemic and who made conducting this study a true pleasure.

Abstract:

Economic games have been useful to study decision making during cooperation in humans and animals. Traditionally, focus lay on the outcome and not on the animals' understanding of these games. In this study, we tested 11 pet dog dyads in the Assurance Game using string-pulling tables. The game has two choices: Hare (low-value reward; achieved alone) or Stag (high-value reward; achieved cooperatively). We assessed if the dogs considered the choice of their partner when making their own choice. If they understood the game, they should coordinate on Stag to obtain the high-value reward. Then we tested if higher cooperation in the game increased the level of tolerance within the dyad after the game. This was expected if the dogs understood that the results of the game were influenced by their partners actions. Lastly, we explored the individual strategies of every dyad in the game. Our results showed that the partner's choice influenced the subject's choice in the Assurance Game. Three of our dyads reliably coordinated on the high payoff, while side bias and local enhancement might account for the matching of choices in the remaining dyads. Further, our dogs did not change their levels of tolerance depending on the level of cooperation. Therefore, despite the dogs being able to match their choices, they might not be aware of the role of their partners' choices in achieving their common goal. Our research opens up possibilities for future follow-up studies and reinterpretation of previous studies with primates.

Zusammenfassung:

Ökonomische Spiele haben sich als nützlich erwiesen, um Entscheidungsfindungen in Tieren und Menschen während Kooperationen zu untersuchen. Traditionell lag der Fokus dabei auf dem Ergebnis und nicht auf dem Verständnis der Tiere für das Spiel. In dieser Studie benutzten wir Seilzieh-Tische um 11 Haushund-Dyaden im Assurance Game zu testen. Dabei gibt es zwei Wahlmöglichkeiten: Hase (Belohnung mit niedrigem Wert; allein erreichbar) und Hirsch (Belohnung mit hohem Wert; kooperativ erreichbar). Wir untersuchten, ob die Hunde die Wahl ihres Partners bei ihrer eigenen Wahl berücksichtigten. Wenn sie das Spiel verstanden haben, sollten sie gemeinsam Hirsch wählen und so die hochwertige Belohnung erhalten. Außerdem testeten wir, ob mehr Kooperation im Spiel das Toleranzniveau innerhalb der Dyade nach dem Spiel erhöhte. Dies war zu erwarten, wenn die Hunde verstanden, dass die Ergebnisse des Spiels von den Handlungen ihres Partners beeinflusst wurden. Schließlich begutachteten wir die individuellen Strategien der Dyaden während des Spiels. Unsere Ergebnisse zeigten, dass die Wahl des Partners einen Einfluss auf die Wahl des Subjekts im Assurance Game hatte. Drei der Dyaden koordinierten zuverlässig mit der Wahl Hirsch, während bei den übrigen Dyaden eine Seitenpräferenz und lokale Verstärkung für übereinstimmende Entscheidungen verantwortlich sein könnten. Außerdem fanden wir keine Veränderung des Toleranzniveaus in Abhängigkeit vom Grad der Kooperation. Obwohl die Hunde in der Lage waren, ihre Wahl im Spiel aneinander anzupassen, waren sie sich der Rolle des Partners beim Erreichen des gemeinsamen Ziels möglicherweise nicht bewusst. Unsere Forschung eröffnet Möglichkeiten für zukünftige Folgestudien und regt zur Neuinterpretation früherer Studien mit Primaten an.

1. Background of the study:

All animals must make decisions about how to interact with their biotic environment (McFarland, 1977): they need to decide when to fight and when to flee, when to feed, sleep or mate. For example, individuals may change their foraging strategy depending on the presence of a predator (e.g., Mitchell & Lima, 2002). Decision making is especially important in the context of cooperation, a form of interaction between two or more individuals that usually results in a higher payoff by working together than working alone (Noë, 2006). Social animals constantly face the challenge of deciding whether to cooperate with other individuals to obtain a common reward, or, on the contrary, defect and act alone (Smith et al. 2019).

From the evolutionary point of view, cooperation is no longer a mystery. Different theories shed light on how cooperation can be beneficial for the fitness of the individual (i.e., through kin selection or reciprocal altruism, Hamilton, 1964; Trivers, 1971). However, we know very little about the cognitive abilities underlying the decisions that lead to cooperation (Brosnan et al., 2011). Yet, cognitively complex cooperation might enable species to interact more flexibly. It has been proposed that cognitively complex cooperative abilities are behind human's extraordinary ability to flexibly cooperate with large groups. Specifically, humans understand the role their cooperative partners play in the achievement of the goal (Tomasello et al., 2005; Duguid & Melis, 2020). Whether and to which extent other non-human animals understand the role of their partners is still an open question. One of the first efforts towards an answer was made by Boesch and Boesch (1989). They defined four levels of cooperation according to the subject's degree of understanding of their partner's role: similarity, synchrony, coordination, and cooperation. *Similarity* describes individuals acting similarly but without any active coordination between them; *synchrony* requires temporal coordination (i.e., individuals adjust the speed of their actions to each other); *coordination* demands tempo-spatial coordination (i.e., individuals adjust the speed and the spatial orientation of their actions); and *collaboration* requires individuals to take up complementary roles in a flexible manner. When two individuals collaborate, they are supposed to understand not only that they need a partner to act in a certain way, but also the specific role of the partner and its importance for the achievement of the common goal. Observational studies of hunting behavior suggest that chimpanzees (Boesch & Boesch, 1989; Boesch, 2002) and some coral reef fishes (Vail et al., 2005) are able to flexibly adopt different roles (i.e., collaboration)

when pursuing prey. However, when presented with more controlled experimental setups, the results are often mixed.

The most widely used experimental paradigm to study whether non-human animals understand the role of their partner is the cooperative pulling table (Crawford, 1937; Hirata & Fuwa, 2007), in which individuals need to work together by simultaneously pulling two ropes (or handles) to gain access to food. To test the extent to which subjects understand the role of their partners, researchers introduced an experimental condition in which one of the animals is delayed (Melis et al., 2006; Hirata & Fuwa, 2007). Studies using this paradigm showed that some animals (e.g., elephants: Plotnik et al., 2011; chimpanzees: Melis et al., 2006; wolves: Marshall-Pescini et al., 2017), but not others (e.g., otters: Schmelz et al., 2017; parrots: Péron et al., 2011; Tassin de Montaigu et al., 2020), would wait for their delayed partner before pulling the rope. While the basic cooperative pulling table can be solved by each animal repeatedly pulling and succeeding by chance (i.e., similarity), the delay condition seems to show that animals are aware that they need a partner to perform the same action as them at the same time and place (i.e., coordination) to solve the task. However, it is possible that they are not paying attention to their partner, but instead following non-social cues (e.g., the movement of the rope caused by the partner pulling on the other side). Also, failing the delay task is no proof for a lack of understanding, as it could simply show bad inhibitory control (Albiach-Serrano, 2015). Even if we disregard all these points, the delay task would not be sufficient to measure collaboration, as there is no room for flexibility or different roles. To account for this, researchers have used role reversal tasks, in which two individuals are trained to perform complementary actions to obtain a reward. After successful training, the roles of the individuals are reversed without any additional training (Mason & Hollis, 1962). Animals successfully performing the complementary role would indicate that they were paying attention to the actions of their partner and an understanding of their importance for the cooperative tasks. This paradigm has been used only with primates, again leading to mixed results (Mason & Hollis, 1962; Povinelli et al., 1992a; Povinelli et al., 1992b). Similar to the cooperative pulling table, one disadvantage of the role reversal paradigm is its reliance on the physical understanding of the apparatus. This, together with the fact that some species might not be as skilled as others in the context of object manipulation, makes it difficult to interpret the results in a comparative way between species. Additionally, in the mentioned string-pulling and role reversal tasks, the subjects lack a variety of behavioral options. Animals are presented with one apparatus and the choice to manipulate it (cooperate) to obtain a reward or do nothing. To better evaluate decision making in the context of

cooperation, we need tasks that provide the animals with an alternative to cooperation that will also result in a reward, thus testing for the subjects' propensity to cooperate.

An approach that can help to investigate decision making and cooperation is economic games (Brosnan et al., 2011). Experimental economics represents complex social situations in the form of a simple two-choice task (Jensen et al., 2007; Smith et al., 2019). The way in which these choices are presented can be adapted to different species (e.g., selecting an icon on the screen: Brosnan et al., 2012, choosing one of two tokens: Brosnan et al., 2011). One of these economic games is the Stag-Hunt or Assurance Game (Skyrms, 2004). This is a two-player game in which two individuals have the choice to go for a low value reward (Hare, LVR, e.g., 1 unit of reward), that can be achieved alone, or to go for a high value reward, (Stag, HVR; e.g., 3 units of reward) that can only be achieved cooperatively. In the Assurance Game, Stag is both the high reward but also high-risk option: If individual one was to choose Stag but individual two chose Hare, individual one would get no reward (i.e., 0 units of reward). Therefore, the game has two Nash-Equilibria (i.e., a strategy that maximizes the payoff for both partners regarding the strategy of the other). Both partners choosing Stag would be the payoff-dominant Nash-Equilibrium, while both choosing Hare would be the risk-dominant Nash-Equilibrium. The temptation to defect here is very low, as both individuals' interests should be aligned on getting the high reward (i.e., choosing Stag) (Parrish et al., 2014). This is useful because it helps to test for basic coordination abilities while minimizing possible understanding problems. All of this makes the Assurance Game an ideal game to explore cooperative coordination. If the subjects understand the game, they should monitor their partners' behavior and adjust their own behavior accordingly.

Previous research has shown that different primate species tend to coordinate with their partners in the Assurance Game after some training (see review in Massen et al, 2019), and many of them do so by coordinating in the payoff-dominant choice (Sag-Stag). However, the strategies behind the choices differ between species (Smith et al., 2019). For example, chimpanzees sometimes show a high degree of coordination (Brosnan et al., 2011; Bullinger et al., 2011; Duguid et al., 2014), while in other studies they seem to play randomly (Hall et al., 2019). Rhesus macaques and squirrel monkeys develop an individual bias for the stag choice over trials (Brosnan et al., 2012; Parrish et al., 2014; Vale et al., 2019), while capuchin monkeys play the Assurance Game by matching their partners' choices. However, capuchin monkeys only do so if they can see what their partner plays (Brosnan et al., 2011; 2012; Smith et al., 2019).

It is not possible to infer the cognitive skills behind those results by solely inspecting the animals' outcomes in the game. It still remains unclear if the tested animals actually understood that their partner was also making decisions that had consequences in the game. Devising methods to investigate the level of understanding of the partner's role during cooperation can be challenging (Brosnan et al., 2010). One possibility to explore that question would be to add behavioral tests measuring whether animals change their attitudes (e.g., social tolerance) towards their partners depending on their performance in the game. This approach was successfully used by Brucks et al. (2016) in an inequity aversion test, showing that dogs co-fed less and showed reduced proximity to the experimenter after receiving unequal payoffs for the same action. This effect was present even if during the test the dogs did not change their behavior between conditions (i.e., they continued giving the paw after the experimenter's command). That means that the dogs somehow associated what happened during the task with their partners in the task.

Additionally, all the species tested in the Assurance Game until now are primates. Non-primate species should be tested to better clarify how animals coordinate with each other across different taxa, as this helps to shed light on the origins and adaptations of the cognitive abilities used during cooperation.

