Department of Interdisciplinary life sciences

University of Veterinary Medicine Vienna

Comparative Cognition Unit, Messerli Research Institute

Head: Univ.Prof. Mag. Dr.rer.nat Ludwig Huber

Investigating weight discrimination in the Goffin`s cockatoo (Cacatua goffiniana)

Diploma thesis

University of Veterinary Medicine Vienna

submitted by

Antonia Rippel-Rachle

Vienna, July 2021

Supervisor: Priv-Doz.Dr. Alice Auersperg

Comparative Cognition Unit, Messerli Research Institute

Department of Interdisciplinary life sciences

University of Veterinary Medicine Vienna

INDEX

1.Introduction **1**

2.Materials and methods **7**

2.1 Subjects and housing *7*

2.1.1 Pre-experience 8

2.2 Ethical note *8*

2.3 Materials *9*

2.3.1 Training 9

2.4 Testing Procedure *9*

3. Analysis **13**

4.Results **15**

5.Discussion **18**

6.Summary **22**

7.Zusammenfassung **23**

8.Acknowledgement **24**

9. References **25**

10. Appendix **28**

1.Introduction

Weight can be considered an 'invisible' property of an object. It can only be perceived by physically handling an object – holding it, lifting it or moving it - in order to gain weight-related proprioceptive information. Investigating how different species perceive and use weight information, and how this might relate to their behavior and environment, can help us to understand the evolution of weight-related cognition in animals. For example, a question of particular interest here is how readily species perceive weight and whether an individual is able to predict an object's physical impact on the environment (Povinelli, 2011). Such a higher-order (cognitive) representation of weight has so far not been concluded for most animal species. However, there is a spectrum between this level of representation and the ability to perceive weight only in a putative sensory-motor way (effort-while-lifting) (Povinelli, 2011). Although the physical perception of weight is not a cognitive skill, investigating the way and speed with which a species learns how to use weight-information will help build a better understanding of shared weight-related cognitive processes in animals. Species that are more sensitive to the weight of an object, and more readily learn to use this information in a task, might be more attentive to weight and force related signals within their natural behaviours. This attention could have evolved to modulate varied actions on different objects (Povinelli, 2011) relevant to the species' lifestyle and environment – for example birds who transport nesting materials.

When focusing on behaviours such as extractive foraging, there is evidence that these actions are influenced and enhanced by the selective use of weighted objects. For example, captive capuchin monkeys discriminate full nuts from empty ones before going through the costly process of opening them (Visalberghi & Neel, 2003). In an experimental setup the animals had to distinguish between nuts differing either in weight and sound or both. When weight was the only discriminative the subjects performed high above chance level. These findings strongly suggest that at least some primate species have a kinesthetic judgement of weight (Visalberghi & Neel, 2003). Many bird species are also extractive foragers. Two studies have shown that birds (Mexican Jays, Black Capped Chickadees and Red-breasted Nuthatches) can discriminate food by mass. Mexican Jays *(Aphelocoma ultramarina*) demonstrated a sensitivity of the correlation between size and weight of nuts. When confronted with peanuts that differed in weight and size, the birds discarded lighter nuts at a higher rate (Jablonski et al., 2015) This bird species therefore may have a concept of how much different nuts should weight, according to their sizes and might have the ability to relate to the higher nutritional value of heavier nuts

(Jablonski et al., 2015). A similar behaviour was observed in Black Capped Chickadees and Red-breasted Nuthatches, who when confronted with normal and plaster of paris filled seeds, rejected the normal ones at a higher rate (Heinrich, 1997).

Some species are even able to use weight information to optimize instrumental problem solving and tool-use. For example, capuchin monkeys (*Cebus spp.)* use tools (stones) to crack open nuts (Liu et al., 2009), and they will develop a weight-based preference relating to the effectiveness of the stone tool (Visalberghi et al., 2009). For Capuchins the efficacy of a tool plays an important role, as the act of nut-cracking is energetically demanding. (Schrauf et al., 2012). When wild bearded capuchins had to choose between stones with different features (either same size and different weight, or light and large, heavy and small and vice versa) they used and transported the heavier stone (functional tool) more often than expected by chance. One can say that wild capuchins consistently and immediately selected the functional tool in any given case and even outperformed captive ones (Visalberghi et al., 2009). In a lab-based experiment, Capuchin monkeys were able to learn to determine objects regarding their success as nut cracking tools when weight was the only feature that differed (Schrauf et al., 2008). The same has been found for chimpanzees (Schrauf et al., 2012). However, it should be noted that similarly for Capuchins, experience in nut cracking affected the attentiveness to the relevant properties for solving the task (Schrauf et al., 2012).

Although species from different groups have demonstrated that they can use weight information, the study of weight discrimination has largely focused on primates. In classical discrimination learning set ups, several primate species have been required to distinguish visually identical objects according to their weight. However, they have shown surprising difficulties. When asking chimpanzees to select the heavier of two objects (an 80g weight and either a 480g or a 680g weight) they took an average of 1179 trials to reach a steady performance (McCulloch, 1941). In a more recent study less than half of a group of five orangutans, three gorillas and five bonobos where able to reach criterion in a weight discrimination task (Schrauf & Call, 2009). Here the subjects needed to return 6 out of 12 correct objects with a certain feature (heavy or light differing by 300 grams) to the experimenter and needed approximately 30 sessions, and a median of 331 exchanges to reach the criterion of 6 out of 6 correct exchanges in the weight discrimination task. When the objects differed in colour instead of weight, they reached criterion in 64 exchanges, but still only around half of subjects met criterion (Schrauf & Call, 2009). In a weight sorting task chimpanzees had to sort two visually identical objects into trays of different colour, according to the object's weight. 6

out of 7 chimpanzees (all which had previous experience with sorting objects and weight related tasks) reached criterion, which was a stringent criterion of 45/50, i.e. 90% correct over 5 sessions. The subjects which reached criterion took an average of 895,2 trials. (Povinelli, 2011). In comparison, 3 to 5 year old children learned to sort heavy and light objects immediately when tested under the same predictions, with 98.1% correct responses across all children and object types (Povinelli, 2011). The high number of trials that the chimpanzees required suggests that weight sorting is a cognitively highly demanding process for this species.

