universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

A Bird's Eye View of Weight: Do kea rely on visual cues to infer weight?

verfasst von | submitted by

Tessa Marie Parker

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

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 013
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Joint-Degree Middle European interdisc. master prog. in Cognitive Science
Betreut von | Supervisor:	Mag. Dr. Ludwig Huber

Abstract

Humans utilize an assortment of clues to pre-determine the weight of an object, including but not limited to size, material, apparent density, and quantity. Previously believed to be a purely human capability, studies with species such as capuchins have shown that animals may also be capable of inferring weight from the visual qualities of an object. To date, this phenomenon remains relatively unexamined, however, and few studies have explored the extent to which other species attend to the visual signals regarding the weight of an object. In this study, we tested to see whether kea (Nestor notabilis), known for their curiosity and physical problem-solving skills, can use the visual cues of size, material, volume, and quantity (i.e., how full a container is) to determine the weight of an object prior to directly interacting with it.

The kea were previously trained to drop either a light (16g) or heavy (116g) object into a box. In this study, the kea were presented with pairs of objects – one light and one heavy – from the four different categories and were rewarded for dropping the object of their target weight into the box. We measured i) which object the birds first contacted, ii) whether they switched after handling (and feeling the weight of) incorrect objects, and iii) which object the birds dropped into the box.

The results from this study indicate that, while highly successful at choosing the correct object based on weight, kea do not appear to use visual markers such as size, material, volume, or quantity to determine how heavy an object is prior to interacting with it.

Follow-up studies currently in the works include determining the minimum weight difference kea are sensitive to and investigating whether kea rely on relative or absolute rules for weight.

Keywords

Visual cues, inferring weight, weight of an object, problem-solving, human capability, understanding of causality

Zusammenfassung

Der Mensch nutzt eine Reihe von Anhaltspunkten, um das Gewicht eines Objekts zu bestimmen, darunter Größe, Material, scheinbare Dichte und Menge. Bisher wurde angenommen, dass es sich um eine rein menschliche Fähigkeit handelt. Studien mit Tierarten wie Kapuzineraffen haben jedoch gezeigt, dass auch Tiere in der Lage sein können, aus den visuellen Eigenschaften eines Objekts auf dessen Gewicht zu schließen. Bislang ist dieses Phänomen jedoch noch relativ unerforscht und nur wenige Studien haben untersucht, inwieweit andere Tierarten auf die visuellen Signale bezüglich des Gewichts eines Objekts achten. In dieser Studie wurde untersucht, ob Keas (Nestor notabilis), die für ihre Neugier und ihre physischen Problemlösungsfähigkeiten bekannt sind, die visuellen Signale von Größe, Material, Volumen und Menge (d. h. wie voll ein Behälter ist) nutzen können, um das Gewicht eines Objekts zu bestimmen, bevor sie direkt mit ihm interagieren.

Die Keas wurden zuvor darauf trainiert, entweder ein leichtes (16 g) oder ein schweres (116 g) Objekt in eine Schachtel fallen zu lassen. In dieser Studie wurden den Keas Paare von Objekten - ein leichtes und ein schweres - aus den vier verschiedenen Kategorien vorgelegt und sie wurden belohnt, wenn sie das Objekt mit dem Zielgewicht in die Kiste fallen ließen. Wir haben erhoben, i) welches Objekt die Vögel zuerst berührten, ii) ob sie nach dem Hantieren mit (und dem Erfühlen des Gewichts von) falschen Objekten wechselten, und iii) welches Objekt die Vögel in die Box fallen ließen.

Die Ergebnisse dieser Studie deuten darauf hin, dass Keas zwar sehr erfolgreich das richtige Objekt anhand des Gewichts auswählen, aber offenbar keine visuellen Marker wie Größe, Material, Volumen oder Menge verwenden, um zu bestimmen, wie schwer ein Objekt ist, bevor sie mit ihm interagieren.
Zu den derzeit in Arbeit befindlichen Folgestudien gehören die Bestimmung des Mindestgewichtsunterschieds, auf den Keas reagieren, und die Untersuchung, ob Keas sich auf relative oder absolute Regeln für das Gewicht verlassen.

Schlüsselwörter

Visuelle Anhaltspunkte, Rückschlüsse auf das Gewicht, Gewicht eines Objekts, Problemlösung, menschliche Fähigkeiten, Verständnis von Kausalität

Acknowledgments

I owe a huge thanks to my supervisor and mentor, Megan Lambert, for bringing me onto this project and for all her help with the statistical analysis, support with the experimental design, and willingness to read multiple drafts of this thesis. Without her help I most certainly would never have finished this thesis, and I am so grateful for all her support and encouragement along the way. I am also deeply grateful to Sarah Jelbert at the University of Bristol for all her input throughout as to the design of the project and how best to realize it. Her contributions were invaluable and greatly appreciated. I would also like to thank the keepers at Haidlhof for all they do every day for the kea and for making this kind of research possible. I'd also like to give a special thanks to Denise Piringer for proofreading and correcting the German abstract – thanks for making my German look much better than it is! Finally, many thanks to the coordinators of the MEi:CogSci program, Elisabeth Zimmermann and Elisabet Delgado Mas, for helping me navigate my way.

To my 'flock' of friends and family: thanks for sticking with me even when all I'd want to talk about were kea parrots and object weight.

Table of Contents

1. Introduction

1.1. The Importance of Weight

Weight is a pervasive, invisible property that both humans and nonhuman animals must reckon with as they navigate their physical reality. While not directly observable, the property of weight (i.e., the 'mass' of an object) can be witnessed indirectly in a myriad of ways and plays a key role in planning and executing any action. It is the property of weight that explains why rocks are chosen to break windows rather than Styrofoam, or why someone taking the stairs to an 8^{th} floor apartment might choose to carry a single volume rather than the entire set of the *Encyclopedia Britannica* at once. Judging the weight of a single book versus a set of books requires not only the appropriate musculature, but the ability to represent weight in some perceptual (and potentially mental) way. While animals may never have to decide between hauling a single book or several hundred kilograms of encyclopedias up eight flights of stairs, they must also navigate a physical environment dominated by the property of weight.

1.2 Representing Weight as a Human

The human body is remarkably good at tracking and adjusting to the perceived weight (mass) of an object. As an object is lifted, the body tracks the amount of force it takes to overcome the gravitational pull on that object and to hold it in place against gravity. During this process, the body builds and stores an internal "sensorimotor representation" [Povinelli, 2012, p. 13] of the force required to lift that object. This representation is, essentially, an internal model of the bodily effort required to lift and hold different things and is built over time through practical experience with a myriad of physical objects. Nor is this representation static; upon physical contact with the object in question the representation is quickly updated to include any haptic information that was gained from contact with the object in order to improve weight prediction and refine the internal representation [Gordon et al., 1993].

Described separately as a long-term "force profile" by Daniel Povinelli [2012, p. 38] and as an "internal model" by Krakauer and Shadmehr [2006, p. 4], the internal representation formed by experience allows humans to subconsciously judge the weight of an object and to pre-program the appropriate amount of effort required to manipulate it prior to even touching it [Povinelli, 2012]. After such a representation has been formed it can then be called upon, even years later, via visual appraisal of an object with similar properties (for a visual depiction of the ways in which humans represent weight, see Table 1) [Povinelli, 2012]. That is not to say that such a method is foolproof. Few of us have managed to make it through life without spilling a beverage on ourselves from exerting more force than is strictly necessary when trying to pick up what we thought was a full cup. When this happens, we refer to it as embarrassing, but it is more technically known as a "positional overshoot" [Povinelli, 2012, p. 38]. Positional overshoot occurs when, after visually identifying an object, the object is determined to be of a type the sensorimotor system has encountered before. The sensorimotor system then applies the amount of force that it has 'stored' as necessary for lifting this type of object. In this instance, however, the internal representation does not match the actual weight of the object, and the object is lifted with greater force than necessary resulting in overshooting the mark. While it can be fooled, the sensorimotor system is remarkably adaptive. In the 1980's, Johannson and Westling demonstrated that our sensorimotor systems are capable of developing an "internal representation" for the force needed to lift an unfamiliar object after only a single trial (with a single trial being made up of several lifts) [Johannson & Westling, 1988, p. 83.]

Humans are not the only species to experience positional overshoot as will be shown further below; nor are we likely alone in the ability of our sensorimotor systems to develop 'pre-set force defaults' for interacting with familiar objects and objects similar enough to feel familiar. As a core property of the physical environment both humans and nonhuman animals inhabit, weight is something all species must be capable of coping with.

Table. 1. A table depicting how the human body and brain represents weight. In this thesis, we are focusing on level 6; the representation of "weight" as a causal mechanism across perceptually distinct categories [Text and image credit Daniel Povinelli, 2012].

Anti-gravity muscles	The mass of many muscles in the body quickly atrophy in response to weightlessness
Effort required to lift object	The sensory-motor system tracks the force used to lift a given object
Effort experienced in lifting object	Humans (and other animals) have phenomenological experiences of effort lifting and moving objects
Feed forward model of effort required to lift object	The sensory-motor system represents the predicted effort needed to lift a given object
Causal representation of a relation involving "weight"	The perceptual system represents the causal connections between object and outcome (organisms may use size, weight, shape, effort sensation, or any cluster of these to learn to predict an "effect" from a "cause")
Effects of object weight across contexts, or *f*(WEIGHT)	The human mind represents "weight" as a causal mechanism common across perceptually distinct categories
Linguistic representation	Humans generate linguistic (symbolic) labels for the causal mechanism of weight ("weight", "heavy", "light", etc.)
Metaphorical expression	Humans generate linguistic (symbolic) descriptions that connect both the phenomenological aspects of weight and *f*(WEIGHT) to nonliteral contexts

1.3 Visual Properties of Weight

Before even picking up an object, humans make several inferences regarding its properties. These inferences stem from prior experience and are built on the associations between certain visual qualities of the object and its properties such as weight [Adelson, 2001; Hirsiger, 2012]. Properties such as the size, color, material, and shape of an object are all visually identified. This process occurs effortlessly and automatically but yields a wealth of information [Adelson, 2001]. These inferences allow humans to determine how to use the object; if it is a large, flat stone it might be used to crush a nut or hold down an errant piece of paper. If it is a sturdy, wooden stick it might be used to move something closer that is just out of reach (such as the cookie jar at the back of the shelf). The relationship between the visual properties of an object and its weight is learned early on in human beings. At 11 months, human infants can already use visual information such as the material composition of an object to determine its weight [Paulus & Hauf, 2011]. By the age of 3, children can use the property of size to predict the weight of an object [Gordon et al., 1992]. After enough experience, novel objects can be correctly categorized by humans as "light" versus "heavy" based on their visual qualities alone.

In concert with the visual-system and the internal long-term "force profile" [Povinelli, 2012, p.38], humans also develop additional 'felt sense' representations of weight. These representations of weight include the subjective experience of weight (i.e., how heavy or light something feels) as well as the causal linking of the weight, size, or density of an object and the sound it makes or the size of the dent it creates when it hits the ground or strikes a metal door. Alongside these sense-based representations are representations mediated at a conscious, cognitive level [Flanagan & Beltzner, 2000; Ellis & Lederman, 1998]. This cognitive-based system of weight perception operates independently of the sensory-based system and is hypothesized to account for the various 'weight illusions' humans (and potentially other animals) experience including the size-weight illusion [Flanagan & Beltzner, 2000].

In the size-weight illusion, participants are presented with two objects that are identical apart from size. Although the objects weigh the same amount, a lifetime of experience associating certain visual cues and object weight creates strong expectations which, when violated, create the perceptual illusion for the participants that the larger object is lighter than the smaller one[1] [Charpentier, 1891, as cited by Murray, D.J., 1999]. The size-weight illusion has been shown to be cognitively impenetrable so that even when the participants have been informed that the objects are the same weight, they still report that the larger object feels lighter than the smaller one [Flournoy, 1894, as cited by Buckingham, 2014]. Similarly to the size-weight illusion, the material-weight illusion also causes participants to perceive one object as being lighter than the other despite the two objects having identical weights [Wolfe, 1898]. In this case, however, it is due not to a size difference but to a difference in surface material. In instances where two objects are identical apart from surface material, participants will perceive the object that appears to be made from a heavier material (such as metal) as being lighter than the one that appears to be made from a lighter material (such as polystyrene) [Wolfe, 1898; Seashore, 1899, as cited by Saccone et al., 2019; Harshfield & DeHardt, 1970; Buckingham et al., 2009]. In the case of both illusions, although the participants' motor-systems adjust over time and begin to apply the correct amount of lift force, the participants continue to report the larger object or the heavier-looking object as being the lighter of the two [Buckingham, 2014; Buckingham et al., 2009].

[1] Interestingly, it has even been shown that the size-weight illusion can be caused by social expectations around weight as in the 'Barbie Experiment' [Dijker, 2008].

These illusions not only highlight the deeply ingrained connection between weight and certain visual cues in humans, but the apparent co-existence of at least two internal representations of weight: sensorimotor and cognitive [Povinelli, 2012]. While the sensorimotor system primarily concerns itself with the 'veridical weight' of individual objects encountered in the physical environment, the cognitive system is more concerned with the learned associations between object properties (such as size) and weight – allowing it to make judgments regarding entire "families" of objects [Flanagan et al., 2008, p. 1744].

At this cognitive level, the brain not only makes determinations regarding families of objects, but also ties together disparate perceptual experiences of weight into a unified theory about how weight works. Representing weight as a "causal variable" [Povinelli, 2012, p. 14] allows the human mind to see the immediate connection not only between the heavy object hitting the ground and the loud thud, but also the connection between the heavy object hitting a packed surface, that large thud, and the sudden gouged appearance of the dirt. This ability to perceive the connection between disparate perceptual experiences and the weight of a single object leads to an understanding of weight as a persistent, causal property of physical objects; it is also generally held to be a uniquely human capability [Povinelli, 2012].

1.4 Weight Representation in Non-human Animals

Given that humans are not the only animals that must make judgments about objects to successfully navigate their physical world, it is likely that humans are also not alone in the ability to form these types of sensorimotor and cognitive representations.[2] While it is generally accepted that nonhuman animals likely rely on sense-based representations, it is less clear whether they are also capable of forming more cognitively sophisticated representations that rely on an understanding of weight as a causal variable and an inherent property of physical objects [Povinelli, 2012]. Although currently there are no conclusive studies confirming that nonhuman animals represent weight cognitively, some evidence (as discussed below) exists to support the idea. Most of this existing evidence comes from

[2] N.B.: The word "representation" is used throughout this thesis to indicate an "internal model"; in the case of the term "sensorimotor representations," the word "representation" refers to the internal pre-set patterns of lift that the body applies prior to lifting objects it recognizes. In the case of "cognitive representations," the term "representation" refers to an understanding of the property of weight as a causal variable and property of physical objects [Povinelli, 2012].

studies done with primates as, until recently, weight understanding in other nonhuman animals remained almost entirely unexplored.

1.5 Povinelli's Experiments

In 1999, a team of researchers headed by Daniel Povinelli (cited above) attempted one of the first comprehensive studies on weight perception in nonhuman animals and their understanding of weight as a causal property of physical objects. At his lab in New Iberia, Povinelli focused his efforts on humanity's closest relative: the chimpanzee. Povinelli set out to systematically explore, not only the chimpanzees' ability to form sensorimotor representations of "effort-to-lift" [Povinelli, 2012, p. 13], but their ability to deduce causal relationships between the weight of an object and its impact upon the world.

One of the early experiments Povinelli and his team conducted was a 'discrimination' experiment which investigated the chimpanzees' ability to discriminate between two visually identical, but differently weighted objects on the basis of weight alone [Experiment 9; Povinelli, 2012, p. 92]. To test this ability, the researchers constructed two identical objects with one weighing approximately 3155g while the other weighed only 330g. The researchers then presented the chimpanzees with the objects one at a time while instructing them to place it in the correct sorting tray based on its weight. While the chimpanzees proved capable of sorting between the two objects, it took them an average of 895.2 trials to reach criterion [Povinelli, 2012]. This high number surprised the researchers who had assumed the task would be easy for the chimpanzees; interestingly, as will be shown further on in the paper, other animals seem to have a much easier time with sorting visually identical, but differently weighted objects.

