universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Compound tool use in the Goffin's Cockatoo (Cacatua goffiniana): Exploring the capacities and limitations of associative tool use in a non-primate species

verfasst von | submitted by

Anna Geisler BSc

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

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 878
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Verhaltens-, Neuro- und Kognitionsbiologie
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Thomas Bugnyar

Table of Contents

Abstract

Complex tool use involves forms of associative tool use – the manipulation of more than one tool – such as tool set use, sequential tool use, composite and compound tool use. The ability to execute associative tool use is not an exclusively human trait, as was initially assumed. In the thesis, I investigate compound tool use, which involves combining separate objects into a single functional tool, in the Goffin's cockatoo (*Cacatua goffiniana*). This foraging generalist is native to the Indonesian Tanimbar Islands. Goffin cockatoos have already demonstrated different forms of associative tool use and therefore represent an excellent non-primate model for studying compound tool innovations. During a problem-solving task involving a Plexiglas apparatus with a food reward only accessible with a sufficiently long tool, they were tested for their ability to innovate a compound tool made of two cone-like elements. Thereby, the birds were given the choice between cone-like and boat-like objects. They did not succeed in creating a compound tool of either object type. However, some individuals consistently touched two separate tools simultaneously or attempted object-object combinations during the test condition requiring more than one tool. Cone-to-cone combination attempts, representing the correct tool configuration, were the most frequent attempts. In a subsequent test the elements were pre-assembled by the experimenter. During this test, the cockatoos chose the functional tool made of two cones significantly more often and succeeded in the vast majority of trials. The results of this experiment reveal a strong preference for the functional tool. This could suggest that Goffin cockatoos respond to a task they cannot solve with a single tool by engaging in object-object combination attempts. However, limitations related to the shape of the tools may have contributed to their difficulty in successfully creating a compound tool. Given the complexity of compound tool use, the birds may face both cognitive and anatomical challenges and limitations in the execution of the sequential steps required to solve this task. Nevertheless, I also discuss alternative explanations for the observed performances.

Zusammenfassung

Komplexer Werkzeuggebrauch umfasst Formen der assoziativen Werkzeugnutzung –die Verwendung von mehr als einem Werkzeug – beispielsweise Werkzeug-Sets, sequenzieller, zusammengesetzer Werkzeuggebrauch und kombinierte Werkzeuge. Diese Art an Werkzeuggebrauch ist kein ausschließlich menschliches Merkmal, wie lange angenommen wurde. Ich untersuche die Kombination einzelner Objekte zu einem funktionalen Werkzeug ("Compound Tool-use") beim Goffin-Kakadu (*Cacatua goffiniana*) – einem Nahrungsgeneralisten, der auf den indonesischen Tanimbar-Inseln heimisch ist. Goffin-Kakadus haben bereits verschiedene Formen assoziativer Werkzeugnuzung demonstriert und eignen sich somit als hervorragendes Modellsystem für die Untersuchung von Compound Tool-use. Das vorliegende Experiment inkludiert eine Plexiglasapparatur mit einer Cashew als Belohnung, die nur mit einem ausreichend langen Werkzeug erreichbar ist. Untersucht wurde, ob die Kakadus ein aus zwei kegelförmigen Objekten zusammengesetztes Werkzeug herstellen können. Sie hatten die Wahl zwischen 3D-gedruckten kegelförmigen und bootsförmigen Elementen. Es gelang ihnen nicht, ein funktionales, kombiniertes Werkzeug aus einem der beiden Objekttypen herzustellen. Einige Individuen berührten jedoch wiederholt zwei separate Werkzeuge gleichzeitig oder versuchten, Objekte miteinander zu berühren, wenn mehr als ein Werkzeug erforderlich war. Kombinationsversuche zwischen zwei Kegel Elementen, die eine korrekte Werkzeugkonfiguration darstellen, traten dabei am häufigsten auf. In einem nachfolgenden Test, in dem die Objekte vorab zusammengesetzt wurden, wählten die Kakadus das funktionale Werkzeug aus zwei Kegeln signifikant häufiger und waren im Großteil aller Durchgänge erfolgreich.

Die Ergebnisse dieses Experiments zeigen eine deutliche Präferenz für das funktionale Werkzeug und deuten darauf hin, dass Goffin-Kakadus auf Aufgaben, die sie mit einem einzelnen Werkzeug nicht lösen können, mit Objekt Kombinationsversuchen reagieren. Einschränkungen hinsichtlich der Form der Werkzeuge können potentiell zu der Hürde beigetragen haben, ein funktionales Werkzeug erfolgreich zu kombinieren. Angesichts der Komplexität dieser Form von Werkzeuggebrauch und der Durchführung der notwendigen Einzelschritte, die zum Erfolg führen, sind die Vögel möglicher Weise sowohl mit kognitiven als auch anatomischen Herausforderungen und Einschränkungen konfrontiert. Folglich berücksichtige ich jedoch auch alternative Erklärungen für die beobachteten Ergebnisse.

Introduction

Historical background

Tool use was long thought to be an exclusively human trait, that helped to define our species. Oldowan tools are the first known simple tools made by early humans during the Early Stone Age. Around the Middle Paleolithic Period the complexity of tools increased to stone spears, a functional tool made of composite parts (Smithsonian National Museum of Natural History, 2024, "Stone tools"). The diversity of tools increased further during the Later Stone Age, reflecting the broader range of human cumulative cultures (Baune, 2009; Smithsonian National Museum of Natural History, 2024, "Stone tools").

Over the past half-century the belief of tool use being an exclusively human trait has changed drastically. One of the first and most well-known cases of tool use has been documented in captive chimpanzees (*Pan troglodytes*) by Köhler (Köhler, 2025).

Apart from the Woodpecker Finch, *Cactospiza pallida* (Tebbich et al., 2002), the New Caledonian Crow (*Corvus moneduloides*) was one of the first bird species to be observed using tools and showing innovative behavior when crafting new tools, including hooks and tapered tools made of natural materials (Hunt, 1996). These habitual tool users show the capacity to spontaneously solve novel problems and infer upon physical relationship between objects (Chappell and Kacelnik, 2002; Kenward et al., 2005; Kenward et al., 2006; Taylor et al., 2007; Bayern, 2018). Other than specialized tool use, there are also non-specialized tool using bird species, like the Northern Blue Jay (*Cyanocitta cristata*) (Jones and Kamil, 1973), the kea (Auersperg et al., 2011; Auersperg, Huber and Gajdon, 2011; Goodman, Hayward and Hunt, 2018), and the Goffin's Cockatoo (*Cacatua goffiniana*) (Auersperg et al., 2012; Auersperg et al., 2016, Auersperg et al., 2017).

Compound tool use was extensively studied in primates (Köhler, 1925; Lethmate, 1982; Bania et al., 2009, Price et al., 2009; Shumaker, Walkup & Beck, 2024), However, within the clade Aves compound tool use is primarily documented in New Caledonian crows. This highlights the importance of expanding comparative research.
The divergence of mammals and birds happened approximately 319 million years ago (timetree.org, 2025), suggesting that the capacity to innovate tool use evolved

independently, due to the phylogenetic distance between crows and apes (Rutz & Clair, 2012). A similar conclusion can be drawn for New Caledonian crows and Goffin cockatoos, which diverged around 66 million years ago (timetree.org, 2025).

Tooling and different forms of tool use

There has been a high number of definitions of tool use with different criteria for its definition. Since then, stricter definitions limit tool use to a narrower range of species, which mainly includes primates and birds (Colbourne, 2021). A fresh approach was suggested by Fragaszy and Mangalam (2018). It refers to the behavior as "tooling" by "transforming the body into the body-plus-object system" and also taking into account the complexity underlying the behavioral sequence (Fragaszy & Mangalam, 2018). The new approach does not include some behaviors, that were previously referred to as tool use and as a result, a considerably smaller number of species exhibited "tooling" when taking into account the revised definition (Colbourne et al., 2021). It is important to mention, that the cognitive complexity of a species or individual is not dependent on whether a behavior qualifies as "tooling". Within tooling the difficulty can vary to quite some extent, for example: a) A higher number of spatial relations is more difficult than a lower number, b) Simultanous relations are more difficult than sequential, c) Dynamic relations are more difficult than static, d) object or surface directed tooling is more difficult than self-directed tooling (Fragaszy & Mangalam, 2018; Colbourne, 2021).

The most complex types of tooling include associative tool use, which is referred to as the use of "two or more tools in combination" (Reindl et al., 2022). It is so far documented in macaques (Hihara et al., 2002), capuchin monkeys (Mannu & Ottoni, 2009), apes (Köhler, 1925; Lethmate, 1982; Bania et al., 2009, Price et al., 2009), cockatoos (O'Hara et al., 2021; Osuna-Mascaro et al., 2022; Osuna-Mascaro et al., 2023) and corvids (Wimpenny et al., 2009; Bayern et al., 2018). Shumaker et al. (2011) distinguishes between tool composite use, tool set use, sequential tool use and secondary tool use.

Tool set use is defined as the use of multiple tools, each serving a different function to achieve a single goal (Osuna-Mascaro et al., 2023). Previously observed use of tool sets also represents a highly complex form of tool use in an allocentric, dynamic, specific and sequential manner (Brewer & McGrew; O'Hara et al., 2021; Mioduszewska et al., 2021; O'Hara et al., 2023). When tools are used in a specific sequence with intermediate steps—

such as retrieving a second tool by using the first one to reach a goal—this is referred to as sequential tool use, which has previously been observed in Great Apes (Martin-Ordas, Schumacher and Call, 2012), New Caledonian Crows and Goffin Cockatoos (Hihara et al., 2002; Wimpenny et al., 2009; Bugnyar, 2019; Gruber et al., 2019; Rössler et al., in prep.). Concurrent tool use, considered the most complex form of tooling, has so far only been documented in humans (Colbourne, Auersperg & Beck, 2024) Additionally, secondary tool use—the use of one tool to create another—has only been observed in the bonobo Kanzi (Schick, Toth & Garufi, 1998).

Compound tool use, a subset of composite tool use, refers to "combining complementary different parts into firmly connected units" (Bayern et al., 2018). In composite tool use, the different parts are not connected, while compound tools are connected to another. Combinatorial actions during tooling impose higher cognitive demands due to the necessity of an intermediate step before achieving the intended goal (Price et al., 2009).

In animals, compound tool use is one of the rarest forms of tool use. There are very few documented examples in chimpanzees, one of the most well-known being Köhler's famous chimpanzee, Sultan. After multiple unsuccessful attempts to reach a banana by jumping, Sultan stacked boxes on top of each other to achieve his goal. In a similar experiment, after several failed attempts, Sultan constructed a compound tool by inserting a smaller tube into a larger one, which he then used to retrieve an out-of-reach banana (Köhler, 1925). Other chimpanzees were also able to combine tools spontaneously or after the observation of a con-specific, in order to make a longer tool (Price et al., 2009). Even in the wild, there is an anecdotal example of an orangutan manufacturing what could be interpreted as a compound tool by bundling *Erythrina* sp. leaves together to form a type of footpad as a protection against the thorny plant (Fox & bin'Muhammad, 2002). Studies of great apes also later inspired similar experiments with other species.

Compound tool use in New Caledonian crows

A study of compound tool use in adult New Caledonian crows was conducted at the Avian Cognition Research Station of the University of Oxford, UK (Bayern et al., 2018).
While various factors influence the complex acquisition of tool use in wild populations, studying these birds in captivity provides a controlled environment that allows for targeted manipulations of these factors. Since New Caledonian are known to be habitual

tool users and flexibly construct and manipulate tools (Hunt, 1996; Taylor et al., 2007; Taylor et al., 2010; Auersperg et al., 2011; Bluff et al., 2010; Hund & Gray, 2004; Rutz & Clair, 2012; Clair & Rutz, 2013; Jacobs, Bayern & Osvath, 2016) they were well-suited subjects for investigating compound tool use in birds, inspired by Köhler's experiments. Unlike Köhler's chimpanzees, however, the crows were naïve to experimental tasks and reinforcing feedback.