A species that has recently drawn the attention of researchers in the field of cooperation is the domestic dog. Dogs (*Canis lupus familiaris)* are well known for their ability to cooperate with humans (e.g., Naderi et al., 2001; Ruusila & Pesonen, 2004; Koster 2008). Further, they have been shown to adjust their behavior according to the context-specific behavior of their human partners (Horn et al., 2012). It has been suggested that dogs' cooperative skills have evolved due to the close contact with humans. That is, during domestication dogs have become more socially tolerant towards humans and conspecifics than wolves (Domestication hypothesis, Hare et al., 2012). In contrast to this, another hypothesis proposes that intraspecific cooperation in wolves has been the basis for the evolution of cooperation between humans and domestic dogs, making additional selection for social tolerance in dogs during domestication unnecessary (i.e., Canine Cooperation hypothesis: Range & Virányi, 2015). Yet another hypothesis revolves around the socioecology of dogs. While wolves rely on cooperative hunting, dogs usually forage alone on human waste. This would suggest that dogs have become less good at cooperating with each other than wolves (i.e., Socioecology hypothesis, Marshall-Pescini et al., 2017).

Experimental studies on conspecific cooperation in dogs have provided mixed results in the past. Some claimed to demonstrate coordination between dog-dog days (Bräuer et al., 2013; 2020), but could not provide any evidence that the individuals actually adjusted their behavior to each other. Using the string-pulling paradigm, it has been shown that pack living dogs seem to understand the need for a partner when that partner is a human (Range et al., 2019a; 2019b), but they perform very poorly when they are paired with conspecifics. This is presumably because of tolerance problems arising when two animals need to manipulate the apparatus at the same time (Marshall-Pescini et al. 2017; 2018). Conversely, in a similar study Ostojić and Clayton (2014) showed that after some training, dog dyads from the same household could coordinate their actions to access a reward and even wait for their partner to access the apparatus for a short time. In this study pet dogs waited for their conspecific partner for an average of 2.2 seconds, which is arguably a very short delay. Furthermore, differences in tolerance might explain the differences between those studies. While the first studies by Marshall-Pescini et al. (2017; 2018) tested pack living dogs and controlled for human influence in the setup, in Ostojić and Clayton (2014) the sample consisted of highly trained pet dogs with their owner present during the tests. This suggests that, when tolerance problems are controlled, dogs are, at least, able to coordinate their actions at the same apparatus. Further research is needed to explore to which extent this coordination is flexible and could be maintained when the subjects can choose between cooperative and individual behaviors, as well as whether dogs understand that the actions of their partners are relevant for the cooperative outcome.

The first aim of this study was to test whether pet dogs are able to coordinate their behavior with a conspecific partner when making a choice in a cooperative task, the Assurance or Stag-Hunt Game. We expected the dogs to coordinate their actions, because there was no payoff in defecting, and we expected them to preferentially choose the option with the highest payoff for both (Stag). The second aim of this study was to investigate whether the dogs understood that the partner was also making decisions in the Assurance Game. Therefore, we measured the level of social tolerance between two individuals before and after participating in the task. If the dogs understood that their partner was making decisions that had consequences for the cooperative outcome, the level of social tolerance should increase after cooperation in the Assurance Game. As tolerance is a factor that can facilitate cooperation, we further expected the dyads that were more tolerant towards each other to coordinate more on Stag than the ones that were less tolerant towards each other.

2. Methods:

2.1. Ethical approval:

This study was discussed and approved by the "Ethik und Tierschutzkomission" of the University of Veterinary Medicine Vienna (Approval number: 142/07/2019) and owners signed a consent form prior to participation. All procedures were non-invasive.

2.2. Subjects:

The test subjects were 22 medium to large sized pet dogs (13 Females, 9 males, mean age = 4.8 years , age range = 1-11 years, see Table 1, Supplementary S1). The subjects were food motivated and showed no signs of extreme resource guarding towards humans or conspecifics. This prerequisite was chosen to avoid conflicts over food during the test. The study took place indoors at the Clever Dog Lab (CDL) at the University of Veterinary Medicine in Vienna, Austria. The subjects were tested in 11 dyads, and to mitigate social stress due to any tolerance issues, dyads were made up of dogs that had been living in the same household for at least 6 months prior to the experiment. One of the dogs was assigned the role of the subject, while the other was the partner (see 2.3. Experimental Design: Assurance Game). This stayed constant within the dyad. Dog owners were recruited via email, using the CDL's database, or through a flyer posted on the CDL's website and social media sites. The owners had to visit the CDL up to ten times in total, once a week if possible. However, due to several lockdowns because of the Corona Virus, where testing at the CDL was stopped, participation in the study stretched over several months (from 2 to 8 months).

Table 1: Subject List

Dyad	ID	Role	Age	Sex	Breed	LVR	HVR	Dog Stooge
1	Annie	Partner	8	F	Shetland Sheepdog	dry Food	2 sausages	yes
	Prim	Subject	5	F	Shetland Sheepdog	dry Food	2 sausages	yes
2	Benny	Partner	7	M	Border Collie	dry Food	2 sausages	yes
	Dawin	Subject	5	M	Border Collie	carrot	sausage	yes
3	Crash	Partner	2	M	Australian Sheperd	zuchhini	2 cheese	yes
	Alou	Subject	4	F	Australian Sheperd	salad	sausage	no
4	Gandalf	Partner	1	M	Andalusian Hound	dry Food	sausage	yes
	Lenny	Subject	6	M	Canarian Warren Hound	zuchhini	sausage	yes
5	Helsa	Partner	2	F	Hovawart	dry Food	sausage	yes
	Dora	Subject	9	F	Hovawart	dry Food	2 sausages	yes
6	Josephine	Partner	4	F	Pumi	dry Food	sausage	yes
	Joseph	Subject	3	M	Pumi-Mix	dry Food	sausage	yes
7	Leia	Partner	10	F	Beagle	dry Food	2 sausages	yes
	Eowyn	Subject	3	F	Mix	dry Food	sausage	no
8	Matteo	Partner	1	M	Segugio Italiano	dry Food	sausage	yes
	Hetti	Subject	2	F	Mix	dry Food	sausage	no
9	Mavie	Partner	2	F	Border Collie	sausage	dry Food	yes
	Cheynna	Subject	7	F	Australian Sheperd	dry Food	sausage	no
10	Miley	Partner	11	F	Border Collie	dry Food	2 sausages	yes
	Tiara	Subject	7	F	Border Collie	zuchhini	2 sausages	yes
11	Timo	Partner	2	M	Mix	dry Food	sausage	yes
	Mailo	Subject	5	M	Mix	dry Food	sausage	yes

2.3. Experimental Design:

General design:

We started the study with a food preference test for each dog of the dyad, where we established a high value (HVR) and low value (LVR) reward to be used in the rest of the experiment. After that, the dyads were trained to operate tables representing the choice Hare (Hare training with a string-pulling table they could operate alone) and Stag (Stag training with a cooperative string-pulling table) in the Assurance Game. In the Stag training, they also learned to wait for a human and conspecific partner. During three of the training days, we conducted tolerance tests. In the final training stage, we exposed the dogs to the different reward qualities associated with the Hare (i.e., LVR) and Stag (i.e., HVR) choice. Then the dyads were tested in the Assurance Game on two different days, with a tolerance test after each Assurance Game.

Food preference test:

To represent the Hare and Stag choice of the Assurance Game, we used rewards of different quality. To determine a high (HVR) and low value food (LVR) reward for each individual, they were tested in a food preference test (see Brucks et al., 2016). The test consisted of 12

trials in which the dog had to choose between a proposed low value or high value food type. We always started with dry food as a LVR-candidate and sausage as an HVR-candidate. One piece of dry food and one piece of equally sized sausage were placed on different container lids. At the beginning of each trial, the dog was held by the owner and allowed to sniff the food. Afterwards, the lids were placed on the ground one meter away from the dog and approximately 50 cm apart from each other. Then, the dog was released and allowed to eat the food off of one lid only. To avoid any side bias effect, the side of the LVR and the HVR were exchanged after each trial. A type of food was considered the HVR if the dog chose it in 9 out of 12 trials. Otherwise, we asked the owners what type of food their dog liked less than dry food but would still work for. Then we redid the test with the suggested different types or quantities of food (e.g., one piece of dry food and three pieces of sausage; carrot and cheese). The chosen LVR and HVR of each dog were not the same within the dyad if they had different preferences (for the LVR and HVR of each dog see Table 1).

Assurance Game:

Experimental Set-up:

To represent the Hare and Stag choices, two types of sliding tables were used. The solo apparatus (i.e., Hare; apparatus that could be solved alone) was a table with a sliding platform (see Figure 1b). Attached to the platform was a small wooden square that contained food rewards (i.e., LVR), and a metal ring that had a rope attached to it. One dog alone could operate this apparatus by pulling at the rope, which caused the platform to slide out and made the reward accessible to them. The cooperative apparatus (i.e., Stag) was a bigger table with a sliding platform, similar to the solo apparatus (see Figure 1c). It was placed so that half of it was on the left and the other half on the right side of the barrier. There were two wooden squares for rewards (i.e., HVR), one on the left side of the platform and one on the right. Also, there were two metal rings on the platform and the rope was threaded through it in a way that each dog could access one end of the rope at their side of the barrier. Pulling simultaneously at the ends of the rope by the partners made the platform slide out and brought the reward into reach for both. However, pulling by only one dog at one end of the rope, caused the rope to slide out of the metal rings. Therefore, the platform could no longer be pulled out and the reward became inaccessible to both members of the dyad (see Figure 1, Figure 2).
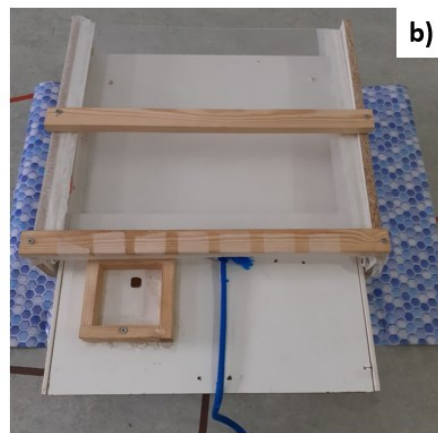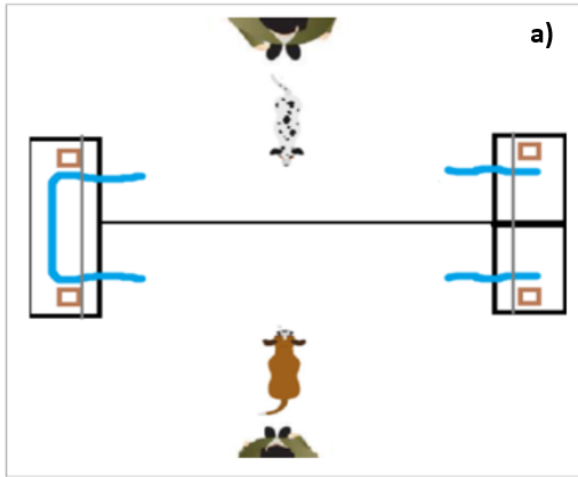
Figure 1: a) Experimental setup with the solo apparatus (Hare) at the bottom and the cooperative apparatus (Stag) at the top. b) Solo apparatus (Hare) baited with food in the unsolved (left) and solved (right) state. c) Cooperative apparatus (Stag) baited with food in the unsolved (left) and solved (right) state.

The dog dyads were in the same room, separated by a see-through fence, and each had access to both types of tables (see Figure 1a). Therefore, each dog was able to decide between choosing Hare or Stag. To make it easier for the subjects to discriminate between the different apparatuses, they were placed on either a white or blue carpet, each table on a different carpet

(counterbalanced between dyads). One member of the dyad was chosen as the partner (i.e., the individual that was released first), while the other one was the subject (i.e., individual that was released second). The owner and experimenter sat on opposing sides of the room at equidistance to the Hare and Stag tables, each on one side of the fence holding a dog. The whole room was recorded by video cameras.