In addition to testing their ability to sort objects based on weight, Povinelli and his team also tested whether the chimpanzees formed stable sensorimotor representations of the effort it took to lift a given object and how rapidly this representation of "effort-to-lift" was formed [Experiment 1; Povinelli, 2012, p. 62]. In one of the first experiments in the series, the chimpanzees were tested using pairs of identical, but differently weighted objects. After a series of lifts, the object was swapped out for its much lighter identical twin, giving researchers the chance to measure how much higher the chimpanzees lifted the new, unexpectedly lighter object. While at first the chimpanzees lifted the new object considerably higher, as they got used to the new weight of the object, the height of the lifts decreased, indicating that the chimpanzees had formed new, stable sensorimotor representations around the weight of the object (Fig. 1). In 2017, this finding was also replicated in a field-based study with wild chimpanzees where the researchers showed that the wild chimpanzees initially experienced positional overshoot (lifting the object with more force than necessary) when lifting artificially hollowed out ‘hammers’ compared to normal ‘hammers’ [Sirianni et al., 2017].

Fig. 1. The experimental situation for Experiment 1 testing the chimpanzees “effort-to-lift” (a) Light object in position on trial 2, (b, c) one of the chimpanzees (Megan) lifting the light box on her fifth trial
(Images and partial text credit Daniel Povinelli, 2012)

1.6 Inferring Weight from Visual Cues

After establishing that the chimpanzees did, in fact, form stable sensorimotor representations about “effort-to-lift,” the researchers pivoted to testing whether the chimpanzees formed *default* sensorimotor representations that would allow the chimpanzees to reliably lift novel objects with the appropriate amount of force [Povinelli, 2012]. It is well-established that humans form default sensorimotor representations based on prior experience lifting various objects [Gordon et al., 1993]. As previously mentioned, the visual features of an object such as its size, texture, shape, and even color, cue humans to the properties of an object, including its potential weight, prior to handling; ongoing experience with the relationship between the visual features of an object and its weight are also what allow the human sensorimotor system to develop “default lift effort” settings for objects based on their visual appearance [Povinelli, 2012, p. 79; Adelson, 2001].

There are solid evolutionary reasons for having such a skill set. The ability to determine the potential weight of an object without handling it first has the twin benefits of limiting risk prior to interacting with the object as well as assisting with planning an action by pre-judging the amount of effort it could take to lift the object [Jelbert et al., 2019]. Given the wide-ranging uses of such an ability, it seems likely that animals as well as humans would share in it from a survival perspective. Despite this, little is currently known about how animals use visual information to determine the weight of an object prior to handling it. A study conducted in 2009 with wild capuchin monkeys shed some light on the subject when it was demonstrated that capuchins are capable of relying on visual information to infer the weight of different stone tools [Visalberghi et al., 2009]. In the case of these wild capuchins, they used visual cues such as size and material to identify the functional stone whenever such cues were both available and reliable. However, to this day, few other studies exist that probe the link between visual cues and weight inference. Povinelli and his team attempted to begin answering this question by presenting the chimpanzees with identically weighted objects of different sizes to test whether the chimpanzees formed default "effort-to-lift" representations based on the visual features of the objects alone.

In the first of a series of three experiments [Experiments 4, 5, and 6; see Povinelli, 2012, p. 79 – 87] exploring this question, the chimpanzees were presented with a small, novel object to lift (Fig. 2a). After several trials, the original object was swapped out for a second object that was identical apart from size. In both cases the lifting height was measured and later compared (Fig. 2b). This study was then replicated a second time with novel cylindrical objects with similar results (Fig. 2c).

Fig. 2a and 2b. Kara preparing to lift the (a) small and (b) large boxes in Experiment 4. Both objects weighed 1.13 kg [Text and image credit Daniel Povinelli, 2012]

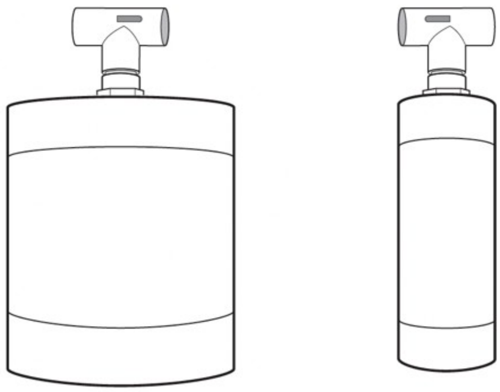


Fig. 2c. The novel cylindrical objects used in Experiment
[Text and image credit Daniel Povinelli, 2012]

In the third and final study in the series, the researchers tested whether the chimpanzees were capable of using cues besides size to infer the weight of an object and tested them using identically weighted rough versus smooth textured boxes.

1.7 Results

The results from the first two studies of the series [Experiments 4 and 5] showed that the chimpanzees were initially lifting the larger box higher than the smaller box despite the equal weight of the boxes – the expected result if the chimpanzees were relying on the visual cue of size to determine how much effort it would take to lift the object. The chimpanzees' overall performance during these two experiments lent evidence to the theory that chimpanzees form default sensorimotor representations that allow them to predict the amount of effort it will take to lift a given object (familiar or novel) and that visual cues such as the size of the object activate these representations [Povinelli, 2012].

The third study in this series [Experiment 6], which investigated whether chimpanzees use visual cues beyond weight to predetermine lift force, presented a somewhat mixed picture, however. In the experiment testing the use of texture as a visual cue, the chimpanzees were first given extended experience handling and moving a variety of both rough and smooth objects with the weight of the objects adjusted so that the smooth objects weighed significantly less than the heavy objects. After familiarizing the chimpanzees with these objects, the researchers tested them using two identically weighted objects with differing textures: one rough and one smooth. Based on the chimpanzees' earlier experiences with the heavy rough objects and the light smooth objects, the researchers predicted that if the chimpanzees relied on default lift force settings that they would initially exhibit positional overshoot and lift the rough test object much higher than the smooth test object. Despite their earlier experience with the heavy rough objects, however, the chimpanzees actually ended up lifting the smooth test object higher than the rough test object. The researchers posited that there were two possible explanations for this result; first, it is possible that the programming of default lift settings is more context dependent for chimpanzees than it is for humans and that familiarizing the chimpanzees with the rough and smooth objects in a different setting than they were tested in had prevented them from transferring their experience with the training objects to the novel test objects. Second, it could also be that the programming of default lift settings in chimpanzees is limited to objects that are visually highly similar, and texture or material are not salient enough cues to for the chimpanzees to have lift expectations for differing texture or material objects [Povinelli, 2012].

Regardless of the reason, this result led the researchers to conclude that, while the chimpanzees did seem to rely on visual cues such as size to preprogram lift force, they were not relying on the use of texture or material to do the same. An additional study looking at whether the chimpanzees could use auditory cues to infer the weight of an object provided similarly mixed results and led the researchers to conclude that there was no evidence to suggest that the chimpanzees could use cues other than size to predict lift force [Povinelli, 2012].

1.8 Indirect Inference Weight Studies

Additional experiments with the chimpanzees in Povinelli's lab continued to provide mixed results regarding the ability of the chimpanzees to infer weight without bodily input and to represent weight cognitively [Povinelli, 2012]. After a series of experiments testing everything from the chimpanzees' ability to infer the weight of an object based on the pulling actions of a human actor [experiment 15] to their ability to successfully predict which structure could support the weight of a plastic block [experiment 27], Povinelli concluded that, while there was evidence to suggest that the chimpanzees could form sensorimotor representations of weight, there was no evidence to indicate that the chimpanzees could reason about weight as an unobservable, constant, causal property of physical reality. In light of a lack of substantive evidence, Povinelli and his team posited that, while nonhuman animals were likely capable of forming sensorimotor based representations of weight, only humans are capable of reasoning about weight as an abstract, unobservable property [Povinelli, 2012].

1.9 Counter Studies

Povinelli's claim that humans alone possess the ability to represent weight cognitively as an abstract, invisible property of objects met with immediate pushback from other researchers, some of whom advocated for a more nuanced approach that took into consideration the possibility that animals can deploy such abilities in certain contexts [Buckner, 2013]. Still others pointed to studies by other research groups that, in contrast to Povinelli's claim, have seemed to find evidence to suggest that animals may be aware of the property of weight and capable of using information about it to solve real-world problems.

1.9. 1 Weight as a Causal Property

In *World Without Weight*, Povinelli makes the claim that nonhuman animals are incapable of reasoning about weight as an invisible property of objects [Povinelli, 2012]. Yet several earlier studies (all of them conducted with apes) seem to provide evidence to the contrary.

A study conducted with great apes tested their ability to use the cue of weight to find the cup baited with food. In this study (which tested 5 bonobos, 6 orangutans, and 1 gorilla) the apes were presented with two opaque cups; one empty and one baited with banana slices. The apes were then allowed to choose one of the cups. Unable to rely on any visual cues to find the food, the apes spontaneously turned to feeling the weight of each cup in turn prior to choosing one. Interestingly, in the arbitrary test condition where the presence of the food was indicated, not by the causal cue of the weight of the food, but by an arbitrary visual cue (in this case different colored triangular and circular stickers), only one of the apes was able to successfully solve the task [Schrauf & Call, 2011]. This study suggests that nonhuman animals are capable of using a relevant, causal cue such as weight to solve a problem – potentially indicating an awareness of the property of weight. As the researchers of the study put it,

> Although both types of cues possess the same predictive value, namely 100% for the correct reward location, only one of them has causal relevance to the task at hand. In

> the present study, both weight and color/shape are indicators for the presence of food. Nevertheless, only the weight cue was caused by the food and thus provided a source of information that is grounded on the laws of physics [Schrauf & Call, 2011, p. 332]

In another study exploring nonhuman understanding of both weight and causality, Hanus and Call demonstrated that chimpanzees seem capable of deducing the presence of a desirable food item using the property of weight. In this particular study the researchers placed two opaque cups on the opposite ends of a balance beam. Under the subject's watchful eyes, a trainer placed a piece of prized banana in one of the cups without giving away which cup they had picked. The weight of the banana naturally tilted the cup and the researchers found that the chimpanzees consistently chose to search for the banana in the lower cup. To test if the chimpanzee were simply attending to the salient downward motion, Hanus and Call also created a 'non-causal' control test where, instead of the weight of the food pushing it down, an experimenter pushed down one side of the balance beam after a few seconds. During this condition, the chimpanzees did not show a preference for the lower side of the balance beam [Hanus & Call, 2008; see Povinelli, 2012 for a critique].

1.9. 2 Weight as a Means to an End

Other studies have also lent support to the possibility that nonhuman animals are both aware of the property of weight and capable of using it to serve their ends. One study conducted with tool-using capuchin monkeys tested, first, whether the capuchins were capable of associating tool weight with the effectiveness of the tool, and second, whether the capuchins would continue to pick the more effective, heavy tool when it was the smaller of the two available tools [Schrauf et al., 2008]. In the first experiment of the study, the capuchins were presented with three identical percussors of varying weights. After incentivizing the capuchins to handle each percussor by coating the bottom with chocolate cream, the capuchins were presented with nuts that were hard enough to require a heavy tool to open. The results from the first experiment indicated that the capuchins were not only capable of differentiating between the tools by weight alone, but that they were also capable of associating weight with increased effectiveness. In the second experiment of the study, one of the capuchins was presented with two percussors: one large and light and the other small and heavy. After a few trials, the capuchin began selecting the small, heavy tool first, leading the researchers to conclude that it was possible the capuchin was relying on his previous experience with the heavy tools being the most effective. The results of the two experiments proved, not only that the capuchin were capable of developing sensorimotor representations of weight, but that they were potentially linking the weight of the object with its probable effectiveness [Schrauf et al., 2008].

In the same theme, another example of support for the idea that nonhuman animals appear cognizant of the property of weight and capable of associating it with real-world objects, comes from an observational study conducted on chimpanzees on the west coast of Africa. The wild chimpanzees in the Taï National Forest in Africa frequently use heavy stones to break open hard *Panda oleosa* tree nuts and softer *Coula edulis* tree nuts. In the early 1980's, it was reported by researchers that these chimpanzees were using the weight of the stones and the distance between the stones and their desired destination to determine which stones to choose. Researchers found that the chimpanzees seemed to choose heavier stones when the distance between the stone and their destination tree was shorter, and lighter stones when the distance between the two was much longer. In addition, the chimpanzees seemed to deliberate between lighter and heavier stones depending on the specific type of nut they needed to crack, indicating at least an awareness of weight if not the ability to cognitively represent it [Boesch & Boesch, 1984].

Taking all these studies into consideration, it appears there is evidence to suggest that (at least some) nonhuman animals may possess an understanding of weight that goes beyond mere perceptual experience and memory. While some researchers such as Daniel Povinelli have dismissed the results from the studies above as merely evidence for extensive sensorimotor memory and capabilities, others argue that it is premature to dismiss such results out of hand when it is likely that the difference between human and nonhuman weight understanding is much more subtle and context reliant [Buckner, 2013]. In either case, this thorny question remains unanswered and, apart from the studies mentioned above, as well as a few other scattered experiments [see Schrauf et al., 2012; Schrauf & Call 2009; and Schrauf & Call, 2011], research into nonhuman weight understanding has remained limited and the question relatively unexplored. When the question has been explored the studies have focused almost exclusively on primates; however, as will be described further below, birds are now emerging as a promising candidate for future research into how nonhuman animals understand weight.

1.10 Avian Cognition

Reasoning about weight in the abstract (i.e., without direct proprioceptive feedback) is a cognitively complex undertaking [Apps, 2006]. Perhaps for this reason, tool-using species such as chimpanzees and capuchins have previously been the primary focus of investigations into nonhuman weight understanding [see Boesch & Boesch, 1984; Povinelli, 2012; Schrauf et al., 2012; Schrauf & Call, 2009; and Schrauf & Call, 2011 among others]. Despite the rigidity and pre-programmed manner in which many species use tools [Biro et al., 2013], tool use is often assumed to be an indicator of the level of cognitive sophistication of a species [Brosnan, 2009; Hayashi, 2015], prompting some researchers to focus specifically on tool-using species for cognitive research. There is evidence, however, to indicate that non-tool using species may possess equivalent cognitive sophistication [Seed & Call, 2010]. In one instance, a group of rooks (*Corvus frugilegus*), a non-tool using species, successfully solved a modified trap-tube test typically used with primates and tool-using birds [Tebbich et al., 2007]. In another instance, yellow-crowned bishops (a species of weaverbird) were tested on their ability to flexibly problem-solve both with and without using tools [Danel et al., 2021]. Although the weaverbirds proved unable to solve the task using the provided tools, they were able to solve the task using their beaks. While studies on weight cognition have previously focused primarily on tool-using primates, the wide range of cognitive abilities demonstrated by both tool-using and non-tool using birds position them as promising candidates for this area of research. In addition, while weight is undoubtedly a prevalent factor for every species, it is likely to be particularly ecologically relevant to birds given their light weight, smaller size, and their tendency to transport objects during flight. Given that birds often fly to get to one place from another, it is also reasonable to assume that birds would need to factor in properties such as weight in order to cope with transporting nesting materials or carrying food long-distances. As Povinelli remarks in his book *World Without Weight*,

> Feathered flight is an energetically extreme form of location and therefore avian species ought to be extremely sensitive to the embodied representation of weight involved in activities such as transporting nesting materials. Thus, it is easy to speculate that evolution may have sculpted the neural systems of birds to track effort-while-lifting for goal-directed purposes [Povinelli, 2012, p. 323, footnotes].