As learning to discriminate objects on the basis of weight seems to be surprisingly difficult for non-human primates it might be the case that, although primate species have evolved various skills relating to tool use and extractive foraging, they might not have evolved specialized systems for readily using weight related signals (Povinelli, 2011). However, it is hard to put the exact abilities of primates regarding their use of weight information into context, due to the lack of similar research with non-primates. Povinelli (2011) suggested the possibility of birds performing exceedingly well on simple weight- sorting tasks. He hypothesized that, due to the fact that birds transport objects of different weights when flying, they might have evolved weight and force-related signals to modulate complex actions on objects in different settings (Povinelli, 2011). Notably, in the single discrimination learning study carried out in a bird species to date, the Goffin´s cockatoo reached criterion (8 out of 8 choices correct, or 9 out of 10) on a weight choice task with visually identical objects in an average of 60.6 trials (Lambert, Stiegler *et al.* in Press). If considering the criterions used in primate studies, these birds met the criterion used in Schrauf & Call (2009) at an average trial number of 69 and 6 of the 7 subjects reached the 45/50 criterion of Povinelli's (2011) study in an average of 101,7 trials (Lambert, Stiegler *et al.* in Press). This is in stark contrast to the hundreds of trials primate subjects in these studies required.

Indeed, there are studies which show that large-brained bird species can also discriminate between weighted objects in the context of a problem-solving task. In a study investigating what subjects learnt about objects during exploration, New Caledonian Crows *(Corvus moneduloides*), and Kea *(Nestor notabilis)* were provided with a selection of objects differing in weight and colour (Lambert et al., 2017). The results of the study show that several individuals performed above chance on test trials with a platform task when they were given objects whose physical properties they had already explored (Lambert et al., 2017). A study by Jelbert and colleagues showed that New Caledonian crows *(Corvus moneduloides)* were able to infer the weight of objects by their movement in a breeze. The setup of the experiment consisted of a fan and two different objects (a heavy one and a light one) mounted in front of the fan. When the fan was on, it created a slight wind current that caused the light object to start moving while the heavier one stayed in place (Jelbert et al., 2019). They chose the correct object (heavy or light) in 73 per cent of all attempts (Jelbert et al., 2019). In fact, in the training phase prior to this test, which had a set up similar to a classical discrimination learning task, the birds learned very quickly to distinguish between objects according to their weight, when they also all differed from each other in colour and shape (Jelbert et al., 2019).

Given the performance in the Goffin´s choice study with visually identical objects, it appears that at least some bird species can learn to discriminate by weight much more quickly than primate species. Additionally, the findings of Jelbert et al. (2019) where the crows could choose correctly between a light and heavy object dependent upon their relative movements in a breeze, suggest that this bird species has a higher-order representation of weight than has been found, so far, in any primate species (Povinelli, 2011). Additionally, there is evidence from the Goffin choice study that the birds not only possessed a sensory-motor representation (effort-while-lifting) of weight but might also have had an abstract concept about `heavy` and `light`: they returned directly to the correct object 80.1% of the time (Lambert & Stiegler et al. in Press), suggesting that they remembered which was the rewarded item without needing to simultaneously experience the physical sensation of moving the object. Povinelli (2011) did not find this type of representation in the chimpanzees he studied. A picture is beginning to emerge that there might be a distinct difference between large-brained bird species and primates with regards to weight-related cognition.

However, we cannot be sure that birds do possess a higher degree of weight-related cognition until we have tested species in more comparable manners. By virtue of the differing methodologies in studies that have tested primates and bird species on the discrimination of weight, it is hard to draw robust comparisons between their abilities. For example, Povinelli's sorting task differs from the other weight discrimination tasks in important ways. In a classical *choice task* (McCulloch, 1941; Schrauf & Call, 2009) subjects will always have one positive, rewarded stimulus, but in the *sorting study* each stimulus is rewarded equally if the correct choice is made. For example, using the rule 'put the light object in the red tray and the heavy object in the blue tray' (sorting) could be more challenging than 'always choose the light object' or 'always choose the heavy object'(choice). Given the potential effect of a sorting task compared to a choice task, we tested the Goffin´s cockatoo in a sorting task like the ones given to chimpanzees.

The Goffin´s cockatoo (*Cacatua goffininania)* serves the ideal avian model for investigating weight discrimination tasks only relying on proprioceptive feedback. In their natural habitat, the Tanimbar Archipelago, a small island group in the Pacific Ocean, the birds are known to acquire a major part of their diet through extractive foraging (Mioduszewska et al., 2019; O'Hara et al., 2019). To reach the edible parts of the cassava plant *(Manihot esculenta)*, the Goffin´s cockatoo excavates the nutritious roots which lie several centimeters underground, by biting at the bottom of the plant and shoveling soil with their beaks. Young coconuts are also part of the diet of the Goffin´s cockatoo (O'Hara et al., 2019). In order to get access to the fruit water and pulp the birds search for preexisting incisions in the shell and begin to tear it off. This process may take the Goffin several hours or up to days (O'Hara et al., 2019; Osuna-Mascaró & Auersperg, 2018). Additionally, captive Goffin Cockatoos have demonstrated sophisticated tool use and manufacturing skills (Auersperg et al., 2013, 2016; Laumer et al., 2016). They were also able to learn to bend hooks from straight wire in order to achieve food from vertical tubes and unbend wire to retrieve food from horizontal tubes (Laumer et al., 2017).

To enable meaningful comparisons between the learning speed of the Goffin´s cockatoo and chimpanzees on a weight discrimination task, we tested the Goffin´s with a similar experimental set-up as in Povinelli's Experiment 9 (Povinelli, 2011). Our subjects had to place either a heavy or a light visually identical object into trays of different colour. Some modifications to the original methodologies were made to allow for a better comparison to the weight choice task. Our subjects were not given orientation sessions to indicate the rules before training began and to avoid creating a side bias, we decided to switch the location of the trays during every trial.

We asked, firstly, whether a sorting task would be more cognitively demanding than a choice task. We predicted that the Goffin´s would need longer to learn to consistently sort objects on the basis of weight than for learning which object was rewarded in the choice study. Taking into account that in the previous weight choice study (Lambert and Stiegler et al. in Press) the Goffin´s learned to discriminate the weighted objects much more quickly than the chimpanzees in the sorting task (Povinelli, 2011), we predicted that the Goffin´s would surpass the chimpanzee's performance in a sorting task. Considering that this is the second weight-based study the Goffin´s have participated in, their performance will likely be enhanced by experience. However, compared to the set-up of Povinelli's task, there is no cue redundancy in our sorting study (the coloured trays do not remain on the same side between trials, as they did for the chimpanzees), which would make our version of the task more difficult. Nevertheless, if the Goffin´s markedly outperform the chimpanzees we could be more confident of a distinct difference between primates and birds regarding their weight cognition. Therefore, we predicted that the birds would outperform the chimps when comparing the responses of the Goffin´s across 15 sessions of the sorting task with the data from the first 15 sessions of Experiment 9 (Povinelli, 2011), which was kindly provided by Daniel Povinelli. We gave the Goffin´s 15 session in this task, because in a previous study (Lambert & Stiegler et al. in Press) the longest any of the subjects took to reach the criterion 45/50 (Povinelli, 2011) were 14 session. Therefore we suggested that if the Goffin´s did not manage to reach the criterion used in (Povinelli, 2011) within 15 session in the sorting study, we could be confident of the sorting task being more difficult than the choice task. In our study we did not test subjects till a criterion, because of the risk of multiple testing issues associated with this approach. Alternatively, analyzing data with Generalized Linear Mixed Models (GLMMs) increases the strength of support we have for any conclusion about the presence, absence or degree of cognitive abilities within a species.