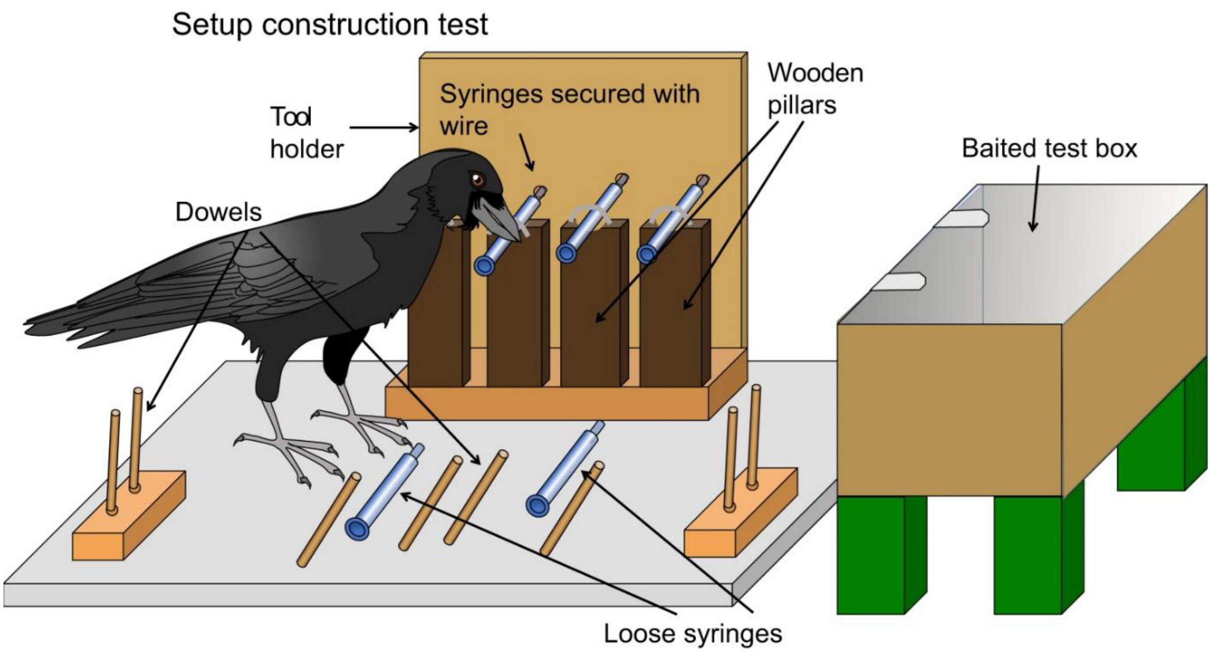

Figure 1: Setup from the construction test conducted on New Caledonian crows. Adapted from "Compound tool construction by New Caledonian crows" by Bayern, A. M. P. von, Danel, S., Auersperg, A. M. I., Mioduszewska, B., & Kacelnik, A., 2018, Scientific Reports, 8(1), p. 2 (https://doi.org/10.1038/s41598-018-33458-z). © The Authors 2018, licensed under CC BY 4.0.

New Caledonian Crows spontaneously combined objects to create functional compound tools (Bayern et al., 2018). Not only did they construct compound tools, but the construction was also dependent on whether it was necessary for the task.

In the first phase the birds were presented with a transparent box containing a treat. Additionally, wooden dowels, long enough to reach the treat, were provided to the birds. By using these dowels to move the treat along a track inside the box, the crows could push the food through an opening to retrieve it.

In addition to the baited test apparatus, the first construction test introduced a tool holder with drilled holes facing the crows and four wooden pillars below. The treat was placed at an increased distance, and only short, cylindrical elements were provided. These could potentially be combined into a longer tool by inserting a syringe plunger or wooden dowel into a syringe barrel. The objects were presented both on the wooden pillars and scattered on the ground.

In the third phase pipe cleaners, wooden dowels and drinking straws were placed in different positions to test whether the shape and configuration influenced compound tool construction. The fourth phase tested whether the construction of a compound tool was goal-directed testing whether the birds would construct tools only when necessary. A "close track" was introduced, where the treat could be reached with a single uncombined tool. In the final phase, the successful birds were provided with even shorter elements, requiring the use of more than two components to reach the food.

During the first phase, all of the birds used to long dowels to reach the treat inside the apparatus. Half of the crows combined two elements during the first construction test and one individual combined up to four during the second construction test. Furthermore, they recognized the need for a compound tool. After initially attempting to retrieve the food with an uncombined tool, they proceeded to construct a longer one when necessary. This suggests that their compound tool use was goal-directed. The birds were more likely to use uncombined elements when sufficient and combined them only when the distant treat required a longer tool.

Nevertheless, the story of compound tool use in New Caledonian crows suggests that tool construction may have initially occurred unintentionally: The syringe barrel was already set up in a horizontal, elevated position on the tool holder, with the insertion hole being at a convenient insertion height. As they poked at the opening with a second element, it naturally became lodged, resulting in the two objects being firmly connected. Later, after initially inserting elements into pre-positioned syringe barrels, the crows successfully combined tools scattered on the ground as well.

Since New Caledonian crows have a predisposition to insert sticks into openings (Auersperg et al., 2014; Kenward et al., 2006), their initial tool combinations may have occurred by chance rather than as a result of problem representation. Various species of the *Psittaciformes* and *Corvids* engage in self-rewarding object play, which is not necessarily goal-directed. While New Caledonian crows primarily exhibit this behavior at a young age, Goffin's cockatoos do not show significant age-related differences in their combinatory actions (Auersperg et al., 2014). Thus, the innovative strength required to reach a solution may not be comparable to the chimpanzees. Nevertheless, combining different objects still implies innovative behavior and the ability to plan.

Tool use in Goffin cockatoos

Goffin cockatoos have shown multiple examples of innovative and associative tool use (Auersperg et al., 2012; Auersperg et al., 2016; Auersperg et al., 2017; Laumer et al., 2017; O'Hara et al., 2021; Osuna-Mascaro et al., 2022; Osuna-Mascaro et al., 2023) as well as social transmission of tool use (Auersperg et al., 2014). Their ecological role as opportunistic foragers also suggests a capacity for innovating extreme strategies to access hard-to-reach food (Mioduszewska, Auersperg and O'Hara, 2022).

The Goffin's cockatoo is a foraging generalist, native to the Indonesian Tanimbar Islands, with a diverse diet, including embedded foods that are difficult for other species to access. Unlike corvids, Goffin's cockatoos do not build nests and therefore do not appear to inherit tool-use potential from predispositions for specific object-related actions (Auersperg et al., 2016; Mioduszewska, Auersperg, & O'Hara, 2022).

Despite not having the same predispositions as New Caledonian crows, such as food caching (Kenward et al., 2011; Jacobs et al., 2014) and nest building (Hansell and Ruxton, 2008), which foster a natural inclination to insert sticks into holes (Auersperg et al., 2014; Kenward et al., 2006), Goffin cockatoos have succeeded in tasks involving highly complex forms of tool use and show innovative tendencies. Hansell and Ruxton (2008) argue that tool use is more likely in birds that build nests, as nest building may require higher levels of cognition, similar to tool use. Food cashing has similarities to tool use in New Caledonian crows and can offer an ultimate explanation in achieving proficiency, for instance, in tool use (Jacobs et al., 2014). But Goffin cockatoos are cavity nesting birds, that do not require weaving sticks together in a complex manner to build a nest, nor are they known to cache food or objects.

They can spontaneously innovate various types of tool use, including stick tools for poking and raking, hook tools for hooking and stone tools for dropping (Auersperg et al., 2012; Laumer et al., 2017; Osuna-Mascaro et al., 2022). Different materials, such as cardboard, wood, wire and twigs have been used to innovate suitable tools (Auersperg et al., 2016). They are even capable of highly complex forms of tool use, which include composite tools and tool sets, both in captivity and in the wild (O'Hara et al., 2021; Osuna-Mascaro et al., 2022; Osuna-Mascaro et al., 2023).

A new and insightful example of tool set use is provided by a previous study in 2022 (O'Hara & Mioduszewska et al., 2022), observing wild Goffin cockatoos that were temporarily housed in captivity. The birds successfully used a set of three different tools to access seeds inside the hard endocarp of the *Cerbera manghas* fruit, also known as "Wawai" (Tanimbarese name). This was observed in 2/15 individuals, which extracted and modified the tools from a branch. The types of tools could be distinguished by their shape and sturdiness, and they corresponded with different motions for inserting, wedging and extracting (Mioduszewska et al., 2021). Goffin's cockatoos can also transport a set of two different tools when faced with a task requiring more than one tool (Osuna-Mascaro et al., 2023). One apparatus used in this test featured a membrane behind the front opening that had to be torn through with a short, rigid stick. Subsequently, a longer, more flexible tool was required to reach the food. Most Goffins successfully innovated the necessary tool set. The sequence of two alternating apparatuses was then randomized to observe whether the tools are used in a goal-oriented manner, and later required movement to reach the apparatus. Not only did they use the tools goal-directedly, but three individuals transported both tools recurrently. One individual even inserted the short tool into the longer one, temporarily creating a compound tool to facilitate transport (Osuna-Mascaro et al., 2023).

It is suggested, that "their foraging ecology could strongly influence the evolutionary emergence of tool use" (Mioduszweska, Auersperg & O'Hara, 2022). This is an example of a very complex type of "tooling", since it is allocentric (directed at an object), dynamic and specific (Mioduszewska et al., 2021). The tendency of manipulating objects in a playful manner might still play a role in the development of tool use behavior. The playful manipulation of objects could also result in combining two or more objects (Auersperg et al., 2014).

Goffin cockatoos previously solved a Golf Course Task, demonstrating composite tool use. This task required the combination of a ball, pushed by a stick-like tool, onto a collapsible platform containing a treat. Both of the objects had to be used in a specific sequence and moved dynamically, controlling the position and momentum (Osuna-Mascaro et al., 2022). This is considered a highly complex form of tool use.

It is challenging to assess the role innovation plays during combinatorial actions in the New Caledonian crows due to their inherent predisposition for tool use.
Therefore, Goffin cockatoos serve as excellent models for comparing cognitive processes with primates, as the same counterargument against innovation does not apply to them.
Inspired by previous examples of compound tool use, this experiment focuses on the cognitive ability to perform object combinations to form a compound tool. In order to do so, two or three identical, 3D-printed cone elements can be combined by inserting the pointed end of one cone into the opening of another. However, there is a deliberate deviation from the design used with the New Caledonian crows, as no element is preemptively set up in the apparatus. To innovate this compound tool, the Goffin cockatoos are expected to insert one of the cone elements into another identical cone, thereby creating a longer tool to reach a treat placed on a platform within the apparatus, without receiving any hints. As a control I also provided a 3D-printed boat-like structure, which can be combined by inserting the "hull" of the boat object into the hollow "deck" of another boat object, resulting in an element with enlarged diameter instead of an elongated shape.

My research question focuses on the Goffin cockatoos' ability to create a compound tool by combining separate elements into one functioning tool. Based on the findings of the study on New Caledonian crows (Bayern et al., 2018), my main hypothesis (H1) is, that Goffin cockatoos will independently innovate a compound tool in a problem-solving task.

My predictions are as follows: 1) Some cockatoos will spontaneously innovate the combination of the provided objects; and 2) they will do so productively, being more likely to combine two cone-elements instead of two boat-elements, resulting in a longer tool, enabling them to reach the food inside an apparatus.

The second hypothesis (H2) is, that social transmission can enhance the ability of otherwise unsuccessful subjects to create a compound tool. This will be done by including demonstration sessions for one bird if all subjects fail the original task and at least one bird has previously attempted to use more than one of the provided elements simultaneously. If that bird is successful, it will be used as a demonstrator.
It was shown, that other social animals, like chimpanzees were able to combine tools significantly better once they observed an unfamiliar chimpanzee performing the task (Price et al., 2009). While primates are the most studied animals in regards to social transmission of tool use and tool acquisition, it was also observed that social transmission can enhance the ability of use tools in New Caledonian crows and Goffin cockatoos (Auersperg et al., 2014; Kenward et al., 2005). Though, in New Caledonian crows certain inherited predispositions seems to play a crucial role as well (Kenward et al., 2005). In a previous study on New Caledonian Crows teaching sessions conducted by a human experimenter were included in the procedure. In the presence of an observing crow, the experimenter would pick up a twig and insert it into a block hole to retrieve food (Kenward et al., 2006).
The prediction of my second hypothesis is as follows: After observing the correct object combination demonstrated by a cockatoo that previously attempted to use more than one tool simultaneously, the observer bird will successfully combine two cone elements into a longer, functioning tool.

During a previous investigation of Laumer, Bugnyar & Auersperg (2016) on the Goffin cockatoos tool choice relative to the quality of the food reward in two different apparatuses, the Goffins chose the functional tool type significantly above chance level. When faced with either one of two different apparatuses, requiring either a ball or a stick tool to retrieve a food reward during the tool selection test, the birds adjusted their tool choice flexibly depending on the apparatus used during each trial (Laumer, Bugnyar & Auersperg, 2016). The findings of this study suggest that Goffin cockatoos can adjust their

tool choice according to the task at hand, leading to my third hypothesis (H3). H3 suggests that the cockatoos are able to recognize functional aspects of the combined tools. Should the attempt of social transmission be unsuccessful too, the task of combining the objects will no longer be required. The elements, consisting of either two cones (CC) or two boats (BB), will be pre-combined by the experimenter ahead of a trial and out of sight of the birds. Since only the cone-to-cone combinations will result in a functioning tool of sufficient length to reach the food inside of an apparatus, the birds should only use the pre-combined cone-tool instead of the two combined boat-elements to solve the task.
Here, my prediction is: The cockatoos will have a preference of handling and using the CC tool to push the food off the platform inside the apparatus.

