Department of comparative cognition
University of Veterinary Medicine Vienna

Head: Univ.-Prof. Dr. rer. nat. Huber

Weight discrimination in dogs

Masterthesis

University of Veterinary Medicine Vienna

Submitted from
Linda Elisabeth Kalb

Vienna, July 2023

The thesis was supervised by

Univ.- Prof. Dr. rer. nat. Huber Ludwig

Dr. rer. nat. Christoph Völter

Megan Lambert, PhD

External Reviewer:

Sarah Marshall Pescini, PhD

Statement

I hereby declare that this thesis was conducted only with the named literature and resources. All work was done by myself, and the work of others is marked as such.
Further, the present study was written by myself, and was not published anywhere else.

Vienna, 10.08.23 Linda Kalb

Acknowledgement

I am profoundly grateful to everyone who has played a role in the completion of this master's thesis. Their support, guidance, and encouragement have been invaluable throughout this journey.

First and foremost, I extend my heartfelt gratitude to my dedicated supervisors Univ.- Prof. Dr. rer. nat. Huber Ludwig, Dr. rer. nat. Christoph Völter and Megan Lambert (PhD). Your unwavering guidance, insightful feedback, and continuous encouragement have been instrumental in shaping the direction of this research. I also want to thank the whole unit of Comparative Cognition at the Messerli Research Institute of the Vetmeduni Vienna, which provided support with administration and IT. The expertise and willingness to invest time in my academic growth are truly appreciated.

I am also indebted to my parents, Margot and Dieter, who not only provided unending patience and belief in my abilities, but also played a pivotal role in my pursuit of this master's degree. Their insistence on my enrollment in this program was the initial push that set me on this rewarding academic journey.

To my loving boyfriend Philip, your unwavering support during the ups and downs of this journey have been my anchor. Your ability to bear with me at my lowest points and uplift me with your positivity have been my driving force.

A special mention goes to Karin Bayer, the esteemed lab manager of the Clever Dog Lab. Your expertise, patience, and willingness to share your knowledge have been pivotal in the successful execution of this research. Your guidance has been invaluable.

Heartfelt thanks to Gabi & Thomas. Your support provided me with the resources needed to carry out this research, but beyond that, your expertise and insightful conversations have added a unique dimension to my academic pursuits.

I would also like to extend my gratitude to all the research participants, both the diligent owners and their canine companions. Your willingness to participate and contribute to this study has enriched its findings and implications. Without your involvement, this research would not have been possible.

Lastly, I acknowledge the enduring support of my friends, well-wishers, and all those who stood by me during this challenging yet rewarding phase of my academic journey.

In conclusion, I am humbled by the collective effort and encouragement I have received, which have enabled me to bring this master's thesis to fruition.

Thank you all.

Linda Elisabeth Kalb

Abstract

Visual discrimination learning in dogs has been the subject of much research in recent years. However, little is known about the role of non-visual discrimination tasks or the underlying absolute or relative discrimination strategies. In this study, we conducted two experiments. The first experiment investigated the ability of dogs to discriminate weight, with a particular focus on disambiguating between absolute and relative learning strategies. The second experiment investigated the sensorimotor memory of dogs using an overshoot task which explores how individuals perceive and judge the weight of objects, specifically focusing on the phenomenon of overestimating the weight of objects. In Experiment 1, despite initial challenges, there was evidence of potential learning, as indicated by improved performance after excluding the first two sessions. However, within the time frame of this study, none of the dogs reached the predetermined learning criterion of 80% correct. Experiment 2 explored sensorimotor memory by investigating whether dogs adjusted lifting forces based on interactions with differently weighted objects. Dogs showed a significant tendency to lift their heads higher when interacting with lighter objects, suggesting an overshoot effect. Although the expected significant difference between the last heavy trial and the subsequent light trial was not observed, the results partially align with the prediction that dogs would lift their heads higher when lifting a lighter object than a heavier one. The results suggest that although dogs demonstrate some weight discrimination ability and sensorimotor adaptations, challenges remain. Recognizing limitations such as side bias, intrinsic motivation and experimental design, future research could refine methodologies by extending training periods, giving the dogs haptic pre-experience with the weights and increasing participant diversity by including dogs without former dummy training, to shed further light on the cognitive abilities of dogs in weight discrimination tasks.

Zusammenfassung:

Das visuelle Diskriminationslernen bei Hunden war in den letzten Jahren Gegenstand zahlreicher Untersuchungen. Wenig ist jedoch über die Rolle nicht-visueller Diskriminationsaufgaben oder die zugrundeliegenden absoluten oder relativen Diskriminationsstrategien bekannt. In dieser Studie wurden zwei Experimente durchgeführt. Das erste Experiment untersuchte die Fähigkeit von Hunden, Gewicht zu diskriminieren, wobei der Schwerpunkt auf der Unterscheidung zwischen absoluten und relativen Lernstrategien lag. Das zweite Experiment untersuchte das sensomotorische Gedächtnis von Hunden mit Hilfe einer Overshoot-Aufgabe, die untersucht, wie Individuen das Gewicht von Objekten wahrnehmen und einschätzen, wobei der Schwerpunkt auf dem Phänomen der Überschätzung des Gewichts von Objekten lag. In Experiment 1 gab es trotz initialer Schwierigkeiten Anzeichen für potenzielles Lernen, wie die verbesserte Leistung nach Ausschluss der ersten beiden Sitzungen zeigte. Allerdings erreichte keiner der Hunde während des Untersuchungszeitraums das vorgegebene Lernkriterium von 80 % Richtigkeit. Experiment 2 untersuchte das sensomotorische Gedächtnis, indem es untersuchte, ob die Hunde ihre Hebekräfte aufgrund von Interaktionen mit unterschiedlich schweren Objekten anpassten. Die Hunde zeigten eine signifikante Tendenz, ihren Kopf höher zu heben, wenn sie mit leichteren Objekten interagierten, was auf einen Overshoot-Effekt hindeutet. Obwohl der erwartete signifikante Unterschied zwischen dem letzten schweren und dem darauffolgenden leichten Versuch nicht beobachtet wurde, stimmen die Ergebnisse teilweise mit der Vorhersage überein, dass Hunde ihren Kopf höher heben, wenn sie ein leichteres Objekt anheben als ein schwereres. Die Ergebnisse deuten darauf hin, dass Hunde zwar eine gewisse Fähigkeit zur Gewichtsdiskrimination und sensomotorischen Anpassung zeigen, aber Herausforderungen bestehen bleiben. Unter Berücksichtigung von Einschränkungen wie Seitenpräferenzen, intrinsischer Motivation und Versuchsaufbau könnten zukünftige Forschungsarbeiten die Methoden verfeinern, indem sie die Trainingszeit verlängern, den Hunden eine vorherige haptische Erfahrung mit Gewichten anbieten und die Vielfalt der Teilnehmer durch die Einbeziehung von Hunden ohne vorheriges Dummy-Training erhöhen, um die kognitiven Fähigkeiten von Hunden bei Gewichtsunterscheidungsaufgaben weiter zu untersuchen.

Contents

1. Introduction

1.1 Weight in general

Humans have an advanced ability to causally understand the coherences of their physical world. These competencies, such as reasoning about invisible properties like weight, enable us to solve complex problems. Humans can use a variety of cues to infer weight without prior handling: for example, using prior experiences with objects of different weights as a reference (a feather is lighter than a piece of wood, or observing others interacting with objects of differing weights (Povinelli, 2012). When it comes to lifting objects, the lifting force applied must be scaled according to the objects' weight (van Polanen & Davare, 2019a). While the weight of an object is not perceivable before the object is lifted, humans are able to estimate the weight and the needed force to lift the object, by using cues such as those mentioned before (Povinelli, 2012). While humans undoubtedly have sophisticated abilities to infer the weight of objects, little is known about whether and how animals can discriminate different weights in the absence of any visual cues.