	Partner B "Hare"	Partner B "Stag"
Partner A "Hare"	A: LVR B: LVR	A: LVR B: Nothing
Partner A "Stag"	A: HVR B: HVR	A: Nothing B: LVR

Figure 2: Payoff-Matrix for the Assurance Game

Training:

The dogs were trained in several phases to teach them how to use the solo and the cooperative apparatus and to give them an idea of the principles of the Assurance Game (see Figure 3). In the first phase of the training, the dog was taught to pull the rope of the solo apparatus to obtain food (Hare training). In the next phase, it needed to learn the principles of the cooperative pulling table (Stag training). Finally, it was made familiar with the different rewards assigned to the tables (Final Training Stage). During the training, both apparatuses were in similar positions as in the test, however only the table the dog was trained on was baited with LVR.

Hare Training (Solo apparatus):

1) Pulling the rope to access the reward: In this phase the dog learned that pulling the rope of the Hare table made the food reward accessible. If the dog did not pull on the rope spontaneously, the experimenter established a command ("Pull!") and trained them via shaping and positive reinforcement. The training was considered successful, when the dog pulled the rope five times in a row to get access to the reward without a prior command. On average, the Hare Training took 2.5 training sessions (range = 1 - 8 sessions).

Stag Training (Cooperative Apparatus):

1) Familiarization: In this phase the dog was familiarized with the cooperative table and the presence of a partner working at the same table. The dog was allowed to move freely while

13

the experimenter held one end of the rope. When the dog approached the apparatus and pulled on the other end of the rope, the platform slid out and made the reward (LVR) accessible to the dog. If the dog did not pull spontaneously, we used the established command. This training phase ended after the dog pulled the rope in three trials in a row without receiving a prior command. On average, it took the dogs 6 trials to reach the criterion (range = 3 - 12 trials).

2) Delay with a human stooge partner: In this phase, the dog experienced the need to wait for a human stooge partner. The experimenter only approached the apparatus and pulled the rope after a delay of five seconds, in which the dog could move freely in the room. If the dog did not manage to wait before pulling the rope, the trial was considered unsuccessful, and the five seconds started again. If the dog managed to wait and pull on the rope simultaneously with the experimenter in three consecutive trials, the delay was expanded to ten seconds. On average, it took the dogs 39 trials to reach the criterion to move onto the ten second delay (range = 7 - 123 trials). After three successful trials in a row with the ten second delay, we moved on to the next phase. This took on average 19 trials (range = 3 - 58 trials).

3) Delay with a conspecific stooge partner: In this phase, the dog experienced working on the apparatus with a conspecific for the first time. For this we used a stooge partner that was trained to pull the rope reliably. The partner was held back by the experimenter for five seconds and then released to pull on the rope, while the subject could move freely. The trial was unsuccessful if the dog pulled the rope before the delay ended. If the subject behaved aggressively towards the stooge partner around the table, but not towards the partner in the dyad, we skipped this part of the training for this individual (see Table 1). If this was the case for both dogs in the dyad, or one of the dogs started showing signs of aggression towards their dyad partner, they were excluded from the study. If the dog managed to pull on the rope simultaneously with the partner two times in a row, the delay was increased to ten seconds. On average, it took the dogs 17 trials to reach the criterion to move onto the ten second delay (range = 3 - 56 trials). After two successful consecutive trials with the 10 second delay, the Stag training was considered finished. This took on average 11 trials (range = 2 - 33 trials).

Final training stage:

Here, the dogs were exposed to the different reward qualities associated with the different apparatuses and the Hare/Stag choice. There were four trials with the conspecific stooge partner. In two of the trials, only the Stag table was baited with the HVR (e.g., sausage) and

the dogs were allowed to operate this apparatus and access the HVR. In the other two trials, the Stag table was made inoperable by removing the rope, upon which the dogs turned to the Hare tables, baited with the LVR (e.g., dry food). This was done to make sure that the dogs associated the different tables with the different reward qualities, as they had only been trained with the LVR. The order of the trial conditions was counterbalanced between individuals.

Assurance Game Test:

There were two testing days with 40 trials each and a 10-minute break after 20 trials. Between testing days, the position of the solo and the cooperative apparatus were switched to control for any bias in the results due to side preference. To ensure that the dogs were aware of the position of the tables, there were four enforced trials before the start of both sessions (i.e., two times Stag, two times Hare, with the other table being inoperable because of missing rope; counterbalanced between dyads). Before each session trial, both the Hare and the Stag table were baited (the order in which the experimenter baited the tables was counterbalanced). During the baiting, one dog was held by the owner and the other one was given the command "Stay!" if possible, or otherwise also held by the owner. At the beginning of each session trial, each dog was held by either their owner or the experimenter on one side of the barrier at the same distance from both tables (see Figure 1 a). During the test, the owner closed their eyes until the dogs made their choice to avoid giving any cues to their dogs. In order to react quickly to any problems if necessary, the experimenter did not have their eyes closed, but looked straight ahead until the dogs made their choice. After the signal (i.e., the spoken word "ok") from the experimenter, the partner was released and was able to choose one of the tables. One second after the release of the partner, the subject was released to make their own choice. The delay was introduced so that the subject could see the behavior of the partner before deciding on a table. A choice was considered the pulling of the rope. After a dog made their choice, it was not allowed to make a second one. The trial ended when both partner and subject chose a table or after 40 seconds. After that, the chosen tables were re-baited, and the unchosen ones were pretended to be baited in order to avoid bias due to local enhancement through the experimenter.

Tolerance Test:

The tolerance test (see Brucks et al., 2016) was conducted to investigate the effect of the animals' choices on their sharing of resources. We presented the dogs with a bowl of food and

measured how long they fed in proximity to their partner before and after the Assurance Game. The tolerance test was done on five different days, three times during training and once after each testing day (see Figure 3). During the Assurance Game, the subjects had the possibility to earn HVR food, while they were only rewarded with LVR food during training days. We wanted to make sure that any effect in the tolerance test was not due to different satiety of the dogs during training and after the testing days of the Assurance Game. Therefore, we fed each subject a variable amount of HVRs equivalent to what they could have earned during the Assurance game, previous to the tolerance test conducted on training days. One of the tolerance tests was done with no extra food before the test, one with 20 pieces of HVRs before the test and one with 40 pieces of HVRs before the test (order counterbalanced between dyads).

To familiarize the animals with the procedure, we presented them with a bowl of food that was placed under a cart box in the middle of the testing room on the first day. The box was attached to a rope that was threaded through a hook in the ceiling. The experimenter was standing on one side of the room holding the free end of the rope. When the dogs approached the bowl, the box was lifted by the experimenter by pulling on the rope. If the dogs did not show any fear of the movement of the box, we moved on to the tolerance test. In the test, the dyad moved freely in the testing room. In the middle of the room, there was the bowl filled with food placed under a cart box. The box was lifted by the experimenter when both dogs were approaching the bowl and were approximately 10cms away. The dogs were both allowed to feed from the bowl. The test ended when both dogs were more than one body length away from the bowl or the bowl was empty.
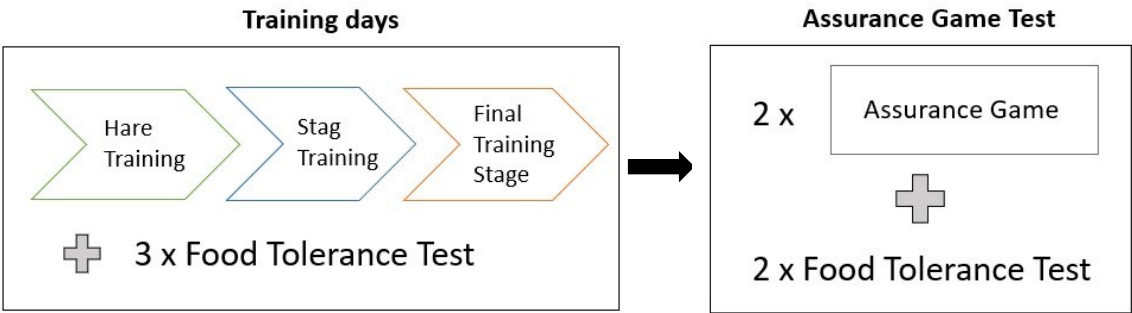

Figure 3: Summary of the procedure after the Food Preference Test.

2.4. Video coding and analysis:

Both the Assurance Game sessions and the tolerance tests were video recorded. This was done by four video cameras attached to the walls that recorded the whole room from different angles. In the Assurance Game, we coded the choice "Hare", "Stag" (i.e., pulling on the rope of the respective table) or "No Choice" (i.e., no pulling on the rope within 40 seconds) and in the tolerance test, we coded the duration of co-feeding, feeding alone and of proximity to the bowl (i.e., being within 1m of the bowl) and the frequency of displacements and growling. Video analysis was done with the software Loopy (Klöckner, 2014). A second coder, blind to the predictions of this study, coded 20% of the videos for reliability checking, revealing good consistency between coders (Intra-class correlation coefficients (ICC) for continuous variables: all ICC > 0.85; Cohen´s kappa for dog´s choices -dichotomous variable-: all kappa = 1)

2.5. Statistical Analysis:

Assurance Game:

We analyzed the results of the Assurance Game with generalized linear mixed models (GLMM, Baayen, 2008) with the lme4 package (Bates et al., 2015) with the optimizer "bobyqa" in R (R Core Team, 2014). To see whether the dyads coordinated their choices, we fitted two GLMMs with binomial error structure and logit link function (McCullagh & Nelder, 1989). In the first model we analyzed if the partner's (i.e., the dog that was released first) choice influenced the subject's (i.e., the dog that was released second) choice. Our response variable was the subject choice, and the predictor was partner choice. To control for possible learning effects and side bias (the position of the Stag table was changed between sessions), we included session number and trial number as fixed factors, as well as the three-way interaction between partner choice, trial number and session number, and all the lower order interactions. In the model, we only used the trials where both dogs made a choice (847 trials). As the subject identity could lead to pseudoreplication, we included it as a random intercept and all theoretically identifiable random slopes (three-way interaction between partner choice, session and trial). For the fixed effects, trial number was z-transformed (to a mean of zero and a standard deviation of one). For the random slopes, trial number was z-transformed, and partner choice and session were dummy coded (i.e., the two levels of the categorical predictors were replaced by either 0 or 1) and centered to the mean. We did this to better the convergence of the model. Because every subject only had one partner, we did not include partner identity in the model. We fitted two versions of the model, one with all the

correlations between intercept and slopes and one with none. Because the correlations were all almost unidentifiable (i.e.. close to one), we chose the simpler model with no correlations to reduce complexity. This caused a modest decrease in model fit (model without correlations: logLik = - 154.1041 (df = 16), model with correlations logLik = -137.6968 (df = 44)).