In case there is no successful construction of a compound tool, the fourth hypothesis (H4) states, that there are attempts of tool combinations. The frequency and duration of the birds touching two tools at the same time as well as attempted tool combinations will be recorded, to assess, whether the birds attempt the correct (cone-to-cone) combination.
During object play, Goffin cockatoos, subdivided into subadults and adults, engaged in free object-to-object and object-to-substrate combinations and combined at least two elements and up to three elements in some instances, a habit they shared especially with young New Caledonian crows (Auersperg et al., 2014). Goffin cockatoos also tend to combine objects when faced with a problem they cannot solve, such as combining plant materials with a food source (Osuna-Mascaro & Auersperg, 2018; O'Hara, unpublished observation). For the first time in the wild up until 2018, a wild individual on Tanimbar was observed combining stiff stalks and several flimsy leaves from the same plant with a coconut, possibly indicating an attempt to open the coconut. (Osuna-Mascaro & Auersperg, 2018). Due to these previous attempts of object-to-object and object-to-substrate combinations, my predictions for the fourth hypothesis include: 1) The cockatoos will touch two tools simultaneously more often and for longer durations during the test condition requiring more than one tool; 2) They will attempt to combine two tools with each other during tool-to-tool contact behavior, regardless of their success; 3) they are more likely to attempt tool combinations that include the cone-element, rather than the boat-element, and; 4) they will attempt C-to-C combinations most frequently and for longer durations compared to other element-to-element attempts.

As a result of recognizing the functional aspects of the different elements – as indicated by a previous study on the Goffins flexibility in tool choice (Laumer, Bugnyar & Auersperg, 2016) – the preference for the correct tool type (cone) is expected to be more pronounced during the test condition requiring more than one tool, as only cone-to-cone combinations can result in a successful trial. This forms the content of my fifth hypothesis (H5), with the following predictions: 1) The preference for either tool type will be approximately equal during both the initial preference test and the test condition requiring only one tool, and; 2) during the test condition requiring more than one tool, the preference will shift to the cone-element.

Methods

Subjects and Housing

In the beginning of the experiment thirteen captive Goffin cockatoos (see Appendix for the list of participants) were included in the experiment. Two of them remained very nervous around the setup and one individual was absent for a prolonged time, therefore, only ten individuals participated eventually, six males and four females. All of the participating animals had been part of previous cognitive problem-solving tasks.

The Goldegg Goffin Lab consists of an outdoor aviary (185m2, 3-5m high), as well as indoor aviary (45m2, 3-6m high), that are connected through a door. Both parts include enrichment, such as toys, bath tubs and toys. Enrichment is restored, suitable food and water is supplied and the aviary is being kept clean on a daily basis during maintenance. As long as temperatures didn't sink below -10°C, the birds were allowed to spend time in the outside aviary every day. An additional testing room is connected to the inside aviary, which is typically only accessed by individual birds during test sessions and habituation. For testing special treats – like cashews and pecan nuts – are provided as a reward.

All participating cockatoos have previous experience with stick tools (Auersperg et al., 2012; Auersperg et al. 2014; Auersperg et al., 2016; Laumer, Bugnyar & Auersperg, 2016; Auersperg et al., 2017; O'Hara et al., 2021; Mioduszewska, Auersperg & O'Hara, 2022; Osuna-Mascaro et al., 2022; Osuna-Mascaro et al., 2023) and basic tool use tasks. Eleven of the thirteen birds – Figaro, Fini, Pipin, Konrad, Dolittle, Heidi, Muppet, Kiwi and Zozo – participated in a composite tool use tasks in the subsequent golf club experiment (Osuna-Mascaro et al., 2022). Ten of the participants have also participated in a task requiring tool set us (Osuna-Mascaro et al., 2023). However, all experiments involving other forms of associative tool use were conducted after collecting data on compound tool use for the experiment at hand. Therefore, the experience gained by the birds from studies on composite tool use and tool set use could not influence this study on compound tool use.

Ethical Statement

All the Goffins from the Goldegg Goffin Lab have CITES certificates and were hand-raised. Additionally, they are registered according Austrian Animal Protection Act (§ 25 - TschG. BGBl. I Nr. 118/2004 Art. 2. 118) at the district's administrative animal welfare bureau (Bezirkshauptmannschaft St. Pölten Schmiedgasse 4-6, 3100; St. Pölten, Austria) and live in housing agreeing with the Austrian Animal Protection Act (§ 24 Abs. 1 Z 1 and 2; § 25 Abs. 3). The experiment was considered a non-animal experiment under the Austrian Animal Experiments Act (§ 2. Federal Law Gazette No. 501/1989), because of its non-invasive procedure. All the animals participated voluntarily in the experiment by being called in individually or by being perched voluntarily on the experimenter.

Materials

Apparatus:

The basic apparatus (Figure 2Figure 2), used for all conditions, consisted of a rectangular plexiglass structure, supported by wooden poles at the corners. The front side of the plexiglass featured a hole for inserting the tools and a food exit at the bottom of the plexiglass wall. Inside the apparatus was a wooden slide leading to the food exit. Glued to the top of the slide was a wooden platform for the food reward, which was a small piece of a cashew nut. By pushing the treat off the platform it would slide towards to food exit. The apparatus could be adapted for three different conditions: the short-distance setup, the medium-distance setup and the long-distance setup. The key difference between the conditions was the distance from the tool opening to the food platform. However, due to the kind of progress made, the long-distance setup became redundant, and only short- and medium-distance setups were included.

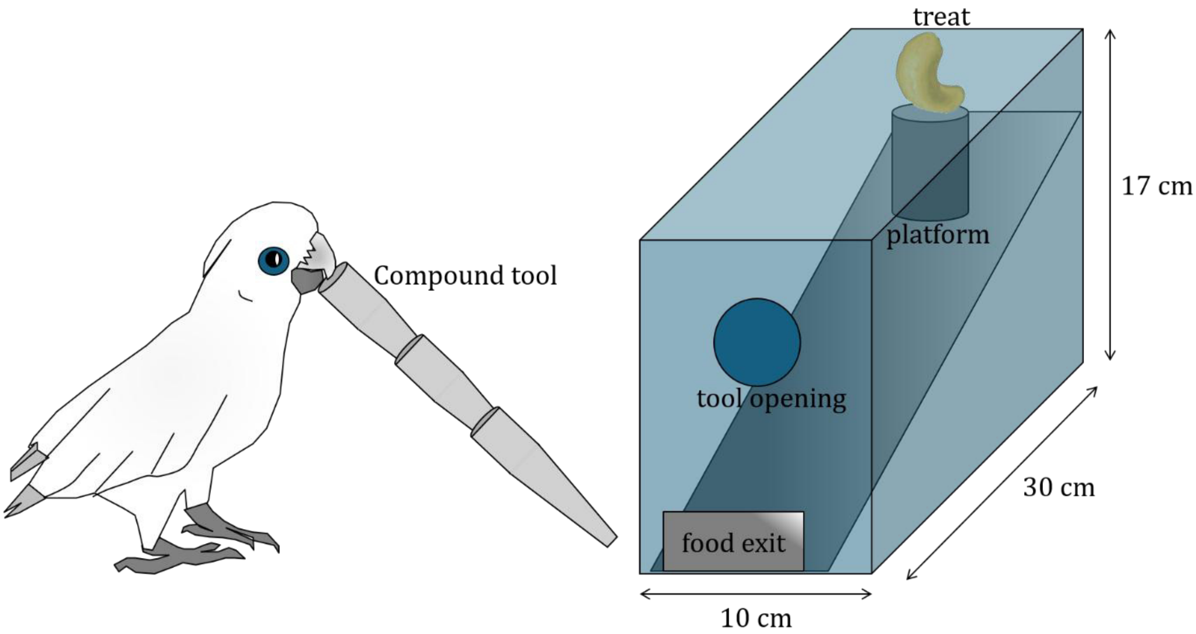

Figure 2: Basic apparatus with 17 cm height, 30 cm depth and 10 cm length (adapted from figure provided by Assoc. Prof. Priv.-Doz. Dr. rer. nat. Alice Auersperg). The cashew nut as a food reward is provided on top of the platform within the apparatus and can be reached with an elongated compound tool. The reward can be pushed off the platform with the tool and slides down towards the food exit.

Tools:

Two distinct elements of different shapes were used as tools in the experiment (Figure 3). The 3D-printed tools were designed by Peter Füreder at the Veterinary University of Vienna. For each tool type a set of three was printed, resulting in six tools overall. The two tool types consisted of a cone-shaped element and a boat-like element, which will be referred to as “cone” and “boat” from here on. The cone measured 90mm in length, was hollow on the inside, and had a diameter of 17mm. The boat was also 90mm long and 17mm wide. The cones were designed to fit into each other by inserting the pointed end of one cone into the open end of another, resulting in an overlap of 40 mm (Figure 3). Similarly, the boats could be combined by inserting the “hull” of one boat into the hollow “deck” of another (Figure 3).

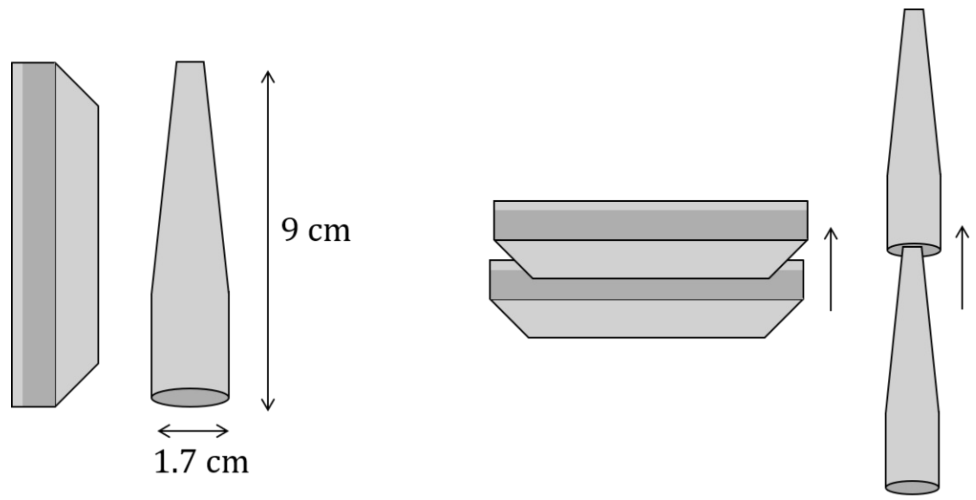

Figure 3: 3D-printed boat and cone tools. The height of both tools is 9 cm with a diameter of 1.7 cm (author's own illustration). The boats can be combined by inserting the "hull" into the "deck" of a second boat. The cones are combined by inserting the pointed end into the opening of the second cone.

Testing procedure

The experiment followed a sequence of steps: habituation to the tools, a preference test, a short-distance setup, a medium-distance setup, and potentially a long-distance setup. Each step included a similar procedure: Three boats and the three cones were placed on the table of the test room in randomized order, determined by using random.org.

Similar to the study on compound tool use in New Caledonian crows (Bayern et al., 2018), the Goffin cockatoos were first presented with a task in which a single tool was sufficient to reach the treat. Once they completed this, the setup was adjusted so that the combination of two cones became necessary for a successful trial. If the birds completed this condition successfully, the experiment would progress to the long-distance setup. If the birds would not be able to complete the medium-distance setup, the experiment would continue with pre-combining the tools for the medium-distance setup.

Each single bird was brought or called into the test room and told to wait on the chair in front of the test table with the apparatus, before each trial began. Tool placement started ahead of each trial. Pushing the food reward off the platform inside the apparatus in the short-distance, medium-distance, long-distance and pre-combined setup would be considered a successful trial, which in turn ended the trial. If the birds were not able to push the food reward off the platform, the trial would continue for the set time and end in an unsuccessful trial.

Habituation:

Before testing, twelve birds were habituated to the tools. Object neophobia was previously investigated in Goffin cockatoos by observing the latency to approach different sizes, shapes, colours and reflectiveness of novel objects. Especially when confronted with elongated objects, the Goffins showed bigger hesitation in approaching, often accompanied by behavioral signs of fear (Cespedes-Gonzalez et al., 2024). These results might influence the success of the experiment.

Ahead of each session a small piece of a cashew nut was put into each tool in front of the bird. During this process the bird was seated on a chairback in front of the table. If an individual showed signs of fear, a second, more confident bird was introduced to the same room. It was necessary to individually respond to each bird's comfort level because they expressed varying intensities of fear. Some were encouraged to go near the tools, if they would not go near any of the tools. In these cases, the experimenter presented both tool types including the treat on her hands to the more hesitant birds.