Believed to have last shared a common ancestor with primates around 324 million years ago [Dos Reis et al., 2015], certain species of birds such as corvids and parrots have performed at levels on-par with primates on a broad range of cognitive tasks [Lambert et al., 2019; Emery 2004, Emery & Clayton, 2004; Wimpenny et al., 2009; Tebbich et al., 2007]. Large-brained, long-lived and highly social, parrots and corvids occupy a similar socioecological niche to primates and were likely subjected to similar evolutionary pressures resulting in some of the shared abilities seen today [Baciadonna et al., 2021]. Both parrots and corvids have likewise proved themselves to be capable problem-solvers. In one experiment, New Caledonian crows were shown to drop objects into tubes containing water to raise the level to reach a floating reward [Jelbert et al., 2014]. In another, wild Goffin Cockatoos proved capable of using tools to access seed matter from tropical fruit [O'Hara et al., 2021]. Cognitive flexibility, the ability to modify one's strategy depending on the problem at hand, is perceived to be a key component of intelligence [Cantwell et al., 2022]. Both parrots and corvids have shown an aptitude for flexible problem-solving [Auersperg et al., 2011] with some species of corvid, such as the New Caledonian crow, proving to also be advanced tool users [Bayern et al., 2018]. Indeed, in some cases, initial research even suggests that corvids and parrots are quicker to solve tasks that require them to use weight as a discriminating factor or to use causal cues to infer weight. In contrast to apes, Goffin's cockatoos (7/8 subjects) reached criterion in a similar weight discrimination task after only 60.6 trials on average compared to the nearly 895.2 trials that the chimpanzees had required [Lambert et al., 2021; Povinelli, 2012]. Similarly, while chimpanzees have previously failed to solve tasks that required them to infer the weight of an object [Povinelli, 2012], New Caledonian crows have proved quite capable of inferring object weight. In one study, the crows proved capable of using the causal cue of a light object blowing in a breeze while the heavy object remained still to later correctly choose between the light and heavy object [Jelbert et al., 2019; see Henley, 2020 for a critique]. Causal reasoning about the physical property of weight has long been thought to belong to humans alone, however, these recent studies with birds are challenging these assumptions and raising the possibility that species outside of humans are in fact capable of relying on causal cues to infer weight. Despite initial promise, there are still many unknowns about what nonhuman animals understand about the concept of weight. While studies have attempted to explore weight understanding in primates, to this day almost nothing is known about how other species, including birds, perceive or infer weight. Currently, even less is known about the visual properties other species might use to make inferences about weight. The present study seeks to address several of these gaps by exploring how a habitually non-tool using parrot, the kea, might use visual cues to infer weight.

1.10. 1 Why Kea?

A highly playful and insatiably curious parrot with a penchant for problem-solving, the kea (*Nestor notabilis*) present themselves as ideal candidates for further research into what nonhuman animals understand about weight (Fig. 3). From an ecological standpoint the kea seem well placed to learn about the property of weight. As generalist, extractive foragers, the kea are highly experienced at digging, pulling, and otherwise manipulating objects in their environments, providing them with ample direct experience with object weight [Diamond & Bond, 1999]. Living in a changing, seasonal environment with few natural predators, the kea are also remarkably neophilic and eager to explore and handle any new object they come across [Diamond & Bond, 1999]. This carefree approach to novel situations and objects means that the kea are likely to have extensive experience with a variety of materials, textures, and weights. Equipped with remarkable problem-solving abilities, the kea have also proven to be cognitively flexible and capable of innovating tool-use both in the wild and in captivity. When baited trap-boxes were installed in New Zealand to catch non-native stoats, the kea, despite having never before been seen using tools in the wild, quickly found ways to set off the traps and access the food inside using sticks. Even after the boxes had been adapted to prevent the kea from accessing them, the kea continued to find ways to remove the food [Goodman et al., 2018]. In another example of their ferocious curiosity and exploratory bend, the kea found multiple creative ways to solve the puzzle of a Multi-Access Box and were quick to switch to new methods whenever a previously successful method had been blocked [Auersperg et al., 2011].

Fig. 3. Juvenile kea at play (Haidlhof, 2023)

The kea have likewise proven capable of apparently spontaneously learning about object properties through exploration. In a study conducted with captive kea, the kea were given the opportunity to thoroughly explore several different objects (some of which could be used as tools) prior to being tested. Subsequently, during testing, the kea were then presented with those same objects and allowed to choose from among them to solve a task. When presented with the objects during testing, six out of the eight kea selected the object that would allow them to solve the task. Of particular relevance to the current study, several of the kea correctly used a heavy object on a weight-oriented portion of the task, potentially indicating an understanding of the property of weight. This result led the researchers to surmise that, during the exploratory phase, the kea had learned about the different properties of the objects, including weight [Lambert et al., 2017].

Another indication of the keas' suitability for studies on weight understanding is the speed at which they can learn to discriminate between visually identical light and heavy objects. While it took Povinelli's chimpanzees an average of 895.2 trials to discriminate based on object weight [Povinelli, 2012], a more recent pilot study conducted with kea showed that the kea can learn to discriminate between objects of different weights after an average of only 30 trials [Lambert, unpublished data]. Although a different study conducted with Great Apes (bonobos, orangutangs, and gorillas) showed that at least some apes may be capable of discriminating between objects with differing weights after a much shorter duration of trials (5 out of 12 apes reached criterion after 30 sessions), birds still seem to grasp the concept of weight much quicker overall than their featherless counterparts. [Schrauf & Call, 2009; Lambert et al., 2021].

These reasons position the kea as likely candidates for possessing an understanding of weight beyond simple sensorimotor representation. As studies regarding weight understanding have remained particularly limited in species outside of the primates, studying how the kea perceive and understand weight is likely to be particularly fruitful and provide a more comprehensive picture of how nonhuman animals navigate a world dominated by the property of weight.

1.11 Hypotheses and Predictions

Broadly, the aim of this study is to investigate what a species outside of the primates understand about weight. More specifically, this study is meant to explore whether a species of parrot, the kea, is aware of and capable of inferring object weight from the visual properties of an object prior to handling it. To do this, the kea will be presented with one pair of objects at a time from four different conditions: a **size** condition, a **material** condition, a **quantity** condition, and a **volume** condition. The objects within each of these conditions will have fixed weights (16 grams or 116 grams) and the kea will be trained to select the object that matches the weight they were trained to attend to. During each condition, the kea will have the opportunity to choose an object in exchange for a food reward.

Our hypothesis is that the kea use visual cues such as the size, material, quantity, and volume of an object to infer its weight prior to directly interacting with it. If this hypothesis is correct, then we predict that the kea will first contact the object of their target weight significantly more often than is expected by chance.

1.12 Interdisciplinarity

A note as to the nature of this study. The study of cognition in nonhuman animals is fundamentally an interdisciplinary undertaking. Rooted in cognitive biology, it necessarily draws from fields such as philosophy, ethology and psychology to answer a perennial question: are we alone in how we understand the world?

Examining such a broad topic requires grounding it by addressing a specific question from the outset. Our particular question, "do kea rely on visual cues to infer weight?" attempts to compare our human perspective on a core understanding of the physical world with that of an alien mind. As with other core knowledge systems that rely on physical input, weight perception and understanding is facilitated by both the body and the mind. This shared responsibility for shaping the way an organism interacts with and understands its environment situates the topic of weight solidly in the realm of predictive processing, phenomenology, and embodied cognition; all core concepts spanning the disciplines of cognitive biology, philosophy, and psychology, and united within the framework of cognitive science.

Perception (at least for our uses) is defined as, "the processing of sensory information that occurs at several levels" [Montemayor & Haladjian, 2017, p. 1]. By its very nature, weight perception is dependent upon the presence of a physical body; perceiving the weight of an object requires both activation of the perceptual and sensorimotor systems as well as the use of cognitive functions that allow for the linkage between the physical force required to lift the object and hold it against gravity and the subjective 'feeling' of weight. In turn, cognitively representing and understanding weight relies on this body-based perception. Without the sensorimotor memory of the 'felt' weight of an object it would be impossible to establish a cognitive representation of weight as a causal entity. Weight understanding is therefore embodied in the sense that cognitively representing weight is impossible without the perceptual input from the body.

Studying weight understanding is likewise an inherently phenomenological task as 'felt' weight is an innate part of the human and (it seems likely) the animal experience. It is clear that, to at least some extent, weight is subjectively experienced [Buckingham & Donald, 2019]. Indeed, the fact that the size-weight illusion is cognitively impenetrable and continues even after our sensorimotor system has adjusted to the actual weight of the object is a sign of the phenomenological character of weight.

It is well-established that the subjective experience of weight is influenced by the order in which a series of weighted objects is picked up [Polanen & Davare, 2015]. If you start with the heaviest object, you are more likely to perceive all the objects following that first object as lighter than they actually are. This can be taken as evidence of the predictive nature of our relationship with weight. Predictive processing concerns how we perceive the world. From the viewpoint of predictive processing, we do not experience the world directly, but instead we experience a version of the world co-created by the filter of our senses and our brain's predictions [Clark, 2023]. As with phenomenology, illusions such as the size-weight illusion lend support to the position that we experience the world not as it is, but as our brain *predicts* it is.

We can only indirectly theorize what nonhuman animals perceive and understand about weight. While we can directly ask another person what their experience is of lifting a surprisingly light object, we do not have the same ability to gain first-person insight into the subjective experience of a nonhuman animal. This lack of access necessitates that we approach nonhuman cognition from an interdisciplinary perspective. Occupying a diverse range of environments and body plans, nonhuman animals are unlikely to experience or understand weight in exactly the way we do. Only by approaching the concept of weight in the multitude of ways it can be perceived and understood can we begin to understand how an alien mind might understand weight.

2. Material and Methods

2.1 Test Subjects and Housing Conditions

The subjects participating in this study were captive kea (*Nestor notabilis*) drawn from a population housed at the Haidlhof research station in Bad Vöslau, Austria and jointly maintained by the Department of Cognitive Biology of the University of Vienna, the Messerli Research Institute of the University of Veterinary Medicine Vienna, the Medical University Vienna, and the University of Vienna. The population at Haidlhof is a blend of hand-raised and parent-raised kea of varying ages. The kea housed at Haidlhof are frequently used in behavioral and cognitive studies and participate voluntarily in testing. The birds are kept together in an enriched aviary measuring 52m x 10m x 4 and provided with a daily diet containing seeds, protein, fruit, and vegetables (Fig. 4). Throughout the study the kea were fed their normal diet and had free access to water, including during testing. To prevent the kea from learning about the study prior to participating in it, subjects were tested in one of two separate experimental compartments that were attached but could be visually closed off from the rest of the aviary.

Fig. 4. The kea aviary at the Haidlhof Research Station in Bad Vöslau
(Image credit by Haidlhof Research Station)

The kea kept at Haidlhof are frequently involved in different types of studies; up until this point, they had already participated in several weight related studies, all of which used similarly sized clay objects of different shapes and colors. The previous studies focused on weight discrimination (i.e., sorting visually identical objects based on weight) and remembering the weight of objects they had interacted with earlier. Some of the kea also participated in a replication of an experiment done with New Caledonian crows [Jelbert et al., 2019] where the objects are placed in front of a fan and the kea had to infer the weight of the object based on how it moved in the breeze. Prior to the beginning of the current study, however, none of the kea had participated in studies that related weight and visual cues such as the size, material, quantity, or volume of an object.

Before running the study, a pilot was conducted using a single male kea, Paul, as the subject. Paul had a target weight of 116 grams and was selected in part because he generally performs well during tests. The pilot study was run with the primary objective of figuring out if the newly constructed objects we wanted to use for testing (specifically certain previously untried objects such as the transparent globes used in the Quantity and Material categories) were appropriate for use with the kea. During the initial pilot, Paul successfully contacted his target weight object first in 8/10 trials and chose the correct object in 10/10 trials. Taking Paul's promising performance as an indicator that the test objects were appropriate, the main study proceeded with a total of 16 kea: 7 females and 9 males (see Table 2 below). Of the original 16 kea selected, only 14 completed the study. Two of the kea, Willy and Skipper, were dropped from the study partway through; Willy was dropped due to a lack of interest during testing and Skipper was dropped due to his habit of destroying every object he was given.

Table 2: List of kea participating in the main study

Name:	DOB:	Sex:	Rearing:	Weight Assignment:
Anu	2007	Male	Hand	Heavy
Coco	2007	Female	Hand	Heavy
Diana	2017	Female	Hand	Heavy
Frowin	2004	Male	Parent	Heavy
Kermit	2004	Male	Hand	Heavy

Lilly	2007	Female	Hand	Heavy
Odo	2015	Male	Parent	Heavy
Paul*	2010	Male	Parent	Heavy
Skipper**	2017	Male	Hand	Heavy
Willy**	2007	Female	Parent	Heavy
Fay	2016	Female	Hand	Light
Jean Luc	2015	Male	Hand	Light
Kiri	2015	Female	Parent	Light
Pancake	2017	Male	Hand	Light
Pick	2004	Male	Hand	Light
Plume	2007	Female	Hand	Light
Roku	2008	Male	Parent	Light
Sunny	2007	Female	Hand	Light

**Pilot study; not used in main study*
*** Dropped from the study due to lack of motivation*

2.2 Prior Training and Test Apparatus

Prior to beginning the study, the kea were randomly assigned a target weight of 16 grams (light) or of 116 grams (heavy) and trained to discriminate between light and heavy objects. These weights were chosen because they were both within the range of what the kea could comfortably lift and the difference between 16 grams and 116 grams was significant enough to allow for easy discrimination. The birds were trained to select the object of their target weight and drop it into a remote-controlled reward box (see Fig. 5) which would then dispense a 1/4 of a peanut as a reward. The objects used (see Fig. 6) were visually identical and made from non-toxic grey fimo modelling clay. To achieve the desired weight of either 16 grams or 116 grams, the objects were filled with paper or with small, non-toxic fishing weights depending on the weight needed. The objects were approximately 5 cm x 2.5 cm.

Fig. 6. The visually identical objects used during training, the pilot, and the refamiliarization and the pre-test phases of the study

The initial training with the objects was conducted over a series of several sessions; first, the light or heavy object (depending on the bird's target weight) was placed on the rim of the reward box and the bird was rewarded for nudging the object into the box. Once the bird was consistently nudging the object into the box, the object was placed in front of the box at increasing distances and the bird was rewarded only when it had picked up and successfully dropped the object into the box. After the kea had successfully completed training, they were presented with both the heavy and light objects simultaneously and only rewarded for dropping their target weight object into the box (Fig. 7). The kea were considered as having successfully passed once they had selected the object of their target weight in at least 8 out of 10 trials for two consecutive sessions (which all the subjects ultimately did in about 3 sessions).

Fig. 5. The remote-controlled reward box used during training, refamiliarization, pre-tests, and testing

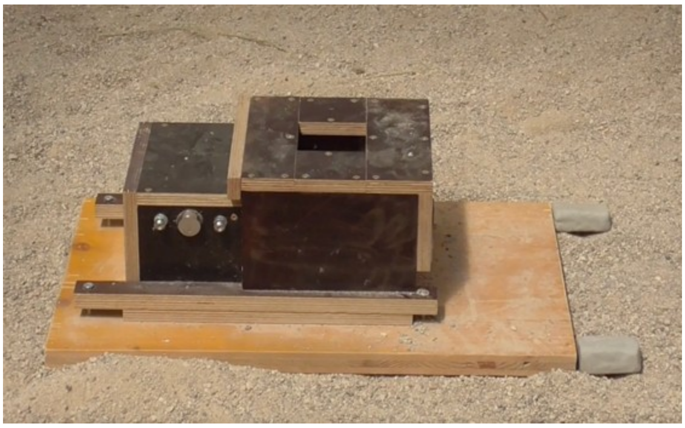


Fig. 7. Set-up for training and testing with original objects

2.3 Stimuli

In addition to the objects used during both training and the re-familiarization phase (see: **Testing Procedure**), a set of 64 objects in total were used as the test stimuli (Fig(s). 8, 9, 10, 11). These objects were chosen to reflect materials that the kea had likely had at least passing experience with over the course of their lifetime. This was done to limit object novelty and ensure that the kea were only asked to infer weight based on visual properties they had previous exposure to and were likely to have formed weight associations with. The objects used in the testing phase of the experiment fell into one of four categories:

Size: 8 object pairs made from fimo modelling clay and varying in color and shape (Fig. 8). Within a pair, there was a light (approx. 16g) and heavy (approx. 116g) version with the heavier object approximately 60% bigger than the lighter object. This particular size difference was chosen based on the results from research done during another student's master's thesis which investigated the ability of kea to establish a relational concept of size [Köck, unpublished data]. Objects within the same pair were of the same color and shape. To achieve the proper weight, objects were filled either with wooden beads (light) or small fishing weights (heavy). For pair 8 (the blue globes) both objects had a zip tie added to it when it was discovered that the kea struggled to grip and lift the larger object.