Our next model was to test if the subject's previous choice had an influence on the partner's choice. The model contained partner choice as the response variable, and previous subject choice as the key predictor. The other fixed factors were session, trial and the three-way interaction between previous subject choice, trial number and session number, and all the lower order interactions. In the model, we only used the trials where both dogs made a choice (825 trials). Partner identity was included as a random intercept, plus all theoretically identifiable random slopes (three-way interaction between previous subject choice, session and trial). Trial was z-transformed before including it in the fixed effects and random slopes and previous subject choice and session were dummy coded and centered to the mean for the random slopes. We then again compared two versions of the model, one with all correlations between random slopes and intercept and one with none. As the correlations were all close to one, we chose the simpler model without correlation. This led to a tolerable decrease in model fit (model without correlations logLik = - 170.8596, (df = 16), model with correlations logLik = - 157.0974 (df = 44)). In order not to increase the type I error by multiple testing (Fostmeier and Schielzeth, 2011), the significance of the full models (models with all variables) was compared with null models (identical models to the full models, but without the key predictor partner choice or subject choice). We did this with likelihood ratio tests (R function anova with argument test set to "Chisq"; Dobson & Barnett, 2018). We then tested the significance of the highest order interactions with likelihood ratio tests (Barr et al., 2013; R function drop1). In case the p-value was above 0.1, we removed the interaction (only one term was removed at a time) and refitted the models. This was done until the two models contained only main terms and interactions with p-values above 0.1.

To assess model stability, we compared the estimates from the working models with the estimates from models that had excluded one subject at a time (Niewenhuis et al., 2012). The model for the influence of partner choice on subject choice was quite stable, with the exception of session number. This was probably due to the fact that some dogs showed side bias between sessions while others did not. Visual inspections of the BLUPs (best linear unbiased predictors, Harrison, 2018) showed deviation from normal distribution for session as well, further indicating the individual variability, possibly because some dogs exhibited a side

bias. The other BLUPs were small or approximately normally distributed. The model for the influence of previous subject choice on partner choice was stable as well. Upon visual inspection of the BLUPs, the partner identity intercept showed considerable deviation from normality, which probably stemmed from the individual variability of the dogs. The interaction between previous partner choice and session also deviated from normality, further showing that some subjects were more side-biased while others were not. To check for collinearity, we assessed the Variance Inflation Factors (VIF, Field, 2009) for standard linear models similar to the main models but lacking interactions and random effects. The maximum VIF was 1.03, so we could exclude problems in the models due to collinearity. Confidence intervals were calculated with the function bootMer, using 1000 parametric bootstraps. To get more reliable values, we obtained the p-values of the fixed effects using likelihood ratio tests (Barr et al., 2013; function mixed of the package afex, Singman et al, 2016).

To see if the dyads coordinated in Stag above chance level, we used chi-square goodness-of-fit tests. The Assurance Game has four possible outcomes, so the chance for each outcome was 25%. If the analysis was significant, we looked at the standardized residuals (srs, if they are outside +/- 2.58 they were statistically significant; Field, 2009) of the choices.

Tolerance Test:

Changes in tolerance after participation in the Assurance Game were analyzed with a generalized linear mixed model (GLMM, Baayen, 2008) with a beta error distribution and logit link function (McCullagh & Nelder, 1989) with the glmmTMB package (Brooks et al., 2017). The response variable was the proportion of time the dyad spent co-feeding and the predictor was the condition (i.e., either test or training). To control for satiety and the effect of different payoffs in the game, the number of HVRs the dog had eaten before the tolerance test was included as a fixed factor, as well as session number and an interaction between number of HVRs and condition. In total, 550 sessions were analyzed. To avoid pseudoreplication, dyad identity was included as a random intercept, together with all theoretically identifiable random slopes (HVRs, condition and session). The number of HVRs and session were z-transformed for the fixed effect. For the random slopes, the number of HVRs and session were z-transformed, and condition was centered to the mean. Two versions of the model were fitted, one with all the correlations between intercept and slopes and one without. As the model with the correlations did not converge, we chose the model without correlations.

Again, we compared the significance of the full model with the null model with a likelihood ratio test (R function anova with argument test set to "Chisq"; Dobson and Barnett, 2018).

To analyze if the level of tolerance predicted the coordination in the Assurance Game, we used a GLMM with binomial error structure and logit link function (McCullagh & Nelder, 1989) with the lme4 package (Bates et al., 2015) with the optimizer "bobyqa". The response variable was the coordination (i.e., coordination on Stag or no coordination) and the predictor was the time the dyad had spent co-feeding during the tolerance test session that was conducted with zero pieces of HVR during training. Session number and trial were also included as fixed factors, as well as the three-way interaction between co-feeding, trial and session number. 880 trials were used for analysis. Subject identity was included as a random intercept together with all the theoretically identifiable random slopes (three-way interaction between co-feeding, trial and session). For both the fixed and random effects, co-feeding and trial number were z-transformed and session was dummy coded and centered to the mean. Two versions of the model were fitted, one with all the correlations between intercept and slopes and one with none. As the correlations were all close to one (i.e., unidentifiable), we chose the model with no correlations. This caused a modest decrease in model fit (model without correlations: logLik = -239.5087 (df = 16), model with correlations logLik = -207.5904 (df = 44)). The significance of the full model (model with all variables) and the null model (identical to the full model, but without the predictor variable co-feeding) was compared with a likelihood ratio test (R function anova with argument test set to: "Chisq"; Dobson & Barnett, 2018).

Model stability in both models was tested by comparing the estimates from the working model with the estimates from a model that excluded one dyad at a time (Niewenhuis et al., 2012). Our model with the predictor variable condition was stable and visual inspection of the BLUPs showed an approximately normal distribution for all variables. The model with co-feeding as a predictor showed some instability, especially for the predictor variable co-feeding. This could be explained by the fact that the co-feeding was very different across dyads, with some co-feeding a lot while others did not co-feed at all. Visual inspection of the BLUPs showed slight deviations from normality. The VIF showed that there was a slight collinearity for session and condition (maximum VIF 4.06) in the model with the predictor condition, however, there was still enough variability to interpret the results. The maximum VIF of the model with co-feeding as a predictor was 2, so there were no problems in the model due to collinearity. We calculated the confidence intervals with the function

20

boot_glmmTMB and used 1000 parametric bootstraps. To get more reliable values, we obtained the p-values of the fixed effects using a likelihood ratio test (Barr et al., 2013; function mixed of the package afex, Singman et al, 2016).

3. Results:

We found that partner choice had a significant influence on subject choice (full-null model comparison: $\chi 2=18.963$, df=4, $p < 0.001$). All the interactions were non-significant, so we dropped them from the model (estimates and p-values of the variables when they were dropped can be found in the Supplementary, S2). Neither trial nor session had a significant influence on the subject choice (see Table 2).
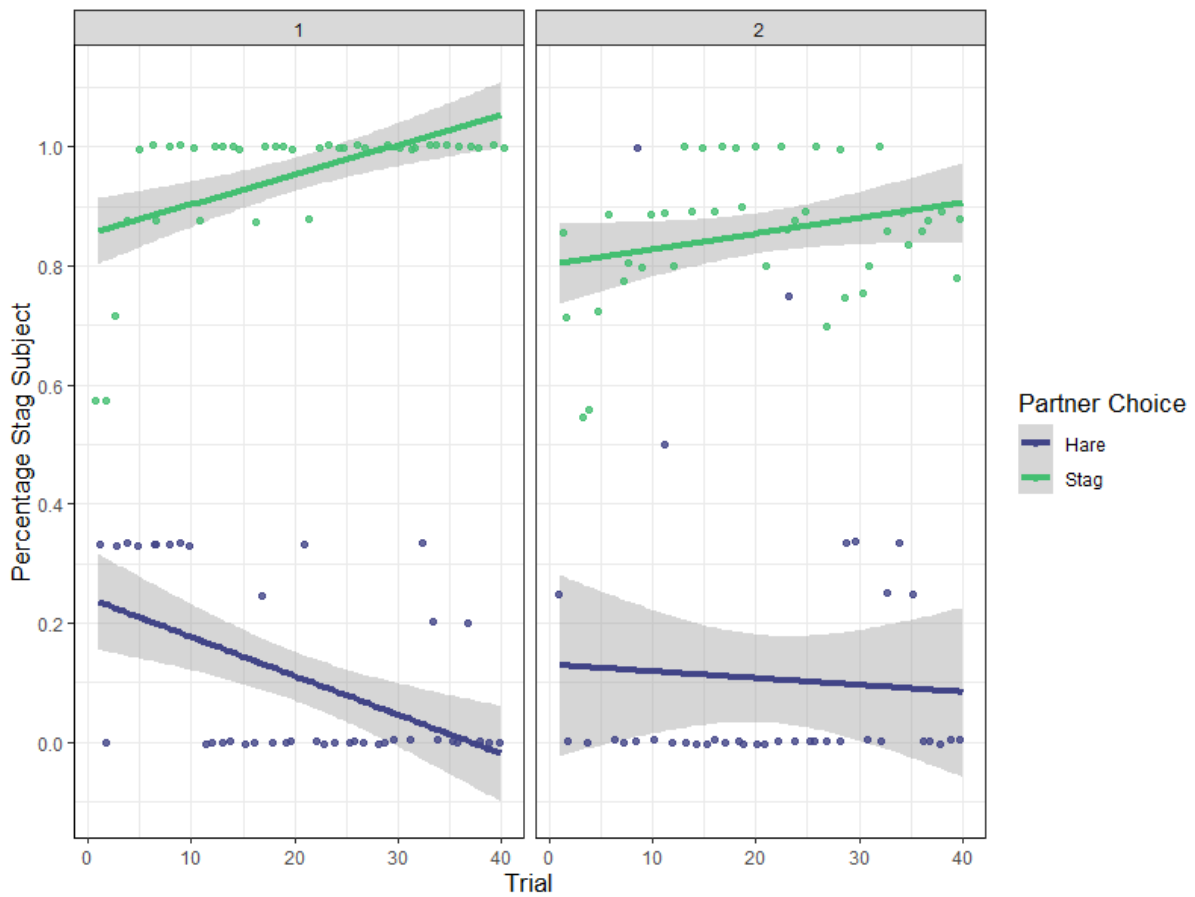
Figure 4: Proportion of the choice "Stag" of the subject depending on the choice of the partner.

Across all trials and sessions, the percentage of choosing Stag by the subject was much lower when the partner chose Hare (mean = 10.66%, SD = ± 19.11%) than when the partner chose Stag (mean = 90.54%, SD = ± 12.06%) (see Figure 4).