2.Materials and methods

2.1 Subjects and housing

Sixteen captive-born and hand-reared Goffin's cockatoos (*Cacatua goffiniana*) participated in this study. The group consisted of 13 adults (aged between ten and 13 years; eight males, four females), and three sub- adults (each three years of age; one male, two females). The subjects were marked with differently coloured aluminium leg bands and housed together at the Goffin Lab Goldegg, in Lower Austria in a large aviary consisting of an indoor and outdoor area (indoors: 45 m2 ground space, 3–6 m high wall to gable; outdoors: 150 m2 ground space, 3–4.5 m high). The aviary is enriched with bathing pools, branches, bark and other wooden toys. The aviary has a 12- hour light-dark cycle. During colder periods the indoor area is kept at a temperature 20°C. The access to the outdoor area is dependent on outside temperatures.

The birds receive fresh food twice a day: in the morning their food consists of a varied range of fruits, plant-based yoghurt, potatoes, noodles mixed with palm oil, fried eggs and cooked grains. The combinations of these foods vary throughout the week. Fresh water and a basic food mixture, to which various types of seeds, dried fruits and different food supplements are added, are supplied ad libitum.

We created two groups with 8 subjects each, divided by sex and age. Group1 was rewarded for placing the heavy object into the blue tray and the light one into the red tray. Group 2 was rewarded for the exact opposite, placing the heavy object into the red tray and the light object into the blue tray. The groups were established to account for possible color preferences in the subjects. Testing was performed in an adjoining compartment to the aviary under visual exclusion from the other subjects.

Group1		Group2	
Olympia	F, A	Irene	F, SA
Jane	F, SA	Heidi	F, A
Pipin	M, A	Moneypenny	F, A
Kiwi	M, A	Muppet	M, A
Fini	F, A	Figaro	M, A
Mayday	F, A	Dolittle	M, A
Zozo	M, A	Muki	M, A
Konrad	M, A	Titus	M, SA

Table1. The sex (M/F) and ages (adult/ sub-adult(A/SA) of subjects in each group.

2.1.1 Pre-experience

All the 16 subjects have previously been involved in various physical problem-solving tasks and have additional experience with a weight-based discrimination (choice) task, were rewarded for offering either the heavy or light object (consistent across trials) from a pair to the experimenter.

2.2 Ethical note

The study was approved by the ethics and animal welfare committee of University of Veterinary Medicine Vienna in accordance with good scientific practice guidelines and national legislation.

All animals had CITES certificates and were registered at the district's administrative animal welfare bureau (Bezirkshauptmannschaft St. Pölten Schmiedgasse 4–6, A-3100; St. Pölten, Austria). These housing conditions comply with the Austrian Federal Act on the Protection of Animals (Animal Protection Act –§ 24 Abs. 1 Z 1and 2: § 25 Abs. 3 – TSchG, BGBl. I Nr. 118/2004 Art. 2). As the study was based on animal cognition and strictly noninvasive, according to the Austrian Animal Experiments Act (TVG 2012) it is classified as non-animal experiment. The subjects participate voluntarily in every experiment: either they were called into the testing compartment by name, or the experimenter asked the subject to step up on their hand to enter the testing compartment.

2.3 Materials

Identical objects of two different weights, two coloured trays and a large cushion (used as a base on top of the testing table) were used in this study. The objects were made of grey Fimo soft® and shaped into small balls with a diameter of 1,5cm. The heavy version had a 20g spherical lead fishing weight inside, while the light version contained a compressed cotton ball of the same shape and size as the fishing weight. The final weight of the heavy objects was 23 gram and the light ones weighed 3 grams. The cushion was made of foam cut to the size of the testing table (70x/70cm) and was covered in white cotton. It could be attached to the table with straps at the corners to prevent it from slipping. The two trays made from plastic were covered in red or blue felt. The use of the cushion and felt was to ensure that the objects would not be visually or acoustically distinguishable.

2.3.1 Training

The birds were already very familiar with the command `Give me´ that was used to ask subjects to pick up an object and place it in a tray. Subjects were not introduced to any other task contingencies – e.g., which weighted object should be placed in which tray – before testing began.

2.4 Testing Procedure

In each trial, a subject was presented with either a heavy or light object. Before starting the trial, the subject was asked to stand on a wooden chair on one side of the testing table, opposite to the experimenter. To ask the subject to stay, one hand is held in front of the bird and the voice command ´bleib´ is given. First the trays were placed onto the cushion and then the weighted object was placed at a marked point between the trays (see Figure 1.) The experimenter then started the trial with the command ´Give me´. Either the blue tray was placed on the left and the red tray on the right, or the opposite (see figure 1). The object (light or heavy) given in the trail was pseudo-randomized across the first 150 trials, as were the positions of the trays (left/right), for every trial but with the caveat that no object type or tray arrangement should occur more than three times in a row.

There were four possible combinations of tray position and objects (e.g., heavy object -blue tray left, heavy object -blue tray right, light object -blue tray left, light object -blue tray right).

The red tray would always go on the opposite side to that specified for the blue tray. Also, the rule of giving each subject the same combination only 3 times in a row for each session was established.

A trial began when the subject picked up the object and finished when the choice was made (defined by any contact between the object and tray) (see Figure 2.). If the placement (or first contact) was correct, they were rewarded with a small piece of cashew. If it was incorrect, this was signaled with a 'Nein'. In either case the birds were asked to return to the starting point (wooden chair) to wait for the next trial. The object (heavy or light) was removed after each trial and the tray positions were changed if required. After a short pause an object was placed again for beginning the next trial. During the whole process the experimenter wore mirrored sunglasses, faced centrally between the two trays and made as minimal movements as possible to avoid cuing the birds.

Figure1. Experimental set-up. The lines show the measurements of the testing table and the distances between the trays and the object Move figure 1 into the testing procedure section

Subjects were given one session per testing day. Sessions consisted of 10 consecutive trials. A trial was a maximum of 10 minutes long (if subjects had not made a choice within this time no response would be recorded for that trial) Subjects were given one session per testing day.

To ensure that each combination of weight and tray positions would occur (nearly) the same number of times for each subject in the 15 sessions, each subject was given 37 of each combination (heavy object -blue tray left, heavy object -blue tray right, light object -blue tray left, light object -blue tray right) (148 trials total). These combinations were pseudo-randomized by random.org for each bird individually. To achieve the 150 trials each subject was given an extra of two of the combinations, there were six possible pairs of combinations. Four pairs were used three times and two pairs two times. After subjects completed 15 sessions, 11 birds were given another 5 sessions, as a post-hoc consideration to see if performance could improve. The other 5 subjects dropped out of the study in earlier sessions.