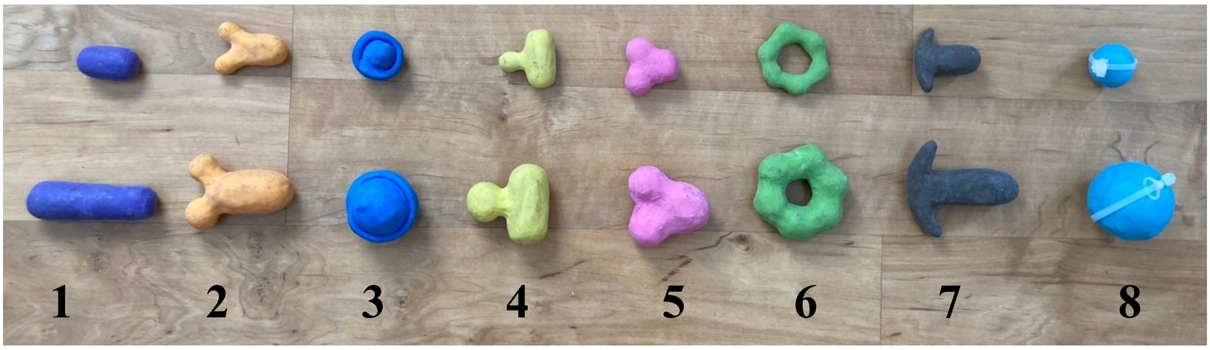


Fig. 8. Objects from the Size category (heavy objects on bottom row)

Quantity: 8 object pairs composed of fillable acrylic globes (Fig. 9). Each pair was made up of a light (approx. 16g) and heavy (approx. 116g) version. The globes were filled with varying amounts of sand, rocks, water, and clay with the light versions filled with Elmer's glue or smaller amounts of material and the heavier versions supplemented with small fishing weights or tungsten mud as well as the aforementioned material.

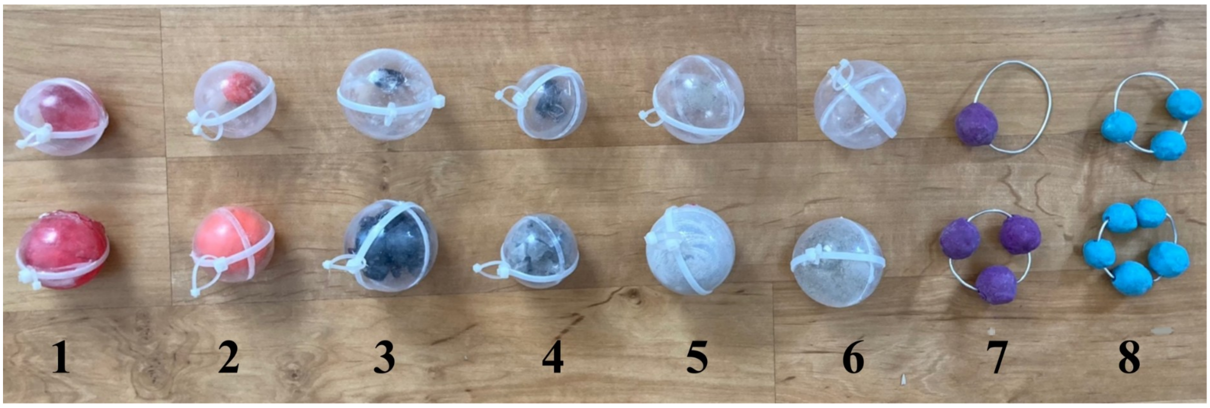


Fig. 9. Objects from the Quantity category (heavy objects on bottom row)

Volume: 8 object pairs comprised of objects with and without holes (Fig. 10). The solid object (no holes) weighed 116 grams while the hollow object (with holes) weighed 16 grams. Objects within a pair were made of the same material with the light object made obviously hollow or less dense. The materials used included rubber, wood, clay, metal, string, and cork. Light objects had their weight supplemented with glue while heavy objects had their weight supplemented with glue, small fishing weights, clay, and tungsten mud.

Fig. 10. The objects from the Volume category (heavy objects on bottom row)

Material: 8 object pairs of varying shape, size, and materials (Fig. 11). Each pair was made up of two objects of different materials with roughly the same size and shape. As with the other categories, each pair was made up of a light object (16g) and a heavy object (116g). The object pairs included (from left to right): a circle of wooden rings and a circle of metal rings, a sock ball and a piece of brick, two connected balls of paper and two connected metal balls, a rock and a sponge, a cardboard 'disc' and a metal 'disc', an acrylic globe filled with feathers and an acrylic globe filled with bolts, a grass filled acrylic globe and a sand filled acrylic globe, and a mini cardboard 'pot' and a mini terracotta pot.

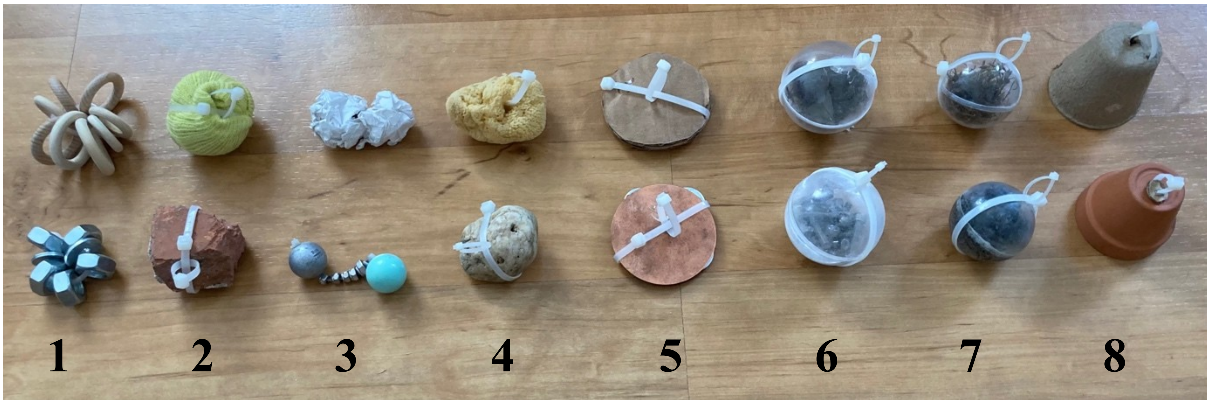


Fig. 11. The objects from the Material category (heavy objects on bottom row)

In addition to the above-mentioned objects, the reward box successfully used in prior weight studies and during training was also utilized during the current study. The reward box was remote controlled and would dispense a peanut when a bird dropped the target object into the rectangular hole at the top.

3. General Testing Procedure

3.1 Test Set-Up

The study was conducted in one of the two experimental compartments in the aviary (see Fig 12). The experimental compartment was visually closed off from the rest of the aviary by movable opaque walls and could be divided in two by a wire gate, creating a 'waiting room' for the kea. Testing took place over approximately 2 months beginning on May 5th, 2023, with each bird being tested once daily several times per week for a total of 8 sessions per bird.

3.2 Refamiliarization

Prior to formally beginning the tests, the kea were each given 10 trials with the visually identical training objects to ensure they still remembered which target weight to select. After they had successfully passed this phase by correctly choosing their target weight object 8/10 times within a session, they were then moved to the testing phase.

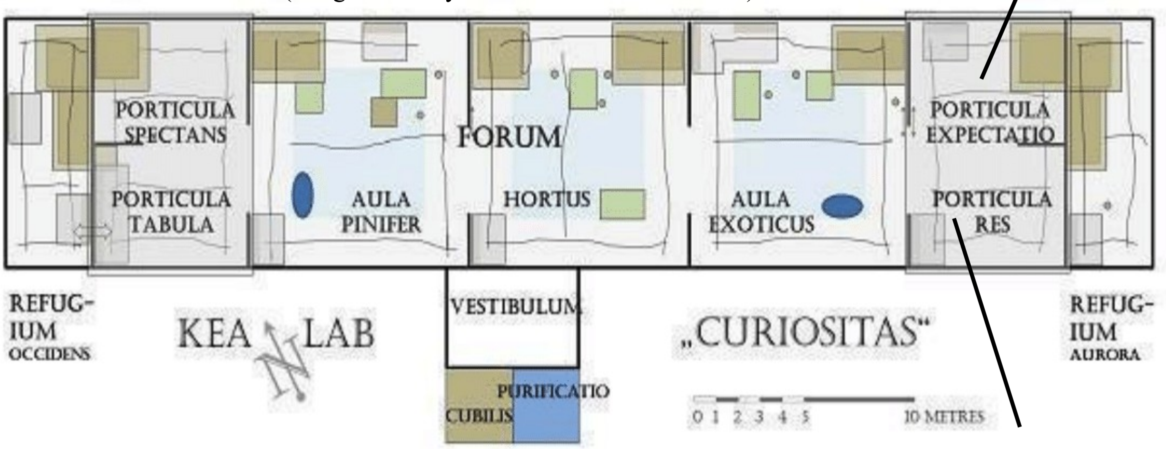


Fig. 12. Map of the kea aviary with experimental compartment and waiting room labeled (Image credit by Haidlhof Research Station)

3.3 Testing Phase

Each test session was comprised of 3 parts: **Pre-Test**, **Observation**, and **Testing**.

Pre-Test: To ensure the kea remembered the procedure and were attending to their target weight, each test session began with 4 'pre-test' trials. During the pre-test, the two visually identical objects used during both training and refamiliarization were placed at a distance from the reward box while the kea watched from the waiting room. The kea was then let out of the waiting room and given the opportunity to select one of the two objects to drop into the box (Fig. 13). If the right object was dropped in, the kea was rewarded with a ¼ of a peanut. If not, the kea was ushered back into the waiting room without a reward. The pre-test consisted of 4 trials and the kea had to correctly drop in their target weight object 3 out of 4 times to move to the next phase of testing. None of the kea had any issues passing the pre-test each time.

Observation: Immediately after successfully completing the pre-test portion, the kea was escorted back into the waiting room for the observation phase (Fig. 14). During the observational phase, the birds were presented with all four pairs of novel objects that would be used in the session (one pair from each category). This was done to reduce the novelty and decrease the likelihood that the bird would choose an object based on curiosity rather than its weight. In this phase, the birds were separated from the objects with a movable fence and allowed to observe, but not interact with, all 8 objects for 1 minute.

Testing: After the observational period had ended the test phase began. Testing was comprised of 4 trials within each session with 8 sessions in total per subject. At this stage, all the objects were removed, and the reward box was placed in the testing compartment along with the first pair of novel objects from a randomized category. The bird was allowed into the testing compartment and given the opportunity to choose one of the two objects and then rewarded with a 1/4 of a peanut if it chose its target weight object (Fig. 15). If the bird chose wrong, it was gently ushered back into the waiting compartment and the test was reset with new objects for the next trial.

After each test trial, the bird was escorted into the waiting room while the test was reset using the next pair of objects. This process was repeated until all four pairs (one pair from each category) had been presented to the bird. The category order was balanced across 8 sessions for each subject and the position of the light and heavy objects both within each trial and across sessions was randomized to prevent the bird from developing a side bias or learning a particular pattern. In addition, the order of the object pairs within each session was also randomized to prevent any kind of pattern from potentially biasing the results.

Fig. 13. Pre-test with training objects

Fig. 14. Observation with objects from the Quantity category

Fig. 15. Test set-up with objects from Material category

3.4 Behavioral Coding

There were several behavioral measures that were assessed during each test trial: First Contact, Choice, and Switching. First Contact referred to the first object the kea touched when it entered the testing compartment and was the primary measure used during the study. First Contact was chosen as the primary measure because we were interested in the birds' spontaneous inferences regarding object weight before they had physically interacted with the objects. If they approached and contacted the correct object first, that was an indicator that they were potentially making the connection between the visual properties of the object and its weight. The second measure, Choice, was chosen because it served as an indicator as to whether the kea understood the task at hand. If they approached the wrong object first, but ultimately dropped in their target weight object, that signified to us that they had understood what was expected of them in the task even if they had not made the correct initial inference. Finally, Switching, referred to whether the kea switched sides upon contacting the wrong (or correct!) object. This measure was chosen because we were curious to see if the kea would recognize if they had contacted the wrong object and whether they would switch. It also served as an indicator for confusion in the kea as the birds would often switch back and forth multiple times when they seemed more uncertain about which object was correct within a particular pair. These measures were all manually noted down in the moment. Additionally, all sessions were recorded with a video camera and reanalyzed later in case of any doubt regarding the behavior of the bird during the session or the outcome of a particular trial.

4. Results

4.1 Methods Used in Analysis

Data were analyzed in R (version 3.6.0). To analyze whether, as a group, the birds performed above chance level (0.5) in any of the four object categories, we performed one sample t-tests on the birds' first contacts, and ultimate choices. We used one-tailed binomial tests to determine whether individuals first contacted or chose the correct object above chance within each object category (i.e., 7/8 or more trials correct). A paired samples t-test was used to compare the number of switches (i.e., contacting one object before dropping in the second) depending on whether subjects first contacted the correct or incorrect object. We also descriptively assessed subjects' side biases by session and overall. This could provide a rough idea of the strategy subjects were using when they first approached the objects; specifically, whether they tried to first choose based on visual information and adjusted as necessary once feeling the weight of the object, or whether they stuck with a single side, ignoring visual information, and adjusted based on the felt weight. To do this, we calculated the total number of times subjects first contacted the Left object in each session as well as overall, using a binomial level of significance (i.e., 24 or more choices on one side, $p < 0.05$) for the overall choice data.

To estimate the effects of Object Category on first contact (binary, correct or incorrect), we fitted a Generalized Linear Mixed Model [GLMM, Baayen, 2008], with a binomial error structure and logit link function using the function lmer of the package lme4 [Bates et al., 2015]. We included Object Category as a fixed effect, as well as Target Weight and Session to control for the effects of the latter two. Subject and object pair were included as random intercepts effects. We included all theoretically identifiable random slopes, specifically: object category (manually dummy coded and centered) and session (z-transformed) within subject, and session and target weight (manually dummy coded and then centered) within object pair.

We began with a maximal model [Barr et al. 2013] which included the correlations among random intercepts and slopes, however, those among the random effect of Object Pair had absolute correlation values near one, suggesting these to be unidentifiable [Matuschek et al., 2017]. Removing these correlations from the model led to a minor reduction in model fit (log-liklihoods, model with correlation parameters included: -294.8193, model without correlation parameters: -295.6457). As an overall test of the effect of Object Category, Target Weight, and Session, we used a likelihood ratio test [Dobson, 2002] to compare the full model to a null model excluding these three effects. Tests of the individual fixed effects were derived using likelihood ratio tests [Barr et al. 2013]; R function drop1 with argument 'test' set to 'ChiSq.'

Model stability was assessed using a function kindly provided by Roger Mundry, which compared the estimates obtained from the model based on all data with those obtained from models with the levels of the random effects excluded one at a time [Nieuwenhuis et al., 2012], revealing the model to be stable (see results). Collinearity, determined for a standard linear model lacking the random effects and the interaction, did not appear to be an issue (variance inflation factors, obtained using the VIF of the package car, were all near one; [Quinn and

Keough, 2002]). Confidence intervals were derived using the confint function of the package lme4 (Bates et al., 2015) using the Wald Method.

The sample for all models encompassed 447 trials (including 248 successes), taken from 14 individuals out of 8 sessions.

4.2 Test Results

Our primary measure for whether subjects made inferences about object weight prior to contacting the object was based on which object the birds first touched in a trial. Overall, 5/14 subjects first contacted the correct object above chance in at least one object category (see Table 3), with one of these subjects (Frowin) performing above chance in two categories (Size and Quantity). All subjects performed at chance levels in the Volume category. As a group, subjects did not first contact the correct object significantly more than would be expected by chance in any of the four object categories: (Size: $\bar{x}$ = 4.5, (t(13)=1.102, p = 0.291; Quantity: $\bar{x}$ = 4.9, (t(13)=1.710, p = 0.111; Volume: $\bar{x}$ = 4.3, t(13)=1.578, p=0.139; Material: $\bar{x}$ = 3.9, t(13)=-0.186, p=0.856).