Our second GLMM was modelled to test whether the previous choice of the subject also had an influence on the choice of the partner. The full-null model comparison showed that the previous subject choice had a significant influence on the partner choice (full-null model comparison: $\chi 2=11.956$, df=4, $p = 0.018$). Specifically, we found a significant

interaction between previous subject choice and trial ($\chi2=5.88$, df=1, $p = 0.015$; see Table 2), showing that over trials, the partners adapted their choice to the previous choice by the subject (see Figure 5). As the interaction between previous subject choice, session and trial, as well as between session and trial were non-significant, we dropped them from the model (see Supplementary, S2).
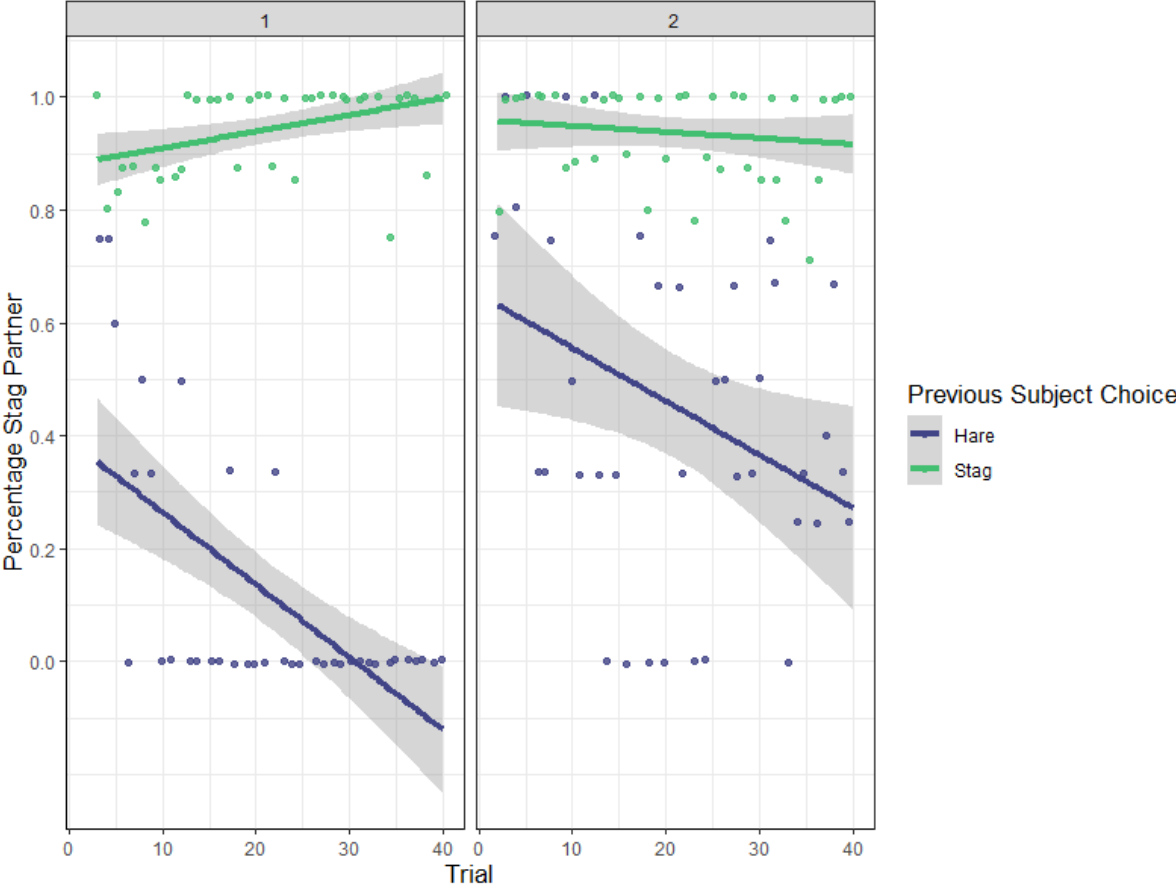

Figure 5: Proportion of the choice "Stag" of the partner depending on the previous choice of the subject.

Table 2: Results of the General Linearized Mixed Models of the Assurance Game

Response Variable	Term	Estimate	SE	LRT	df	p-value	Stability		CI	
							Min	Max	2.50%	97.50%
Subject choice	Intercept	-0.829	1.954				-2.294	0.782	-5.164	3.375
	Partner choice (0, Hare ; 1, Stag)	5.24	1.365	11.72	1	**< 0.001**	4.078	6.489	2.625	9.412
	Session (0, Session 1 ; 1, Session 2)	-0.077	2.922	0	1	0.976	-2.408	1.135	-6.489	6.471
	Session	0.808								
	Trial	0.808	0.673	1.44	1	0.231	0.388	1.087	-0.485	2.266
Partner choice	Intercept	-0.316	1.759				-1.105	1.021	-3.943	3.485
	Previous Subject Choice (0, Hare ; 1, Stag)	4.055	1.554	4.9	1	**0.027**	3.379	4.991	1.016	7.746
	Session (0, Session 1 ; 1, Session 2)	3.2	2.427	0.58	1	0.447	1.769	4.288	-1.819	9.513
	Trial	-1.25	0.458	2.58	1	0.091	-1.653	-0.864	-2.561	-0.343
	Previous Subject Choice:Session	-3.274	2.149	3.23	1	0.072	-4.222	-2.532	-8.859	0.828
	Previous Subject Choice:Trial	1.078	0.426	5.88	1	**0.015**	0.927	1.371	0.123	2.219

*Statistically significant p-values are indicated by bold letters.*

Trial was z-transformed, for response variable subject choice original mean = 20.74, SD = ± 11.54;

for response variable partner choice original mean = 21.27, SD = ± 11.23

24

The previous models showed that both subjects and partners coordinated their choices with each other. Our second aim was to find out if they coordinated in the payoff-dominant strategy (StagStag). We found that all of the dyads' choices significantly differed from chance level (see Supplementary, S3). In total 7 dyads coordinated in Stag, that is, they chose StagStag significantly more often than chance (Annie-Prim: $sr = 5.58$, $p < 0.001$; Benny-Dawin: $sr = 5.46$, $p < 0.001$ Crash-Alou: $sr = 12.08$, $p < 0.001$; Gandalf-Lenny: $sr = 6.78$, $p < 0.001$; Leia-Eowyn: $sr = 11.53$, $p < 0.001$; Miley-Tiara: $sr = 11.53$, $p < 0.001$; Timo-Mailo: $sr = 12.71$, $p < 0.001$). The 4 other dyads also matched their choices (StagStag and HareHare) above chance, but they chose both options equally often (Helsa-Dora: StagStag $sr = 3.51$, HareHare $sr = 3.06$, $p < 0.001$; Joseph-Josephine: StagStag $sr = 2.59$, HareHare $sr = 2.83$, $p < 0.001$; Matteo-Hetti: StagStag $sr = 3.51$, HareHare $sr = 4.42$, $p < 0.001$; Mavie-Cheynna: StagStag $sr = 4.25$, HareHare $sr = 4.47$, $p < 0.001$). This can probably be explained by a side bias, as the dyads that matched on both choices did not do so randomly, but usually matched on Hare in one of the sessions and coordinated on Stag in the other.

When looking at the dyad's individual performance in the Assurance Game (see Figure 6, for a summary see Table 3), we found that Annie-Prim only coordinated on Stag in the last few trials of session one and session two, while they did not match their choices for most of session one. In this dyad the partner was side biased. Benny-Dawin coordinated on Stag in the first session, but coordination and the matching of choices dropped in the second session. Crash-Alou coordinated on Stag in both sessions, in the first session the subject adjusted to the partner choosing Stag by changing their choice from Hare to Stag after the third trial. In the second session both dogs matched on Hare first but coordinated on Stag after a few trials. Gandalf-Lenny coordinated on Stag in the first session, but the subject coordinated less after the partner had growled at it during the game. Helsa-Dora matched on Hare in the first session and coordinated on Stag in the second session. In this dyad, the subject was clearly side biased while the partner also occasionally diverted from that side bias. This was also the case in Joseph-Josephine, who coordinated on Stag in the first session and matched on Hare in the second session. Leia-Eowyn coordinated perfectly on Stag in the first session, in the second half of session two coordination on Stag dropped as the subject occasionally chose Hare. Matteo-Hetti were side biased and matched on Hare in session one and coordinated on Stag in session two. In the end of session two, both dogs changed their choice from Stag to Hare and matched on Hare. Mavie-Cheynna were clearly side-biased and matched on Hare in session
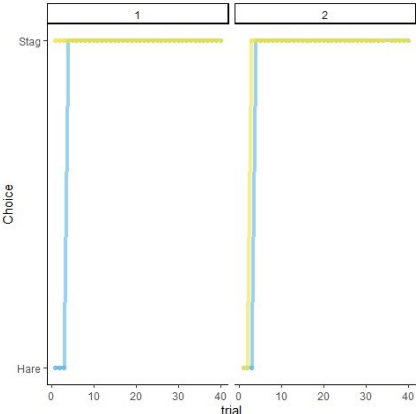
25

one and coordinated on Stag in session two. Miley-Tiara and Timo-Mailo mostly coordinated on Stag in both sessions, with a few unmatched choices.
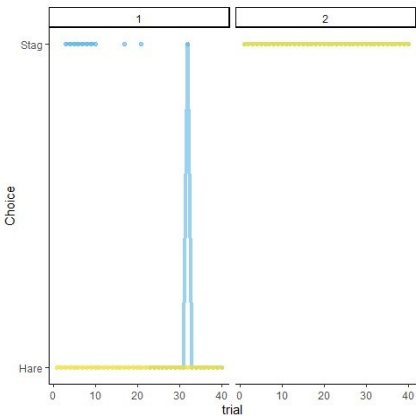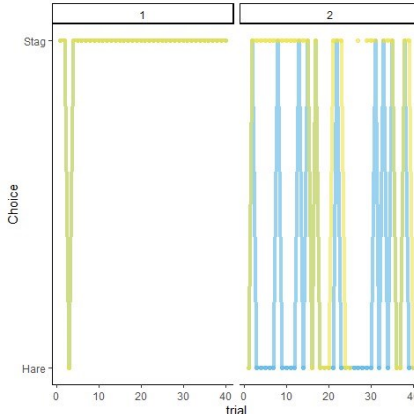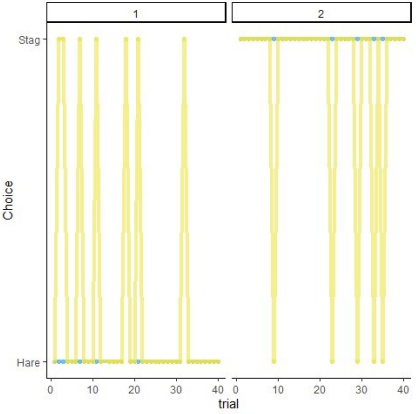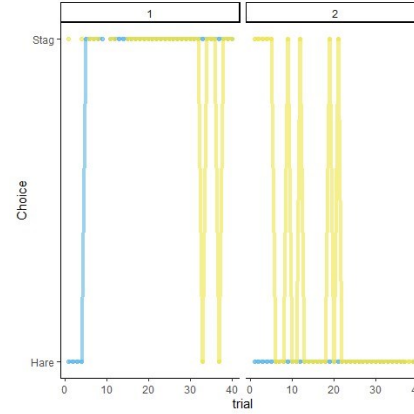