1)

2)

3)

Figure 2. A subject carrying out a trial 1) Initial set up, with the subject sitting on the chair and the weighted object situated in between the two trays of different colour.2) The subject picking up the weighted object (either heavy or light). 3) The subject choosing a tray (before contact is shown here).

3. Analysis

We ran four Generalised Linear Mixed Models (GLMMs) with the data (success: 0/1) from sessions 1 to 15 of Povinelli's (2011) Experiment 9 and sessions 1 to 20 of our sorting task. Two of these models addressed our main research questions, whilst two addressed post-hoc questions (whether the Goffins would perform above chance levels with further testing (session 16-20), and whether the performance of the 7 chimpanzees of Povinelli's (2011) study was significantly above chance levels over the first 15 sessions of the task). We ran our GLMMs with a binomial error structure and logit link function. For greater interpretability of the results, we z-transformed trial number and session number. A sigmoidal-shaped learning curve was a hypothetically likely scenario, so we included the session term as session + $session^2$ (in an interaction with trial number), to account for this possibility. To avoid pseudo-replication, we included a random intercept of subject and the combined factor of session and subject ('dayID') to account for day specific variation (e.g., motivation/mood). We included all theoretically identifiable random slopes in all of our model and we removed correlations of random slopes and random intercepts if the majority of correlations appeared unidentifiable (Matuschek et al., 2017) or if any resulted in a non-numerical output; we did this for the random slopes within dayID in all models.

After constructing our full models, we explored whether any of assumptions associated with GLMMs with a binomial error structure were violated: we calculated the dispersion parameter and Variance Inflation Factor (Field, 2009) and we inspected whether 'Best Linear Unbiased Predictors' were (approximately) normally distributed. After testing the significance of a full-null model comparison, we examined the significance of the estimate associated with each term in the full model and followed a process of dropping each non-significant highest order interaction to derive the final version of the model (to gain the most reliable estimates). We assessed the stability of the final models by comparing the estimates of the model using all data versus estimates derived when levels of the random effects (subject, and day ID) were excluded one by one (Nieuwenhuis, 2012).

To answer our first research question, whether a weight sorting task is more difficult for the Goffin´s than a weight choice task, we examined whether the intercept of a model using (Goffin) data from sessions 1 – 15 was significantly different from zero (an intercept of zero represents chance level), given that we knew the intercept estimate for a similar model exploring the performance of Goffin subjects in the choice task. We did not choose to do a within-subject comparison between the tasks as the testing structure differed between studies. In this model (model 1), we controlled for an effect of trial number, session number, group and the weight (heavy or light) given on each particular trial. The factors of ‘group’ and ‘weight’ were dummy coded and centered, to provide a meaningful interpretation of the intercept.

We hypothesized that the learning speed could differ between group 1 and group 2. Additionally, we hypothesized that it could be easier to learn the correct tray colour for one type of object (heavy or light) over the other. Therefore, this model was comprised of a four-way interaction between these terms. As a post-hoc exploration, we ran this same model structure with data from Goffin sessions 16–20 (model 1a). We then asked the same post-hoc question for the data from chimpanzees (model 3): here the model was comprised of a three-way interaction between trial number, session number and weight (as the chimpanzee subjects were not divided into two groups).

To answer our second research question, whether the performance of Goffin´s and chimpanzees significantly differed from each other in (the first) 15 sessions of a weight sorting task, we ran a model (model 2) with the same structure as for model 1, but with a term ‘group-species’ rather than group (which was comprised of three levels: chimpanzee, Goffin 1 and Goffin 2). In the full model the four-way interaction allowed for any difference in learning speed for heavy and light objects to depend upon the level of group-species: an aspect of the particular absolute or relative weights of the objects, compared to subject body weight and strength, or given species morphology, could lead to greater differences between heavy and light learning speeds in one species over the other. In this model the term of particular interest was ‘group-species’, so our null model included only a three-way interaction between trial number, session number and weight.

4.Results

Overall, the performance of Goffin subjects did not notably improve over 15 sessions. In session 1-5 (15 individuals), subjects chose correctly in 47,6% of trials, in sessions 6–10 (13 individuals), 53,8%, and in session 11-15 (11 individuals), 50,2%. Taking the responses from all individuals over these 15 sessions, we found that subject performance was not significantly different from chance level (model 1; intercept estimate: 0,08, SE ± 0,099, z = 0,804, p = 0,42). At a group level, subjects had an estimated probability of being correct of 52,1% (see Table 1 in the Appendix for the estimated probability of being correct at the subject level). Both groups performed similarly (group 1: 50,3% correct overall in sessions 1-15, group 2: 50,4% correct overall in sessions 1-15); the effect of group on success was not significant (model 1, group.code estimate: p=0,873). Looking at the additional sessions (16-20; 11 individuals), subjects were successful on 52,0% of trials. Performance over these sessions was also not significantly different from chance (model 1a; intercept estimate: 0,083, SE ± 0,101, z = 0,819, p = 0,41). Again, the effect of group upon success was not significant (model 1a; group.code estimate: p=0,599). The success rate between heavy and light trials differed (see Table 2 for a full breakdown of success rate for heavy and light trials). Over all 20 sessions, Goffin subjects were successful 53,0% of trials with a heavy object, whilst they were successful in 48,6% of trials with a light object.

The interaction between session^2, trial and weight (entered into the model as a dummy-coded and centred variable, weight.code) was significant (model 1; z.session^2:z.trial:weight.code estimate: p= 0.01). We found a greater difference in success rate between heavy and light trials in Group 1 (Heavy on Blue) (54,1% correct heavy trials and 47,9% correct light trials) than in Group 2 (Heavy on Red) (51,5% correct heavy and 49,3% correct light). However, the group term was not involved in a significant interaction with weight, session and trial in model 1 or 1a. See Tables 2 and 3 in the Appendix for the estimates from these models. When comparing the performance of Goffin´s and chimpanzees, there was a significant effect of ‘group-species’ (model 2; full-null model comparison: χ^2 = 37,95, df = 24, p = 0,035). The four-way interaction between trial number, session number, the group-species term and weight was significant (model 2; z.trial:z.session:groupspeciesgoffin1/2:weightlight estimates: p_s<0,03). Figure 3 shows the estimated probability of success for subjects in each level of group-species and for both types of weighted object, across sessions and trials (see Table 4 in the Appendix for the value of each estimate from model 2).