Subject (Heavy)	Category	Correct 1st Contact	Correct Choice
Anu			
	Size	7	8
	Volume	6	8
	Quantity	5	8
	Material	3	5
Coco			
	Size	4	8
	Volume	4	7*
	Quantity	4	8
	Material	7	6*
Diana			
	Size	6	8
	Volume	5	8
	Quantity	6	8
	Material	5	7
Frowin			
	Size	7	8
	Volume	4	7*
	Quantity	8	8
	Material	2	8
Kermit			
	Size	7	8
	Volume	4	8
	Quantity	6	8
	Material	3	7
Odo			
	Size	3	7
	Volume	4	8
	Quantity	6	8
	Material	4	7
Lilly			
	Size	4	8
	Volume	4	8
	Quantity	8	8
	Material	3	4

Subject (Light)	Category	Correct 1st Contact	Correct Choice
Pick			
	Size	5	8
	Volume	5	7*
	Quantity	3	6
	Material	4	6
Plume			
	Size	4	5
	Volume	5	7
	Quantity	5	6
	Material	4	6
Roku			
	Size	5	8
	Volume	5	7
	Quantity	3	5*
	Material	6	8
Sunny			
	Size	2	8
	Volume	4	6
	Quantity	2	5
	Material	2	5
Fay			
	Size	3	8
	Volume	2	5*
	Quantity	4	6*
	Material	5	7
Pancake			
	Size	3	7
	Volume	4	8
	Quantity	4	8
	Material	3	5
Kiri			
	Size	3	8
	Volume	6	5*
	Quantity	5	6*
	Material	4	6

Table 3. Breakdown of successful first contacts and choices per object category for each subject. Subjects with a "Heavy" designation had a target weight of 116 grams while subjects with the "Light" designation had a target weight of 16 grams. The * mark indicates that the results from one or more sessions were counted as "0" because the subject did not make a choice (see **Results** for more details)

In terms of which object was ultimately dropped into the box, the birds showed very high accuracy dropping in the correct object significantly more than would be expected by chance in all categories (see Fig. 16): (Size: $\bar{x}$ = 7.6, $t(13)=16.19$, $p < 0.001$; Quantity: $\bar{x}$ = 6.9, $t(13)=8.251$, $p < 0.001$; Volume : $\bar{x}$ = 7.1, $t(13)=10.724$, $p < 0.001$; Material: $\bar{x}$ = 6.2, $t(13)=6.972$, $p < 0.001$). In certain trials, the kea failed to choose an object to drop into the box.

In these cases, the trial was reset, and the kea were given a second opportunity to select an object to drop into the box. In the case that they did drop an object into the box this was denoted with either a "1" to indicate a correct choice or a "0" to indicate an incorrect choice. In instances where the kea failed once again to make a choice, this was also indicated with a "0." This is denoted in Table 3 with an asterix mark next to the number to indicate that in one or more trials in a given category the kea failed to make a final object choice, and the result was recorded as "0."

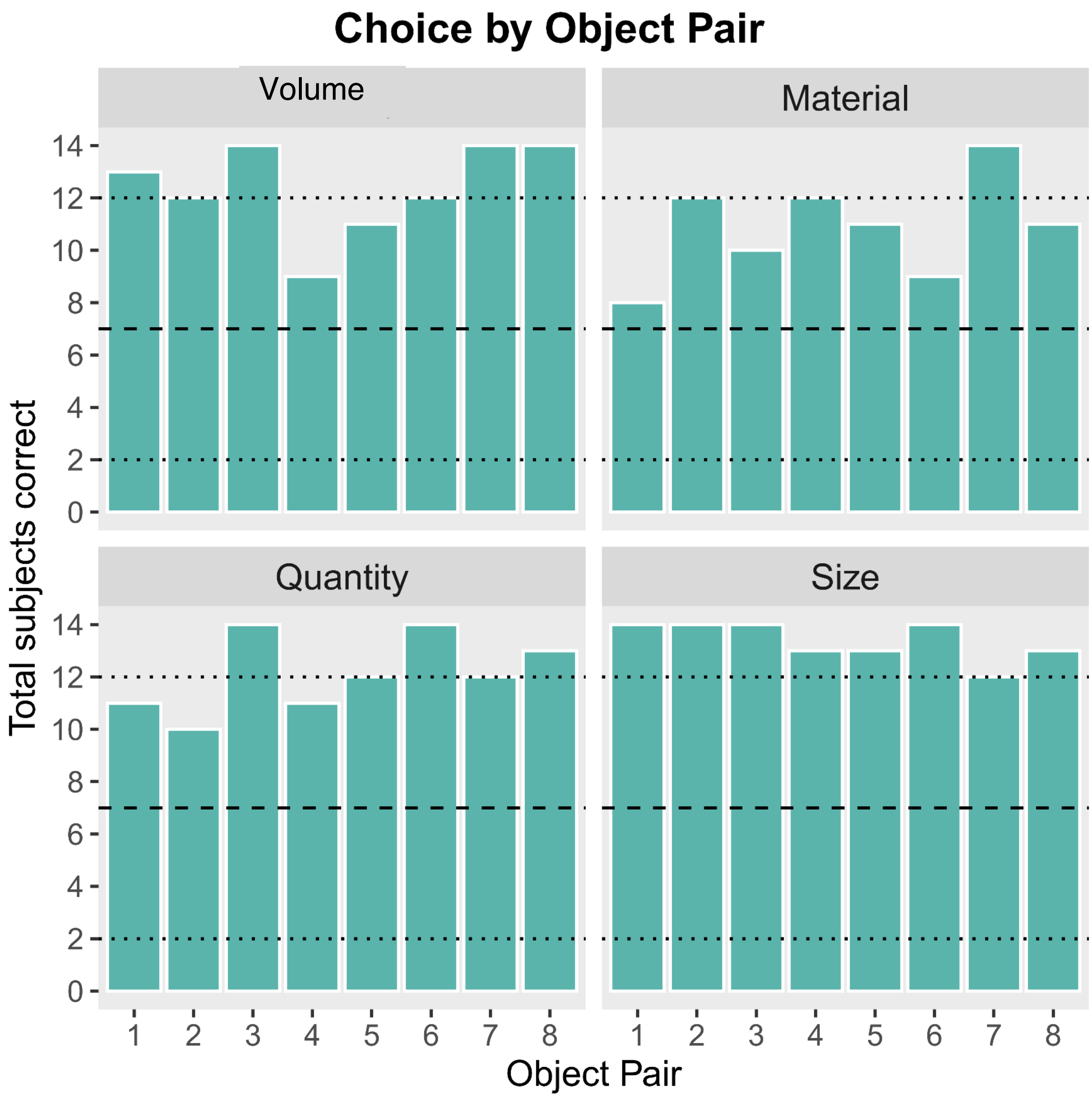


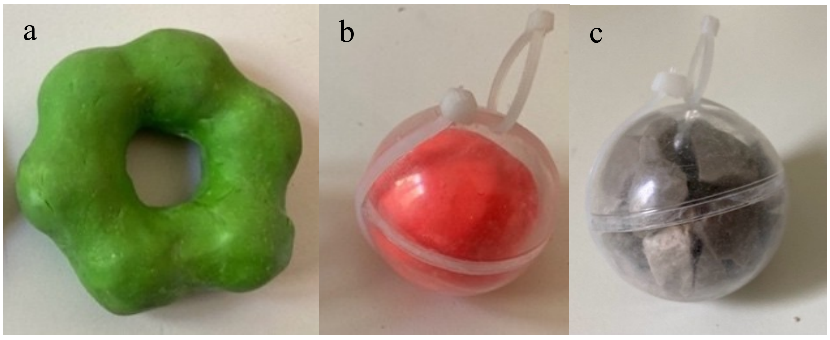
Fig. 16. Plot depicting the final object choice of each subject over the 8 sessions. Dashed line represents the median, dotted lines represent upper and lower chance limits, such that bars at or above upper dotted line would reflect a significant group-level tendency to choose the correct object of that pair, and bars at or below lower dotted line would reflect group level tendency to choose the incorrect object of the pair

Table 4. Results of GLMM of effects of Object Category, Target Weight, and Session on probability to first contact correct object in each trial. Includes estimates, standard errors (SE), confidence intervals (CI), test results and minimum and maximum of model estimates after dropping levels of individual random effects sequentially

Term	Estimate	SE	Lower CI	Upper CI	x^2	Df	P	Min	Max
Intercept	0.525	0.246	0.042	1.008				0.399	0.665
Object Category – Material	-0.268	0.303	-0.861	0.326	2.093	3	0.553	-0.431	-0.123
Object Category – Quantity	0.308	0.329	-0.337	0.953				0.144	0.452
Object Category – Size	0.053	0.291	-0.518	0.624				-0.068	0.199
TargetWeight – Light (1)	-0.588	0.282	-1.141	-0.034	4.033	1	0.045	-0.699	-0.454
Session(2)	0.183	0.112	-0.036	0.401	2.576	1	0.109	0.142	0.228

(1) Dummy coded with 'heavy' being the reference category
(2) Z-transformed to a standard deviation of zero and a mean of one

There was no significant effect of Object Category (see Table 4 above) (x^2 =2.093, df=3, p = 0.553) or session (x^2 =2.576, df=1, p = 0.109) on whether subjects *first contacted* the correct object, however, there was a significant effect of Target Weight x^2 =4.033, df=1, p = 0.045). Specifically, subjects with the heavy target weight were more likely to first contact the correct object. Visual inspection of the plots revealed the difference between target weight categories to be most pronounced in the Size and Quantity categories (Fig. 17 next page). Post Hoc binomial tests revealed that the birds showed strong, group-level preferences for the heavier object (regardless of their target weight) in three object pairs (Size: Pair 6, and Quantity: Pairs 2 & 4, see Fig(s). 18a, b, c).



Fig(s). 18a, b, c. The heavy objects from: pair 6 in Size category (left image), pair 2 (middle image) and pair 4 (right) from the Quantity category (to see the light version of these objects refer to Fig(s). 8 and 9)

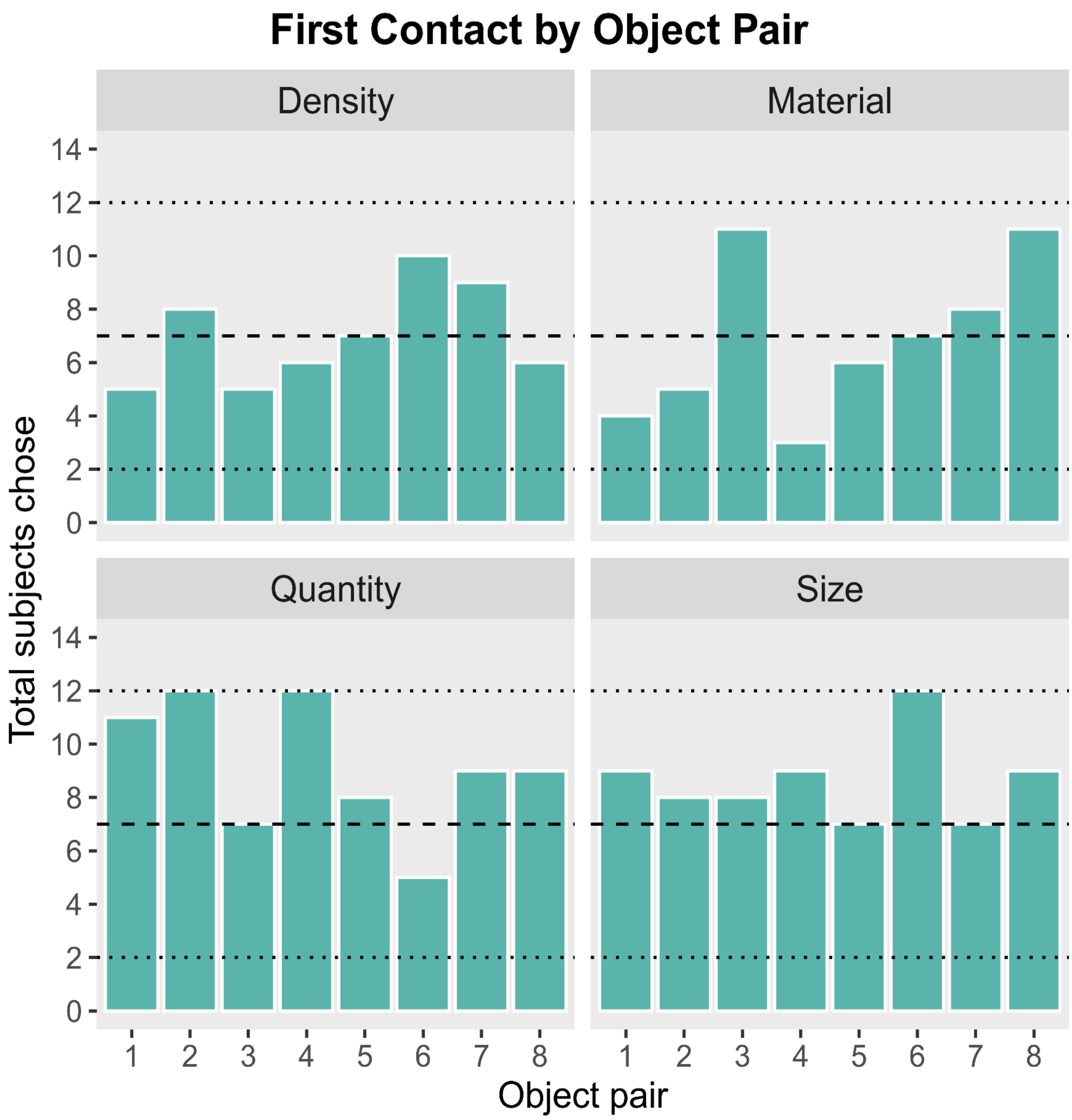


Fig. 17. The total number of subjects (out of 14) in either target weight category that first contacted the heavy object within a pair. Dashed line represents the median, dotted lines represent upper and lower chance limits such that bars at or above upper dotted line would reflect a significant group-level tendency to first contact the heavy object of the pair, and bars at or below lower dotted line would reflect group level tendency to first contact the light object of the pair

As subjects were allowed to contact both the heavy and the light objects during each trial, we assessed whether their tendency to switch (i.e., contact the second object after contacting the first object) depended on whether they first contacted the correct or incorrect object. A necessary precondition for selecting the correct object was switching when subjects first contacted the incorrect object. Indeed, subjects switched an average of 11.64 times (range: 6 – 15) throughout their 8 sessions (Fig. 19), and significantly more often when they had first contacted the incorrect object (paired samples t-test: ($t(13)=16.111$, $p < 0.001$).

The subjects' switching and choice performance indicates that they clearly understood the task (and that side biases were not a problem for choice). When assessing whether subjects exhibited side biases when *first contacting* an object, we found that 12/14 subjects exhibited a side bias of either 100% right or left first contacts in at least one session (see Fig. 20), but that, overall, only 4/14 subjects (all males, 2 in the light group, 2 in the heavy group) first contacted the Left object significantly more than would be expected by chance (i.e., 24 or more times out of 32 trials).