Instead, most studies to date have focused on a more sophisticated understanding of properties such as weight in animals, and the results are not conclusive. Research in this field is needed because it can shed light on how different animals understand their physical environment. An example of this is the study of Schrauf et al. (2008), in which capuchins were presented with three tools differing only in weight, to examine whether they would learn to select the more effective tool (the heavier one) to crack up nuts. For this study, the capuchins were offered 3 differently weighted aluminum boxes as tools, to crack nuts and gain access to the kernels. As the nuts were selected such that they were too hard to be cracked by their teeth, the capuchins needed to find the most effective (the heaviest) tool to crack the nuts. To assess the weight of the different tools the experimenter needed the capuchins to lift every single tool prior to making a choice. To encourage the capuchins to lift all the boxes before scoring a choice, chocolate cream was smeared on them to make sure they always lift all three tools, before selecting. The results suggest, that capuchins are not only able to discriminate weight and select the most effective tool to help them foraging, but also that this selectivity helps them to increase their tool using effectiveness (Schrauf et al., 2008). Another study by Hanus and Call (2008) investigated if chimpanzees can use causal cues like the effect differently weighted objects can have on a balance scale (heavier objects are lower). For this purpose, the researchers baited two opaque cups on a balance scale with a slice of a banana. The experimenter showed a piece of banana to the subject. After placing the piece of banana behind the experimenter's back and from hand to hand in a kind of blind shell game, the experimenter placed both closed hands into the opaque cups. After opening both hands, the balance falls to the side with the piece of banana inside, cueing an attentive viewer where the reward is hidden. In Hanus and Call's (2008) experiment, the chimpanzees selected the baited cups significantly over chance. Additionally, experimenters supplemented their experimental set-up with an edition in which the scales were tilted on one side and could not be moved. When repeating the experiment now (with the only change in procedure that the bait was always placed in the lowered cup of the tilted balance scale), the experimenters found out, that this condition was way harder for them to learn. Hanus and Call (2008) set up an additional control condition to find out if chimpanzees have an intrinsic preference for the cups moving downward, by introducing a delayed balance condition in which the set up was exactly the same as in the beginning except the balance scale was manipulated so that it would not tilt when releasing the bait in one of the cups, but instead after 4 seconds the experimenter would mechanically tilt the scale on the baited side. The chimpanzee's performance decreased, suggesting that the chimpanzees were using causal effects of weight as a cue, using the tilt of the scale or cup position as a cue only when the weight of the food was the cause of the tilt or cup position (Hanus & Call, 2008).

Among birds, studies reveal mixed results within species. Neilands et al. (2016) aimed to study problem-solving by insight in New Caledonian crows and presented them with an apparatus, where a heavy object (stone) needed to be dropped onto a platform, to receive a food reward. Even though the crows received initial object handling training, the results indicate that the subjects did not selectively drop in heavier objects to collapse the platform and could therefore not pass the task. As humans are capable of acquiring information about objects through observing their physical interactions in the world, another project by Jelbert et al. (2019) aimed to find out if wild-caught New Caledonian crows can distinguish between heavy or light objects, moved by an electric fan without handling the objects directly. In the first part of the experiment, the crows were trained to pick up and drop objects into a tube. While the one group was rewarded for dropping a light object into the tube, the other groups was rewarded for picking up and dropping a heavy object into it. After successful training phase, the crows were presented with two new objects (a heavy and a light one), tied to a string and out of reach. The experimenter then turned on a fan, causing the lighter object to move in the breeze. Afterwards, the objects got untied and presented to the New Caledonian crows to make a choice which object they will choose to drop in the tube. The crows were rewarded in line with the rule they learned in the training phase (heavy or light). Results showed, 11 of 12 birds touched the correct object first in the experimental condition and performed at chance in a control condition (with the fan off, and objects unmoving), leading the authors to argue for a higher-order casual reasoning about weight in New Caledonian crows (Jelbert et al., 2019).

These studies provide a mixed picture about whether animals possess a more sophisticated understanding of weight; however, before asking further questions along these lines, it is important to examine the basic ability to discriminate weight. Few studies on this topic exist, and those that do focus largely on primates, suggesting that even weight discrimination alone is difficult. A project by McCulloch, (1941) aimed to find out if chimpanzees can discriminate the heavier box in a set of two, and what the smallest noticeable weight difference would be. In the preliminary training phase, five chimpanzees first learned to put a box on a shelf and then to select the heavier of two boxes. In the beginning, the boxes had a large weight difference (80 vs. 480/640 grams) but as soon as they reached criterion, the weight of the heavier box was decreased to see what the smallest noticeable difference for them would be (average differential limen = 7.3 grams). To reach criterion, the chimpanzees needed a median of 1179 training trials to accomplish the initial weight discrimination training. In a study by Schrauf & Call (2009), researchers investigated and directly compared the results of achromatic color and weight discrimination tasks in bonobos, gorillas and orangutans. To verify their hypothesis, the superiority of visual over kinesthetic information, they set up an exchange paradigm in which apes had to swap the correct objects with a human experimenter to obtain a reward. They tested 12 subjects, from which 5 individuals performed above chance level in the weight discrimination task and 6 individuals in the color discrimination task. Supporting their hypothesis, the apes needed about 30 sessions to reach criterion (6 consecutive choices correct) in weight discrimination tasks, but only a median of 6.5 session were needed to reach criterion in the color discrimination task. Aligning with McCulloch (1941) and Schrauf & Call (2009), it took chimpanzees nearly 900 trials in a sorting task, to learn to place visually identical objects in colored trays according to weight (Povinelli, 2012). While weight discrimination tasks seem to be surprisingly difficult for primates, a more recent project from Lambert et al. (2021) investigated how Goffin's Cockatoos perform in weight discrimination tasks. The researchers trained the Cockatoos to pick up two small balls differing either in color & weight or solely in weight and afterwards form a decision by using their prior proprioceptive experience with the balls and drop the assigned ball into a specific tray. Results showed that Cockatoos learn weight discrimination faster when additional color cues are used, but they can also learn to discriminate between visually identical objects based on weight only, faster than primates. The Cockatoos without prior handling experience with the colored balls needed a median of 60.6 trials to reach criterion. The Cockatoos with prior experience only needed a median of 40.8 trials to meet criterion in the color weight task. These findings suggest that weight discrimination is an ability possessed not only by primates, but also by other taxa. In fact, these results also show that cockatoos learn weight discrimination maybe more quickly and more easily (even though methodological differences in the study designs preclude direct comparisons to the previous primate studies). Therefore, it seems worthwhile to look at other species as well.

1.2 Absolute and relative learning concepts in weight discrimination

In addition to the question whether animals can discriminate objects based on weight, a related (and thus far unexplored) question is what rules they use to make these discriminations. While there is not much research about weight discrimination, the ability to discriminate quantities has been a focus of several research projects (Bortot et al., 2019; Honig & Stewart, 1989; Miletto Petrazzini et al., 2016). Discriminating quantities is thought to be beneficial for many species, to decide whether or not to engage in an intergroup conflict (Range et al., 2014), or which branch has more ripe berries on it. Studies show that animals can be trained to discriminate between two quantities, and while it is still under debate which decision criteria the animals use to do so, it seems like there is a shared system for quantity discrimination across all vertebrate species (Bortot et al., 2019). To discriminate two quantities a subject can either select a specific (absolute) number of items or by applying a more general (relative) numerosity rule (Miletto Petrazzini et al., 2016; Bortot et al., 2019). Another contributing feature to the shared system across vertebrates is the ratio-dependent discrimination of numerosity, which describes the fact, that relative discrimination performance decreases as soon as the numerical ratio between the smaller and the larger quantity increases. Humans, for example, are better at discriminating 9 vs. 18 (0.5 ratio) than 7 vs. 11 (0.63 ratio) (Agrillo et al., 2015). This ratio can be traced back to Ernst Weber, founder of the field psychophysics, who discovered humans' ability to notice that an increase of a stimulus' magnitude is proportional to the change in the original stimulus's magnitude (Povinelli 2012). For humans and most other vertebrates, it seems that there is a preference towards selecting a relative number of items rather than absolute numerosity. In comparison, Bortot et al. (2019) investigated if this preference can also be found in free flying honeybees and aimed to find out if honeybees rather use absolute or relative numerosity for quantity discrimination. In their experimental set up, Bortot et al. (2019) trained 24 bees to choose variable images containing three dots. First, the bees were familiarized with the Y-maze and learned to search for and collect a sucrose solution as a reward. In the training phase, while one group of bees learned to distinguish 3 from 2 (group "larger") the other group was trained to discriminate 3 from 4 dots (group "smaller"). Each bee had 32 consecutive trials in which only the 3 dots were rewarded while the alternative stimulus was punished by offering quinine solution. The test phase consisted of two conditions: In the training-like condition the bees were presented with the same numerical discrimination as in the training condition but with new item configuration. In the numerical condition, bees from group “larger” (trained to discriminate 3 from 2) had to discriminate 3 from 4 and bees from group “smaller” (trained to discriminate 3 from 4) had to discriminate 3 from 2. Results showed, bees were more likely to use an absolute numerosity rule because in the numerical tests, both bee groups always preferred stimuli with three items, regardless of whether they had been trained to choose the smaller or the larger quantity (Bortot et al., 2019). Therefore, the researchers concluded, that bees differ from vertebrates in the usage of absolute rather than relative numerical values. While the concepts of absolute and relative learning are well researched around quantities, so far, no studies investigated absolute and relative weight discrimination learning.