Preference test:

The goal of the preference test was to identify any individual preferences for interacting with a specific tool type. This was done by measuring the number of times and the duration (s) the birds handled each of the tools. In order to do so three cones and three boats were placed on the table in randomized order. Each bird was individually introduced to the test room. Three sessions of each five minutes were conducted per bird. No additional treat was provided during the preference test.

Short-distance test:

Starting with the short distance setup, the apparatus was placed on the table in the test room. The platform for a small piece of cashew nut as a reward was at a 9 cm distance from the tool opening (Figure 4.a). Only one tool – either a boat or a cone – was sufficient to reach the food reward inside of the apparatus. Each of the two sessions per bird contained ten trials. For the birds that did not participate by touching any of the objects or try solving the apparatus the number of sessions was increased to nine. Ahead of each trial, three boats and three cones were again placed on the table in randomized order in front of the apparatus. To prevent the birds from simply choosing the last tool placed, the experimenter hovered her hands over all six tools after placement. Each trial began with the vocal cue “los geht’s” (“let’s go”) and lasted five minutes each, ending with an alarm set by the experimenter.

Medium-distance test:

In the subsequent step, the distance to the treat inside the apparatus was increased to 14 cm from the tool opening (Figure 4.b). In order to reach it, the cockatoos would need to combine two cones into an elongated compound tool. This required the handling and coordination of two tools simultaneously, inserting one cone into another in a way they stay firmly connected, inserting the resulting compound tool into the apparatus and push the food reward from the platform in order to be successful.

Long-distance test:

The final condition in the long-distance setup required the combination of three cone elements to reach the food reward, now at a distance of 20 cm from the tool opening (Figure 4.c).

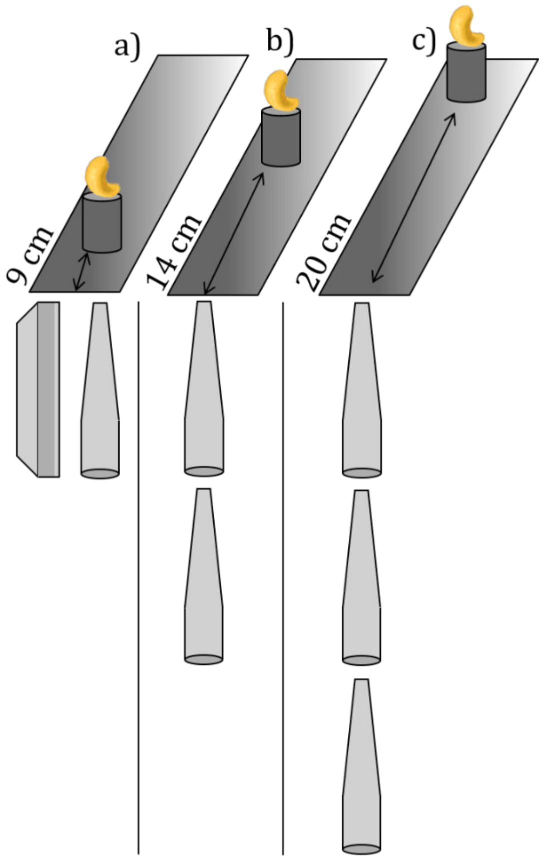

Figure 4: Three different conditions are distinguished by the distance of the food platform to the tool opening (adapted from figure provided by Assoc. Prof. Priv.-Doz. Dr. rer. nat. Alice Auersperg): a) for the short-distance setup the distance to the opening was 9 cm. Both tool types could potentially be utilized to retrieve to reward; b) increasing the distance to 14 cm, it became necessary to combine two cone elements into a longer tool to retrieve to food; and c) a final increase to 20 cm would require three cone elements to be combined into the longest compound tool.

Demonstration:

In case the cockatoos were not able to combine the cone tools by themselves, without receiving any hints, five sessions with a human demonstrator were potentially included as an additional step. Since past experiments have shown, that there can be social transmission of tool use in Goffin cockatoos (Auersperg et al., 2014), it was attempted to demonstrate the necessary task of combining the cone elements into an elongated compound tool to one bird, that had at least attempted to use more than one element simultaneously.

In the beginning of each session, one medium-distance trial was conducted in the usual manner. This was followed by six demonstration trials. In each of the trials the experimenter combined either two cones or two boats in front of the bird and continued to attempt inserting the respective compound tool into the apparatus. This was only possible when with a cone-to-cone compound tool. Subsequently, the treat was also consumed by the experimenter in front of the bird. If two boats were combined, the experimenter would try to insert them and fail, after which the trial was stopped.

Pre-combined tools:

For the alternative step of the pre-combined tool test, the medium-distance setup (Figure 4.b) was used. During the pre-combined test setup the birds were presented with two pre-assembled boats (Figure 5.a) and two pre-assembled cones (Figure 5.b), prepared out of view. Each was placed in front of the apparatus, while the birds waited on the chair in front of the table. To retrieve the cashew nut, the birds had to choose the compound tool with the correct shape and length for the apparatus. Only the combined cones were able to fit the tool opening and reach the cashew nut on the platform.

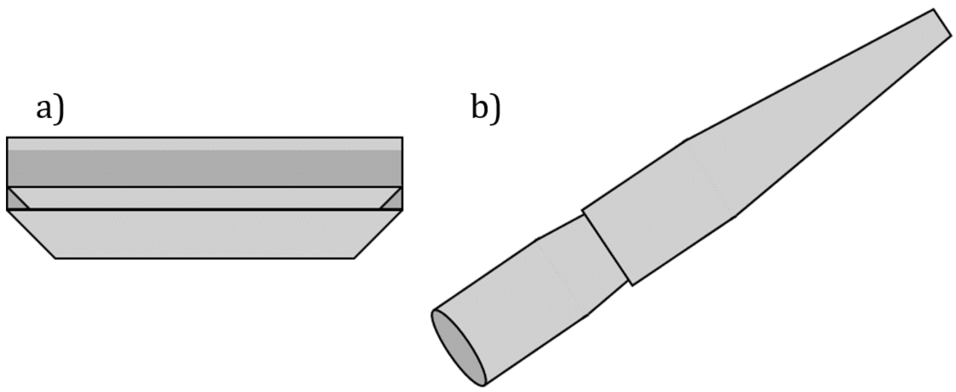

Figure 5: The two compound tools were pre-assembled ahead of each trial (author's own illustration): a) shows boat-to-boat; and b) shows cone-to-cone

In case the pre-combined tools would get disassembled – by accident or by the birds separating them – the experimenter instructed the bird to return to the chair and reassembled them into a compound tool out-of-view of the bird before continuing.

Behavioural Coding

The coding software BORIS (v. 8.27.1) was used to record point and state events.

Touching cone or boat elements as well as touching the pre-combined tools with their claws or beak are coded as state events, as well as touching the apparatus with either tool. Touching two tools simultaneously – usually one tool was held in the beak, while another was held with their claws – was recorded as a separate state event. On occasion, birds would hold two tools with their beak. Each tool-to-tool contact, including cone-to-cone, cone-to-boat, boat-to-boat and boat-to-cone was recorded separately as state events.

Due to several instances of the pre-combined tool falling apart when the birds handled it in a rough manner or separated them intentionally, they were recorded as a point events.

A time-out to rearrange the tools on the table marks the moment the birds had either thrown at least two tools of one type off the table or the tools getting stuck in the apparatus. During this time-out, the behaviors were not recorded and the session timer was not stopped.

Each trial was timed from the moment a vocal cue was given by the experimenter and ended either with the success or the time-out of the trial. Whether a trial was successful was coded with either yes (y) or no (n) as point events, as soon as the bird pushed the cashew nut off the platform with a tool.

Statistics

For the analysis I worked with R version 4.4.2 (2024-10-31 ucrt). To read and structure the data, I used the packages "readxl" (Wickham & Bryan, 2023), "dplyr" (Wickham et al., 2023) and "tidyr" (Wickham, Vaughan & Girlich, 2024). The statistical analysis and models were carried out with the packages "lme4" (Bates et al., 2015), "glmmTMB" (McGillycuddy et al., 2025), "MASS" (Veneables & Ripley, 2002), "Matrix" (Bates, Maechler & Jagan, 2025) "AER" (Kleiber & Zeileis, 2008), "car" (Fox & Weisberg, 2019), "emmeans" (Lenth, 2025) and "DHARMa" (Hartig, 2024). Finally, for the visualisations I used "ggplot2" (Wickham, 2016).

For each test in the experimental sequence both the duration and frequency of tool touches was averaged and the standard errors (SE) were calculated. Furthermore, it was determined which tool was touched first during each session. The number of successful trials was summed up for the short-distance test, the medium-distance test, as well as the pre-combined test. For tests, which resulted in any successful trials (short-distance and pre-combined), the time to success was calculated by determining the difference between trial success and trial start for each individual, and a mean was measured across trials.

The statistical analysis was conducted using Generalized Linear Mixed Models (GLMMs) and one Linear Mixed Model (LMM), including subject as random effect for each model. I fitted GLMMs with a Tweedie distribution to analyze the duration of tool handling as response variable. For the preference test, zero-inflation was modelled with a constant intercept. Tool type (cone and boat), session and sex were included as fixed effects in all full models. For the frequency of tool touches negative-binomial GLMMs were applied, due to overdispersion observed in model comparisons with Poisson GLMMs. Investigating the frequency of tool choice during the pre-combined test I additionally specified a random intercept and a random slope for session by subject using „||“, assuming no correlation between random intercepts and slopes. This was done to increase model fit. Model selection was based on Akaike Information Criterion (AIC), along with an inspection of residuals and an evaluation for potential overdispersion.

Analyzing the durations of the combination attempts and the duration of touching two tools simultaneously during the medium-distance task, I used zero-inflated models. For the frequency I used negative-binomial models. Four potential tool-to-tool combination attempts, as well as session and sex were included as fixed effects in the full models. To model the number of successful and unsuccessful trials during the pre-combined task I used binomial GLMMs and added session and sex as predictors. Furthermore, the time until success during the pre-combined test was investigated using an LMM with session and sex as fixed effects. Finally, assessing the cone-preference across different tests I used Beta GLMMs for ratios for both frequency and duration and included test and sex as predictors in the full model.

All full models were compared to null models, excluding all predictors and checked for overdispersion by calculating the overdispersion ratio, defined as the ratio of the sum-of-squared Pearson residuals to the residual degrees of freedom. A ratio substantially greater than 1 was considered indicative of overdispersion. Model diagnostics were assessed using the „DHARMa“ package (Hartig, 2024), including residual simulation plots and zero-inflation tests to evaluate potential deviations from distributional assumptions. Finally, additionl checks for collinearity were conducted using vif() function from the „car“

package (Fox & Weisberg, 2019). Variance inflation factors (VIF) greater than 5 were indicative of potential collinearity issues.

Tool-to-tool contact:

For the medium-distance test, the duration and frequency of each possible tool-to-tool contact was calculated, which include cone-to-cone (C-to-C), cone-to-boat (C-to-B), boat-to-cone (B-to-C) and boat-to-boat (B-to-B). It is necessary to note, that it often is unclear in the observations, whether the attempts of tool-to-tool contact were intentional. Although no successful construction of a compound tool occurred, some attempts were made to handle two tools simultaneously and touch one tool with another. An individual touching one tool with another was coded in a consistent manner across all trials and regardless of intention, and will hereafter be referred to "combination attempt".

For the analysis of the duration and frequency of combination attempts in the medium-distance test, I utilized the glmer() function from the "lme4" package (Bates et al., 2015), as well as the glmmTMB() function from the "glmmTMB" package (Brooks et al., 2017). I assigned combination attempts, sex and session as fixed effects and, since there are repeated measures of each subject, included subject as a random effect. After model comparison and model selection based on AIC and likelihood ratio tests, I used Tukey's estimated marginal means for pairwise comparisons, using the "emmeans" package (Lenth, 2025). Due to high zero-inflation and overdispersion for continuous duration data of combination attempts, I applied a zero-inflated GLMM for the tool-to-tool duration as a response variable. For investigating frequencies of combination attempts I applied a GLMM with a poisson-distribution and a log-link function.

Touching two tools simultaneously:

The frequency and duration of two tools simultaneously was calculated for the medium-distance test, as well as the short-distance test. This was done for the full data sets of each test as well as separately for each individual. To analyze the duration of touching two tools simultaneously in the medium-distance test, I fitted a zero-inflated GLMM using the "glmmTMB" package (Brooks et al., 2017) including session and sex as fixed effects. A Tweedie distribution with a log link was specified to deal with the right-skewed distribution of the duration data. For the frequency data I applied a negative-binomial GLMM with a log link from the same package.