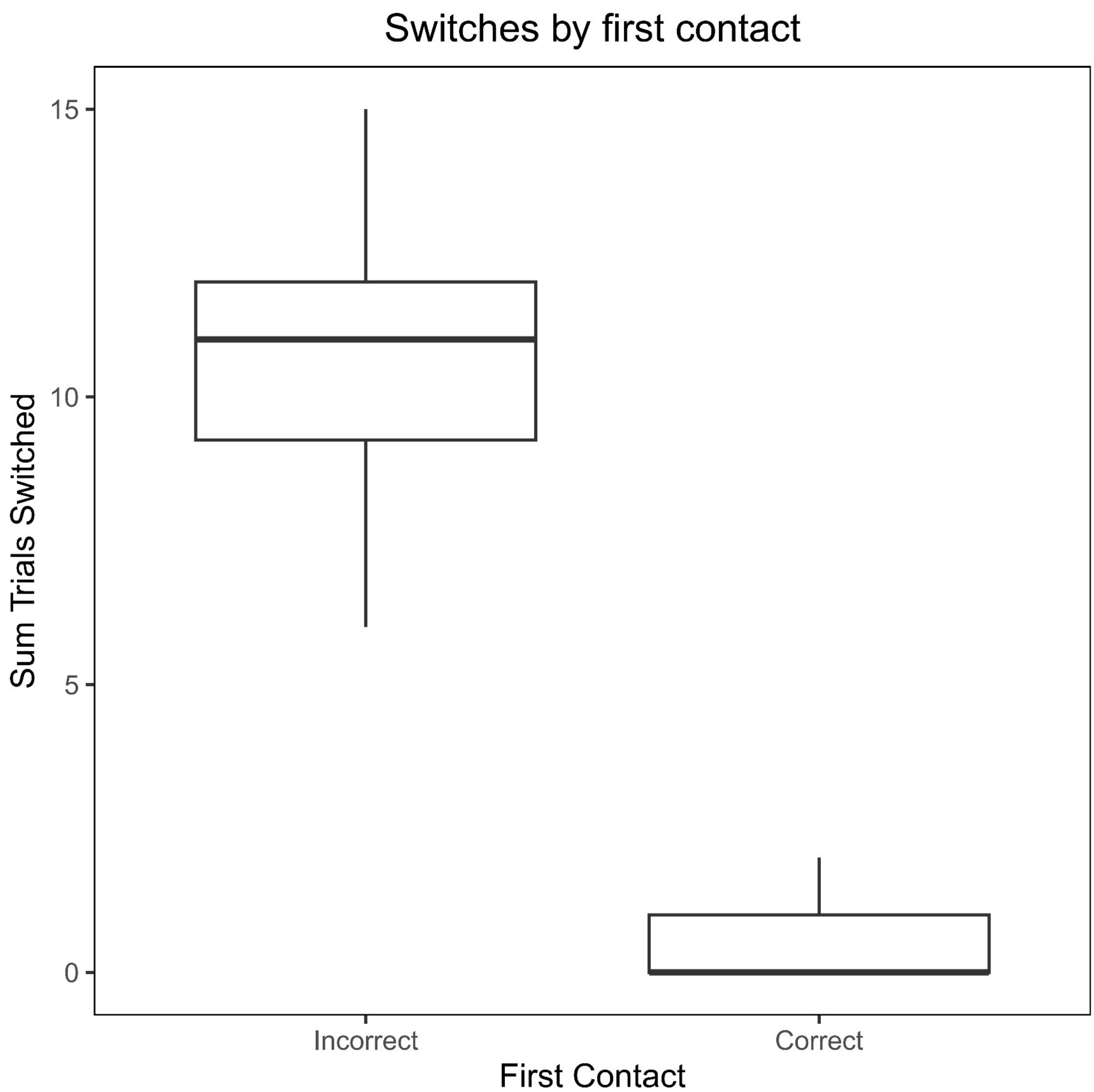


Fig. 19. Plot showing subjects' total trials in which they switched after first contacting the correct or incorrect object. Thick black line denotes median, box denotes interquartile range between 25th and 75th percentile, whiskers show range of data within IQR*1.5

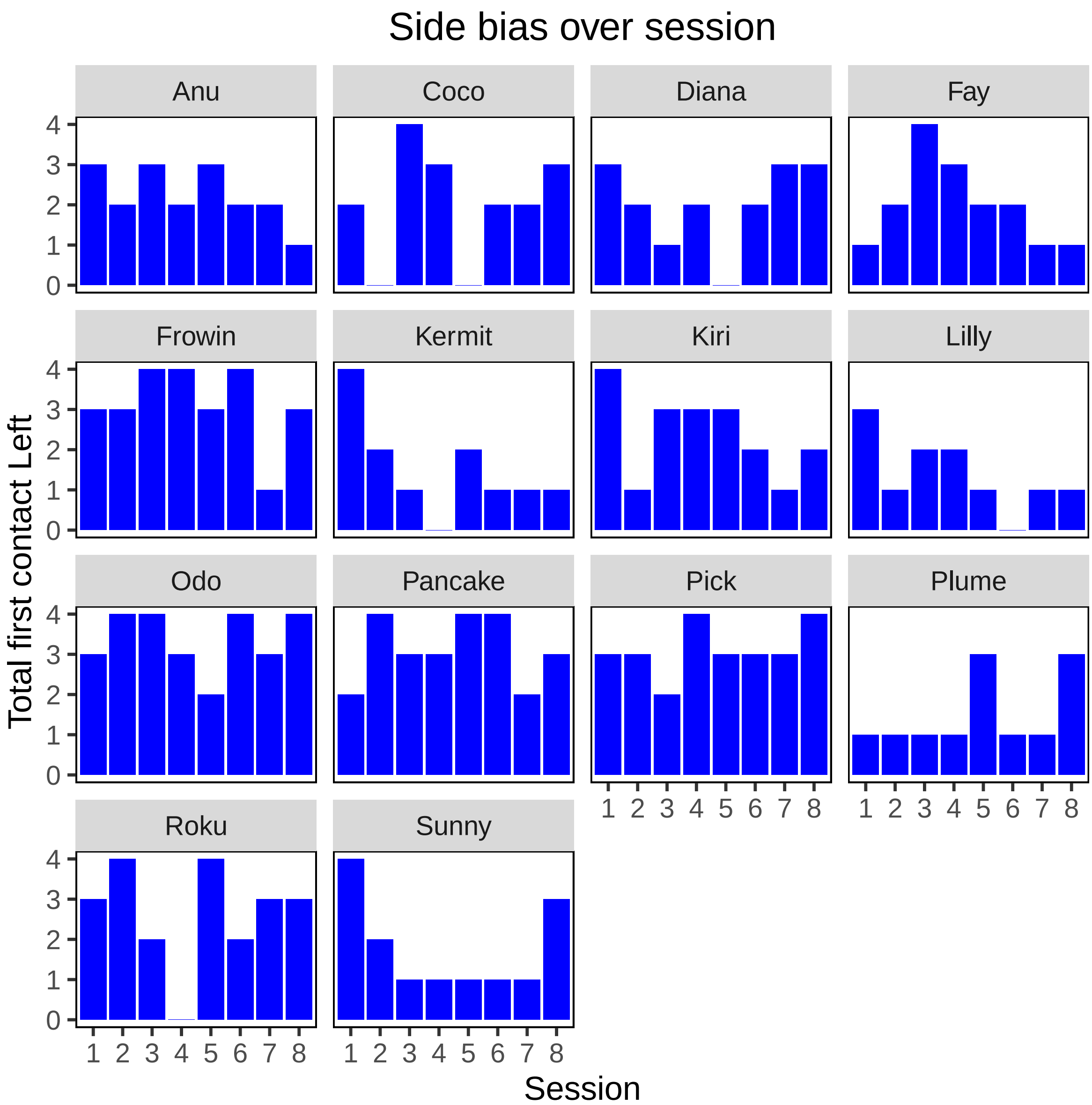


Fig. 20. Total number of trials within each session in which the subject first contacted the Left object. Lower bars (or values of zero) reflect a tendency to select the object on the right side

As we had no *a priori* reason to believe there would be sex differences in performance, and to avoid overfitting our model, we did not include sex as an effect in our analysis. Descriptively speaking, on average, males did (very) slightly better than females at both selecting the correct object and at first contacting the correct object. Males correctly chose the correct object an average of 29.1 times and correctly contacted the correct object first an average of 18.4 times. Females chose the correct object an average of 26.7 times and contacted the correct object first an average of 17.1 times. We did not have a sufficient distribution of age categories to examine its effect on the keas' success during the task.

5. Discussion

5.1 Overall Performance

In this study we attempted to answer the question, "Do kea use visual cues to infer the weight of an object prior to interacting with it?" According to the results, the initial answer to this question appears to be that they do not or, at least, that there is no evidence to suggest that they do. Over the course of the study, less than half of the kea (5 out of 14) first contacted the correct object above chance in at least one category. Only one of the kea, Frowin, contacted the correct object first above chance in more than one category (Size and Quantity). Overall, at the group level, the kea did not contact the correct object first significantly more often than chance in any category, potentially indicating that kea do not rely on visual cues such as size or material to infer the weight of an object prior to interacting with it.

While the kea did not contact the correct object first significantly above chance, they did, however, ultimately *choose* the correct object significantly above chance (Fig. 16). Additionally, the kea frequently switched sides when first contacting the incorrect object (refer to Fig. 19 in **Results**), indicating that they understood what was required of them in this task. These results further suggest that while kea in general may not be relying on visual cues such as size, volume, material, or quantity to discern the weight of an object prior to touching it, they are highly sensitive to differences in weight upon handling an object.

There are several potential explanations for why the kea may not have first contacted the correct object significantly above chance in this study. Each of the following reasons is broken down in its own section below.

5.2 Kea Do Not Rely on Visual Cues?

An initial possibility worth considering is that kea truly do not rely on visual cues to infer weight prior to physically interacting with an object. The few studies exploring whether different animal species use visual cues to determine object weight have previously met with mixed results. Avid tool users, capuchin monkeys use stone tools to break open nuts and show a preference for heavier, more effective stones [Spagnoletti et al., 2011]. A study conducted in 2009 with wild capuchin monkeys established that capuchins are able to rely on visual information to infer the weight of different stones tools [Visalberghi et al., 2009]. In the case of these wild capuchins, they relied on visual cues about the properties of stones such as size and material to identify the functional stone whenever such cues were reliable and readily available. Without such cues, the capuchins would turn to exploratory behaviors much as the kea have been shown to do in this study; tapping, moving, and/or lifting the different experimental stones in turn [Visalberghi et al., 2009]. Similarly, it has been shown that captive chimpanzees appear to rely on the visual cue of size to decide how much effort to apply in lifting an object [see Experiment 4; Povinelli, 2012, p. 79]. In contrast to the capuchins, however, these same captive chimpanzees seemed not to rely on other visual cues such as material and texture to infer the weight of the objects [see Experiment 6; Povinelli, 2012, p. 83]. One explanation for the success of the wild capuchins over the captive chimpanzees in this regard was put forward by Visalberghi and her colleagues when they wrote that,

> Wild capuchins' recognition of the affordances of objects relied on their lifelong experience with a wide variety of nuts and stones. In captive nonhuman primates, training involving a larger and more varied set of stimuli allows better generalization of abstract concepts [Visalberghi et al., 2009, p. 215].

Presumably, while the captive primates' varied experiences enhanced their ability to grasp abstract concepts, it is the repeated experience with a set of particular stimuli that gives the wild capuchins the ability to discern from sight alone whether a stone will be a heavy, effective tool or not. Indeed, one of the primary critiques of the comprehensive weight studies Povinelli and colleagues undertook is based on the fact that they used captive chimpanzees in those studies and their inability to solve the various tasks presented to them may have more to do with their atypical upbringing than with their cognitive abilities [Buckner, 2013].

Unlike kea, both chimpanzees and capuchins are habitual tool users in the wild [Spagnoletti et al., 2011] potentially making the connection between certain visual cues and the weight of an object more salient for them. As Visalberghi and colleagues wrote in their 2009 paper, the "repetitive effective use of a tool induces a plastic modification of the body representation in the brain" [p. 215]. Wild capuchin and chimpanzees routinely use heavy stones to break open nuts, potentially creating a perceptual link between the visual qualities and the heft of certain stone types and their effectiveness. While kea clearly rely on visual information (and have innovated tool-use under certain circumstances, see Goodman et al., 2018), it is possible that they do not link visual cues and weight because they do not regularly rely on or use objects as tools in a manner that would make knowing the weight of an object prior to physically interacting with it important. Instead, kea seem to rely heavily on haptic exploration and will go out of their way to engage directly with novel objects in their environment.

Known for being both intensely curious and intensely playful, kea often interact with objects simply to satiate their curiosity [Lambert et al., 2017]. As foraging generalists living in a harsh, seasonal environment with low predation, kea occupy an ecological niche that tends to produce neophilic, exploratory species [O'Hara et al., 2016]. This exploratory 'touch first, ask questions later' approach may have an adaptive role; foraging generalists such as kea living in changing environments typically have more to gain from adopting a neophilic attitude as it exposes them to a wider range of potential food sources [Greenberg & Mettke-Hofmann, 2001; Diamond & Bond, 1999]. The benefits of such an approach can also be seen in the literature; in a study on flexibility in problem-solving and tool-use in kea and New Caledonian crows using a food baited Multi-Access Box, the kea were both more individualistic in their solutions and quicker to come up with multiple solutions while solving the box than the more neophobic crows [Auersperg et al., 2011]. The researchers speculated that the balance between neophilia and neophobia in the kea and the crows, as well as their different exploratory behaviors, were likely responsible for the difference in their approaches and levels of success. As the researchers put it, "The kea showed more haptic exploration while the crows, probably due to their higher level of neophobia, seemed to explore more in a visually guided manner" [Auersperg et al., 2011, p. 5].

While kea clearly rely on visual information, the urge to haptically explore objects seems to commonly override any dependence on visual information. Whereas other, more neophobic bird species such as New Caledonian crows seem to rely more heavily on visual information [see Auersperg et al., 2011] kea seem to prefer haptic exploration above all else. This was further demonstrated in a 2011 paper that showed that while captive kea could more quickly learn a new skill when it was demonstrated to them by another kea, they soon reverted to indiscriminate haptic exploration to solve the task rather than relying on the visual demonstration provided by the other kea [Gajdon et al., 2011].

In a way, the keas' propensity to physically explore everything they come across makes the results of this current study all the more surprising as it would seem that the kea would have ample opportunity to form perceptual associations between object properties and weight. However, one of the major benefits associated with being able to infer the weight of an object from its visual properties is that it decreases risk; it is, after all, often much safer to assess a situation at a distance. For species that occupy habitats with variable or high-predation levels, it is safer to be neophobic and to rely more heavily on visual information gathered from a cautious distance [Brown et al., 2013; Greenberg et al., 2001]. In the case of the kea, a low-predation level means that there is usually minimal risk to approaching and exploring an object directly. For this and other reasons, it is possible that visual discrimination is simply not an evolutionarily or culturally favored means of information gathering for the kea.

5.3 Potential Methodological Limitations

As with any study, it may be that limitations within the study design itself influenced the success of the kea. While care was taken to eliminate any potential design flaws that could influence the keas' behavior, it is still possible that the kea were affected by a previously unnoticed issue. In this section, we will explore possible limiting factors.

Prior to testing the kea, a pilot study with one male subject, Paul, was completed. During the pilot, Paul successfully contacted the correct object first in 8 out of 10 trials and his success seemed to indicate that the methodology used was appropriate. However, there are several potential limitations to the test set-up that might have affected the ability of the kea to infer the weight of the objects based on visual cues alone.

The key parameter used to measure success during the study was whether the kea contacted their target weight object first or not [i.e. the 'correct' object]. Regardless of which object the kea contacted first, they were then able to freely choose an object to drop into the box. One possibility is that rewarding the kea for dropping the correct object into the box, regardless of which object they contacted first, disincentivized them from immediately contacting the correct object. The kea are known to rely heavily on extensive search behavior, although whether this search behavior is primarily oriented towards gaining information or as a form of play is still under debate [Diamond & Bond, 1999]. In Schloegl et al.'s 2009 study they used a foraging task (fashioned after an original task developed by Call and Carpenter

[2001]) in which both kea and ravens were simultaneously presented with both straight and bent tubes, one of which was always baited with a food reward. In the study, the ravens quickly disposed of the empty, straight tubes when the opening was facing them, relying on visual information to verify the absence of food. The kea, however, studied both ends of the empty, straight tubes in nearly a third of the trials before finally picking the tube with the food reward.

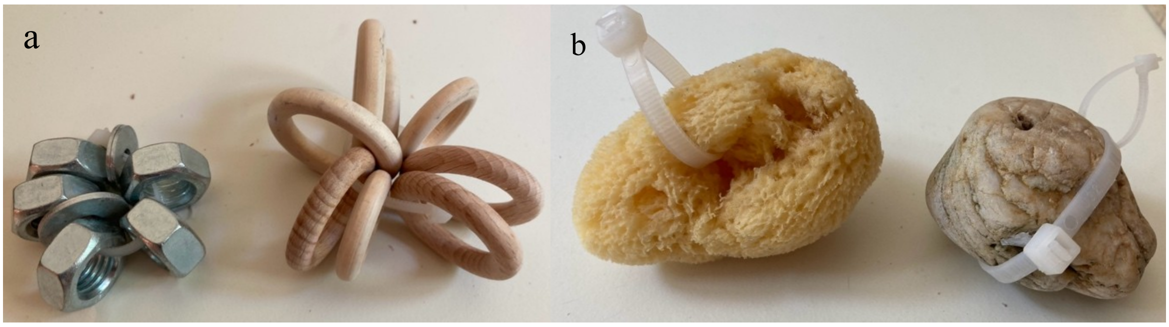
This preference for extensively exploring a novel object or situation, even if clear visual information exists to direct the kea immediately to the correct answer [as existed in Schloegl et al.'s, 2009 study], could easily influence testing results involving a choice between novel objects by compelling the kea to explore all the objects prior to completing the task. In their 2012 paper, O'Hara and colleagues made the point that, "the kea seem to frequently have other priorities than recovering the reward as fast as possible at the beginning of a new task." [O'Hara et al., 2012, p. 28] Without a penalty for initially choosing incorrectly in the current study, there was no incentive for the kea to limit their explorations to the correct object. Instead, they could opt to physically explore both of the novel objects prior to making a choice, starting with the object that appeared the most novel and/or interesting.