1.3 The overshoot effect

Outside of weight discrimination, studies with both humans and great apes have examined whether individuals form sensorimotor expectations about the weight of objects after repeated interaction with them. When humans manipulate objects, we use visual cues to recall an internal representation of the object, which helps to predict the weight of an object and therefore adjust the effort to lift it. After some repeated experience with similar or the same objects, a “long-term force profile” (Povinelli, 2012) is built, from which expectations about an object’s weight are derived. This long-term force profile helps improve the manipulative performances and helps to adjust the effort to lift an object (Sirianni et al., 2018). Gordon et al. (1993) could show that humans are able to form “(…) internal motor commands that are scaled to the weight of familiar objects (soda cans, telephone books, etc.).” This internal motor commands derive from the visual identification of the object and are stable across time. Nevertheless, Gordon and colleagues (1993) found an easy way to manipulate this force-profile by increasing the density of a common every-day object the participant is already familiar with. When the object's weight differs from the internal expectation, human tend to not apply enough force-to-lift and fail to pick the object up within the first try. This works with an object denser than expected, but also with an object which is lighter than expected. Johansson & Westling (1988) for example showed in another experiment, that humans adjust their effort-to-lift also with unfamiliar objects that are lighter than expected leading to the so-called overshoot effect. The experimenters asked the participants to lift an unfamiliar object several times and then exchanged the object with a visually identical, but significantly lighter object. Results showed that the participants used the information from the previous lifts (when the object was heavier) to adjust their lifting force, a positional overshoot effect is created. While humans quickly adapt to the unexpected weight of an object and adjust their force-to-lift for the next trial, those experiments show that the lifting effort is planned before feedback of the proprioceptive system is obtained.

While most of the experiments concerning internal representations about weight focused on humans, Povinelli (2012) conducted an experiment in which he trained 7 chimpanzees to vertically lift a 7.0 kg box from a platform and measured the maximum height to which the apes lifted the objects. After 4 trials with the heavy box, the experimenter exchanged the box with a lighter one (1.5 kg) and asked the chimpanzees again to lift it. He measured the maximum height of the lifted box. Results showed that subjects lifted the boxes higher in trial 5 than in trials 1-4, indicating an acceleration overshoot effect in trial 5 and suggesting stable sensorimotor representations of the effort needed to lift an object based on information received from previous experience (Povinelli 2012). Sirianni et al. (2018) for example, aimed to find out if chimpanzees of the Taï forest, known for using tools to crack nuts, use long-term force profiles to anticipate the weight of a nut-cracking hammer from its size. They tested this by presenting the chimpanzees a natural wooden hammer and an artificially hollowed, lighter hammer with the same visual appearance. The researchers then compared the initial lifting accelerations for the hollowed and natural hammers, finding out that the chimpanzees apply a similar force when lifting both hammer types for the first time, which leads to a higher acceleration for the hollowed hammers (overshoot effect) but also, that the overshoot effect will fade away after repeated interaction with it (Sirianni et al., 2018).

1.4 Weight perception in dogs

While weight discrimination can be seen as a logical precondition in some species, for dogs the evolutionary benefit is not easy to spot. But since using long-term force profiles has been proposed to help "ensuring a stable grasp, determining speed and precision in execution and producing a smooth lift" (Sirianni et al., 2018) there are possible ecological circumstances which explain why weight discrimination might also be important for dogs.

Canine predatory sequence consists of several motor pattern such as for example : stalk – chase – grab-bite – kill-bite – dissect – consume (Spady & Ostrander, 2008). Grabbing and carrying prey (over potentially long distances) is already part of the natural canine behavioral repertoire, which could have modified their ability to pay attention to the weight of an object. Additionally, domestication and special breeding strategies produced modern hunting dogs which exhibit refined traits adapted for different aspects of hunting like pointing, chasing but also retrieving (Spady & Ostrander, 2008).

In our pilot study from 2020, we used the innate trait "retrieving" in an experimental set-up, in which dogs from the FCI group "Retriever" retrieved objects of different weights. For this experiment we used a similar exchange paradigm like Schrauf and Call (2009), training dogs to retrieve either a heavy or a light object (i.e., dummy) in order to receive a food reward. Like Schrauf and Call (2009), we also compared weight discrimination with achromatic color discrimination. For both tasks, the dummies of the same shape and size were used. In the weight discrimination group, dummies only differed in weight, but had the same color (white), shape, and size. The light dummy weighed around 120 grams while the heavy one weighed about 500 grams and was filled up with water. For the achromatic color discrimination group, the same dummies were used but they differed in color (black and white & white). Out of 11 dogs, there was only one individual in the weight discrimination group that performed significantly above chance, indicating that she might have learned to discriminate the dummies based upon their weights. This individual was also the only one to show increasing dummy switches across session, which are necessary to assess the weight in the first place in order to make a choice. In the color discrimination group there was also one individual which performed above chance level, but due to limited sample size and restrictions due to COVID-19, the results were inconclusive and further investigations are needed.

The aim of the current study was to evaluate and compare how dogs perform in weight discrimination tasks and, more specifically, which learning strategy they use to do so. For this purpose, we followed again the methodology of Schrauf & Call (2009) using an exchange paradigm, but modified it by setting up a retrieval exercise. The dogs were divided into two groups and trained to only retrieve their target dummy (group heavy: heavy target = 364g vs. 200g, group light: light target = 661g vs. 1202g). After training phase, we planned to test all dogs in the same condition, offering them both heavy and light target objects (364g vs. 661g). Now in this test phase, the heavier target from training group heavy is the lighter one in the test condition and the other way round. The results of the test phase might be indicative about whether they use absolute or relative concepts to discriminate the targets. If dogs assigned to group *heavy* fetch again the heavier dummy and ignore their former absolute target weight in the test condition, we can assume they used a relative discrimination rule. If dogs from group *heavy* fetch the lighter one in the test condition, it can be assumed, that the dogs use absolute discrimination learning because the dogs stick to their original test weight. The same rules apply also for the dogs from group *light*. If dogs from training group *light* decide to stick with their absolute target weight in the test condition, they use absolute learning strategies. We hypothesized that dogs can learn to discriminate weight and, based tentatively on studies of absolute versus relative discrimination in vertebrates, that they represent weight in relative terms.

Since all our dogs were well-trained dummy dogs, various difficulties arose. For example, the dogs learn during their training that under no circumstances should they drop an already selected dummy to pick up another one. This learning experience made it almost impossible for us to get the dogs to pick up the two different dummies to get a feel for the weight. In addition, it became apparent that for some dogs it is crucial which side of their handler they are on when they are asked to retrieve one of the two dummies. For example, if a dog is sitting on the left side of the handler and the handler asks the dog to retrieve a dummy, the dog is likely to retrieve the left dummy. Unfortunately, these subtleties have too much to do with the individual learning history to make them controllable, so we have come up with an alternative set up where the dogs do not have to switch dummies.

We conducted another, even more basic experiment, investigating if dogs form implicit expectations about the weight of an object. Therefore, we built an apparatus in which we placed the heaviest dummy form experiment 1 and trained the dogs to lift the dummy off the apparatus (see Figure 5). After several trials with the heavy object, we exchanged it with a significantly lighter one and recorded the reaction of the dogs and measured the maximum lifting height in every trial. We hypothesized, that dogs form implicit expectations about weight and plan their lifting force based on the previous object, which will lead to an overshoot effect. Accordingly, we predicted that dogs would lift their head higher in light, than in heavy trials.

2. Experiment 1: absolute and relative rules in weight discrimination learning

The aim of the current study was to evaluate and compare how dogs perform in weight discrimination tasks and if they use an absolute or relative rule to discriminate weight. We adapted an exchange paradigm (Schrauf and Call, 2009) for use with dogs, in which they had to discriminate a dummy with a specific weight against another one and retrieve it, in order to gain a reward.

2.1 Methods

2.1.1 Subjects

We originally planned to test 24 dogs of different breeds, living in and around Vienna. Participants were recruited using social media (Facebook and Instagram) as well as the Clever Dog Lab's database. Due to the ongoing pandemic situation, participation and recruiting was more difficult than usual, therefore only 14 dogs participated in Experiment 1. Like in our pilot study with there were 8 males and 6 females with ages ranging from one to eleven years (mean age = 5.5 years). All of our subjects lived indoors with their caretakers. Some of the dogs had already participated in other experiments at the Clever Dog Lab but none had ever been tested in a weight discrimination task. The dogs were chosen based on their preexisting ability to retrieve objects for example dummies in exchange for rewards which helped to minimize time for prior training as this would go beyond the scope of this project. The training status of the subjects varied but 9 out of 14 dogs were professionally trained in dummy work. All dogs participating were required to be above a shoulder height of 35 cm to ensure that they were physically capable of lifting the heaviest (1202g) dummy.