Success during pre-combined test:

For the pre-combined test, I analysed predictors for both successful and unsuccessful trials using binomial GLMMs. First, I defined the number of unsuccessful trials as a response variable and as fixed variables I assigned a) the binomial variable of BB touched first during a trial, b) session, and c) sex, and added the individual as a random factor. Second, for the number of successful trials as a response variable I assigned a) first touched tool, b) session, and c) sex as fixed variables, again, including individual as random factor.

Cone preference:

Across all tests, excluding the habituation, the frequency and duration of tool touches were compared using a Spearman rank correlation for each tool type.

Subsequently, the cone preference was established for the preference test, short-distance test and medium-distance test, by calculating the cone-ratio (cone-touches dividing by all tool touches). The preference for either cone or boat in the short-distance test and the medium-distance test was compared to the initial preference test to identify any potential changes in preference and determine whether the tool of choice depended on the task at hand. The ratio of cone choice to boat choice was calculated across the tests (preference test, short-distance, medium-distance). Before applying a Beta GLMM using the glmmTMB() function from the "glmmTMB" package (Brooks et al., 2017) with the cone ratio as the response variable, a small constant of 0.00001 was added to 0-values and a constant of 0.99998 was subtracted from 1-values, as the model cannot handle exact 0 or 1 values. Both test and sex were included as fixed effects, and subject was added as a random effect. The same procedure was applied to both frequency and the continuous duration data.

Time to success:

In addition to the main question, the time to success calculated as the difference between trial success and trial start for each individual during the pre-combined test was investigated, assigning session and sex as potential predictors. I applied a linear mixed model with the lmer() function (Bates et al., 2015). The log-transformed time-to-success data for each session was used as a response variable.

Results

None of the birds successfully combined either of the different element types during any test of the experiment, with only one exception. During the first session of the medium-distance test Fini first inserted a single cone into the apparatus. This cone came to lie inside the apparatus with the open end pointing at the tool opening of the apparatus. She thereafter picked up a second cone and inserted it through the tool opening into the opening of the first cone. The combined construct deteriorated before the food could be retrieved. While one female bird, Mayday, handled two tools at the same time, she wasn't successful in combining them into a compound. Due to her near success, Mayday received five human demonstration sessions. The demonstration sessions had no effect on her performance.

Habituation

Across the habituation phase, the birds touched the cone on average 66.1% (SE ± 6.04) of the time while the boat was touched in 33.9% (SE ± 6.04) of cases. Muki and Olympia, however, continued showing signs of fear or tension, expressed by raising their crest, spreading their tail feathers or swaying from side to side and were therefore excluded prior to the medium-distance test.

Preference test

During the preference test, the birds touched the cone 134 times (38.5%, SE ± 7.12%) and the boat 188 times (53.8%, SE ± 7.77%) in total. No significant difference between the frequency of cone and boat touches could be established, following the negative-binomial GLMM (Figure 6; Appendix 1B for detailed model output). A null-model comparison (AIC = 333.80, log-likelihood = -163.90, deviance = 327.80, df = 3) revealed no improvement of the full-model (AIC = 335.65, log-likelihood = -160.83, deviance = 321.65) with a $\chi^2(7)$ of 6.144 and a p-value of 0.189. However, on average, they spent 54.6% (SE ± 9.76%) of tool manipulation time with the cone and only 37.7% (SE ± 9.19%) with the boat, resulting in a significantly longer cone manipulation time ($\beta = 0.717$, SE = 0.360, z = 1.993, p = 0.046), as indicated by the zero-inflated GLMM (Figure 6). Additionally, both session ($\beta = 0.437$, SE = 0.216, z = 2.029, p = 0.042) and sex ($\beta = 1.232$, SE = 0.397, z = 3.102, p = 0.002) had significant effects on manipulation time, with duration increasing across sessions and males spending more time handling tools in total (Appendix 1A includes model output). Across sessions the boat was touched first 19 times, compared to the cone being touched first only eight times.

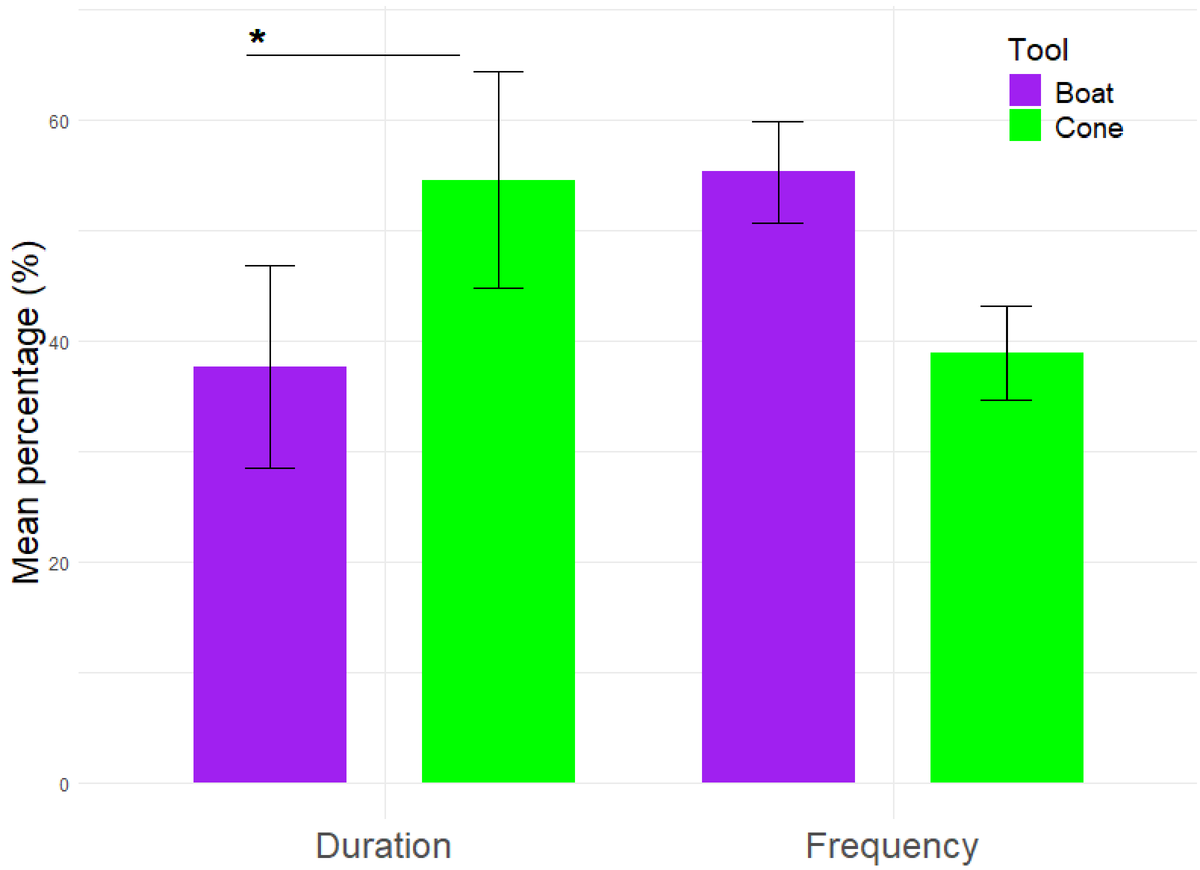

Figure 6: Mean percentages of both duration and frequency of touching cone (green) or boat (purple) for the preference test. The difference in duration is significant between cone and boat touches, but there is no significant difference in their frequency.

Short-distance test

During the short-distance test, the birds spent on average 63.3% (SE ± 10.5%) of their total tool-handling time using the cone. Inspection of the Tweedie GLMM results (see Appendix 2A) with a log-link confirm that subjects spend significantly more time with the cone ($\beta = 1.027$, SE = 0.2910, $z = 3.529$, $p < 0.001$) than compared to the boat (Intercept: $\beta = 2.6085$, SE = 0.514, $z = 5.079$, $p < 0.001$), as visualized in Figure 7. Increasing number of sessions also had a positive effect on the manipulation time ($\beta = 0.3101$, SE = 0.1513, $z = 2.049$, $p = 0.040$), but sex did not ($\beta = 0.371$, SE = 0.449, $z = 0.828$, $p = 0.408$). There were 510 incidences of cone touches, accounting for an average of 64.6% (SE ± 2.78%), resulting in a significant difference between the frequency of cone touches ($\beta = 0.788$, $SE = 0.200$, $z = 3.945$, $p < 0.001$) compared to boat touches (Intercept: $\beta = 2.380$, $SE = 0.356$, $z = 6.683$, $p < 0.001$). Across sessions the cone was touched first 64.6% (SE ± 10.4%) of times on average, however this did not significantly affect the frequency of tool touches ($\beta = -0.118$, $SE = 0.225$, $z = -0.526$, $p = 0.599$) – see Appendix 2B for detailed model output.

The short-distance test resulted in 228 successful trials and only 12 unsuccessful trials in which the birds were not able to retrieve the food reward with either the boat or the cone. Only two individuals, Muki and Olympia, never succeeded during this test. Notably they both continued to show signs of neophobia and were subsequently excluded from the test. The duration from trial start to trial success ranged from 8.91 s to 31.10 s across trials and between individuals, with a mean of 17.37 s.

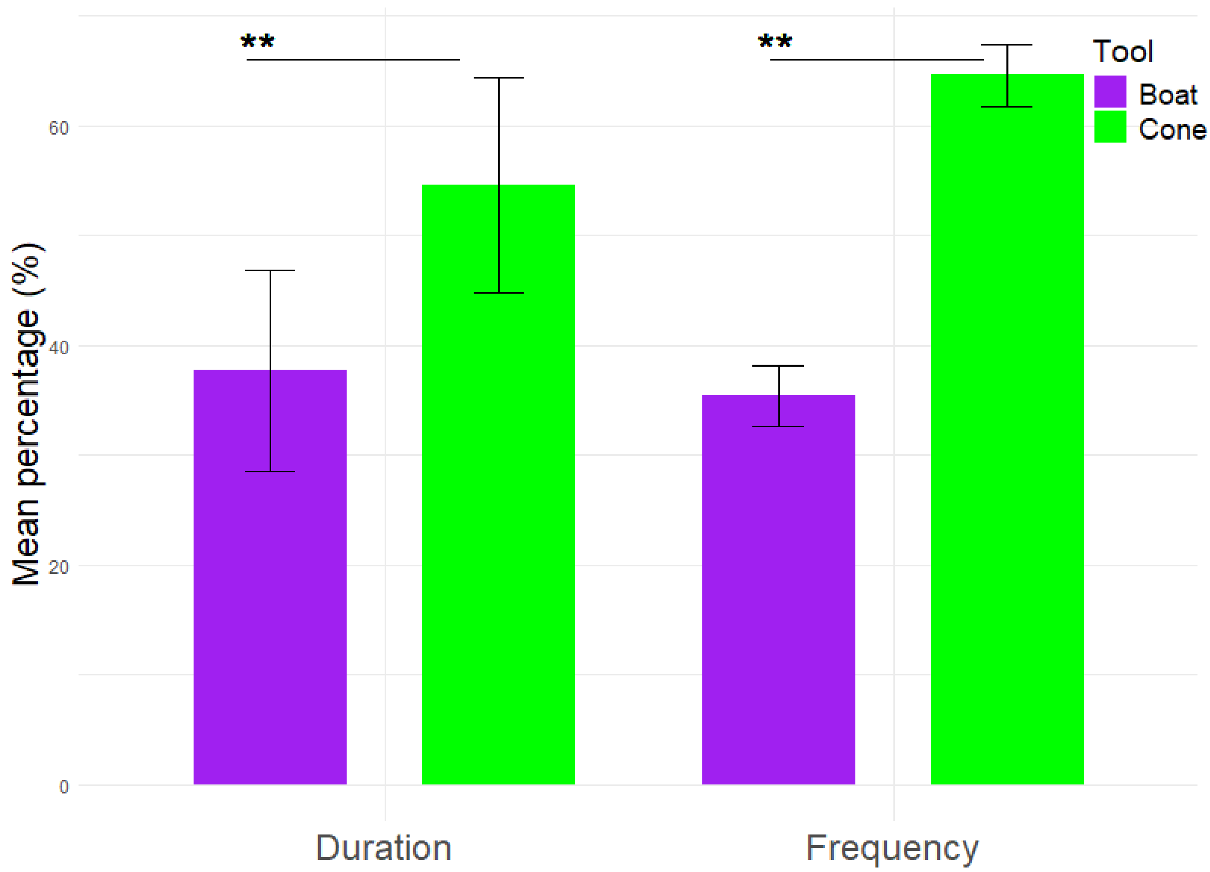

Figure 7: Mean percentages of tool touches during the short-distance test show significant differences in both duration and frequency.