Lastly, to answer a post-hoc question about chimpanzee performance, we found that in the first 15 sessions of their sorting task, chimpanzees did not perform significantly differently from chance levels (model 3; intercept estimate intercept estimate: -0,257, SE = 0,319, z= -0,807, p = 0,42). See Table 5 in the Appendix for estimates from model 3.

		Session Block			
	Object	**1-5**	**6-10**	**11-15**	**16-20**
Group 1	**Heavy**	51.8%	62.6%	52.2%	51.7%
	Light	44.8%	40.8%	48.2%	55.0%
Group 2	**Heavy**	48.1%	57.0%	54.3%	49.2%
	Light	46.7%	55.3%	43.0%	51.6%

Table 2. Percentage of successful trials for Goffin subjects in each group, according to session block and the object type (heavy or light) given per trial.

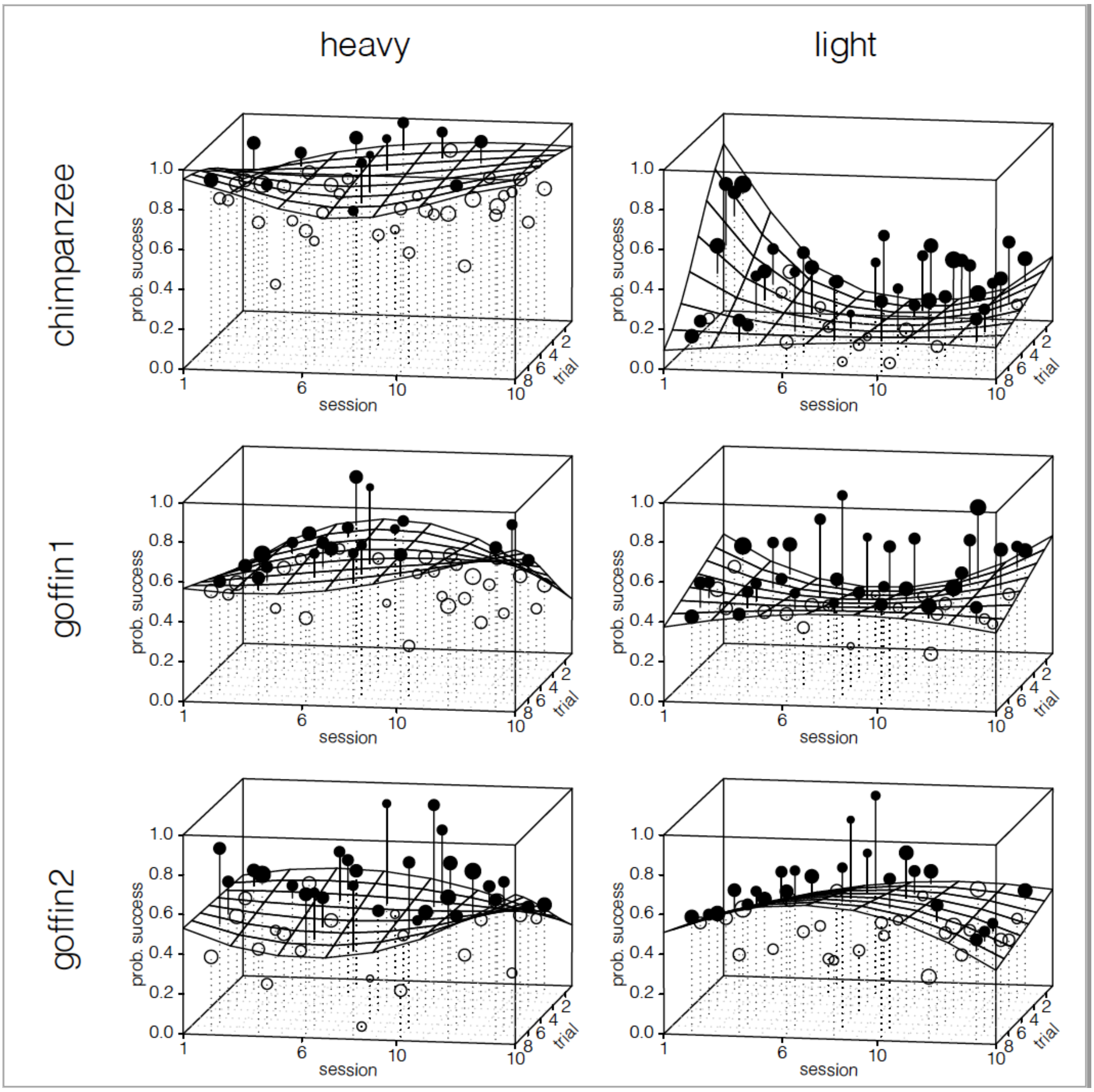

Figure 3. The effect of trial number and session number on the probability of success, depicted separately for trials with heavy objects and trials with lights objects, for subjects belonging to each level of 'group-species'. Depicted are the fitted model (surface) and the success rate (averaged over all subjects) per cell (filled/open points). Values lying above the fitted model are shown as filled points (connected to the surface by a solid line) and values lying below the fitted model as open points (connected to the surface with a dotted line). The 'volume' of the points corresponds to the number of samples per cell.

5.Discussion

Our findings support our hypothesis that Goffin´s perform better in a weight choice task than in a weight sorting task. In the choice study the Goffin´s were able to reach a consistently high performance in the weight-only task (i.e., with visually identical objects) at an average of 60,6 trials (Lambert & Stiegler et al. in Press). Furthermore, in five weight-only sessions given to subjects after they reached criterion, subjects performed well above chance levels. In contrast, in the weight sorting study subjects' performance did not differ from chance level over sessions 1-15 or sessions 16-20 (200 trials in total) .

We hypothesized that a weight sorting task would be more cognitively demanding than a choice task. In the weight choice task, the subjects needed to learn a single association (reward for either heavy or light object), while in the sorting study they needed to remember two rules (heavy goes in red tray, light into blue tray or vice versa). We can therefore consider the weight choice task as involving successive discrimination (only one stimulus that effects the response (Axe, 2008)), while weight sorting can be considered to involve conditional discrimination (a response underlies the operant control of one stimulus whilst the presence of another one (Axe, 2008)). Our results are in line with the fact that conditional discrimination is more multifold than successive discrimination; a difference in the nature of discrimination could be one possible explanation for a weight sorting task being more difficult than weight choice task for the Goffin's cockatoo.

However, there are alternative reasons that might also explain our results. For example, animals in general seem to face difficulties when learning arbitrary relationships. Whether a task is arbitrary or causal outlines major differences in performance (Schrauf & Call, 2011). Primates have proven that mastering a weight discrimination task in the absence of causalities is difficult and takes a lot of time (McCulloch, 1941; Povinelli, 2011; Schrauf & Call, 2009).