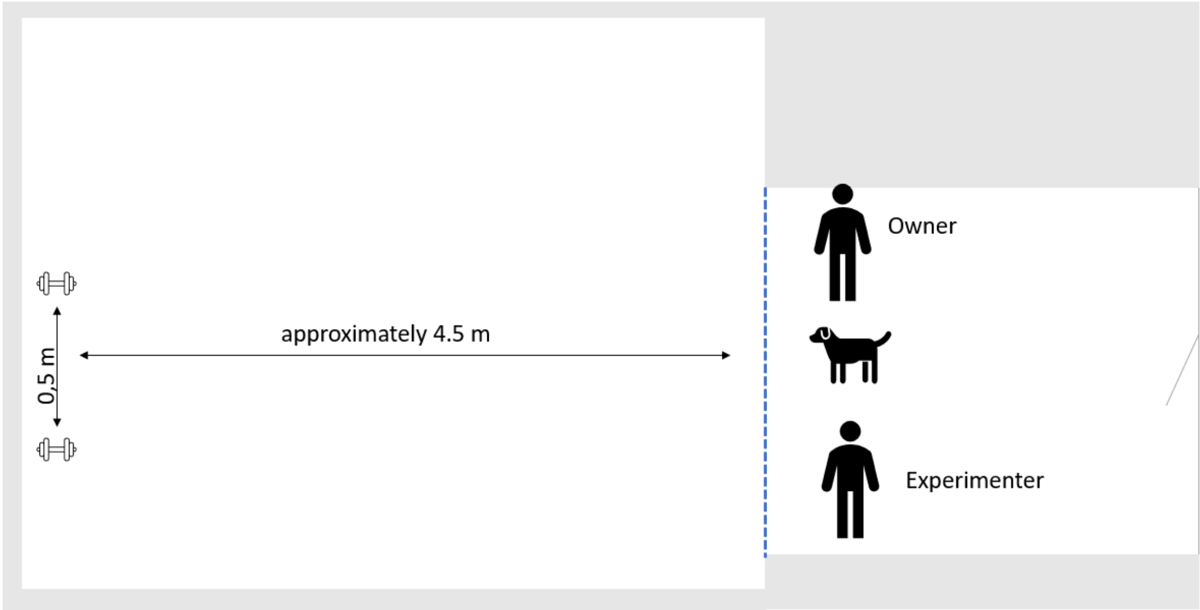

Figure 1: Experimental set up. The experimenter, the owner and the dog are at the right side of the room. After every retrieval, the dog is asked to return to his position between experimenter and owner and wait until the next set of dummies is set up by the experimenter. A choice is scored when the dog passes the dashed blue line, carrying a dummy.

2.1.2 Set-up

The experiment was carried out in a test room of the Clever Dog Lab, which was equipped with a camera system. At the beginning of each test session, the dog was positioned between the experimenter and the owner on a marked position (see Figure 1). To prevent the owners from giving social cues (like gazing or pointing towards the target object) they were instructed to look towards the wall.

Figure 2: Dummies used for experiment 1 & 2.

2.1.3 Objects

All dogs were offered visually identical dummies that only differed in weight, but had the same (white) color, shape, and size (see Figure 2). In order to achieve the desired weight, the dummies were manipulated by cutting and filling them with a mixture of memory foam and lead shot covered by a PVC pipe.

According to the Weber-Fechner-Law, which helps quantifying the perception of a change in a given stimulus, humans rather attune to relative differences than to absolute differences (Povinelli 2012). We decided on a ratio of 11:20 for the weights in both groups of Experiment 1. The ratio was chosen due to the restrictions imposed on the setup by the weights of the lightest dummy (i.e., dummy only filled with foam = 200g) and of the heaviest dummy (1202g) that the dogs could carry without problems. The weights of the dummies used in our experiment were **200g, 364g, 661g, 1202g**.

2.1.4 Procedure and design

2.1.4.1 Training phase Procedure A

Prior to testing, all 14 dogs underwent a training phase to evaluate whether they could reliably discriminate the target object based on its weight within the maximum number of 12 sessions with 30 trials each. The dogs were pseudo-randomly allocated into two different groups: group “heavy” and group “light”. In the first part of the experiment, for dogs from group heavy, the target object was the heavier dummy (**364g** vs 200g); for group light the target object was the lighter dummy (**661g** vs 1202g). We planned to run a maximum of 12 training sessions with both groups or until they met criterion and reliably discriminated the target weight. The criterion included that the dogs’ performance had improved significantly across session and that they reached a final performance of ≥80% correct within a session. If dogs had not met criterion after 12 sessions, they would have been excluded. Before every session we conducted 2 warm-up trials with a non-target toy to get the dogs familiarized with the experimenter and the set-up, but also to establish a fetching game. After 2 successful warm-up trials (requiring object retrieval within 30 seconds of the release), the training trials were administered. In the training condition, the experimenter placed the two visually identical dummies on the other side of the room (approximately 4.5 meters away from the dog) on prior marked positions (0.5m apart of each other) and returned to her position afterwards (see Figure 1). Subsequently, the dog was asked to retrieve one of the two objects by using an individual, verbal command which was already known to the dog. As soon as the dog crossed the dashed blue line with an object, a choice was scored. As the experimental set up was designed as an exchange paradigm, whenever the dog retrieved the target object, the dog was rewarded with verbal praise and treats. We used dry food (Belcando®) for rewarding the dogs whenever they made a correct choice. Whenever the dog fetched the wrong object or did not cross the score line, the experimenter did not provide any direct feedback to the dog but just took the object and started another trial. The testing was carried out between April and December 2021 in an indoor testing room of the Clever Dog Lab in Vienna.

2.1.4.2 Training phase procedure B

As none of the dogs met criterion or improved their performance over session and trial number, we decided to modify the setup of our experiment slightly to help the dogs understand the task. The professional trained dummy dogs (retriever type dogs) in particular struggled to score correct choices, so we administered a change in the routine for the dogs by adding two sessions in which the dogs were offered only the target dummy for ten trials and rewarding them for retrieving it. After ten trials with the target dummy, we went back to the normal routine, offering the dog both dummies, expecting them to perform better due to the successful ten trials with target weight beforehand.

2.1.4.3 Test phase: absolute and relative learning rules

Every dog that would have reached criterion in the training phase was meant to be tested in an absolute and relative weight discrimination task. In this phase, all dogs from both groups would have been tested under the same condition. In the test phase, two visually identical dummies with the target weights from both training groups (**364g** vs**. 661g**) would have been offered to the dogs in the same way as in the training. We predicted if dogs with the heavy training target (**364g** vs. 200g) choose the heavier dummy in the test session (364g vs. 661g), that they use relative learning rules. If dogs with the heavy training target would choose the lighter dummy in the test phase (364g vs. 661g), they were using absolute learning rules. If the dogs assigned to the lighter training target (**661g** vs. 1202g) preferred and retrieved the lighter one in the test session, they were using relative learning rules. If the dogs with the lighter training target would choose the heavier target, they use absolute learning rules.

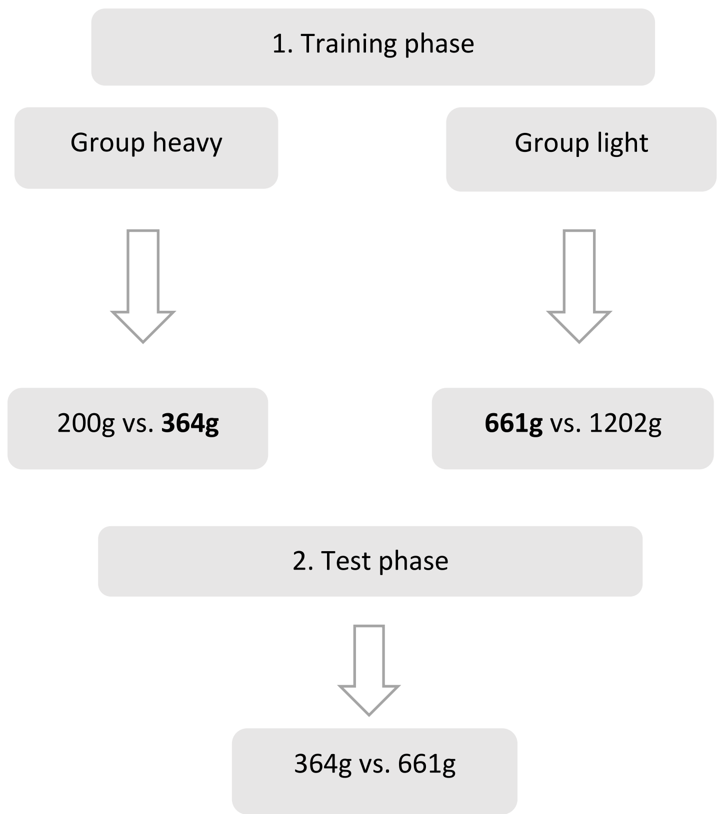

Figure 3: Group composition. In the trainings phase, the dogs were pseudo randomly allocated to either group heavy or group light. Dogs from group heavy were only rewarded for retrieving the 364g dummy and dogs from group light were only rewarded for retrieving the 661g dummy. In the test phase all dogs were offered their former target dummies.

2.1.5 Analysis

As none of the dogs met criterion in the training phase, we analyzed the dogs' training performance to determine whether this improved across sessions. We fitted a generalized linear mixed model (GLMM) with binomial error structure. We analyzed whether the dogs chose the correct or incorrect dummy (coded as 1 and 0). We included the session number and trial number as continuous predictor variables. Before including the predictor variables, we centered them to a mean of 0 and scaled the variable to a standard deviation of 1. We also included the random effect of subject ID as well as the random slopes of session number and trial number within subject ID. Inferences with respect to the fixed effects were drawn using likelihood ratio tests using the R function drop1. The GLMM was fitted with the r function glmer of the package lme4 (Bates et al., 2015). We used parametric bootstraps to calculate the 95% confidence intervals. We checked the model stability by removing one subject at the time and by inspecting how much the estimates varied. This procedure showed that the model was stable with respect to the fixed effects.