The potential for this to be an issue was identified at the beginning of the study and there are two primary counterarguments to it. The first is that in similar, earlier weight studies, we find that the kea do not typically waste time approaching the wrong object or tool if they already recognize which one is needed [Lambert et al., 2017]. While they are quite curious, in the observed cases (including in the pilot with Paul), it did not seem to override their desire to receive a piece of prized peanut as soon as possible. Second, an attempt was made to reduce the novelty of the objects in the current study by both choosing materials that the kea were at least passingly familiar with, and by including an "observation period" of 1 minute during testing to allow the kea to see (but not interact with) the objects. While this was done under the assumption that it would take the 'edge off' the keas' curiosity, it is possible that these measures were insufficient at eliminating object novelty and reducing the keas' urge to thoroughly, and haptically, explore both objects.

One potential way to mitigate these concerns in the future would be to test the kea using an ephemeral choice test where the kea are presented with two (or perhaps more) object choices but are only rewarded based on the object they touch first (the other untouched objects being removed as soon as the kea has first touched one). By making it costly to first contact the wrong object, this could eliminate, or at least reduce, the keas' interest in exploring multiple objects prior to making their final choice. In addition, a simple way to make contacting the wrong object less appealing would be to place the two objects much further apart so that the kea would have to exert more effort to contact an object simply for the sake of exploring it. Finally, to help reduce novelty, the objects used in the test could be placed in the keas' enclosure for an extended period of time prior to testing (perhaps a day or more). During this 'exposure time' the objects would be enclosed or cordoned off in a way that would allow the kea to see them but not physically interact with them (to avoid allowing the kea to develop potential weight associations with the objects). This prolonged visual exposure would help address the concern that the observation period used during testing was simply too short to overcome the kea's curiosity about novel objects.

5.4 Object Preferences

Another potential consideration is that the objects themselves influenced the choices of the kea. Some of the objects such as the heavy object in pair 4 from the Material category and the heavy object in pair 1 from the Material category (see Fig(s). 21a, b) performed particularly poorly with the kea. At a group level, the kea seemed to favor the light object from pair 4 (Material) in terms of which object they contacted first. The heavy object in pair 1 (Material) likewise seemed to fare poorly at a group level with the kea in both assessment categories: First Contact and Choice, with the kea showing a preference for the light object in both.



Fig(s). 21a, b. Heavy object from pair 1 (far left) and heavy object from pair 4 (far right) from the Material category with their lightweight counterparts

Other objects that performed poorly with the kea included the light objects from pairs 2 and 4 in the Quantity category (see Fig(s). 26a, b on p. 55). The kea showed a strong group level preference for the heavy objects in these two pairs in the First Contact category and the kea were also the least successful at choosing the correct object to drop into the box when choosing between the light and heavy objects in pair 2. While the upcoming section will focus on pairs 1 and 4 from the Material category, a potential explanation for the keas' performance with pairs 2 and 4 from the Quantity category will be discussed further down.

The kea were originally trained to distinguish between weights using rectangular, grey cubes (see Fig. 6 in **Methods**). The objects used in testing (as shown above in Fig(s). 21a, b, and below in Fig(s). 22 and 23) were frequently quite different looking from the objects that the kea were familiar with from training. While the kea did have a 'pre-test' session with the original grey objects (see **Methods**) prior to every testing session in hopes that it would remind the kea of what was expected of them, it is possible that the differences between the familiar objects and the testing objects occasionally threw them off. In the case of the heavy objects in pairs 1 and 4 in the Material category, the objects may just have been too unfamiliar to the kea for them to recognize what was expected of them.

Another possibility is that some of these objects performed poorly with the kea – not because they were too unfamiliar – but because they were too interesting. One of the objects (see Fig. 22), a heavy clay ball wrapped in fronds, proved to be a favorite object for the kea to unleash their destructive energy upon. When this happened, the kea would need to be ushered out of the testing compartment without a reward and the test would be reset. If the kea once again attempted to destroy the object, testing would continue using the next set of objects and the kea would receive a '0' in regard to object Choice. It is possible that having interacted with either the heavy or light object the kea then associated that weight object (even if it was the correctly weighted one) with not receiving a reward and were therefore more reluctant to choose it again the second time, biasing the results. In these cases, first contact was counted only from the first trial while a second trial was used to determine which object the kea would ultimately select. In all cases, care was taken to indicate the results were from a second trial.

Addressing both of the above mentioned issues in future testing requires a similar approach as previously mentioned: limiting object novelty. By using objects that the kea are more visually familiar with and avoiding using materials that the kea associate with food or with play (such as the object in Fig. 24 which we later discovered was similar to something the animal keepers occasionally gave the kea as enrichment), the kea would be less curious about the object and more likely to interact with it appropriately. While the enrichment association with the object in Fig. 24 (light, pair 8, Material) did not seem, in this case, to strongly bias the kea towards contacting the object first (if anything, the kea actually showed a slight preference for the heavy object in the pair), it could have very easily done so.

Fig. 22. The heavy object of pair 4 from the Volume category

Fig. 24. The light object in pair 8 in the Material category

A final explanation for the apparent object preferences exhibited during testing concerns the prior associations the kea may have with those objects. Certain objects may have been initially investigated, or even chosen by the kea, because the kea already associated them with food rewards from prior testing experiences. The captive kea at Haidlhof have been a part of ongoing cognitive tests for many years, and it is possible that some of the object types used in this experiment were used previously. For instance, prior studies investigating the keas' understanding of weight also used different colored clay objects [Lambert et al., unpublished data]. Although care was taken to ensure none of the objects used in the current study were identical to previously used clay objects, it could be that the kea have formed positive associations with certain colors or shapes over multiple rounds of testing and that certain objects were then chosen during this study due simply to their apparent similarity. This could perhaps explain why the heavy object in pair 6 from the Size category was strongly preferred by the kea at a group level in the First Contact category (Fig. 25). However, any clay object pairs used in this study were matched in terms of shape and color and varied only in terms of size and volume which should have addressed this issue to some extent.

Fig. 25. Pair 6 (Size). Heavy object is on the right

5.5 Side Preferences

Another aspect to consider in discussing the study results are side preferences. While overall the subjects did not contact one object first over the other at a rate significantly higher than chance, 4 of the 14 subjects (all males, 2 in the light group and 2 in the heavy group) did exhibit a strong Left side bias (i.e., the keas' Left side) when first contacting an object. Crucially though, even among the 4 males with a Left side bias, the kea had a high rate of switching when they contacted the incorrect object first and successfully chose the correct object to drop into the reward box at a rate significantly higher than chance. These results indicate both that the kea understood the task and that they are sensitive to differences in weight when directly handling objects. The fact that the kea chose the correct object more often than not further indicates that side biases did not ultimately influence their choice.

One possible reason the kea showed side preferences during testing is that they may not have known how to interpret the provided visual cues and so simply chose a side at random and then switched (or not) depending on the 'felt' weight of the object on that side. Since the position of the correct object was counterbalanced throughout the sessions, the kea would have been rewarded on a semi-regular basis regardless of which side they started out on (assuming, of course, that they chose the correct object to drop into the box). Given the frequency that they were rewarded on any given side and the lack of consequences if they initially chose wrong, it is possible that the kea frequently started on a particular side out of habit [Jensen & Altschul, 2015]. This habit of frequently starting on one side (even if it is not always the side with the correct object and, therefore, the *rewarded* side) could potentially be adaptive. As extractive foragers living in a changeable, alpine environment [Greer et al., 2015], the kea will often expend more energy exploring different food resources and are likely to return to food sources that do not immediately, or always, reward their efforts [Winkler & Leisler, 1999]. In this way, there could be an adaptive component to the kea returning again and again to the same side despite the fact that each side is only occasionally rewarded.

While few studies have specifically looked at the visuospatial perception of kea, another point to consider is that kea, like many bird species, may have a left-sided visuospatial bias and will naturally gravitate towards the left side without a compelling reason (such as a clear visual cue) to choose otherwise [Diekamp et al., 2005]. If the kea do not, in fact, rely on visual cues to distinguish between objects of different weights then it makes sense that they may instead choose a starting side based on natural preference and only switch sides if, after they have interacted with the object, they realize it is the wrong weight. In this case, only 4 of the 14 birds in the study showed a strong Left side bias, but this could potentially explain their actions in this study.

5.6 Target Weight

Interestingly, there seemed to be an effect of the assigned target weight on how successful each kea was at contacting the correct object first. The kea assigned the heavy target weight (i.e., 116 grams) were significantly more likely than the kea assigned a light target weight (i.e., 16 grams) to contact the correct object first, especially in the Size and Quantity categories. This effect, however, can be explained by group level preferences among the kea for contacting the heavy object first in three particular object pairs: Pair 6 in the Size category, and Pairs 2 and 4 in the Quantity category (see Fig. 18a, b, c in **Results**).

For the lightweight versions of objects 2 and 4 in the Quantity category (see Fig. 26a, b) it is possible that the kea simply did not register the objects at first. During testing, the objects were pushed slightly into the sand to hide any obvious visual imperfections, and given that objects 2 and 4 were transparent globes (containing a small lump of orange clay in the case of obj. 2 and partially filled with pebbles in the case of obj. 4), it is possible that they blended into the sand enough that the kea did not notice them until they had already approached the other object. While there is limited information available at present on kea visual acuity, two studies found that ground foraging bird species were less likely to have a high level of visual acuity [Caves et al., 2024; Dolan et al., 2010]. If this is the case for kea, it is possible it impacted their ability to use visual cues to determine the weight of the objects prior to touching them. Anecdotally, during testing, the kea would occasionally run up and get very close to an object to inspect it without touching it before either choosing it or switching to the other object. It could be that the kea simply do not have the visual acuity to perceive transparent or very small objects from a distance of roughly a meter (the distance from where the bird waited in the observation compartment to the testing site). An easy fix for this would be to use larger, opaque objects, and to place them on an even, flat surface such as a wooden board to prevent one object from being more visible than the other.

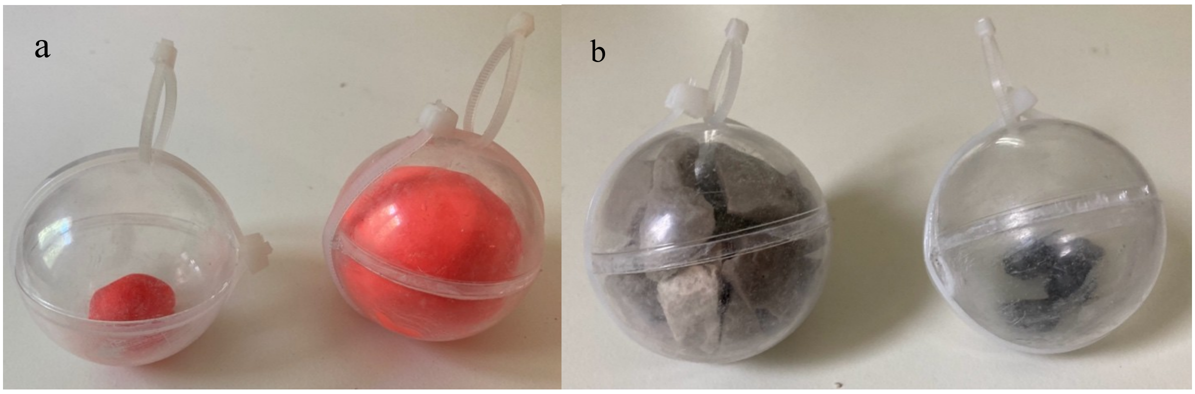


Fig. 26a, b. Two of the objects the kea struggled with: Light object from pair 2 (far left) and the light object from pair 4 (far right). Both from the Quantity category with the heavy weight counterparts

Another possible answer for the target weight preference, however, is that kea, like other bird species, may have more experience and familiarity with heavy, solid objects and therefore have an object bias towards objects that appear heavier and more solid. In the Aesop's Fable paradigm (dropping objects into water filled tubes to bring floating food closer), it was shown that certain species of corvids have a strong bias towards solid objects over hollow ones, a preference the authors of the study explained by stating that, "heavy and solid objects are likely to be more familiar than light polystyrene or hollow metal structures to many birds, given that these objects are more similar to the normal stones that birds experience outside of a test context" [Miller et al., 2016, p. 3]. In a similar vein, if the kea show a default preference for objects that are solid and heavy (or at least appear "heavy"), the kea with a heavy target weight may have been, overall, more successful at contacting the correct object first because they are at an advantage.

5.7 Age/Sex

Given the uneven distribution of age ranges it is hard to make any inference regarding the influence of age on success, however, two of the most successful kea in terms of ultimately selecting the correct object to drop into the box (Kermit and Frowin) were also two of the oldest kea in the group. Frowin was also the only kea to perform above chance at contacting his target weight object first in two categories (Size and Quantity).

In terms of the effect of sex, the males seemed to perform very slightly better than the females at both contacting the correct object first and dropping the correct object into the box. This is only a descriptive result, however, as since we had no reason to believe that sex would affect success, we did not analyze the data with this effect in mind.

While it is difficult to make any claims regarding the influence of age or sex on the success of the subjects in this study, it presents a potentially interesting avenue for future exploration.

6. Conclusion and Directions for Future Research

This study explored the ability of captive kea to infer the weights of objects based on visual cues alone. Although the results from this study provided no evidence to support the hypothesis that kea use visual cues to infer weight prior to contacting an object, the study did provide further, clear evidence that kea are highly sensitive to differences in object weight and are capable of discriminating between objects based on their felt weight. This result is broadly in line with research conducted on primates by both Daniel Povinelli, who showed that captive chimpanzees are capable of learning to sort objects based solely on weight, and with research done by Schrauf and colleagues, who revealed that wild capuchin monkeys are capable of discriminating between ineffective light stone hammers and effective heavy stone hammers using only the cue of weight [Schrauf et al., 2008]. It also reinforces more recent research establishing birds as potential contenders for further weight research by highlighting their seemingly heightened sensitivity to weight [Lambert et al., 2021]. Perhaps unsurprisingly, given the apparent convergent evolution of physical reasoning in primates and large-brained birds such as parrots, both groups have been shown capable of discriminating between objects using weight as the sole discriminating factor [Povinelli, 2012; Schrauf et al., 2008; Lambert et al., 2021].What is perhaps more surprising is that while different species of primate have proven to rely on at least certain visual object cues such as size to pre-determine weight, there is no indication, from this study at least, that kea parrots rely on visual cues at all.

However, while the results from this study provided no clear indication that kea rely on visual cues to infer object weight it is possible that this was due, in part, to the use of novel objects and a binary choice model during testing that rewarded the kea for choosing the correct object despite first contacting the incorrect (but potentially more interesting) object. To correct for this, the kea could be tested in the future using an ephemeral choice test where the untouched object is removed after the kea has made first contact and the kea are only rewarded for contacting the correct object first. Depending on the value of the reward, this could prevent the kea from first contacting an object based on curiosity alone. Other possible considerations for mitigating any testing biases caused by curiosity could include placing novel test objects within the aviary for extended periods of time to allow the kea to observe (but not interact with) them and using materials (where possible) that more closely mimic the training materials initially used with the kea.

Future planned studies in this area include exploring how sensitive the kea are to changes in weight and establishing the average minimum gradation of weight that the kea can detect. Over the course of this study the kea have proved to be highly attuned to object weight despite not seeming to infer weight from visual cues. Testing their sensitivity to changes in object weight would provide a deeper insight into how the kea view and interact with the physical world and the properties that govern it. It would also help establish a framework for future studies exploring weight understanding in kea by ensuring that the objects used are appropriately weighted for the species. Another area of interest for future study includes looking at whether the kea rely on absolute or relative rules using the concept of weight. Such a study would grant us a greater understanding of how the kea make choices and solve physical reasoning problems in captivity and in the wild. An ever-present question in animal cognition seems to be “how are they like us?”, and lately, and arguably more importantly, “how are they different?” Studying what aspects of physical reasoning we share (or may not share) with kea and other species ultimately helps us answer those questions.