In an experiment 5 out of 9 chimps were able to master a causal condition while only one of the 9 subjects was able to learn the arbitrary one (Schrauf & Call, 2011). The arbitrary nature of the sorting task could explain the Goffin´s poor performance.

On the other hand, the Goffin´s performance in the similarly arbitrary weight choice task was very high (Lambert and Stiegler et al. in Press). This can be attributed to the fact that, although both tasks can be considered arbitrary, they considerably differ in methodology. In the weight choice task, the Goffin´s were able to directly compare both objects by having to lift both prior to the testing situation (Lambert & Stiegler et al. in Press). In contrast, in this study they were presented with either heavy or light objects on each trial. Therefore, in the choice study they only had to make a real-time comparison, but not remember how heavy or light objects felt. Being able to lift an object and immediately compare it to another one could have helped the Goffin´s in the choice task (Lambert & Stiegler et al. in Press), by allowing them to keep track of the weight difference. Tackling this question would until require further investigations such as giving subjects the opportunity to compare both weights before placing them in the trays.

Counter to our prediction, the Goffin´s did not appear outperform the chimpanzees in the weight sorting task, when considering overall success rate in the first 15 sessions for both species Experiment 9 (Povinelli, 2011). During Orientation session 1-15 chimpanzees had a success rate of 55,1% in session 1-5, But their performance decreased within the following sessions with a success rate of 51,1% (Session 6-10) and then increased again to 53,1% (Session 11-15). Their overall performance within these 15 sessions was not statistically different to chance levels (as was the case for the Goffin's). The Goffin´s had a success rate of 47,6% (Session 1-5), 53,8% (Sessions 6-10), and 50,2% (Session 11-15).

Methodologically, the tasks given to the subjects (chimpanzees and Goffin´s) are similar, but they do differ with regards to the cue redundancy (or lack of) associated with the location each type of object should be placed. In Povinelli's task the trays were always located on the same side, while in our study we changed the location of the trays within every trial. This was in order to prevent a possible side bias that could develop as a result. It is likely that cue redundancy made the task easier for the chimpanzees. Therefore, it remains possible that if the Goffin´s could use tray location as a stable cue in addition to colour, their performance would increase.

We found a significant effect of the type of weighted object on the learning speed of the Goffins. We also we found a significant interaction between trial number, session number (the interaction between these two accounts for learning), weight and group-species (i.e., chimpanzees, and the two groups of Goffin subjects).

The difference in relative success rates between light and heavy objects we saw in the two Goffin groups is not statistically significant, so the significant four-way interaction in the Goffin-chimpanzee comparison is driven by the difference between chimpanzees and Goffin´s in the differences they have in success rate for heavy and light trials (there was a greater difference in success – i.e. rate of rule-learning – for heavy and light trials in chimpanzees). In the first 15 sessions of their sorting task, the chimpanzees were successful in 30,3% of light trials and 76,0 %) of heavy trials. The difference in success rate (both between the two Goffin groups themselves, and between these groups compared to the chimpanzees) could be accounted for by a particular salience of the heavy weight and the colour blue, in particular for chimpanzees (Group 1 of the Goffin´s and the chimpanzees were trained with the rule ‘heavy on blue’). One can say that, with regards to food items, heavy objects tend to be more valuable for animals than lighter ones, therefore they might pay more attention to a heavier cue (Heinrich, 1997; Schrauf et al., 2008). Additionally, a heavier object might provide a more salient cue due to greater stimulation of the muscular/nervous system, which could aid in association learning.

When considering the differences between these species, the weight difference of the objects compared to the subject’s body weight and strength should be considered. Although the weight difference in both studies were at an almost 10:1 ratio, the ratio according to the body weight of the subjects differed (6-8% of subjects body weight in the Goffin´s; 0,0007% for chimpanzees (Lambert & Stiegler et al. in Press). Strength in relation to body weight is difficult to measure, therefore chimpanzees might have perceived the weight difference in a divergent way. It is possible the chimpanzees perceived their light version of the objects (about 330g) as much ‘lighter’ than the Goffin´s did with their light weight (about 3g). It is also possible that Povinelli’s chimpanzees developed a colour preference during the orientation sessions where they already learned the association between heavy and the colour blue (Povinelli, 2011).

Our findings show that, as for apes who struggled to learn about the absolute weight of different objects (Schrauf & Call, 2011), Goffin´s seem to face difficulties when learning about weight in a relative sense. A food caching avian species might have found the task easier. It is important to consider the ecology of a species, and how ecological challenges have shaped cognition. In general, while there might be a need for parrots like Goffin´s to differentiate between weights in a decision-making situation, there is little need to remember the weight of an object. We show that the methodology used in Povinelli (2011) notably different to methodologies used in other weight discrimination tasks to date and that differences in methodology can have a huge impact on performance. Therefore, the conclusions we make about the cognitive abilities of the species tested should be cautious.

This is corroborated by the potential influences small sample sizes compared with methodological differences might have on results (Schrauf & Call, 2009). Due to our relatively large sample size, we can be confident of our conclusion that a sorting task is more difficult for Goffin´s to learn than a weight choice task. Previous work suggested a stark difference in the speed at which Goffin´s could master a weight discrimination task compared to primates (Lambert and Stiegler et al.in Press). The findings of our study don´t support this conclusion.

However, we cannot yet say if birds possess a higher order weight-related cognition than primates or not. Further studies will be needed to create a distinct concept about the species-specific differences regarding their concept of weight. To enable meaningful comparisons to be drawn, the number of trials that it would take the Goffin´s to learn the weight sorting task should be evaluated. Furthermore, Goffin´s should be tested on their causal understanding of weight. Additionally, primates should be tested on a weight choice task that is comparable to the one the Goffin´s participated in. Those findings could support our conclusions about the presence and the degree of weight related cognitive abilities within the species tested.

6.Summary

The discrimination of weight in the animal kingdom is less understood than discrimination within other senses. However, paying attention to weight influences and improves natural behaviours like extractive foraging and tool use. Whilst the basic perception of weight is not a cognitive ability, the way and speed with which a species can master a weight discrimination task reveals underlying cognitive processes.

Several studies have been carried out with multiple primate species, investigating learning on weight-based discrimination tasks. To generate a better understanding on how cognitive abilities pertaining to weight evolved across the animal kingdom, it is important to address these questions in different species in a comparative manner.

We tested the Goffin´s cockatoo *(Cacatua goffinana*), an extractive-foraging parrot species which has demonstrated sophisticated tool use and manufacturing skills, on a weight discrimination task. This weight sorting task replicated a study conducted with chimpanzees. Our findings show that this type of task is more difficult for the Goffin´s than the weight choice task they were tested with previously. Additionally, the Goffin´s performance on the sorting task did not meaningfully differ from the performance of chimpanzees. Both findings are important in shaping our growing understanding of how weight-related cognition might differ between birds and primates, given the different ecological challenges the groups experience. Lastly, the findings highlight that a consideration of methodological differences between tasks, and how these might relate to species-specific factors, is of upmost importance when drawing comparisons between species' cognitive abilities.