To analyze whether the dogs overall performed significantly different from the chance level of 0.5 we conducted one-sample t-tests based on aggregated data across all sessions but also after excluding the first two sessions.

2.2 Results

Due to the restrictions imposed by the Sars-Cov2 pandemic our maximum number of 12 training sessions per subject was not reached. The dogs completed on average 4 sessions (range: 2-

9) None of the dogs reached criterion (final performance of ≥80%) in the training phase. Overall, the dogs' proportion of correct choices was 0.53 +/- 0.02 (mean +/- SE). Their performance did not significantly deviate from the hypothetical chance level (t (13)=1.46, p=0.169). When we excluded the first two sessions, the dogs' mean performance was significantly above chance (t = 3.08, df = 6, p= 0.022).

To examine whether the dogs' performance improved across trials within session and across session we fitted a GLMM with binomial error structure and we included the predictor variables session number and trial number. Neither session number (χ^2=1.42, df=1, p=0.233) nor trial number (χ^2=0.49, df=1, p=0.484) had a significant effect on the dogs' performance.

Table 1 Results of the GLMM.

	Estimate	**SE**	**Lower CI**	**Upper CI**	**X^2**	**df**	**p**
Intercept	0.16	0.07	0.03	0.3			
Session	0.18	0.1	-0.04	0.39	1.42	1	0.233
Trial	-0.05	0.07	-0. 19	0.09	0. 49	1	0.484

For procedure B, we conducted binominal tests at the individual level to check if it helped to improve the performance of the dogs but only one individual performed significantly above chance: (25 correct choices out of 32 trials, p= 0.002).

A logical precondition for discriminating objects of different weights is to lift the objects before making a choice. Due to the setup, it is only possible for the dogs to assess one weight at the time. Therefore, the number of dummy lifts is a crucial aspect of the task. We predicted that dogs which show multiple dummy lifts per trial have better performance than dogs which do not switch between and lift both dummies. But when comparing dogs that lifted the dummies multiple times in at least 5 trials to dogs that lifted multiple dummies less often there is no significant difference in performance: independent-samples t-test: $t = -0.17$, $df = 5.86$, $p = 0.875$. When comparing the dogs' performance (only considering dogs that switched dummies in at least 5 trials (N=6) in trials in which they lifted multiple dummies to trials in which they lifted only one dummy, there was no significant difference either: $t = -1.11$, $df = 5$, $p = 0.318$

None of the dogs met criterion but the group-level results when excluding the first two sessions suggested that some individuals started learning to discriminate the weights. Since none of the subjects passed the training phase, the second phase (testing absolute and relative learning rules) of the experiment could not be carried out. We therefore decided to conduct another experiment with the dogs to examine implicit expectations about weight.

3. Experiment 2: overshoot experiment

Due to the fact, that none of the dogs in Experiment 1 met criterion, we decided to conduct another experiment. The overshoot experiment examines whether dogs have implicit expectations about the weight of objects. In contrast to Experiment 1, the dog did not have to learn to differentiate between objects. This experiment is primarily investigating whether dogs can acquire certain expectations concerning the weight of an object. For this purpose, we orientate ourselves on overshoot experiments from human psychology, whereby humans adjust their lifting efforts to the known weight of an object, such that when the object is suddenly lighter than expected, they lift it with greater force than necessary.

3.1 Methods

3.1.1 Subjects

In Experiment 2 our sample consisted of the same 14 individuals of Experiment 1 but two individuals were excluded, resulting in 12 individuals. The dogs were excluded because they could not lift the dummy out of the apparatus in any session. The other dog (6 males and 6 females) aged between one and eleven years (mean age: 5.5 years).

3.1.2 Setup

The overshoot experiment was conducted in the same room as Experiment 1 but used a different set-up. For Experiment 2, we used an apparatus which held the dummy at a specific height and position. The apparatus (Fig. 4B) was placed between the wall of the room and a fence to allow the dog to approach the dummy only from one side. We used this apparatus so that the dogs could easily lift the dummy starting at a predetermined height. The experimenter sat opposite from the dog and the apparatus but behind the fence (Fig. 4 A). The heaviest (1202g) and lightest (200g) dummy of experiment 1 were used in this experiment. The camera, as well as the chair, the fence and the apparatus stood in predetermined positions. The fence was about 1.08m high and 1.52m long and was positioned 0.7m from the wall. The apparatus itself was 0.5m high and was screwed on a wooden platform (0.5m x 0.5m) to make sure, that the dogs could not move it.

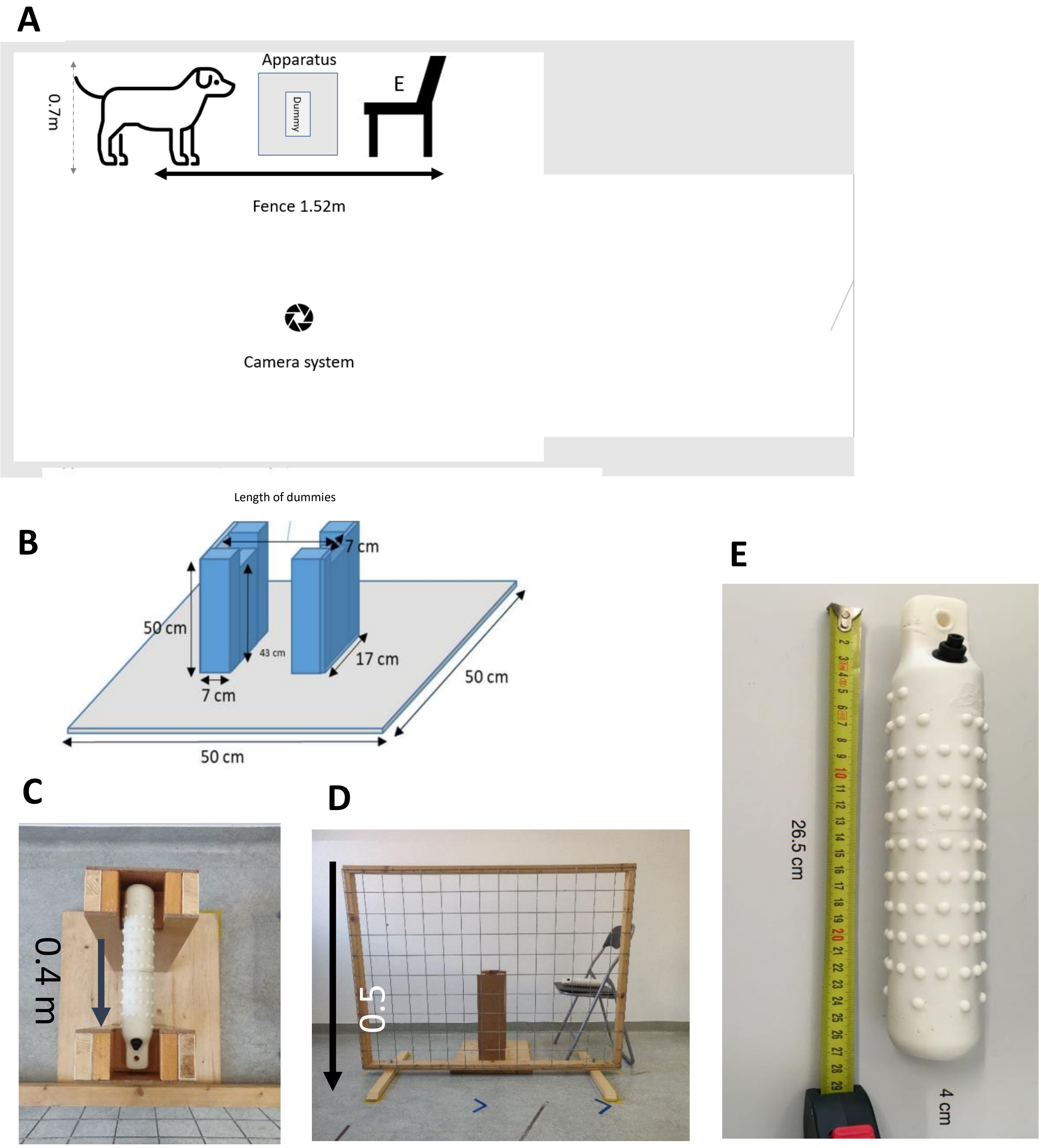

Figure 4 ***A:*** *As the overshoot experiment is conducted in the same room as experiment 1, we decided to place the apparatus in a way that it is between two barriers (wall and fence) in order to help the dogs to approach the dummy only from one side. The experimenter sits opposite to the apparatus and the dog, between the fence and the wall.* ***B****: The apparatus built for the overshoot experiment is a 50 x 50 x 50 cm wooden platform with blocks holding the dummy, allowing the dogs to lift the dummy only in a specific way.* ***C:*** *Close up of the wooden blocks holding the dummy.* ***D:*** *The apparatus and the experimenter are placed behind a fence which allows the dogs to approach the dummy only from the left side. Additionally, the fence grid helps as a reference for the maximal lifting height.* ***E:*** *The dummies we used were the same as in Experiment 1, with a length of 26.5 cm and a width of 4 cm*

3.1.3 Procedure

Each dog had two sessions in which they completed two different conditions (condition A and B) of 15 trials each. The dogs were divided into two groups to counterbalance the order of conditions across dogs. In both conditions, the dogs had to lift the dummies a total of 15 times. In condition A, the dogs had to lift the heavy dummy (**1202g**) out of the apparatus *9* times, then a lighter (but visually identical) one (**200g**) 5 times and finally the heavier one again. In condition B the order of the lifts was slightly different. To begin with, the dogs had to lift the heavy dummy *8* times, then the lighter one 5 times and finally the heavier one twice. This procedure was designed in order to ensure that the observed effect is due to the different dummy weights, not the trial sequence.