Medium-distance-test

Similar to the short-distance test, the ratio of time spent with the cone was quite long with an average of 66.1% (SE ± 7.89) during the medium-distance test, resulting in significantly longer manipulation times for the cone ($\beta = 0.806$, SE = 0.126, z = 6.379, $p < 0.001$), compared to the boat (Intercept: $\beta = 4.258$, SE = 0.194, z = 21.944, $p < 0.001$). Additionally, the results of the Tweedie GLMM indicate, a decrease of manipulation time over sessions ($\beta = -0.041$, SE = 0.021, z = -2.022, p = 0.043). Appendix 3A includes the full model outputs.

The frequency of cone touches is also much higher with an average of 65.4% (SE ± 5.68%) than boat touches, and cones were more often touched first during each session (65.2%, SE ± 6.91%). The mean percentages of both duration and frequency are visualized in Figure 8. The negative-binomial GLMM (see Appendix 3B) confirmed a significant difference in number of cone touches ($\beta = 0.565$, SE = 0.096, z = 5.868, $p < 0.001$) compared to boat touches (Intercept: $\beta = 2.56347$, SE = 0.197, z = 13.011, $p < 0.001$).

Neither the tool that was touched first ($\beta = -0.114$, SE = 0.104, z = -1.101, p = 0.271) nor session ($\beta = 0.022$, SE = 0.016, z = -1.319, p = 0.187) or sex ($\beta = 0.140$, SE = 0.201, z = -0.700, p = 0.484) had any significant effects on tool touches.

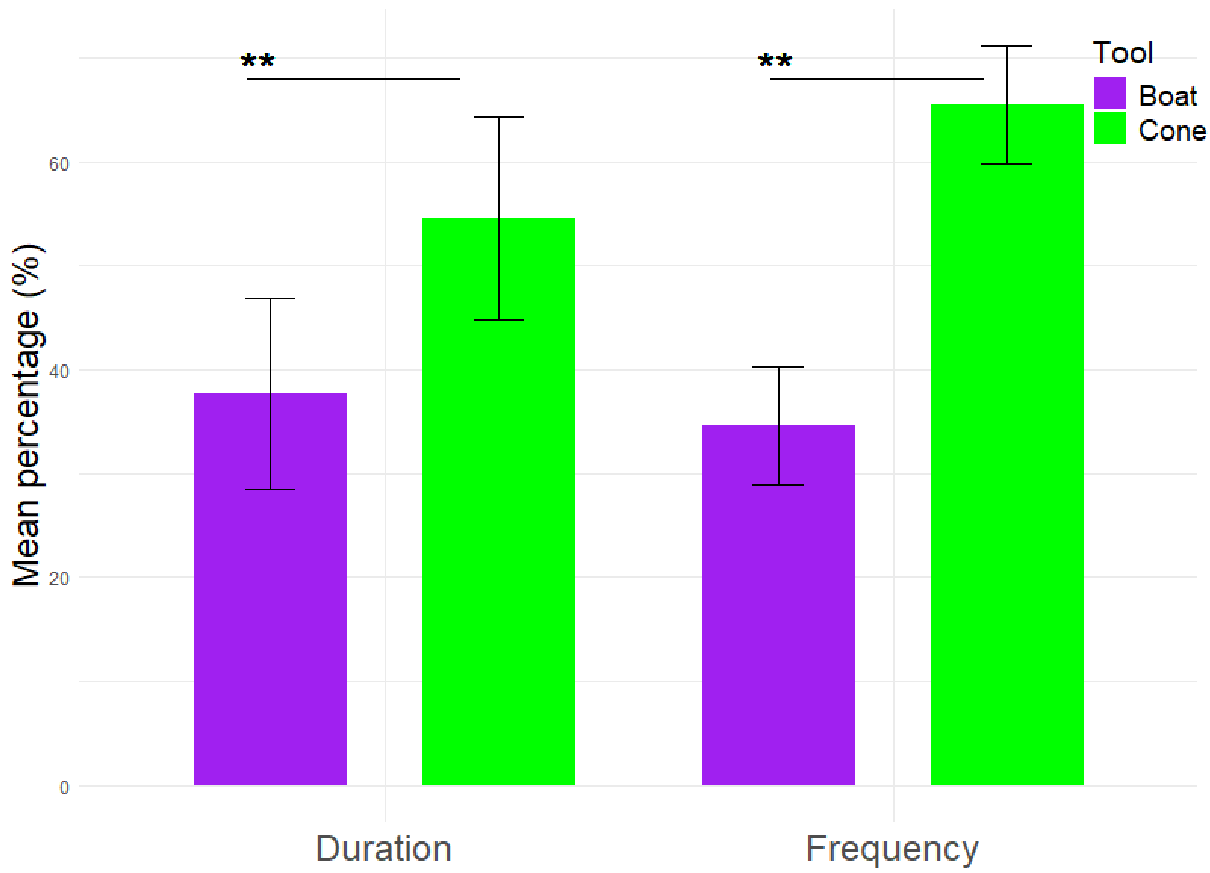

Figure 8: Mean percentages of tool touches during the medium-distance test show significant differences in both duration and frequency.

Among the possible tool-to-tool combination attempts (Table 1), subjects combined boat-to-cone (B-to-C) on average for the longest time (x̄ = 6.98 s, SE ± 1.13 s) and boat-to-boat (B-to-B) for the shortest time (x̄ = 0.639, SE ± 0.129). Cone-to-cone (C-to-C) attempts were the most frequent (x̄ = 8.64, SE ± 1.99), whereas B-to-B attempts were the least frequent (x̄ = 1.55, SE ± 0.578). There were no successful trials during the medium-distance test, indicating that there was no successful construction of a compound tool from two cones in order to reach the food reward. Likewise, the birds did not construct a compound tool from two boats.

With C-to-C as the reference tool-to-tool contact (β = -0.459, SE = 0.733, z = - 0.627, p = 0.029) there were significant differences across all other combination attempts (Figure 9): B-to-B combinations took significantly shorter (β = -0.476, SE = 0.218, z = -2.180, p = 0.029), while B-to-C and C-to-B took significantly longer (both $p < 0.001$). Neither session (β = -0.031, SE = 0.019, z = -1.647, p = 0.100) nor sex (β = -0.218, SE = 0.911, z = -0.240, p = 0.811) had any significant effect on touch durations (see Appendix 3C). The full model showed a significantly better fit (AIC =1522.2, log-likelihood = -751.09, deviance = 1502.2, df = 5, p < 0.001) than the null model (AIC =1694.9, log-likelihood = -842.47, deviance = 1684.9). Only mild overdispersion was detected (overdispersion ratio = 1.469). Residual diagnostics revealed potential patterning at the at the 0.75 quantile, suggesting some unmodeled non-linear effects. However, following model comparisons using a Gamma GLMM and a linear mixed model (lmer), the zero-inflated GLMM demonstrated the best model fit. Post-hoc analysis confirmed, that B-to-C has the highest log mean of 1.137 (Se = 0.462, 95% CI(0.231, 2.043), followed by C-to-B, then C-to-C and finally B-to-B. Tukey's pairwise comparisons (see Appendix 3C.1) reveal significant differences between all possible combination attempts (p < 0.001), except between C-to-C and B-to-B (p = 0.129).

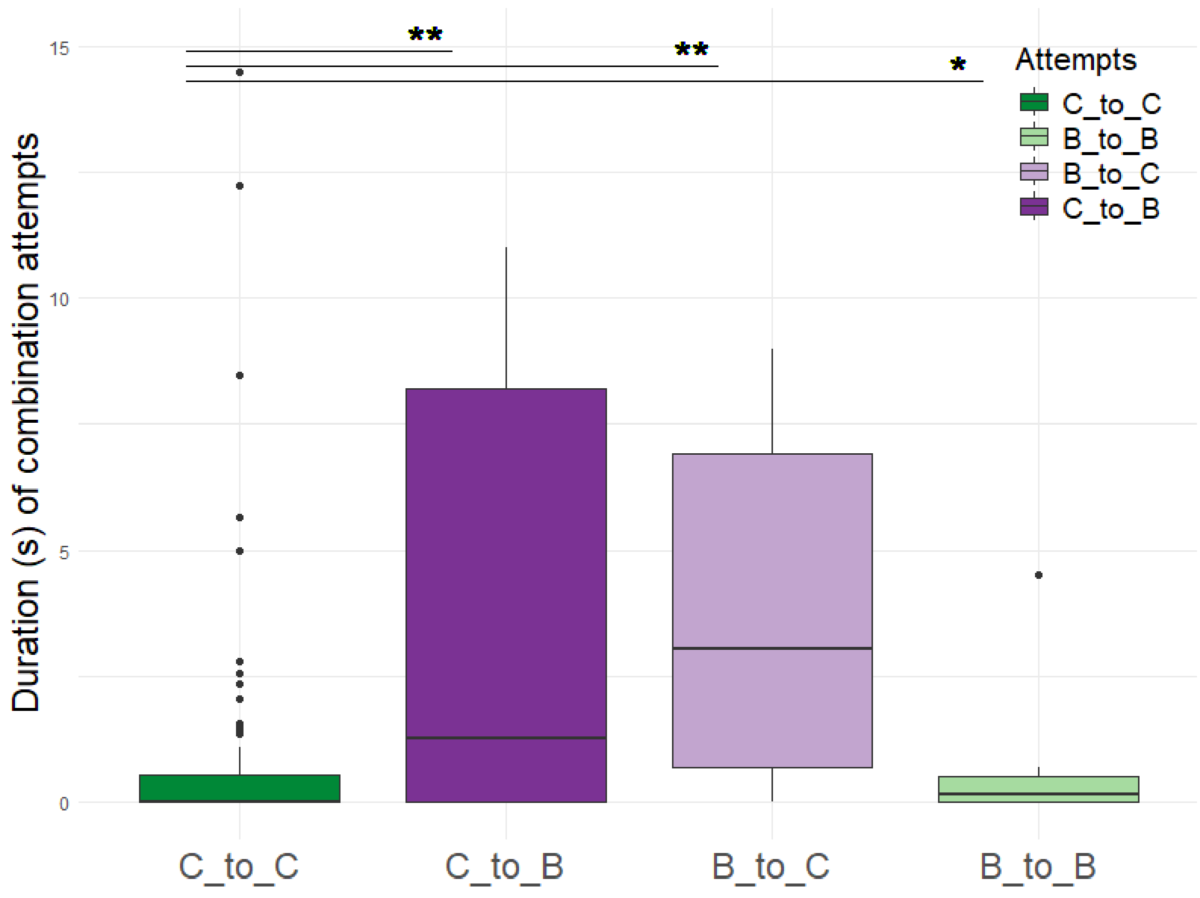

Figure 9: Observed durations of the four potential combination attempts. Compared to C-to-C, the combination attempts C-to-B and B-to-C have significantly longer durations while B-to-B is significantly shorter.

In contrast to the duration, results of the GLMM for the frequency (see Appendix 3D; Figure 10) of combination attempts showed significantly less C-to-B attempts (β = - 0.755, SE = 0.253, z = -2.979, p = 0.003), as well as significantly less B-to-B attempts (β = -1.763, SE = 0.322, z = -5.484, p < 0.001) with C-to-C as reference (intercept: β = -0.135, SE = 0.390, z = -0.345, p = 0.730). B-to-C, however, was not significantly different (β = -0.132, SE = 0.231, z = -0.568, p = 0.570). Neither sex (β = 0.163, SE = 0.420, z = -0.389, p = 697), nor session (β = 0.004, SE = 0.032, z = 0.115, p = 0.909) had any significant effect on the frequencies. Comparing the full model (AIC = 804.87, log-likelihood = -394.44, deviance = 788.87, df = 5) to the null model (AIC = 834.21, log-likelihood = -414.11, deviance = 828.21), the likelihood ratio test indicated a significantly better fit of the full model with p < 0.001. Following Tukey's pairwise comparisons, C-to-C was found to be the most frequent of the four possible behaviours with a log mean of -0.196 (Se =0.250, 95% CI(-0.686, 0.294), followed by B-to-C, then C-to-B and finally B-to-B. In more detail, C-to-C was significantly more frequent than C-to-B (estimate = 1.764, p < 0.001). C-to-B was also significantly more frequent than B-to-B (estimate = 1.632, p < 0.001). No significant difference between C-to-C and B-to-C (estimate = 0.132, p = 0.942), and between C-to-B and B-to-C (estimate = -0.623, p = 0.066) was detected.