7.Zusammenfassung

Die Fähigkeit von Tieren Gewicht zu unterscheiden ist weniger gut erforscht als andere Sinne. Die Beachtung des Gewichts eines Objektes beeinflusst und verbessert jedoch natürliche Verhaltensweisen wie die extraktive Nahrungssuche und den Gebrauch von Werkzeugen. Die grundlegende Wahrnehmung von Gewicht ist zwar keine kognitive Fähigkeit, aber die Art und Geschwindigkeit, mit der eine Tierart eine Aufgabe zur Gewichtsunterscheidung bewältigen kann, lässt auf zugrunde liegende kognitive Prozesse schließen.

Es wurden mehrere Studien mit verschiedenen Primatenarten durchgeführt, in denen das Lernen bei gewichtsbezogenen Unterscheidungsaufgaben untersucht wurde. Um besser zu verstehen, wie sich die kognitiven Fähigkeiten in Bezug auf das Erkennen von Gewicht im gesamten Tierreich entwickelt haben, ist es wichtig, diese Fragen bei verschiedenen Arten vergleichend zu untersuchen.

Wir testeten den Goffin-Kakadu (Cacatua goffinana), eine Papageienart, die auf der Suche nach Nahrung ist und hochentwickelte Fähigkeiten im Umgang mit Werkzeugen und in der Herstellung von Produkten aufweist, bei einer Aufgabe zur Gewichtsunterscheidung. Diese Aufgabe zur Gewichtssortierung wurde in Anlehnung an eine Studie mit Schimpansen durchgeführt. Unsere Ergebnisse zeigen, dass diese Art von Aufgabe für die Goffin-Papageien schwieriger ist als die Aufgabe zur Gewichtsauswahl, mit der sie zuvor getestet wurden. Außerdem unterschied sich die Leistung der Goffins bei der Sortieraufgabe nicht wesentlich von der Leistung der Schimpansen. Beide Ergebnisse sind wichtig für unser wachsendes Verständnis der Frage, wie sich die gewichtsbezogene Wahrnehmung von Vögeln und Primaten angesichts der unterschiedlichen ökologischen Herausforderungen, denen diese Gruppen ausgesetzt sind, unterscheiden könnte. Schließlich unterstreichen die Ergebnisse, dass die Berücksichtigung methodischer Unterschiede zwischen den Aufgaben und die Frage, wie diese mit artspezifischen Faktoren zusammenhängen könnten, von größter Bedeutung sind, wenn man die kognitiven Fähigkeiten verschiedener Arten vergleicht.

8.Acknowledgement

I would like to thank Alice von Auersperg for supervising this Diploma thesis and giving me the opportunity to work with these great animals, as well as Poppy Lambert for all the help and advice she gave me during the whole process and for her statistical support, Daniel Povinelli for kindly providing the chimp data and Florian Kadic for the great illustrations of the Cockatoos.

9. References

Auersperg, A. M. I., Borasinski, S., Laumer, I., & Kacelnik, A. (2016). Goffin's cockatoos make the same tool type from different materials. *Biology Letters*, *12*(11), 20160689.

Auersperg, A. M. I., Laumer, I. B., & Bugnyar, T. (2013). Goffin cockatoos wait for qualitative and quantitative gains but prefer 'better' to 'more'. *Biology Letters*, *9*(3), 20121092.

Axe, J. B. (2008). Conditional Discrimination in the Intraverbal Relation: A Review and Recommendations for Future Research. *The Analysis of Verbal Behavior*, *24*(1), 159–174.

Field, A. P. (2009). *Discovering statistics using SPSS: And sex, drugs and rock 'n' roll* (3rd ed). SAGE Publications.

Heinrich, B. (1997). Black-Capped Chickadees and Red-Breasted Nuthatches 'Weigh' Sunflower Seeds. *The Auk*, *114*(2), 298–299.

Jablonski, P. G., Lee, S., Fuszara, E., Fuszara, M., Jeong, C., & Lee, W. Y. (2015). Proximate mechanisms of detecting nut properties in a wild population of Mexican Jays (Aphelocoma ultramarina). *Journal of Ornithology*, *156*(S1), 163–172.

Jelbert, S. A., Miller, R., Schiestl, M., Boeckle, M., Cheke, L. G., Gray, R. D., Taylor, A. H., & Clayton, N. S. (n.d.). *New Caledonian crows infer the weight of objects from observing their movements in a breeze*.

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.

Lambert, M. L., Schiestl, M., Schwing, R., Taylor, A. H., Gajdon, G. K., Slocombe, K. E., & Seed, A. M. (2017). Function and flexibility of object exploration in kea and New Caledonian crows. *Royal Society Open Science*, *4*(9), 170652.

Laumer, I. B., Bugnyar, T., & Auersperg, A. M. I. (2016). Flexible decision-making relative to reward quality and tool functionality in Goffin cockatoos (Cacatua goffiniana). *Scientific Reports*, *6*(1), 28380.

Laumer, I. B., Bugnyar, T., Reber, S. A., & Auersperg, A. M. I. (2017). Can hook-bending be let off the hook? Bending/unbending of pliant tools by cockatoos. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1862), 20171026.

Liu, Q., Simpson, K., Izar, P., Ottoni, E., Visalberghi, E., & Fragaszy, D. (2009). Kinematics and energetics of nut-cracking in wild capuchin monkeys (*Cebus libidinosus*) in Piauí, Brazil. *American Journal of Physical Anthropology*, *138*(2), 210–220.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

McCulloch, T. L. (1941). Discrimination of lifted weights by chimpanzees. *Journal of Comparative Psychology*, *32*(3), 507–519.

Mioduszewska, B. M., O'Hara, M. C., Haryoko, T., Auersperg, A. M. I., Huber, L., & Prawiradilaga, D. M. (2019). Notes on ecology of wild Goffin's cockatoo in the late dry season with emphasis on feeding ecology. *TREUBIA*, *45*, 85–102.

Nieuwenhuis, R. (2012). *Influence.ME: Tools for Detecting Influential Data in Mixed Effects Models*. *4*, 10.

O'Hara, M., Mioduszewska, B., Haryoko, T., Prawiradilaga, D. M., Huber, L., & Auersperg, A. (2019). Extraction without tooling around—The first comprehensive description of the foraging- and socio-ecology of wild Goffin's cockatoos (Cacatua goffiniana). *Behaviour*, *156*(5–8), 661–690.