At the beginning of each session, the experimenter, the dog owner and the dog entered the room, and the dog could freely examine and get familiar with the room and the set-up (duration: 2 minutes). Afterwards the dog was called to the area between the wall and the fence with a verbal command. Before the first session started, the dog was given a warm-up to also get used to the apparatus and the limited space. The aim of the warm-up was for the dogs to learn, through trial and error how to lift the dummy out of the apparatus. It took most of the dogs only 2-3 trials to learn how they could retrieve the dummy from the apparatus. As soon as the dogs were able to lift the dummy out safely and effortlessly of the apparatus, the test trials were started. After the dogs had lifted the dummy out of the apparatus, they either gave it to the experimenter, dropped it or waited until the experiment had taken the dummy out of the dog's mouth. The experimenter sat on the chair motionless, with the hand (palm up) resting on the knee, in order to avoid influencing the dogs' lifting height. After every successful lift, the dog was rewarded with few pieces of dry food to keep them motivated (Fig. 5). Before each trial, the experimenter took both dummies behind her back and shuffled them so that the dog had no information as to which dummy (heavy or light) was placed next in the apparatus. If individual trials were not carried out properly, e.g., when dogs tried not to lift the dummy in the middle but at the corners of the dummy during a session, tilted their heads too much or tried to push the dummy off the apparatus with their snout, the trial was repeated (Fig. 5D).

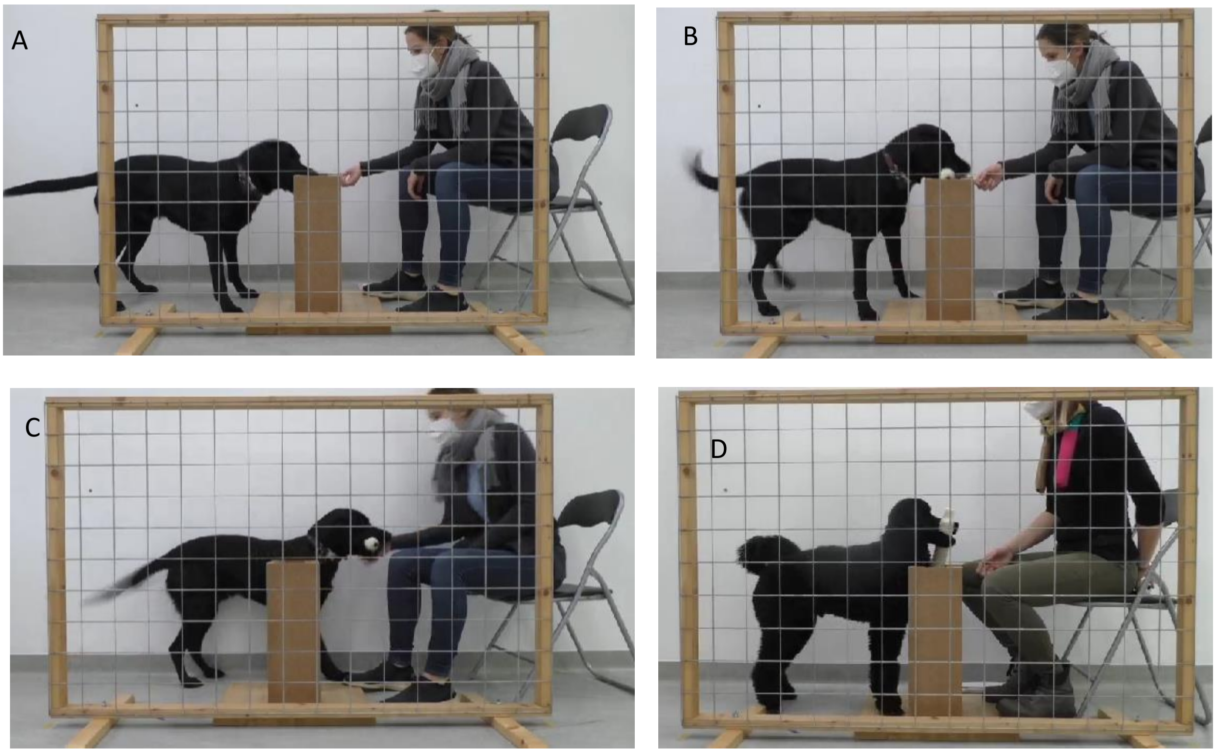

Figure 5: ***A:*** *The dog entered the testing area through the open side on the left. The experimenter rests her hand (palm up) on her knee to avoid influencing the lifting motion of the dog.* ***B***: *The dog successfully lifts the dummy of the apparatus and proceeds to place it in the hand of the experimenter* ***C:*** *The dog finishes one lifting unit by exchanging the dummy for a reward.* ***D:*** *Whenever a dog failed to pick up the dummy in one upward motion, the trial was repeated.*

3.1.4 Coding

In the overshoot experiment we filmed the dogs (N=12) from the side which allowed us to record their profile in the video. After every session we uploaded the videos on Loopy (Loopbio, Vienna). For each trial in which the dogs successfully lifted a dummy out of the apparatus, we froze the video at the precise frame when the tip of the snout was lifted the highest, in a continuous lifting up movement (before moving down again) and set two annotation points: a reference point (on the left outer edge of the apparatus) and a measuring point as a constant in case the video picture was slightly different within or across sessions, for example, when the dogs moved the camera or the apparatus mid-session. The set annotation points allowed us to continuously measure the pixel distance between the apparatus and the tip of the dog's snout. We excluded all non-dummy-lifting motions, such as the dog moving forward to place the dummy in the experimenter's hand or moving backwards in order to sit down (in dummy training the dog is trained to sit and hold the dummy in the mouth). As the aim of this experiment was to check for implicit expectations about weight in dogs, we focused on the transition from heavy to light and from light to heavy trials. We then exported these values from Loopy for analysis.

3.1.5 Analysis

In Experiment 2 we first aggregated the data of dogs to compare overall performance between the conditions heavy and light. We calculated the average difference between the maximal snout height and the reference point for each dog and condition (heavy / light). We conducted paired-samples t-tests on the aggregated data, using the mean values per individual and condition. We did several paired samples t-tests to compare the two test conditions: We compared the two conditions across all trials and specifically the transitions between conditions (heavy to light and light to heavy). To check whether the data matched the assumption that the difference between the two conditions is normally distributed we visually inspected the distribution. This revealed that the data did not deviate from the normality assumption in an obvious way.

3.2 Results

The dogs lifted their heads significantly higher in the light condition than in the heavy condition ($t = -2.86$, $df = 11$, $p = 0.015$). We then specifically looked at the comparison between the first light trial and preceding heavy trial but there was no significant difference between these trials ($t = -1.06$, $df = 11$, $p = 0.312$). We also looked at the heavy to light transition only in the first session. But again, there was no significant difference between these two trials ($t = -0.82$, $df = 11$, $p = 0.432$). Next, we looked at the comparison between the last light and the subsequent heavy trial, which did not reveal a significant difference either ($t = -1.52$, $df = 11$, $p = 0.157$).