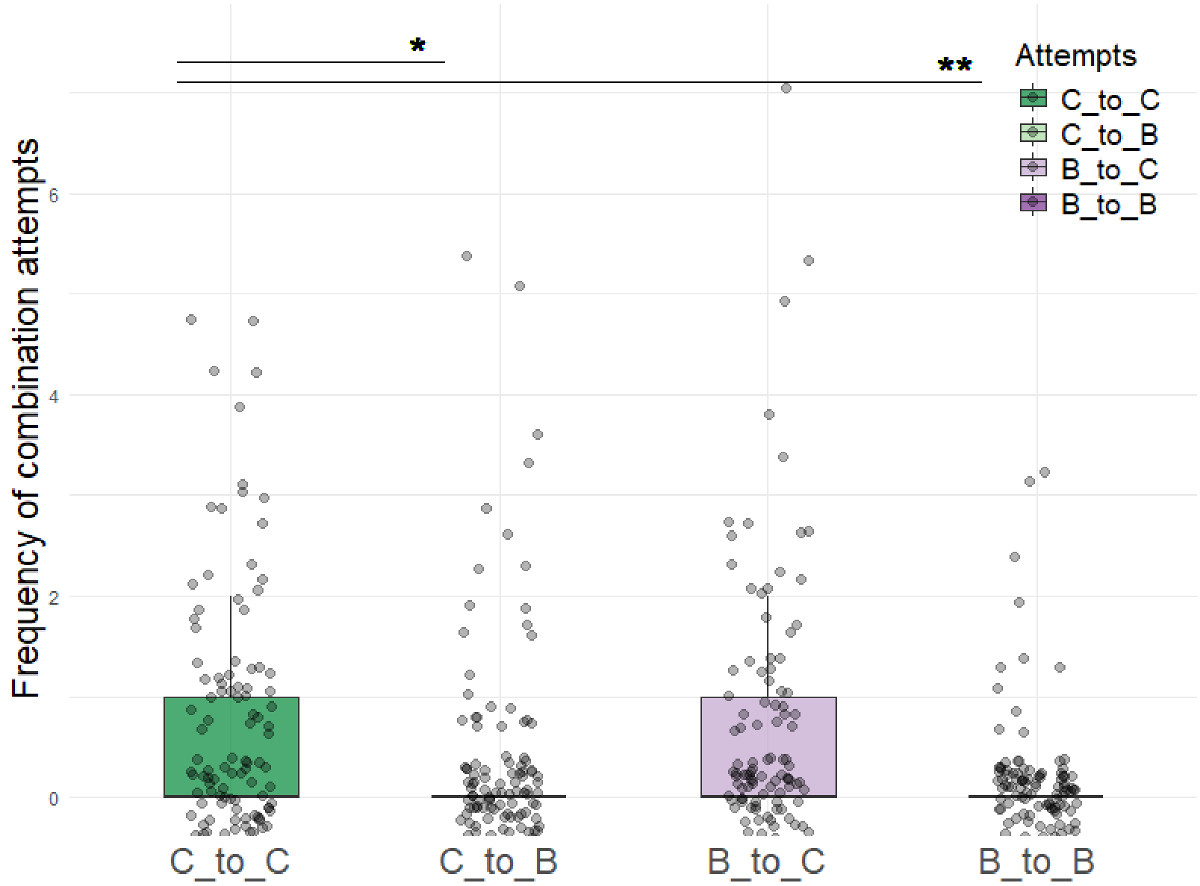

Figure 10: Frequency of four possible combination attempts. Significant differences lie between C-to-C and C-to-B, as well as between C-to-C and B-to-B. Due to multiple data points sharing similar values jitter was added to the boxplot in order to visually separate individual data points.

Two tools were touched simultaneously 163 times in total and 14.8 (SE ± 2.96) times on average per bird during the medium-distance test for a duration of 32.2 (SE ± 13.2), compared to only nine times during the short-distance test in total. Averaged frequency and duration of two tools being touched simultaneously along with four potential combination attempts are shown in Table 1. Every bird, except Konrad, touched two tools simultaneously during the medium-distance test, while only five birds touched two tools at once during the short-distance test. The inclusion of session and sex as fixed effects in the zero-inflated GLMM (Appendix 3E) did not significantly improve model fit (AIC = 485.19, log-likelihood = -235.59, deviance = 471.19, df = 2, p = 0.081) in the likelihood-ratio test, suggesting neither had any significant effect on the duration of touching two tools at the same time. Likewise, the frequency of touching two tools simultaneously (see Appendix 3F for detailed results) was not significantly affected by either session or sex, since a comparison of the full model (AIC = 348.24, log-likelihood = -

169.12, deviance = 338.24, df = 2, p = 0.428) to the null model (AIC = 345.94, log-likelihood = -169.97, deviance = 339.94), did not show significant improvement of model fit.

Table 1: Mean raw counts and percentages of combination attempts and touching two tools simultaneously, with standard errors calculated for both frequency and duration data. While C-to-C accounts for the largest proportion in terms of frequency, there are larger proportions of B-to-C and C-to-B attempts contributing to the duration of combination attempts.

Median-distance	Frequency	Duration	Frequency %	Duration %
C_to_C	8.84, SE ± 1.99	0.770, SE ± 0.216	42.4, SE ± 5.82	8.90, SE ± 1.73
C_to_B	4.45, SE ± 1.54	3.60, SE ± 0.422	16.6, SE ± 4.39	27.8, SE ± 2.45
B_to_C	8, SE ± 2.08	6.98, SE ± 1.13	34.4, SE ± 3.34	57.4, SE ± 3.07
B_to_B	1.55, SE ± 0.578	0.639, SE ± 0.129	6.55, SE ± 2.72	5.87, SE ± 0.878
two tools	14.8, SE ± 2.96	32.2, SE ± 13.2		

Demonstration of compound tool construction

Regardless, none of the birds was successful in creating a compound tool from any tool type. Therefore, one female was chosen for five sessions of being demonstrated the task the combining of two tools into a compound tool by the experimenter. For this, Mayday was picked, since she showed the highest frequency of combination attempts, except for B-to-B and, compared to other birds, long durations of combination attempts for C-to-C (21.2 s) and for B-to-C (40.2 s), as well as the highest frequency (37 of 163 times) and duration (160 s of 365 s) for touching two tools simultaneously.

After having received demonstrations, she touched the cone 23 times (for a duration of 1317 s) and the boat 12 times (for 217 s) across all sessions. Overall, she attempted very few combinatory actions, accounting for two C-to-C attempts (0.24 s), only one of each C-to-B (1.26 s) and B-to-C (0.36 s) and finally no B-to-B (0 s) at all. The boat was touched first one more time than the cone. Due to very low values and no successful trials following the demonstration sessions, statistical analysis was deemed unnecessary for this test.

Since construction of a compound tool has not been successful under given conditions, I proceeded with the pre-combined test, by providing both pre-assembled compound tools, CC (compound tool constructed of two cones) and BB (constructed of two boats) ahead of trial start.

Pre-combined test

For the pre-combined test one more bird, Konrad, had to be excluded, because he was taken ill at the time. The birds showed a high frequency (87.1% SE ± 3.66%) and duration (96.6% SE ± 1.20%) of CC touches. CC was touched first a total of 136 times, on average 92.1% (SE ± 2.96%). The Tweedie GLMM (Appendix 4A) confirmed the significantly longer CC manipulation times (β = 3.700, SE = 0.205, z = 18.050, p < 0.001). Furthermore, a decrease in manipulation time over sessions could be observed (β = -0.060, SE = 0.015, z = -4.075, p < 0.001), however, there was no effect of sex (β = -0.379, SE = 0.365, z = -1.037, p = 0.300). Significantly more CC touches (β = 2.290, SE = 0.181, z = 12.673, p < 0.001) were confirmed by the negative-binomial model (Appendix 4B). The model additionally indicates a significant decrease in expected frequency when CC was touched first (β = -1.007, SE = 0.286, z = -3.522, p < 0.001). As a supplementary finding (Appendix 4E), each additional session was associated with decrease in the frequency of tool touches (β = -0.061, SE = 0.017, z = -3.609, p < 0.001).

The pre-combined test resulted in 136 successful trials and 13 unsuccessful trials, which coincides with the number of CC being touched first across trials. However, no effect of the first touched tool on the number of unsuccessful trials across sessions and individuals could be detected as shown in the binomial GLMM (β = 0.093, SE = 1.027, z = 0.091, p = 0.928) as shown in Appendix 4C. Additionally, sex and session did not appear to influence the number of unsuccessful trials. In a full-model comparison (AIC = 84.896, log-likelihood = -37.448, deviance = 74.896, df = 3) against a null-model without these predictors (AIC =82.465, log-likelihood = -39.232, deviance = 78.465), no significant improvement was assessed (p = 0.312).

The success of a trial (β = 2.894, p = 0.034) was also not influenced by sex (β = -1.381, p = 0.255), session (β = 0.112, p = 0.181) or whether the cone was touched first (β = 0.093, p = 0.928) – shown in detail in Appendix 4D. Compared to the null model (AIC = 88.473, log-likelihood = -39.232, deviance = 78.465) there was no significant improvement of model fit (AIC = 84.896, log-likelihood = -37.448, deviance = 74.896), with a $\chi^2(3)$ of 3.569 (p = 0.312).

The average time across all individuals and sessions from beginning of a trial to success of the trial amounts to 43.8 s, varying from 20.3 s to 76.4 s between individuals. The time to success was significantly affected by the session, as a null-model comparison confirmed ($\chi^2(15) = 41.82$, $p < 0.001$), but not by sex (b = -0.295, SE = 0.265, t = -1.113). Compared to the first session, negative estimates indicate a faster time to success, with Session 8 showing the strongest effect (b = -1.614, SE = 0.351, t = -4.593), relative to the intercept (b = 4.562, SE = 0.310, t = 14.723). Overall, the full model (AIC = 344.85, log-likelihood = -154.42, deviance = 308.85, df = 15, p < 0.001) significantly improved model fit (AIC = 356.67, log-likelihood = -175.33, deviance =350.67).

Cone-preference between tests

Beta GLMMs were applied for both frequency and duration ratios – detailed results are shown in Appendix 5A and 5B. The medium-distance test differed significantly from the initial preference test in both frequency (β = 0.997, SE = 0.282, z = 3.54, $p < 0.001$) and duration (β = 0.622, SE = 0.298, z = 2.089, p = 0.037). In contrast, the short-distance test did not differ significantly from the preference test in either the count data (β = 0.359, SE = 0.324, z = 1.11, p = 0.267), or the duration data (β = 0.148, SE = 0.337, z = 0.438, p = 0.661). Mean proportions of cone and boat manipulations for each test are shown in Table 2.

Additionally, males overall expressed lower cone preference than females in both frequency (β = -0.769, SE = 0.382, z = -2.01, p = 0.044), and duration (β = - 0.916, SE = 0.452, z = 2.026, 0.043). While there was no overdispersion or zero-inflation in the model for the count data, visual inspection of residuals showed a significant deviation in the expected distribution, as indicated by the Kolmogorov-Smirnov (KS) test (p = 0.016), suggesting that not all variability might be explained. No such deviation was detected in the duration data, suggesting a better model fit.

When considering the CC-ratio of the pre-combined test as a cone-ratio, the pre-combined test showed the highest ratio with 87.1% SE ± 3.66% for the count data and even 96.6% SE ± 1.20% for the duration. However, including the pre-combined test in the model resulted in significant within-group deviations from uniformity, as well as a significant deviation from the expected distribution (KS: p = 0) in the count data. Similar issues arose in the time data, which also showed a significant variance in the Levene test for homogeneity.

Table 2: Mean percentages in frequency and duration of tool manipulation for both cone and boat with standard errors. CC (cone compound tool) was included in the cone column and BB (boat compound tool) was included in the boat column. Since the number of subjects wasn't constant across all tests, the number of participating birds in each of the tests was added to the table. Preference for the cone tool was most pronounced in the pre-combined test in both frequency and duration.

Tool Manipulation Frequency	**Nr. of birds**	**% Cone / CC**	**% Boat / BB**
Habituation	13	66.1, SE ± 6.04	33.9, SE ± 6.04
Preference-test	13	38.5, SE ± 7.12	53.8, SE ± 7.77
Short-distance	13	64.6, SE ± 2.78	35.4, SE ± 2.78
Medium-distance	11	65.4, SE ± 5.68	34.6, SE ± 5.68
Pre-combined	10	87.1, SE ± 3.66	12.9, SE ± 3.66
Tool Manipulation Duration	**Nr. of birds**	**% Cone / CC (s)**	**% Boat / BB (s)**
Habituation	13	51.5, SE± 4.34	48.5, SE ± 4.34
Preference-test	13	54.6, SE ± 9.76	37.7, SE ± 9.19
Short-distance	13	63.6, SE ± 10.5	29.0, SE ± 9.34
Medium-distance	11	66.1, SE ± 7.89	33.9, SE ± 7.89
Pre-combined	10	96.6, SE ± 1.20	3.44, SE ± 1.20

Discussion

The aim of this study was to determine whether Goffin cockatoos can combine two elements and construct a functioning compound tool for retrieving a food reward. Throughout the entire test there was no successful construction of a compound tool, with a single exception. Therefore, none of the trials during the medium-distance test were successful, making the long-distance test obsolete. On one occasion, the adult female Fini inserted a cone into another that was already inside the apparatus. Since the open end was pointing at her though the opening, thereby enhancing a successful combination by chance, an intention of combining the cones cannot be inferred from this single incident as in later trials, Fini did not repeat this action. Therefore, H1 must be rejected.