Annie-Prim
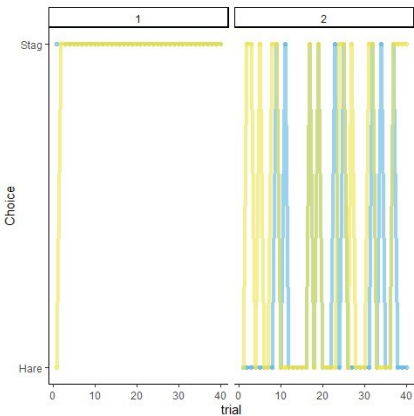
Benny-Dawin

Crash-Alou

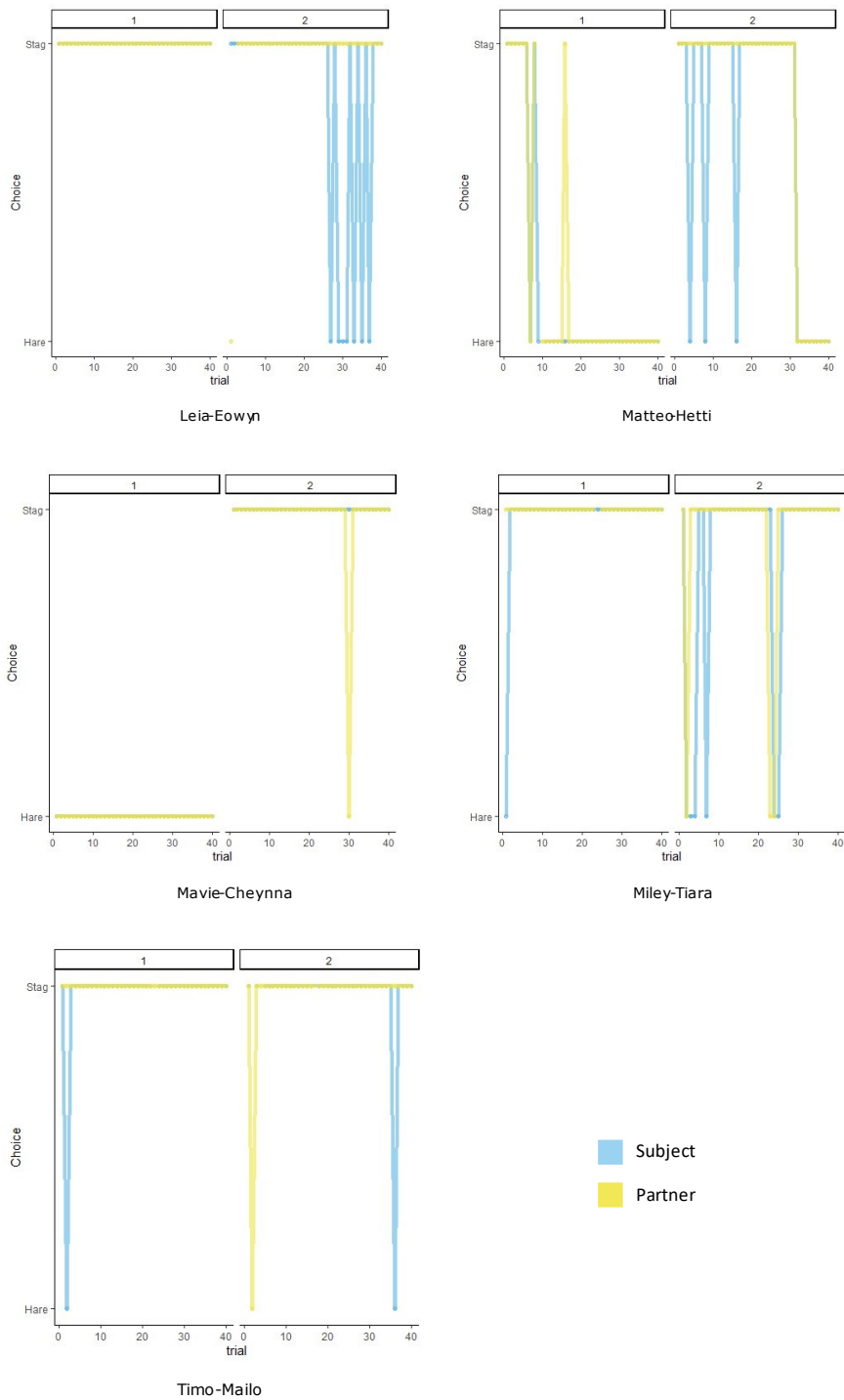Gandalf-Lenny

Helsa-Dora

Josephine-Joseph

Figure 6: Dyads' performances in the Assurance Game

Table 3: Description of the performance of the dyads in the Assurance Game

Dyad	ID	Description of AG
1	Annie Prim	Subject side biased, partner adjusted to it after the first half of session one
2	Benny Dawin	Session 1 almost perfect coordination on Stag; session 2 both Hare and Stag coordination with mixed performance
3	Crash Alou	After first 3 trials subject adjusted to partner's Stag choice in both sessions
4	Gandalf Lenny	Session 1 perfect coordination; session 2 coordination mixed after partner growled at subject
5	Helsa Dora	Subject side biased, partner adjusted to it after a couple of trials with some exceptions
6	Josephine Joseph	Subject side biased, partner adjusted to it after a couple of trials in each session
7	Leia Eowyn	Session 1 perfect coordination; session 2 coordination dropped after first half because subject sometimes chose Hare
8	Matteo Hetti	Coordinated on both choices in both sessions with mixed performance
9	Mavie Cheynna	Almost perfect coordination in both sessions, at least one of them had side bias
10	Miley Tiara	Session 1 almost perfect coordination; session 2 subject adjusted to partner's Stag choice after a few trials
11	Timo Mailo	Almost perfect coordination in both sessions, with a few uncoordinated trials at beginning and end

The results of the tolerance test showed that condition did not have a significant influence on the proportion of time the dyad spent co-feeding (full-null model comparison: $\chi 2=1.164$, df=10, $p < 0.559$). Further, neither the number of HVRs eaten, nor the session number influenced the co-feeding (see Table 4). This means that the level of co-feeding stayed the same within dyads over the course of the tolerance tests. As we included the number of HVRs the dog had eaten before each test into the model, it also showed that the different payoffs during the game did not influence the level of tolerance. We found displays of aggression in only two dyads (growling: once in Lenny-Gandalf during training, displacement: three times in Leia-Eowyn during training and test).

When testing if the level of co-feeding during the tolerance test could predict the level of coordination on Stag during the Assurance Game, we found no significant influence of the co-feeding on the coordination (full-null model comparison: $\chi2=4.371$, df=4, $p < 0.358$). Further, neither trial nor session had a significant influence on the level of coordination (see Table 4). This means that dyads that co-fed more during the tolerance test did not coordinate more on Stag than those who co-fed less.

Table 4: Results of the Generalized Linear Mixed Model of the Tolerance

Response Variable	Term	Estimate	SE	LRT	df	p-value	Stability Min	Stability Max	CI 2.50%	CI 97.50%
Proportion of time co-feeding	Intercept	-2.136	0.601				-2.632	-1.948	-3.365	-1.107
	Condition	-0.489	0.349	1.74	1	0.187	-1.048	-0.102	-1.210	0.236
	HVRs	0.277	0.157	2.33	1	0.127	-0.180	0.415	-0.033	0.582
	Session	-0.266	0.184	2.08	1	0.149	-0.435	-0.030	-0.622	0.063
	Condition:HVRs	-0.218	0.200	1.18	1	0.277	-0.415	0.376	-0.589	0.157
Coordination on Stag	Intercept	0.457	1.598				-4.889	2.569	-2.761	3.678
	Co-feeding	0.335	2.308	0.77	1	0.381	-9.420	3.362	-4.806	5.471
	Trial	-0.922	0.697	1.98	1	0.159	-1.521	0.855	-2.053	9.143
	Session (0, Session 1; 1, Session 2)	2.144	2.336	0.86	1	0.355	0.423	7.151	-2.559	0.041
	Co-feeding:Trial	-1.403	1.420	0.19	1	0.655	-4.059	1.262	-4.077	17.092
	Co-feeding:Session	4.145	2.838	2.17	1	0.141	1.825	16.010	-1.649	0.939
	Session:Trial	0.463	0.699	0.44	1	0.507	-0.978	0.834	-0.866	1.825
	Co-feeding:Session:Trial	1.820	1.325	0.73	1	0.394	0.162	2.561	-1.105	4.286

HVRs, session and co-feeding were z-transformed,
original mean HVR = 23, SD = ± 16.67;
original mean session = 3, SD = ± 1.43; original mean co-feeding = 103.46, SD = ± 174.50

4. Discussion:

In this study we investigated if dogs were able to coordinate their actions in the Assurance Game. In line with our predictions, we found that the study subjects coordinated their actions and most of them preferentially chose the cooperative option (Stag). We also tested if dogs would improve their tolerance towards their partners after cooperation, showing that they understood that their partners' actions were important in obtaining the cooperative rewards. Contrary to our prediction, the tolerance within the dyad was not influenced by coordination or the achieved payoffs in the Assurance Game. Further, we found that the level of tolerance in the dyads during the tolerance test did not predict the level of coordination in the game.

In detail, the choice of the subject was significantly influenced by the choice of the partner, and neither trial nor session had an effect on the subject's choice. That is, if the partner (the dog that was released one second before the other one) chose Stag or Hare, the subject (the dog released second) was significantly more likely to make the same choice. In the model used to analyze this, session was not a stable factor. This means that the estimate was very different for different subjects. This can possibly be explained by the fact that some dogs exhibited a side bias in their choices (the side of the tables was switched between sessions), while others did not. In the dyads in which one or both individuals were side biased, coordination destabilized when the position of the Stag table in the room changed. Further, we not only found an influence of the partner choice on the subject choice, but also that the partner choice was affected by the previous choice the subject had made. The partner's coordination increased with trials, which indicates a learning effect. The partner first was more likely to choose Stag, and if that led to a reward, it stuck with it. When it was not successful because the subject was choosing Hare more often, the partner gradually changed their choice towards the Hare option.

Seven of the dyads coordinated on the payoff-dominant strategy of the game (Stag) as expected, while the other four dyads coordinated on Stag and matched on Hare above chance. Inspection of the individual dyads' performances showed that in several pairs, coordination on Stag could be explained by the fact that the dogs had a general preference for the HVR, so they went to the Stag table first and, if successful, stuck with that choice. Further, the dogs were trained last on the Stag table, and even though there was a final stage and warm-ups with the Hare table, it might have made them more likely to choose Stag. However, even the dyads that coordinated solely on Stag did not do so perfectly, but also had unmatched trials.

This could indicate that it was not solely a preference for the HVR that drove the coordination, as they also explored their other options. There were three dyads that showed signs of true coordination (Crash-Alou, Miley-Tiara, Timo-Mailo), as they coordinated on Stag in both sessions reliably with only very few exceptions. The four dyads that coordinated on Stag and matched on Hare did not do so randomly but made one choice in one session and the other in the next. This means that these dogs exhibited side bias in their choices, going to the same side of the room in both sessions. However, even these dyads matched their choices, which means that either both of the dogs had a side bias for the same side, or one of them adjusted to the other. Further, our model clearly showed a significant influence of the choice of the partner on the individual's own choice over all dyads.

This study in dogs extends the research that has been done in primates on decision making and cooperation during the Assurance Game. Not only does it add another taxon, but it also tested for flexible coordination. Different primate species (chimpanzees: e.g., Bullinger et al., 2011; rhesus macaques: e.g., Brosnan et al., 2012; Parrish et al., 2014; squirrel monkeys: Vale et al., 2019; capuchin monkeys: e.g., Brosnan et al., 2011; 2012) tend to coordinate with their partners by choosing the payoff-dominant strategy (Stag; see review in Massen et al., 2019). The dogs showed a higher level of coordination than capuchin monkeys, who only coordinated on the payoff-dominant strategy after extensive experience with the test procedure (Brosnan et al. 2011, 2012; Smith et al., 2019). Contrary to our dogs, however, the capuchins did not receive any training on the contingencies of the game before testing. In a recent study capuchin monkeys coordinated on Stag, but the choice of the partner did not influence their actions (Robinson et al., 2021). While rhesus macaques and squirrel monkeys developed a bias for the Stag choice over trials (Brosnan et al., 2012; Parrish et al., 2014; Vale et al., 2019), the choice of the subjects in our study was independent of trial. Further, the dogs not only coordinated on Stag, but some dyads also matched on Hare. This is the first study to show that the choice of the animal in the Assurance Game is influenced by their partner's choice. The better performance of the dogs in our study could be due to methodological differences. For the monkeys, the different choices in the Assurance Game were represented by tokens (Brosnan et al., 2011; Vale et al., 2019) or choosing icons on a screen (Brosnan et al., 2012; Parrish et al., 2014; Smith et al., 2019). Using string-pulling tables might have conveyed the contingencies of the game more intuitively for the dogs. Further, contrary to the capuchin monkeys in Brosnan et al. (2011), Smith et al. (2019) and Robinson et al. (2021), rhesus macaques (Brosnan et al. 2012, Parrish et al. 2014) and squirrel monkeys (Vale et al.,

2019), the dogs went through an extensive training before testing to ensure their understanding of the apparatus. The dogs were introduced to the choices of the Assurance Game and only passed into the test if they were able to wait for their partner. Another important aspect in our study was that one dog was released first, so the other could see what the partner had played before making their own choice. Similar setups also seemed to increase the performance of the capuchin and rhesus monkeys (Brosnan et al., 2012; Robinson et al., 2021). In fact, our setup was similar to Bullinger et al. (2011) and Duguid et al. (2014), who tested chimpanzees in the Assurance Game by using a string-pulling paradigm. The chimps were able to coordinate on Stag by applying a leader-follower strategy. In fact, in Bullinger et al. (2011) a loud noise produced by the experimenter announced the availability of the Stag table, which might have acted as a non-social cue for the individuals to approach the Stag regardless of the actions of their partner. This was not the case for our study. However, the coordination of the choices of the chimps and the dogs could be the result of local enhancement (i.e., the presence of another individual draws the attention to that particular location; Thorpe, 1956) or stimulus enhancement (i.e., the presence of another individual at a certain object draws the attention to objects of that type; Spence, 1937). As one of the dogs (the partner) is released first and goes to one side, the attention of the subject could be drawn to that side and lead them to move to the same side. There they pull the rope they see in front of them or interact with the same object type the partner has. This is also a possible explanation for many hunting situations, where simple strategies suffice to be successful in a cooperative hunt (Bailey et al., 2013). To avoid the problem of local and stimulus enhancement, it might be a good idea to devise an apparatus where cooperation is achieved by the partners each manipulating a differently looking half of an apparatus on opposing sides of the room. However, in this setting, a second test would be needed to check that the animals understand they are working together despite the separated halves of the apparatus.