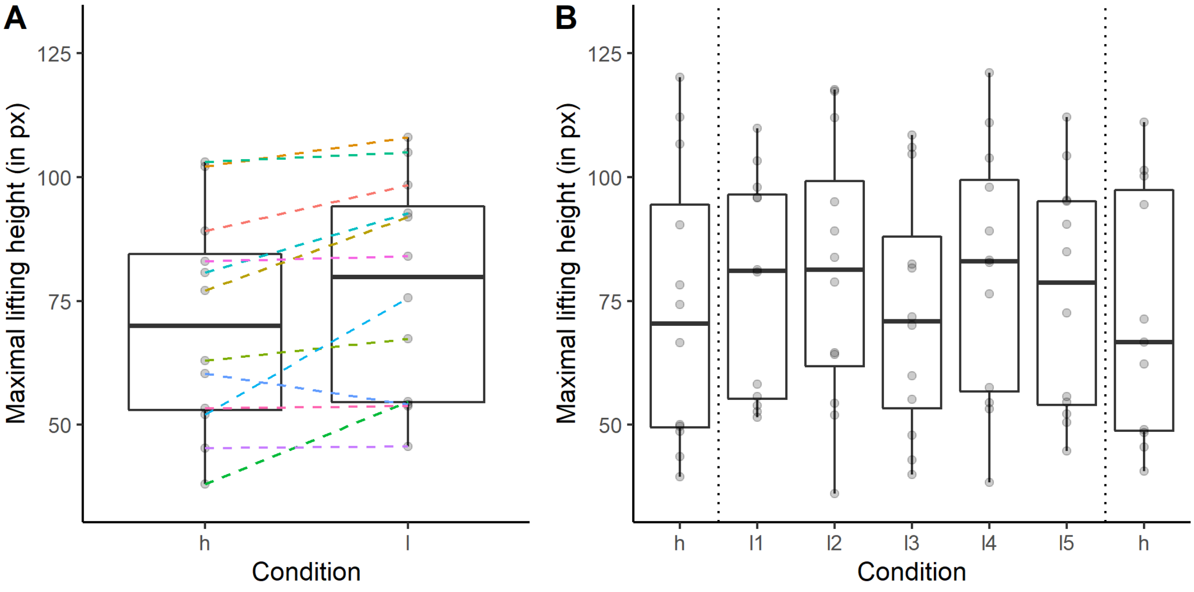

Figure 6: ***A:*** *Box plot based on the average performance across all heavy and light trials. The dots represent the dogs' individual mean values across trials and sessions. The lines connect values of the same individual across the two conditions (h=heavy/l=light).* ***B****: Boxplot combined with dot plot showing the distribution of mean individual performance across the last heavy lift (h) before the transition and then the 5 light trials (l1-l5) and again the subsequent heavy lift. Horizontal lines represent the overall median.*

4. Discussion

4.1 Experiment 1

The objective of our research project was to determine if dogs form implicit expectations about weight. The aim of the first experiment was to show whether dogs use absolute or relative learning rules when learning to discriminate two differently weighed objects. We found that dogs had difficulty discriminating between objects of different weights in a two-choice task but showed some adjustment of their lifting forces as demonstrated by an overshoot effect found in Experiment 2 in which they had to lift differently weighed objects out of a bracket.

In Experiment 1, none of the dogs passed the initial weight discrimination training that would have been necessary to test for absolute and relative learning strategies. When excluding the first two sessions the dogs' performance improved but neither session nor trial number had a significant impact on their performance.

While the results did not align with our hypothesis, that dogs would be able to discriminate objects of different weights in a two-choice task, they might have learnt it, if they would have participated in the full 12 sessions, because the group-level analysis excluding the first two sessions suggests that the dogs learned something. Our prediction that dogs who performed multiple dummy lifts per trial were more likely to be correct was not met. The results of our research suggest that we may not have found the right set up for dogs yet or that it just takes more time for them to learn it.

While apes, corvids or octopuses are tool using species and known to solve problems by object manipulation, this is not true for dogs. As our experimental set up was based on Schrauf and Call's (2009) exchange paradigm with apes, one might need to come up with another set up adapted for dogs. Additionally in discrimination learning tasks, the animals often develop some sort of bias or preference for arbitrary attributes. Former research projects about visual discrimination learning in mice state that the more visually similar two stimuli are, the higher the probability that the animals have an intrinsic side bias (Treviño, 2015). As the objects in our set-up had to be visually identical and only differed in weight, this could be a possible explanation for the occurrence of side biases.

As a limitation for this study, retrieving can be assumed to be intrinsically motivating for most participating dogs and for some dogs, retrieving an object itself could be even considered a greater reward than food, making a food reward and verbal praise for making a correct choice less relevant (Turcsán et al., 2015). This could have influenced the decision making of the dogs: if the task of retrieving was so exciting and rewarding in itself they might have payed less attention to the weight cues. While their training history with fetching objects accelerated the training, it might have biased our results. Since we had 7 retriever type dogs in experiment 1 (6 Labrador Retrievers and 1 Golden Retriever), which all were professionally trained dummy dogs, this could have had a great impact. In addition, professionally trained dummy dogs learn during their training to stick to a decision that has been made (Ordnung für Arbeitsprüfungen mit Dummies für Retriever, 2020). Once the dog has picked a dummy, they are not allowed to exchange the object during training. If the dog tries to replace the dummy it has picked up with one that is still lying on the ground during training, an aversive stimulus follows to make this behavior less likely to occur. Additionally, the starting position of the dogs might have influenced their choice. While in our experiment the dog started from a position between the experimenter and the owner, professionally trained dummy dogs always start sitting to the left side of their handler. We had some cases (only professional trained dummy dogs) where dogs had a very strong side bias and struggled to overcome it, although we administered a correction procedure. The dogs' starting position combined with the professional dummy training could be a contributing factor to their side preferences.

Irrespective of these limitations, previous weight discrimination experiments with apes (McCulloch, 1941; Povinelli 2012; Schrauf and Call, 2009) support the assumption that the ability to reliably discriminate heavy from light objects is hard to train. For instance, McCulloch needed a median of 1179 trials to teach his chimpanzees to discriminate between a heavy and a light box (Povinelli, 2012). As weight discrimination tasks have not been conducted with dogs before and there are indications that some dogs started to improve, they might just need more trials, like the primates. Additionally, the weight difference between the objects was relatively small so it is possible that this was an additional limiting factor which hindered the dogs in mastering the weight discrimination. Future experiments could increase the weight difference of the dummies to facilitate the discrimination. As mentioned in the introduction, we already conducted a pilot study investigating the difference in color vs. weight discrimination learning, using the same set up as in the current experiment. In the pilot study, the weight differences were much greater and even though there were other limitations, one dog started to learn the difference. Unfortunately, both studies took place during the SARS-CoV-2 pandemic, making access to the research facility, but also the personal safety concerns of the participating dog owners, a limiting factor during the data collection.

Therefore, our results do not allow us to conclusively evaluate whether dogs can learn to discriminate weight. More research with longer training protocols and more extreme weight differences will be needed. Also testing more naïve individuals might help given the fact that none of the professionally dummy trained dogs passed our training session.

The test sessions which aimed to examine whether dogs use absolute and relative strategies could not be carried out. Therefore, our second hypothesis that dogs use absolute and relative rules learning to discriminate weight could therefore not be evaluated. For future directions in this field, it could be interesting to use the setup of Lambert et al. (2021) which was designed in a way the animals needed to lift the weights before each trial. This presumably facilitated the task a lot because the animals did not have to seek information about the weight proactively. It could be interesting to test if such a procedure would help dogs to learn the discrimination more easily. Conversely, it could be examined how tool-using species such as the cockatoos would perform in the current weight discrimination tasks, without any direct incentives to lift the objects prior to the discrimination.

4.2 Experiment 2

With our second experiment, the overshoot experiment, we aimed to find out if dogs form implicit expectations of the weight of a common object after interacting with it. For this purpose, we built an apparatus that enabled us to position the dummy in a way, that compels the dogs to lift the dummy in a specific motion. We predicted that dogs form explicit expectations about weight and therefore will lift their head higher in trials with the lighter dummy, after interacting several times with a heavier one, leading to an overshoot effect. Results showed that dogs lifted their head higher in trials with the lighter, 200g dummy, than in the trials with the heavier (1202g) dummy. While this result is promising at the first look, we could not find any significant difference between the last heavy trial and the first following light trial. When comparing the last light trials with the subsequent heavy trials there was also no significant difference. This results partly align with our prediction, that dogs will lift their head higher in lighter conditions, than in heavier conditions. Nevertheless, it is still under question why the transition from the last heavy trial to the first light trial did not show any significant difference in the maximal lifting height.

While overshoot experiments are well researched in human psychology and have been adapted for the use with apes (Povinelli, 2012), until now there is no study investigating the sensorimotor memory of pets like dogs. Until now, it is still under debate which of the two theories concerning sensorimotor memory is valid in the context of weight discrimination. Due to the fact that the lifting process can be broken down in small movement units: loading phase (from object contact until lift-off), lift, hold (the object is held in the air), replacement and unloading, some researchers argue, that the weight of an object can already be estimated as soon as the finger grasp the object and the fingertip force overcomes the objects weight and it starts moving (van Polanen & Davare, 2019). The overshoot effect typically occurs when individuals exert more force or movement than necessary after reaching a target. The argument that the weight of an object can be estimated as soon as the finger grasps it and the force of the fingertip overcomes the weight of the object is consistent with a more accurate and efficient lifting process. This would suggest that the agent would apply the necessary force to lift and hold the object without overshooting or applying excessive force beyond what is required.

But since sensorimotor memory is generally based on the previous handled object, another hypothesis concerning the sensorimotor memory (in line with the overshoot effect) is that holding phase (when the object is held in the air) is more influential than the lifting phase because it provides information for a longer time span and is the final sensory input. While the time duration for a lifting motion is quite short and also can be influenced by “movement induced sensory noise”, holding an object could provide more accurate information about the weight (van Polanen & Davare, 2019).