Osuna-Mascaró, A. J., & Auersperg, A. M. I. (2018). On the brink of tool use? Could object combinations during foraging in a feral Goffin's cockatoo (Cacatua goffiniana) result in tool innovations? *Animal Behavior and Cognition*, *5*(2), 229–234.

Poppy J. Lambert*†1, Alexandra Stiegler*2, Theresa Rössler1,3, Megan L. Lambert1 & Alice M. I. Auersperg1, P. J. L., Alexandra Stiegler*2. (n.d.). *Goffin's cockatoos discriminate objects based on weight alone*.

Povinelli, D. (2011). *World without weightPerspectives on an alien mind*. Oxford University Press.

Schrauf, C., & Call, J. (2009). Great apes' performance in discriminating weight and achromatic color. *Animal Cognition*, *12*(4), 567–574.

Schrauf, C., & Call, J. (2011). Great apes use weight as a cue to find hidden food. *American Journal of Primatology*, *73*(4), 323–334.

Schrauf, C., Call, J., Fuwa, K., & Hirata, S. (2012). Do Chimpanzees Use Weight to Select Hammer Tools? *PLOS ONE*, *7*(7), e41044.

Schrauf, C., Huber, L., & Visalberghi, E. (2008). Do capuchin monkeys use weight to select hammer tools? *Animal Cognition*, *11*(3), 413–422.

Visalberghi, E., Addessi, E., Truppa, V., Spagnoletti, N., Ottoni, E., Izar, P., & Fragaszy, D. (2009). Selection of Effective Stone Tools by Wild Bearded Capuchin Monkeys. *Current Biology*, *19*(3), 213–217.

10. Appendix

Subject	Estimated Probability of Success
Dolittle	56,191
Figaro	47,817
Fini	46,482
Heidi	50,489
Irene	52,559
Jane	49,255
Kiwi	56,965
Konrad	51,69
Mayday	50,975
Moneypenny	53,037
Muki	52,677
Muppet	55,78
Olympia	57,628
Pipin	48,901
Titus	47,578
Zozo	54,564

Table 1. Estimated Probability of Success at Subject Level

Model 1	Estimate	Std. Error	z value	Pr(>|z|)
Intercept	0,08	0,1	0,805	
z.session	0,071	0,057	1,256	0,209
I(z.session^2)	-0,057	0,073	-0,778	
z.trial	0,029	0,081	0,359	
group.code	0,018	0,11	0,159	0,874
weight.code	-0,43	0,4	-1,075	
I(z.session^2):z.trial	0,041	0,062	0,663	
I(z.session^2):weight.code	0,118	0,246	0,48	
z.trial:weight.code	0,349	0,174	2,009	
I(z.session^2):z.trial:weight.code	-0,323	0,126	-2,574	0,01

Table 2. Estimates from model 1 with associated standard errors and z-values. P-values for the highest-order interaction terms are included (the P-values for any interactions or single terms included

in a higher-order interaction are not included due to their limited interpretation). Weight.code and group.code are the dummy-coded and centred variables.

Model 1a	Estimate	Std. Error	z value	Pr(>|z|)
Intercept	0,083	0,102	0,819	
z.session	-0,009	0,109	-0,083	0,934
z.trial	0,049	0,103	0,477	0,633
group.code	-0,105	0,199	-0,525	0,6
weight.code	0,205	0,201	1,022	0,307

Table 3. Estimates from model 1a with associated standard errors and z-values. P-values for the highest-order interaction terms are included (the P-values for any interactions or single terms included in a higher-order interaction are not included due to their limited interpretation). Weight.code and group.code are the dummy-coded and centred variables.

Model 2	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1,585	0,399	3,977	
z.trial	-0,121	0,201	-0,6	
z.session	0,287	0,291	0,987	
I(z.session^2)	0,352	0,195	1,803	
groupspeciesgoffin1	-1,302	0,407	-3,196	
groupspeciesgoffin2	-1,507	0,445	-3,383	
weightlight	-3,578	0,807	-4,435	
z.trial:z.session	-0,162	0,179	-0,902	
z.trial:I(z.session^2)	0,227	0,099	2,284	
z.trial:groupspeciesgoffin1	0,022	0,209	0,104	
z.trial:groupspeciesgoffin2	-0,099	0,222	-0,446	
z.session:groupspeciesgoffin1	-0,211	0,338	-0,623	
z.session:groupspeciesgoffin2	-0,28	0,384	-0,73	
I(z.session^2):groupspeciesgoffin1	-0,4	0,175	-2,283	0,022
I(z.session^2):groupspeciesgoffin2	-0,443	0,188	-2,357	0,018
z.trial:weightlight	0,039	0,274	0,144	
z.session:weightlight	-0,705	0,532	-1,326	
I(z.session^2):weightlight	0,048	0,28	0,172	
groupspeciesgoffin1:weightlight	3,02	0,852	3,547	
groupspeciesgoffin2:weightlight	3,705	0,922	4,019	

z.trial:z.session:groupspeciesgoffin1	0,286	0,202	1,411	
z.trial:z.session:groupspeciesgoffin2	0,311	0,217	1,437	
z.trial:z.session:weightlight	0,503	0,233	2,159	
z.trial:I(z.session^2):weightlight	-0,403	0,133	-3,036	0,002
z.trial:groupspeciesgoffin1:weightlight	0,267	0,273	0,976	
z.trial:groupspeciesgoffin2:weightlight	0,526	0,295	1,782	
z.session:groupspeciesgoffin1:weightlight	0,688	0,635	1,083	
z.session:groupspeciesgoffin2:weightlight	0,784	0,728	1,077	
z.trial:z.session:groupspeciesgoffin1:weightlight	-0,616	0,266	-2,319	0,02
z.trial:z.session:groupspeciesgoffin2:weightlight	-0,74	0,295	-2,511	0,012

Table 4. Estimates from model 2 with associated standard errors and z-values. P-values for the highest-order interaction terms are included (the P-values for any interactions or single terms included in a higher-order interaction are not included due to their limited interpretation). Chimpanzee is the reference level for group-species and heavy is the reference level for weight.

Model 3	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0,257	0,319	-0,807	
z.session	0,19	0,208	0,916	0,36
I(z.session^2)	0,346	0,227	1,527	
z.trial	-0,038	0,255	-0,149	
weight.code	-3,357	1,76	-1,908	
I(z.session^2):z.trial	0,08	0,218	0,365	
I(z.session^2):weight.code	-0,761	0,91	-0,836	
z.trial:weight.code	0,509	0,539	0,944	
I(z.session^2):z.trial:weight.code	-0,937	0,452	-2,074	0,038

Table 5. Estimates from model 3 with associated standard errors and z-values. P-values for the highest-order interaction terms are included (the P-values for any interactions or single terms included in a higher-order interaction are not included due to their limited interpretation). Weight.code and group.code are the dummy-coded and centred variables.