Our results stand in line with the pulling table studies by Ostojić and Clayton (2014) and Marshall-Pescini et al. (2017; 2018), which seemed to indicate that tolerance might be the limiting factor for solving the string-pulling paradigm. Like Ostojić and Clayton (2014), we tested pet dog dyads from the same household, while Marshall-Pescini et al. (2017; 2018) tested pack living dogs. Our dogs were used to living with each other in close contact, and their interactions are usually monitored closely by their owners to mitigate any conflict between them. Their owner was in the room during testing, which might further help to mediate any conflicts. Additionally, we only selected dyads for our study where at least one of

the dogs showed no signs of aggression against conspecifics and there was no aggression within the dyad. This is also shown by the lack of aggression in almost all dyads during the tolerance test. Thus, we created an environment for our experiment where possible conflicts due to tolerance issues were mitigated. Interestingly, in one of our dyads, Lenny-Gandalf, Gandalf stopped coordinating on Stag for some trials after Lenny growled at him in session two.

We showed that dogs are able to flexibly adjust their actions in time and space to a partner (i.e., coordination, Boesch & Boesch, 1989). It is possible that they understood the need for a partner to gain access to the high value reward and that their partner was also making choices in the game that had an influence on their own payoff. Due to the presence of both tables, the attention of the dogs was not inherently directed at one of them from the beginning. It is possible that the dogs coordinated due to local enhancement, even though the study by Marshall-Pescini et al. (2017) would suggest that dogs would rather avoid the same location because of tolerance issues. Yet, from this experiment alone, we cannot infer the level of understanding of the partner's role. When looking at the individual strategies of the dyads across trials, we found that some of the coordination and matching of choices could be explained by external factors (i.e., preference for the HVR, side bias). Only three of the dyads showed signs of truly coordinating during the game (i.e., the ones where the level of coordination on Stag was high and stable in both sessions). Indeed, the tolerance test, conducted to test if the animal understood that their partner was also making decisions during the Assurance Game, showed no significant results. Neither condition (training or test), nor the different payoffs (number of HVRs eaten before the test), had an influence on the time the dyads spent co-feeding. This is in contrast to the study by Brucks et al. (2016), where the duration of co-feeding increased if the dogs had been treated equally before. However, the conditions prior to the tolerance test in this study were different to ours. We did not reward the dogs equally or unequally before the test, instead the dogs could still get food even if they did not make the cooperative choice. Further, knowing that they need to coordinate with a partner does not necessarily include knowledge about the existence of goals and intentions in the partner. It also does not ensure that they reciprocate the help they received in one context by being more tolerant in another context. The tolerance test could have been perceived as an entirely separate interaction from the Assurance Game. Supporting this point is the fact that we found that the level of coordination on Stag in the game could not be predicted by the

tolerance level of the dyad during the tolerance test. This means that more tolerant dyads were not more likely to coordinate on Stag in the game than less tolerant dyad.

While these results do not allow the conclusion that the dogs understood their partners role during the task, it also does not allow the conclusion that they did not. The test could have been influenced by several other factors. The situation of competing over a bowl of food might not be a new situation to the dyad. Previous interactions could have led to a feeding hierarchy, where the subordinate would not try to approach the bowl, thus not allowing the possibility for co-feeding. Indeed, it has been shown that in dogs the dominant individual usually monopolizes the food source (Range et al., 2015) and that the strength of the social bond and the rank distance affect the likelihood of food sharing (Dale et al., 2017). Another important factor is that owners often prevent their dogs from approaching each other and stealing food during feeding time at home. Further, we put the dogs into direct competition in the tolerance test, where sharing with the other necessarily decreased their own payoff, while the contrary was true for the cooperative Assurance Game. With this tolerance test, we tried to find a way to evaluate how much the animals understood about the role of their partner during the task, which has not been done before. It would be important to explore this question further. The food in the tolerance test was not obtained cooperatively, which might have made sharing less likely. Chimpanzees, for example, share the meat of a hunt with the individuals that participated in it (Boesch, 2002). Therefore, one approach could be to measure the tolerance around food that has been gained by cooperation.

In summary, we are the first to show that the choices of an animal playing the Assurance Game depended on the choice of their partner. We found that dogs are able to coordinate their actions in the Assurance Game and that they mostly played the payoff-dominant strategy. Our experiment makes room for interesting follow-up studies. One example would be to exchange the role of the partner and subject within dyads and include control conditions (e.g., a condition with an opaque barrier or no partner on the other side of the fence). In a broader context, it highlights the importance of exploring new techniques to research cooperation and especially ones to demystify the understanding of the partner's role during cooperation in animals.

5. Literature:

Albiach-Serrano, A. (2015). Cooperation in primates: A critical, methodological review. *Interaction Studies*, *16*(3), 361-382.

Baayen, R. H. (2008). *Analyzing Linguistic Data: a Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Bailey, I., Myatt, J. P., & Wilson, A. M. (2013). Group hunting within the Carnivora: physiological, cognitive and environmental influences on strategy and cooperation. *Behavioral Ecology and Sociobiology*, *67*(1), 1-17.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.

Bates, D., Maechler, M., Bolker, B. and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Boesch, C. (2002). Cooperative hunting roles among Tai chimpanzees. *Human Nature*, *13*(1), 27-46.

Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Tai National Park. *American journal of physical anthropology*, *78*(4), 547-573.

Bräuer, J., Bös, M., Call, J., & Tomasello, M. (2013). Domestic dogs (Canis familiaris) coordinate their actions in a problem-solving task. *Animal Cognition*, *16*(2), 273-285.

Bräuer, J., Stenglein, K., & Amici, F. (2020). Dogs (Canis familiaris) and wolves (Canis lupus) coordinate with conspecifics in a social dilemma. *Journal of Comparative Psychology*, *134(2),* 211-221.

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M., Bolker, B.M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378–400

Brosnan, S. F. (2011). What do capuchin monkeys tell us about cooperation?. In *For the greater good of all* (pp. 11-27). Palgrave Macmillan, New York.

Brosnan, S. F., Parrish, A., Beran, M. J., Flemming, T., Heimbauer, L., Talbot, C. F., ... & Wilson, B. J. (2011). Responses to the Assurance game in monkeys, apes, and humans using equivalent procedures. *Proceedings of the National Academy of Sciences*, *108*(8), 3442-3447.

Brosnan, S. F., Salwiczek, L., & Bshary, R. (2010). The interplay of cognition and cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1553), 2699-2710.

Brosnan, S. F., Wilson, B. J., & Beran, M. J. (2012). Old World monkeys are more similar to humans than New World monkeys when playing a coordination game. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1733), 1522-1530.

Brucks, D., Essler, J. L., Marshall-Pescini, S., & Range, F. (2016). Inequity aversion negatively affects tolerance and contact-seeking behaviours towards partner and experimenter. *PLoS One*, *11*(4), e0153799.

Bullinger, A. F., Wyman, E., Melis, A. P., & Tomasello, M. (2011). Coordination of chimpanzees (Pan troglodytes) in a stag hunt game. *International Journal of Primatology*, *32*(6), 1296-1310.

Crawford, M. P. (1937). The cooperative solving of problems by young chimpanzees. *Comparative Psychology Monographs*,*14* , 1-88.

Dale, R., Range, F., Stott, L., Kotrschal, K., & Marshall-Pescini, S. (2017). The influence of social relationship on food tolerance in wolves and dogs. *Behavioral ecology and sociobiology*, *71*(7), 1-14.

Dobson, A. J., & Barnett, A. G. (2018). *An introduction to generalized linear models*. Boca Raton: CRC press.

Duguid, S., & Melis, A. P. (2020). How animals collaborate: Underlying proximate mechanisms. *Wiley Interdisciplinary Reviews: Cognitive Science*, *11*(5), e1529.

Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K., & Tomasello, M. (2014). Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1796), 20141973.

Field, A. (2009). *Discovering Statistics Using SPSS*. London: Sage Publications

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behavioral ecology and sociobiology*, *65*(1), 47-55.

Hall, K., Smith, M., Russell, J. L., Lambeth, S. P., Schapiro, S. J., & Brosnan, S. F. (2019). Chimpanzees rarely settle on consistent patterns of play in the hawk dove, assurance, and prisoner's dilemma games, in a token exchange task. *Animal behavior and cognition*, *6*(1), 48.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of theoretical biology*, *7*(1), 17-52.

Hare, B., Wobber, V., & Wrangham, R. (2012). The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Animal Behaviour*, *83*(3), 573-585.

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Goodwin, C.,Robinson, B., Hodgson, D. J., & Inger, R. (2018). Best practice in mixed effects modelling and multi-model inference in ecology. *PeerJ*, 6:e4794.

Hirata, S., & Fuwa, K. (2007). Chimpanzees (Pan troglodytes) learn to act with other individuals in a cooperative task. *Primates*, *48*(1), 13-21.

Horn, L., Virányi, Z., Miklósi, Á., Huber, L., & Range, F. (2012). Domestic dogs (Canis familiaris) flexibly adjust their human-directed behavior to the actions of their human partners in a problem situation. *Animal cognition*, *15*(1), 57-71.

Jensen, K., Call, J., & Tomasello, M. (2007). Chimpanzees are rational maximizers in an ultimatum game. *Science*, *318*(5847), 107-109.

Klöckner, A. (2014). Loo.py: transformation-based code generation for GPUs and CPUs. *Proceedings of ARRAY '14*: ACM SIGPLAN Workshop on Libraries, Languages, and Compilers for Array Programming. Edinburgh, Scotland.

Koster, J. M. (2008). Hunting with dogs in Nicaragua: an optimal foraging approach. *Current Anthropology*, *49*(5), 935-944.

Marshall-Pescini, S., Basin, C., & Range, F. (2018). A task-experienced partner does not help dogs be as successful as wolves in a cooperative string-pulling task. *Scientific Reports*, *8*(1), 1-12.

Marshall-Pescini, S., Cafazzo, S., Viranyi, Z., & Range, F. (2017). Integrating social ecology in explanations of wolf–dog behavioral differences. *Current Opinion in Behavioral Sciences*, *16*, 80-86.