With respect to the set-up, we decided on, there were some dogs who struggled to lift the dummy off the apparatus, because it required them to first move upwards in a straight line to get the dummy out of the hollows. While some dogs tried to grab the dummy and pull it backwards out of the apparatus, other dogs preferentially grabbed it on the endings resulting in the dummy getting stuck in the apparatus (see Figure 5). Even though we trained all dogs to pick up the dummy, some still had difficulties in the test phase. As it is still under debate which motion (lifting or holding) is influencing the sensorimotor memory and dogs manipulated the dummies several times while trying to lift them of the apparatus, it is retrospectively difficult to determine when and what kind of information was received by the dog and if the lifting heights occurred due to the expectancy violation concerning the weight of the object or the differences in weight *per se*. An overshoot particularly in the first light trial would have provided more direct evidence for the expectancy violation account. However, we found no evidence that the overshoot was particularly pronounced in the first light trial.

Likewise, it is not clear if the dogs that lifted their heads higher in the light conditions had an intrinsic preference for the lighter dummy and therefore lifted their head higher because of excitement. Another possible explanation for the inconclusive results is that the dogs had difficulties keeping the 200g dummy in their snout and tilted their heads backwards/upwards in order to get the dummy to slide deeper in the mouth for a better grip.

5. Conclusion

This is, to our knowledge, one of the first studies investigating dogs' attention to the invisible property of weight. Contrary to our initial hypotheses, the results of the research project indicate that dogs faced challenges in discriminating between objects of different weights in a two-choice task, and did not show a significant overshoot effect that would have indicated they adjust their lifting forces based on sensorimotor representations of weight.

However, there were findings that suggest the dogs might have learned something during the course of the weight discrimination experiment. When excluding the initial sessions, the dogs' performance showed improvement, indicating a potential learning effect over time.

Although we took measures to adapt existing weight discrimination and overshoot tasks for the dogs (e.g., presenting the tasks in a dummy retrieving context), Several limitations should be taken into consideration when interpreting the results. Firstly, the occurrence of side biases among the dogs could have influenced their choices, potentially affecting the discrimination task. Additionally, the intrinsic motivation for retrieving objects, which may have been more rewarding than food or verbal praise for some dogs, could have influenced their decision-making process. The dogs' training history as professionally trained retrievers might have further impacted their performance and decision-making, as they are trained to stick to their initial choice and face some sort of punishment when they pick up more than one object before retrieving it (*changing retrieves*). Moreover, the dogs' starting position and the specific muscles involved in lifting the dummies, as opposed to using hands and fingers like humans and apes, might have caused differences in sensorimotor feedback. Therefore, the experimental set-up, originally designed for humans and monkeys, may not have been fully suited to the capabilities and physical characteristics of dogs.

While the results did not provide conclusive evidence regarding dogs' ability to discriminate weight, they suggest the need for further investigation. Future experiments could involve longer training phases and more extreme weight differences between objects to facilitate discrimination. Additionally, including a larger sample of naive individuals, rather than professionally trained dummy dogs, may provide additional insights into dogs' learning abilities in weight discrimination tasks. Exploring alternative experimental setups, such as the one used by Lambert et al. (2021), which involves lifting the weights before each trial, could be a promising direction for future research. This setup may reduce the need for dogs to proactively seek information about weight and potentially facilitate the learning process. In summary, although the results did not support the initial hypothesis and several limitations were identified, the findings from this research project provide first insights into dogs' ability to discriminate weight.

Further research with adapted experimental setups, longer training protocols, and diverse participant samples is necessary to gain a comprehensive understanding of dogs' cognitive capacities in weight discrimination tasks. These investigations will contribute to the broader understanding of comparative cognition and shed light on the intricacies of dogs' perception and learning mechanisms.

6. References

Agrillo, C., Piffer, L., Bisazza, A., & Butterworth, B. (2015). Ratio dependence in small number discrimination is affected by the experimental procedure. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01649

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

Bensky, M. K., Gosling, S. D., & Sinn, D. L. (2013). The World from a Dog's Point of View. In *Advances in the Study of Behavior* (Bd. 45, S. 209–406). Elsevier. https://doi.org/10.1016/B978-0-12-407186-5.00005-7

Bortot, M., Agrillo, C., Avarguès-Weber, A., Bisazza, A., Miletto Petrazzini, M. E., & Giurfa, M. (2019). Honeybees use absolute rather than relative numerosity in number discrimination. *Biology Letters*, *15*(6), 20190138. https://doi.org/10.1098/rsbl.2019.0138

Deutscher Retriever Club e.V. (2022, September 09). *Ordnung für Arbeitsprüfungen mit Dummies für Retriever (APD/R)* [PDF file]. https://drc.de/retrieverarbeit/pruefungsordnungen

Gordon, A. M., Westling, G., Cole, K. J., & Johansson, R. S. (1993). Memory representations underlying motor commands used during manipulation of common and novel objects. *Journal of Neurophysiology*, *69*(6), 1789–1796. https://doi.org/10.1152/jn.1993.69.6.1789

Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology*, *18*(9), R370–R372. https://doi.org/10.1016/j.cub.2008.02.039

Jelbert, S. A., Miller, R., Schiestl, M., Boeckle, M., Cheke, L. G., Gray, R. D., Taylor, A. H., & Clayton, N. S. (2019). New Caledonian crows infer the weight of objects from observing their movements in a breeze. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1894), 20182332. https://doi.org/10.1098/rspb.2018.2332

Johansson, R. S., & Westling, G. (1988). Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Experimental Brain Research*, *71*(1). https://doi.org/10.1007/BF00247522

Lambert, P. J., Stiegler, A., Rössler, T., Lambert, M. L., & Auersperg, A. M. I. (2021). Goffin's cockatoos discriminate objects based on weight alone. *Biology Letters*, *17*(9), 20210250. https://doi.org/10.1098/rsbl.2021.0250

McCulloch, T. L. (1941). Discrimination of lifted weights by chimpanzees. *Journal of Comparative Psychology*, *32*(3), 507–519. https://doi.org/10.1037/h0058792

Miletto Petrazzini, M. E., Agrillo, C., Izard, V., & Bisazza, A. (2016). Do humans (Homo sapiens) and fish (Pterophyllum scalare) make similar numerosity judgments? *Journal of Comparative Psychology*, *130*(4), 380–390. https://doi.org/10.1037/com0000045

Neilands, P. D., Jelbert, S. A., Breen, A. J., Schiestl, M., & Taylor, A. H. (2016). How Insightful Is 'Insight'? New Caledonian Crows Do Not Attend to Object Weight during Spontaneous Stone Dropping. *PLOS ONE*, *11*(12), e0167419. https://doi.org/10.1371/journal.pone.0167419

Povinelli, D. J. (2012). *World without weight: Perspectives on an alien mind*. Oxford University Press.

Range, F., Jenikejew, J., Schröder, I., & Viranyi, Z. (2014). Difference in quantity discrimination in dogs and wolves. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01299

Schrauf, C., & Call, J. (2009). Great apes' performance in discriminating weight and achromatic color. *Animal Cognition*, *12*(4), 567–574. https://doi.org/10.1007/s10071-009-0216-1

Schrauf, C., Huber, L., & Visalberghi, E. (2008). Do capuchin monkeys use weight to select hammer tools? *Animal Cognition*, *11*(3), 413–422. https://doi.org/10.1007/s10071-007-0131-2

Sirianni, G., Wittig, R. M., Gratton, P., Mundry, R., Schüler, A., & Boesch, C. (2018). Do chimpanzees anticipate an object's weight? A field experiment on the kinematics of hammer-lifting movements in the nut-cracking Taï chimpanzees. *Animal Cognition*, *21*(1), 109–118. https://doi.org/10.1007/s10071-017-1144-0

Spady, T. C., & Ostrander, E. A. (2008). Canine Behavioral Genetics: Pointing Out the Phenotypes and Herding up the Genes. *The American Journal of Human Genetics*, *82*(1), 10–18. https://doi.org/10.1016/j.ajhg.2007.12.001

Treviño, M. (2015). Stimulus similarity determines the prevalence of behavioral laterality in a visual discrimination task for mice. *Scientific Reports*, *4*(1), 7569. https://doi.org/10.1038/srep07569

Turcsán, B., Szánthó, F., Miklósi, Á., & Kubinyi, E. (2015). Fetching what the owner prefers? Dogs recognize disgust and happiness in human behaviour. *Animal Cognition*, *18*(1), 83–94. https://doi.org/10.1007/s10071-014-0779-3

van Polanen, V., & Davare, M. (2019a). Sensorimotor memory for object weight is based on previous experience during lifting, not holding. *Neuropsychologia*, *131*, 306–315. https://doi.org/10.1016/j.neuropsychologia.2019.05.025

Visalberghi, E., & Neel, C. (2003). Tufted Capuchins (Cebus apella) Use Weight and Sound to Choose Between Full and Empty Nuts. *Ecological Psychology*, *15*(3), 215–228. https://doi.org/10.1207/S15326969ECO1503_2