References

Adelson, E. H. (2001). *On seeing stuff: The perception of materials by humans and machines* (B. E. Rogowitz & T. N. Pappas, Eds.; pp. 1–12). https://doi.org/10.1117/12.429489

Antinucci, F., & Visalberghi, E. (1986). Tool use in Cebus apella: A case study. *International Journal of Primatology*, *7*, 351-363.

Apps, J. N. (2008). Abstract thinking. In S. J. Loue & M. Sajatovic (Eds.), *Encyclopedia of Aging and Public Health* (pp. 67–69). Springer US. https://doi.org/10.1007/978-0-387-33754-8_2

Auersperg, A. M., Von Bayern, A. M., Gajdon, G. K., Huber, L., & Kacelnik, A. (2011). Flexibility in problem solving and tool use of kea and New Caledonian crows in a multi access box paradigm. *PLOS One*, *6*(6), e20231.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Baciadonna, L., Cornero, F. M., Emery, N. J., & Clayton, N. S. (2021). Convergent evolution of complex cognition: Insights from the field of avian cognition into the study of self-awareness. *Learning & Behavior*, *49*(1), 9–22. https://doi.org/10.3758/s13420-020-00434-5

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

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

Bayern, A. V., Danel, S., Auersperg, A. M. I., Mioduszewska, B., & Kacelnik, A. (2018). Compound tool construction by New Caledonian crows. *Scientific Reports*, *8*(1), 15676.

Bird, C. D., & Emery, N. J. (2010). Rooks perceive support relations similar to six-month-old babies. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1678), 147–151. https://doi.org/10.1098/rspb.2009.1456

Biro, D., Haslam, M., & Rutz, C. (2013). Tool use as adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1630), 20120408. https://doi.org/10.1098/rstb.2012.0408

Block, N. (2010). Attention and Mental Paint. *Philosophical Issues*, *20*(1), 23–63. https://doi.org/10.1111/j.1533-6077.2010.00177.x

Boesch, C., & Boesch, H. (1984). Mental map in wild chimpanzees: An analysis of hammer transports for nut cracking. *Primates*, *25*(2), 160–170. https://doi.org/10.1007/BF02382388

Brosnan, S.F. (2009). Animal behavior: The right tool for the job. *Current Biology,* 19(3): R124-125. doi: 10.1016/j.cub.2008.12.001

Brown, G. E., Ferrari, M. C. O., Elvidge, C. K., Ramnarine, I., & Chivers, D. P. (2013). Phenotypically plastic neophobia: A response to variable predation risk. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1756), 20122712. https://doi.org/10.1098/rspb.2012.2712

Buckingham, G. (2014). Getting a grip on heaviness perception: A review of weight illusions and their probable causes. *Experimental Brain Research*, *232*(6), 1623–1629. https://doi.org/10.1007/s00221-014-3926-9

Buckingham, G., Cant, J. S., & Goodale, M. A. (2009). Living in A Material World: How Visual Cues to Material Properties Affect the Way That We Lift Objects and Perceive Their Weight. *Journal of Neurophysiology*, *102*(6), 3111–3118. https://doi.org/10.1152/jn.00515.2009

Buckingham, G., & Donald, H. (2019). Move on up: Fingertip forces and felt heaviness are modulated by the goal of the lift. *Attention, Perception, & Psychophysics*, *81*(7), 2384–2390. https://doi.org/10.3758/s13414-019-01703-w

Buckner, C. (2013). In search of balance: A review of Povinelli's world without weight: Daniel J. Povinelli: World without weight: perspectives on an alien mind. Oxford University Press, 2012, 353 pp. *Biology & Philosophy*, *28*(1), 145–152. https://doi.org/10.1007/s10539-012-9352-0

Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen?. *Animal Cognition*, *3*, 207-220.

Cantwell, A., Buckholtz, J. W., Atencia, R., & Rosati, A. G. (2022). The origins of cognitive flexibility in chimpanzees. *Developmental Science*, *25*(5), e13266.

Caves, E. M., Fernández-Juricic, E., & Kelley, L. A. (2024). Ecological and morphological correlates of visual acuity in birds. *Journal of Experimental Biology*, *227*(2), jeb246063.

Clark, A. (2023). *The experience machine: How our minds predict and shape reality*. Pantheon Books.

Danel, S., Rebout, N., von Bayern, A. M., & Osiurak, F. (2021). Complex nests but no use of tools: An investigation of problem solving in weaverbirds (Ploceidae). *Behavioural Processes*, *192*, 104493.

Dautenhahn, K. (1996). *Embodied Cognition in Animals and Artifacts*. AAAI Fall Symposium. https://cdn.aaai.org/Symposia/Fall/1996/FS-96-02/FS96-02-006.pdf

Diekamp, B., Regolin, L., Güntürkün, O., & Vallortigara, G. (2005). A left-sided visuospatial bias in birds. *Current Biology*, *15*(10), R372-R373.

Dijker, A. J. (2008). Why Barbie feels heavier than Ken: The influence of size-based expectancies and social cues on the illusory perception of weight. *Cognition*, *106*(3), 1109-1125.

Diamond, J., & Bond, A. B. (1999). *Kea, bird of paradox: the evolution and behavior of a New Zealand parrot*. Univ of California Press, pp. 40, 57, 122, 148, 149

Dobson, A. J. (2002). *An introduction to generalized linear models*. Chapman & Hall/CRC

Dolan, T., & Fernández-Juricic, E. (2010). Retinal ganglion cell topography of five species of ground-foraging birds. *Brain Behavior and Evolution*, *75*(2), 111-121.

Dos Reis, M., Thawornwattana, Y., Angelis, K., Telford, M. J., Donoghue, P. C., & Yang, Z. (2015). Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. *Current Biology*, *25*(22), 2939-2950.

Ellis, R. R., & Lederman, S. J. (1998). The Golf-Ball Illusion: Evidence for Top-down Processing in Weight Perception. *Perception*, *27*(2), 193–201. https://doi.org/10.1068/p270193

Emery, Nathan. (2004). Are corvids 'feathered apes'? Cognitive evolution in crows, jays, rooks, and jackdaws. *Comparative Analysis of Minds,* 181-213.

Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science*, *306*(5703), 1903-1907.

Fallon, F. T., Ryan, T. J., Krakauer, J. W., & The Rppf Group. (2023). What are we talking about? Clarifying the fuzzy concept of representation in neuroscience and beyond. *The Transmitter*. https://doi.org/10.53053/PFBH5329

Flanagan, J. R., & Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size–weight illusion. *Nature Neuroscience*, *3*(7), 737-741.

Flanagan, J. R., Bittner, J. P., & Johansson, R. S. (2008). Experience Can Change Distinct Size-Weight Priors Engaged in Lifting Objects and Judging their Weights. *Current Biology*, *18*(22), 1742–1747. https://doi.org/10.1016/j.cub.2008.09.042

Gajdon, G. K., Amann, L., & Huber, L. (2011). Keas rely on social information in a tool use task but abandon it in favour of overt exploration. *Interaction Studies*, *12*(2), 304-323.

Goodman, M., Hayward, T., & Hunt, G. R. (2018). Habitual tool use innovated by free-living New Zealand kea. *Scientific Reports*, *8*(1), 13935. https://doi.org/10.1038/s41598-018-32363-9

Gordon, A. M., Forssberg, H., Johansson, R. S., Eliasson, A. C., & Westling, G. (1992). Development of human precision grip: III. Integration of visual size cues during the programming of isometric forces. *Experimental Brain Research*, *90*, 399-403.

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.

Greenberg, R., & Mettke-hofmann, C. (2001). Ecological aspects of neophobia and neophilia in birds. In V. Nolan & C. F. Thompson (Eds.), *Current Ornithology, Volume 16* (pp. 119–178). Springer US. https://doi.org/10.1007/978-1-4615-1211-0_3

Greer, A. L., Gajdon, G. K., & Nelson, X. J. (2015). Intraspecific variation in the foraging ecology of kea, the world's only mountain-and rainforest-dwelling parrot. *New Zealand Journal of Ecology*, *39*(2), 254-261.

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.

Harshfield, S. P., & DeHardt, D. C. (1970). Weight judgment as a function of apparent density of objects. *Psychonomic Science*, *20*(6), 365–366. https://doi.org/10.3758/BF03335692

Hayashi, M. (2015). Perspectives on object manipulation and action grammar for percussive actions in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1682), 20140350. https://doi.org/10.1098/rstb.2014.0350

Henley, T. (2020). Do New Caledonian Crows Represent <Weight>? An Analysis of Jelbert et al. (2019). *Animal Behavior and Cognition*, *7*(4), 661–664. https://doi.org/10.26451/abc.07.04.12.2020

Hirsiger, S., Pickett, K., & Konczak, J. (2012). The integration of size and weight cues for perception and action: Evidence for a weight–size illusion. *Experimental Brain Research*, *223*(1), 137–147. https://doi.org/10.1007/s00221-012-3247-9

Hoffmann, M., & Pfeifer, R. (2011). *The implications of embodiment for behavior and cognition: Animal and robotic case studies*. https://doi.org/10.5167/UZH-72303

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

Jelbert, S. A., Taylor, A. H., Cheke, L. G., Clayton, N. S., & Gray, R. D. (2014). Using the Aesop's fable paradigm to investigate causal understanding of water displacement by New Caledonian crows. *PLOS One*, *9*(3), e92895.

Jensen, G., & Altschul, D. (2015). Two perils of binary categorization: why the study of concepts can't afford true/false testing. *Frontiers in Psychology*, *6*, 168.

Johansson, R. S., & Westling, G. (1988). Programmed and triggered actions to rapid load changes during precision grip. *Experimental Brain Research*, *71*(1). https://doi.org/10.1007/BF00247523

Köck, C. (2014). *Size Discrimination in Kea (Nestor notabilis)* [Master's Thesis, Unpublished]. University of Vienna.

Krakauer, J. W., & Shadmehr, R. (2006). Consolidation of motor memory. *Trends in Neurosciences*, *29*(1), 58–64. https://doi.org/10.1016/j.tins.2005.10.003

Lambert, M., Farrar, B., Garcia-Pelegrin, E., Reber, S., & Miller, R. (2022). ManyBirds: A multi-site collaborative Open Science approach to avian cognition and behavior research. *Animal Behavior and Cognition*, *9*(1), 133–152. https://doi.org/10.26451/abc.09.01.11.2022

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. https://doi.org/10.1098/rsos.170652

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

Lambert, M. L., Jacobs, I., Osvath, M., & von Bayern, A. M. (2019). Birds of a feather? Parrot and corvid cognition compared. *Behaviour*, *156*(5-8), 505-594.

Lycan, W. (n.d.). Representational Theories of Consciousness. In *The Stanford Encyclopedia of Philosophy* (Winter 2023). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2023/entries/consciousness-representational/>

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.

Miller, R., Jelbert, S. A., Taylor, A. H., Cheke, L. G., Gray, R. D., Loissel, E., & Clayton, N. S. (2016). Performance in object-choice Aesop's fable tasks are influenced by object biases in New Caledonian crows but not in human children. *PLOS One*, *11*(12), e0168056. https://doi.org/10.1371/journal.pone.0168056

Montemayor, C., & Haladjian, H. H. (2017). Perception and cognition are largely independent, but still affect each other in systematic ways: Arguments from evolution and the consciousness-attention dissociation. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00040

Murray, D. J., Ellis, R. R., Bandomir, C. A., & Ross, H. E. (1999). Charpentier (1891) on the size—weight illusion. *Perception & Psychophysics*, *61*(8), 1681-1685.

Nieuwenhuis, R., te Grotenhuis, H. F., & Pelzer, B. J. (2012). Influence. ME: Tools for detecting influential data in mixed effects models. *The R Journal*, 4(2):38-47.

Nyssen, R., & Bourdon, J. (1955). Contribution to the study of the size-weight illusion by the method of P. Koseleff. *Acta Psychologica*, *11*, 467–474.

O'Hara, M., Schwing, R., Federspiel, I., Gajdon, G. K., & Huber, L. (2016). Reasoning by exclusion in the kea (Nestor notabilis). *Animal Cognition*, *19*(5), 965–975. https://doi.org/10.1007/s10071-016-0998-x

O'Hara, M., Mioduszewska, B., Mundry, R., Haryoko, T., Rachmatika, R., Prawiradilaga, D. M., ... & Auersperg, A. M. (2021). Wild Goffin's cockatoos flexibly manufacture and use tool sets. *Current Biology*, *31*(20), 4512-4520.

O'Hara, M., Gajdon, G. K., & Huber, L. (2012). Kea logics: how these birds solve difficult problems and outsmart researchers. *Logic and Sensibility*, *5*, 23-38.

Paulus, M., & Hauf, P. (2011). Infants' use of material properties to guide their actions with differently weighted objects. *Infant and Child Development*, *20*(4), 423-436.

Pautz, A. (2010). Why explain visual experience in terms of content? In B. Nanay (Ed.), *Perceiving the World* (pp. 254–309). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195386196.003.0010

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

Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists.* Cambridge University Press.

R Core Team. (2019). *R: a language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria.

Saccone, E. J., Landry, O., & Chouinard, P. A. (2019). A meta-analysis of the size-weight and material-weight illusions. *Psychonomic Bulletin & Review*, *26*(4), 1195–1212. https://doi.org/10.3758/s13423-019-01604-x

Schloegl, C., Dierks, A., Gajdon, G. K., Huber, L., Kotrschal, K., & Bugnyar, T. (2009). What you see is what you get? Exclusion performances in ravens and keas. *PLOS One*, *4*(8), e6368.

Schloegl, C., Schmidt, J., Boeckle, M., Weiß, B. M., & Kotrschal, K. (2012). Grey parrots use inferential reasoning based on acoustic cues alone. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1745), 4135–4142. https://doi.org/10.1098/rspb.2012.1292

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., & Call, J. (2011). Great apes use weight as a cue to find hidden food. *American Journal of Primatology*, *73*(4), 323–334. https://doi.org/10.1002/ajp.20899

Schrauf, C., Call, J., Fuwa, K., & Hirata, S. (2012). Do Chimpanzees Use Weight to Select Hammer Tools? *PLOS One*, *7*(7), e41044. https://doi.org/10.1371/journal.pone.0041044

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

Seed, A. M., & Call, J. (2010). Problem-solving in tool-using and non-tool-using animals. *Encyclopedia of Animal Behavior*, *2*, 778-785.

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

Spagnoletti, N., Visalberghi, E., Ottoni, E., Izar, P., & Fragaszy, D. (2011). Stone tool use by adult wild bearded capuchin monkeys (Cebus libidinosus). Frequency, efficiency and tool selectivity. *Journal of Human Evolution*, *61*(1), 97–107. https://doi.org/10.1016/j.jhevol.2011.02.010

Steiner, P. (2014). Enacting anti-representationalism. The scope and the limits of enactive critiques of representationalism. *AVANT. The Journal of the Philosophical-Interdisciplinary Vanguard*, *V*(2), 43–86. https://doi.org/10.26913/50202014.0109.0003

Tebbich, S., Seed, A. M., Emery, N. J., & Clayton, N. S. (2007). Non-tool-using rooks, Corvus frugilegus, solve the trap-tube problem. *Animal Cognition*, *10*, 225-231.

van Polanen, V., & Davare, M. (2015). Sensorimotor memory biases weight perception during object lifting. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00700

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. https://doi.org/10.1016/jv.cub.2008.11.064

Watzl, S. (2019). Can representationism explain how attention affects appearances? In A. Pautz & D. Stoljar (Eds.), *Blockheads!* (pp. 581–608). The MIT Press. https://doi.org/10.7551/mitpress/9196.003.0036

Wimpenny, J. H., Weir, A. A., Clayton, L., Rutz, C., & Kacelnik, A. (2009). Cognitive processes associated with sequential tool use in New Caledonian crows. *PLOS One*, *4*(8), e6471.

Winkler, H. & Leisler, B. (1999). Exploration and curiosity in birds: Functions and mechanisms. In: Adams, N.J. & Slotow, R.H. (eds) Proc. 22 Int. Ornithol. Congr., Durban: 915-932. Johannesburg: BirdLife South Africa. https://www.internationalornithology.org/PROCEEDINGS_Durban/Symposium/S16/S16.5.htm

Wolfe, H. K. (1898). Some effects of size on judgments of weight. *Psychological Review*, *5*(1), 25–54. https://doi.org/10.1037/h0073342