Marshall-Pescini, S., Schwarz, J. F., Kostelnik, I., Virányi, Z., & Range, F. (2017). Importance of a species' socioecology: Wolves outperform dogs in a conspecific cooperation task. *Proceedings of the National Academy of Sciences*, *114*(44), 11793-11798.

Mason, W. A., & Hollis, J. H. (1962). Communication between young rhesus monkeys. *Animal Behaviour, 10*(3-4), 211-221.

Massen, J. J., Behrens, F., Martin, J. S., Stocker, M., & Brosnan, S. F. (2019). A comparative approach to affect and cooperation. *Neuroscience & Biobehavioral Reviews*, *107*, 370-387.

McCullagh, P. & Nelder J.A. (1989). *Generalized linear models*. London: Chapman and Hall.

McFarland, D. J. (1977). Decision making in animals. *Nature*, *269*(5623), 15-21.

Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science*, *311*(5765), 1297-1300.

Mitchell, W. A., & Lima, S. L. (2002). Predator-prey shell games: large-scale movement and its implications for decision-making by prey. *Oikos*, *99*(2), 249-259.

Naderi, S., Miklósi, Á., Dóka, A., & Csányi, V. (2001). Co-operative interactions between blind persons and their dogs. *Applied Animal Behaviour Science*, *74*(1), 59-80.

Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B. (2012). influence.me: Tools for detecting influential data in mixed effects models. *The R Journal*, 4:38–47

Noë, R. (2006). Cooperation experiments: coordination through communication versus acting apart together. *Animal Behaviour*, *71*(1), 1-18.

Ostojić, L., & Clayton, N. S. (2014). Behavioural coordination of dogs in a cooperative problem-solving task with a conspecific and a human partner. *Animal cognition*, *17*(2), 445-459.

Parrish, A. E., Brosnan, S. F., Wilson, B. J., & Beran, M. J. (2014). Differential responding by rhesus monkeys (Macaca mulatta) and humans (Homo sapiens) to variable outcomes in the assurance game. *Animal Behavior and Cognition*, 1(3), 215-229.

Péron, F., Rat-Fischer, L., Lalot, M., Nagle, L., & Bovet, D. (2011). Cooperative problem solving in African grey parrots (Psittacus erithacus). *Animal cognition*, *14*(4), 545-553.

Plotnik, J. M., Lair, R., Suphachoksahakun, W., & De Waal, F. B. (2011). Elephants know when they need a helping trunk in a cooperative task. *Proceedings of the National Academy of Sciences*, *108*(12), 5116-5121.

Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1992a). Comprehension of role reversal in chimpanzees: Evidence of empathy?. *Animal behaviour*, *43*(4), 633-640.

Povinelli, D. J., Parks, K. A., & Novak, M. A. (1992b). Role reversal by rhesus monkeys, but no evidence of empathy. *Animal Behaviour*, *44*, 269-281.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Range, F., Jenikejew, J., Schröder, I., & Virányi, Z. (2014). Difference in quantity discrimination in dogs and wolves. *Frontiers in Psychology*, *5*, 1299.

Range, F., Kassis, A., Taborsky, M., Boada, M., & Marshall-Pescini, S. (2019a). Wolves and dogs recruit human partners in the cooperative string-pulling task. *Scientific reports*, *9*(1), 1-10.

Range, F., Marshall-Pescini, S., Kratz, C., & Virányi, Z. (2019b). Wolves lead and dogs follow, but they both cooperate with humans. *Scientific reports*, *9*(1), 1-10.

Range, F., Ritter, C., & Virányi, Z. (2015). Testing the myth: tolerant dogs and aggressive wolves. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1807), 20150220.

Range, F., & Virányi, Z. (2015). Tracking the evolutionary origins of dog-human cooperation: the "Canine Cooperation Hypothesis". *Frontiers in psychology*, *5*, 1582.

Robinson, L. M., Martínez, M., Leverett, K. L., Rossettie, M. S., Wilson, B. J., & Brosnan, S. F. (2021). Anything for a cheerio: Brown capuchins (Sapajus [Cebus] apella) consistently coordinate in an Assurance Game for unequal payoffs. *American Journal of Primatology*, e23321.

Ruusila, V., & Pesonen, M. (2004). Interspecific cooperation in human (Homo sapiens) hunting: the benefits of a barking dog (Canis familiaris). In *Annales Zoologici Fennici* (pp. 545-549). Finnish Zoological and Botanical Publishing Board.

Schmelz, M., Duguid, S., Bohn, M., & Völter, C. J. (2017). Cooperative problem solving in giant otters (Pteronura brasiliensis) and Asian small-clawed otters (Aonyx cinerea). *Animal Cognition*, *20*(6), 1107-1114.

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2016). afex: Analysis of factorial experiments. R package version 0.16-1. Available at https://CRAN.R-project.org/package=afex

Smith, M. F., Leverett, K. L., Wilson, B. J., & Brosnan, S. F. (2019). Capuchin monkeys (Sapajus [Cebus] apella) play Nash equilibria in dynamic games, but their decisions are likely not influenced by oxytocin. *American journal of primatology*, *81*(4), e22973.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge University Press.

Spence, K. W. (1937). Experimental studies of learning and the higher mental processes in infra-human primates. *Psychological Bulletin*, *34*(10), 806.

Tassin de Montaigu, C., Durdevic, K., Brucks, D., Krasheninnikova, A., & von Bayern, A. (2020). Blue-throated macaws (Ara glaucogularis) succeed in a cooperative task without coordinating their actions. *Ethology*, *126*(2), 267-277.

Thorpe, W. H. (1956). *Learning and instinct in animals*. London: Methuen.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, *28*(5), 675-691.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly review of biology*, *46*(1), 35-57.

Vail, A. L., Manica, A., & Bshary, R. (2013). Referential gestures in fish collaborative hunting. *Nature communications*, *4*(1), 1-7.

Vale, G. L., Williams, L. E., Schapiro, S. J., Lambeth, S. P., & Brosnan, S. F. (2019). Responses to economic games of cooperation and conflict in squirrel monkeys (Saimiri boliviensis). *Animal behavior and cognition*, *6*(1), 32-47.

6. Supplementary Material:

Table S1: Extended Subject List

Dyad	ID	Role	Age	Sex	Breed	LVR	HVR	Dog Stooge	Carpet Stag	Position Stag First Game	Sausages				
											TT 1	TT 2	TT 3	TT 4	TT 5
1	Annie	Partner	8	F	Shetland Sheepdog	dry Food	2 sausages	yes	white	door	20	0	40	1	42
	Prim	Subject	5	F	Shetland Sheepdog	dry Food	2 sausages	yes							
2	Benny	Partner	7	M	Border Collie	dry Food	2 sausages	yes	patterned	back	0	40	20	40	8
	Dawin	Subject	5	M	Border Collie	carrot	sausage	yes							
3	Crash	Partner	2	M	Australian Sheperd	zuchhini	2 cheese	yes	patterned	back	20	40	0	39	39
	Alou	Subject	4	F	Australian Sheperd	salad	sausage	no							
4	Gandalf	Partner	1	M	Andalusian Hound	dry Food	sausage	yes	patterned	door	0	20	40	41	12
	Lenny	Subject	6	M	Canarian Warren Hound	zuchhini	sausage	yes							
5	Helsa	Partner	2	F	Hovawart	dry Food	sausage	yes	patterned	door	0	30	40	2	37
	Dora	Subject	9	F	Hovawart	dry Food	2 sausages	yes							
6	Josephine	Partner	4	F	Pumi	dry Food	sausage	yes	patterned	back	20	0	40	32	2
	Joseph	Subject	3	M	Pumi-Mix	dry Food	sausage	yes							
7	Leia	Partner	10	F	Beagle	dry Food	2 sausages	yes	white	door	0	40	30	42	33
	Eowyn	Subject	3	F	Mix	dry Food	sausage	no							
8	Matteo	Partner	1	M	Segugio Italiano	dry Food	sausage	yes	white	door	40	0	30	9	29
	Hetti	Subject	2	F	Mix	dry Food	sausage	no							
9	Mavie	Partner	2	F	Border Collie	sausage	dry Food	yes	white	door	40	20	0	2	42
	Cheynna	Subject	7	F	Australian Sheperd	dry Food	sausage	no							
10	Miley	Partner	11	F	Border Collie	dry Food	2 sausages	yes	patterned	back	20	40	0	40	35
	Tiara	Subject	7	F	Border Collie	zuchhini	2 sausages	yes							
11	Timo	Partner	2	M	Mix	dry Food	sausage	yes	white	back	40	20	0	40	38
	Mailo	Subject	5	M	Mix	dry Food	sausage	yes							

Table S2: Results of the General Linearized Mixed Models of the Assurance Game

Response Variable	Term	Estimate	SE	LRT	df	*p*-value	Stability Min	Max	CI 2.50%	97.50%
	Intercept	-0.829	1.954				-2.294	0.782	-5.164	3.375
	Partner choice (0, Hare ; 1, Stag)	5.24	1.365	11.72	1	**< 0.001**	4.078	6.489	2.625	9.412
	Session (0, Session 1 ; 1, Session 2)	-0.077	2.922	0	1	0.976	-2.408	1.135	-6.489	6.471
	Trial	0.808	0.673	1.44	1	0.231	0.388	1.087	-0.485	2.266
Subject choice	**Dropped Interactions**									
	Partner choice: Session: Trial					0.118				
	Session: Trial					0.138				
	Partner choice: Trial					0.213				
	Parter Choice: Session					0.175				
	Intercept	-0.316	1.759				-1.105	1.021	-3.943	3.485
	Previous Subject Choice (0, Hare ; 1, Stag)	4.055	1.554	4.9	1	**0.027**	3.379	4.991	1.016	7.746
	Session (0, Session 1 ; 1, Session 2)	3.2	2.427	0.58	1	0.447	1.769	4.288	-1.819	9.513
	Trial	-1.25	0.458	2.58	1	0.091	-1.653	-0.864	-2.561	-0.343
Partner choice	Previous Subject Choice: Session	-3.274	2.149	3.23	1	0.072	-4.222	-2.532	-8.859	0.828
	Previous Subject Choice: Trial	1.078	0.426	5.88	1	**0.015**	0.927	1.371	0.123	2.219
	Dropped Interactions									
	Previous Subject Choice: Session: Trial					0.164				
	Session: Trial					0.701				

Statistically significant p-values are indicated by bold letters.

Trial was z-transformed, for response variable subject choice original mean = 20.74, SD = ± 11.54;

for response variable partner choice original mean = 21.27, SD = ± 11.23

Table S3: Dyads Choices in the Assurance Game

Dyad	Stag-Stag	Hare-Hare	Hare-Stag	Stag-Hare	$\chi 2$	*p*-value
Annie-Prim	**40**	17	11	0	50.235	**< 0.001**
Benny-Dawin	**44**	21	4	10	47.228	**< 0.001**
Crash-Alou	**74**	2	0	4	194.8	**< 0.001**
Gandalf-Lenny	**49**	11	0	17	69.026	**< 0.001**
Helsa-Dora	**35**	**33**	5	5	43.231	**< 0.001**
Josephine-Joseph	**29**	**30**	2	11	31.66	**< 0.001**
Leia-Eowyn	**71**	0	1	7	178.77	**< 0.001**
Matteo-Hetti	**35**	**39**	0	4	63.641	**< 0.001**
Mavie-Cheynna	**39**	**40**	1	0	76.1	**< 0.001**
Miley-Tiara	**71**	2	1	5	177.76	**< 0.001**
Timo-Mailo	**75**	0	0	2	215.42	**< 0.001**

Choices above chance level are indicated by bold letters.