Although Goffin cockatoos have shown to be dexterous tool users and tool manufacturers (Auersperg et al., 2012; Auersperg et al., 2014, Auersperg et al., 2016; Laumer et al., 2017; Osuna-Mascaro et al., 2022; O'Hara et al., 2021; Osuna-Mascaro et al. 2023), it can be challenging to coordinate even a single tool with their beak, tongue and claws. To not only hold, but also combine two separate elements in a very specific way, is both cognitively (Price et al., 2009) and anatomically challenging, because both the body and the task impose morphological and ergonomic constraints (Mangalam & Fragaszy, 2018). Nevertheless, Goffin cockatoos engaged in combinatorial actions during object play (Auersperg et al., 2016; Zewald et al., in prep.). Different requirements have to be met to successfully create a compound tool with two cones:

1) They need to secure one cone in a way it stays in a fixed position to insert the second cone. To secure the first tool they need to either hold it with their foot by stepping on it, or push it against another structure with the second cone while it is inserted. During object play the Goffins exhibit this behaviour (Zewald et al., in prep.); 2) Both of the tools have to point into the same direction; 3) The cones need to be firmly connected after insertion of the second cone.

Spontaneous innovation of a compound tool to succeed in a problem-solving task may represent an ill-structured problem as it includes an intermediate step of tool inseration which doesn't immediately result in a food reward. Additionally, the information required to combine two elements is not provided initially. Goffin cockatoos previously succeeded in a hook bending experiment (Laumer et al., 2017), representing an ill-structured problem generally defined as a problem which is missing information about its start- or goal state or the required tool transformation (Wood, 1983). Compared to a well-structured problem, an ill-structured problem includes unpredictable steps of the solution. The process of inserting a multi-branched stick into a tube and sequentially breaking off each branch as it become stuck serves as an example of a well-structured or straight-forward problem. Without any additional assistance, hook bending only includes the start state without any initial information regarding the transformation of the tool required for a specific problem-solving task (Cutting et al., 2013). Even children up to eight years old, tested in the UK, have difficulty in spontaneously innovating a hook tool (Beck et al., 2010). Children in the age groups 4-5 and 5-6 performed poorly prior to any kind of demonstration (Cutting et al., 2013). Both in a hook-bending task and in the constructing a compound tool the information of the required transformation of one or more tools is not predictable initially. Consequently, Goffin cockatoos may encounter limitations with certain ill-structured problems despite their previously observed playful combination of objects.

During the previous experiment on New Caledonian crows investigating compound tool construction (Bayern et al., 2018), the first element was already placed on a tool holder in the correct position, facilitating insertion of the second element. Insertion may have been enhanced additionally due to their inherent drive to inspect openings with other objects, therefore, New Caledonian crows may have created their first compound tool by chance (Auersperg et al., 2014). Preparing the first element of the compound tool in a tool holder and the crows inherent drive to inspect openings arguably may have resulted in a less ill-structured problem as compared to the task expected from the Goffin cockatoos. Since no tool holder was provided to the Goffin cockatoos, this limited the possibility for creating a functioning compound tool accidentally. As mentioned above, in a single instance, Fini inserted a cone into another, but since there was no structure stabilizing the first tool inside the apparatus, the two cones could not be firmly connected and fell apart soon after.

Combining two boats involved similar requirements comparable to the cones, however, the construction of a boat compound tool could be simpler, as the larger opening of the boat provides easier access for insertion of the second boat. Nevertheless, the cockatoos also did not construct a boat compound tool. This can be expected due the overwhelming preference of the cone during most of the tests, providing less opportunity for combining two boats. The shape of the boat may offer an alternative explanation for the observed cone preference, as the boat may be more difficult for the birds to manipulate.

Although one individual, Mayday, did show some attempts of touching the two cone tools simultaneously, she was not successful after observing the demonstration sessions conducted by the experimenter. In further consequence, she did not demonstrate the construction of a compound tool to her conspecifics, which contrasts with my prediction of H2.

A human demonstrator might not be sufficient as a model for enhancing successful compound tool construction. Perceiving a conspecific using a tool, may improve the observers' ability to successfully acquire this skill via emulation learning, as shown in a study of Auersperg et al. in 2014. However, this option was not available during the investigation of compound tool use. Only a demonstrator bird was able to "teach" successful tool use in a problem-solving task, while the "ghost-demonstrations" – using only magnets to move the tool or the reward – did not result in the acquisition of successful tool use (Auersperg et al., 2014).

However, in the pre-combined test, the birds apparently recognized the functionality of a pre-constructed compound tool, which is supported by 136 successful trials, compared to only 13 unsuccessful ones. Their choice for the functioning tool during the pre-combined test was expected due to their flexibility in adjusting tool choice depending on the context (Laumer, Bugnyar & Auersperg, 2016). Previously, the Goffins have also shown their capability in adjusting tool length depending on the distance to the food reward. When presented with an apparatus similar to the current one, they constructed significantly longer tools from cardboard when the distance from the tool opening to the food reward was increased (Auersperg et al., 2018). This finding supports the claim that the Goffins are capable of recognizing the necessity for varying tool lengths depending on the distance to the food reward. Here, they show a clear preference for the long CC tool both in frequency (87.1, SE ± 3.66) and duration of use (96.6, SE ± 1.20), supporting the prediction of H3 – preference for handling the CC tool – and suggesting that tool choice is not random. The selection for the more effective tool points towards an ability to evaluate tool utility. It is likely that previous exposure to the short-distance task involving similar tool components enhanced the association of the cone with success, possibly playing a role in the recognition of the correct tool choice.

While no successful construction of a compound tool was performed during the medium-distance test, two tools were touched simultaneously by the birds more often than during the short-distance. However, this may also be attributed to longer trial durations during the middle-distance test, as they had previously succeeded in the short-distance test without exhausting the entire trial time.

Additionally, and regardless of intention, tool-to-tool combination attempts occurred during the middle-distance test, with the highest proportion being C-to-C (42.4, SE ± 5.82), which is the correct combination for a functioning compound tool. As previously mentioned, the cone may be food associated due to their success during the short-distance test. A closer inspection of the duration in the zero-inflated GLMM revealed, that the time spent on C-to-C attempts was significantly shorter compared to both C-to-B and B-to-C, while B-to-B had significantly shorter touch durations than C-to-C. These contrasting results suggest, that there is no clear preference for attempting C-to-C combinations. The high C-to-C ratio for the frequency is not reflected in the duration, since a higher frequency does not necessarily result in longer durations.

These findings suggest that the birds did not consistently engage in the correct combination attempts, despite of frequently making C-to-C contact. The shorter durations of C-to-C contact imply, that the high frequency of C-to-C contact was likely incidental rather than intentional. Nevertheless, the higher frequency and duration of touching two tools simultaneously during the medium-distance test requiring the use of two tools, as compared to the short-distance test requiring only one tool, indicate that the cockatoos adjusted their behaviour in response to the necessity of using more than one tool in the medium-distance task. When Goffin cockatoos reach an impasse during a problem-solving task, they tend to start combining elements with each other (Osuna-Mascaro & Auersperg, 2018; O'Hara, unpublished observation). It is possible that their inability to solve the task, propelled them to try different tool combinations and touch more than one tool at the same time.

These results confirm my first prediction of H4, stating that the cockatoos will touch two tools simultaneously more often and for longer durations during the test condition requiring more than one tool. The second and third predictions – that cockatoos would attempt to touch one tool with another regardless of success during the medium-distance task, and that these attempts would involve the cone element more frequently and for longer durations – were supported. Combination attempts including at least one cone-element were more frequent and had longer durations than B-to-B attempts. This is likely a result of the cone preference, as indicated by the large cone-ratio, during the medium-distance test.

The tool choice in the initial preference-test is not consistent in both frequency in duration, as the calculated 38.5%, SE ± 7.12% cone-ratio in the frequency data and 54.6%, SE ± 9.76% cone-ratio in the duration data indicate. The cone-ratios showed higher consistency between frequency and duration data during both short-distance (64.6%, SE ± 2.78% for frequency, 63.6%, SE ± 10.5% for duration) and medium-distance tests (65.4%, SE ± 5.68% for frequency, 66.1%, SE ± 7.89% for duration). The preference for the cone was more pronounced in the short-distance and significantly higher in the medium-distance task compared to the initial preference-test, as revealed by the beta GLMMs used for both frequency and duration data. The model also included sex as a fixed effect, which showed that males had a significantly smaller cone-ratio compared to the females. While the cone ratio was highest in the short-distance test, as compared to the medium-distance test and the initial preference test, the model results still indicated a significant difference only between the medium-distance and the preference test. This suggests, that there might be higher variability in the short-distance data and other contributing factors, like sex. Female Goffin cockatoos tend to have slightly smaller body size and smaller beaks than males (Grahl, 1990), possibly adding to the difficulty in manipulating certain tool types, such as the boat element.

The cockatoos adjust their preference when they are confronted with an apparatus, providing a food reward, with females being even more likely to adjust their tool choice. The preference for either tool is not consistent in frequency and duration during the preference test, but more so during the short-distance test and the medium-distance test, which does not fully support the first prediction of H5, stating that the preference for either tool type will be approximately equal during the preference test and the short-distance test. The second prediction, however, is supported by the results of the model, indicating a preference-shift towards the cone especially during the medium-distance test.

The birds might shift their preference due to recognizing the necessity of using a specific tool-type to solve this particular task, although, they did not actually succeed. An alternative explanation for the cone preference after the initial preference test is that the cone can be inserted into the tool opening of the apparatus more easily than the boat. On occasion, the edges of the boat may create more resistance during insertion into the tool opening, making it more difficult to use. Therefore, during a problem-solving task involving an apparatus, the cockatoos might simply choose the tool that is easier to handle.

While no successful trials were completed during the medium-distance task, the cockatoos were successful in most cases in both the short-distance test and the pre-combined test. As expected, due to their success in former studies using a similar apparatus (Auersperg et al., 2016; Laumer, Bugnyar & Auersperg, 2016, Auersperg et al., 2018), the birds showed little difficulty in solving the short-distance task requiring only a single tool. They did so within a short mean time across individuals and trials (17.37 s). Even though, any tool would have been sufficient for reaching the food reward inside the apparatus, the Goffins chose to use the cone more often relative to the boat. They have already used stick-type tools, including even different materials, in many previous studies (Auersperg et al., 2012; Auersperg et al. 2014; Auersperg et al., 2016; Laumer, Bugnyar & Auersperg, 2016; Auersperg et al., 2017; O'Hara et al., 2021; Mioduszewska, Auersperg & O'Hara, 2022; Osuna-Mascaro et al., 2022; Osuna-Mascaro et al., 2023), therefore, it comes as no surprise, that they would continue to use another single, elongated element as a tool to push the food reward off a platform. However, the short-distance task only assessed the birds' ability to interact with the apparatus and apply the tools effectively. During the medium-distance test the focus shifted to the additional task of combining two cones. They continued using the cone more often and handling it for a longer time. Many birds would continue using one tool on the apparatus.

Additional findings during this experiment show that the Goffins display different strategies in their execution of the problem-solving task and the handling of their tools. These findings are fully consistent with a previous master thesis on the handling of disbalanced stick-tools by Özge (2024). During my experiment at least four different strategies were observed: 1) Utilizing exclusively the beak, some birds first picked up the tool while holding it at one end between upper and lower mandible and subsequently adjusted it between their tongue and upper mandible, and finally, stepped back and inserting it into the tool opening of the apparatus with the other end. Figaro seemed to prefer this method which was referred to as "precision-grip" (Özge, 2024); 2) After picking up the tool with their beak gasping it between upper and lower mandible, some individuals stepped towards the apparatus and subsequently inserted the end closer to their holding position, sliding the tool into the apparatus bit by bit with their tongue. This technique was called "needling-technique" (Özge, 2024); 3) The technique referred to as "foot insertion" (Özge, 2024) included holding the tool at one end between upper and lower mandible after picking it up with their beak, followed by making direct contact between tool and the apparatus or placing it near the tool opening. Finally, these individuals grasped the tool with one foot in the middle or at the distal end to slide it into the tool opening of the apparatus; 4) Some birds secured the tool at the tool opening at the proximal end after grasping it between upper and lower mandible which was called "entrance securing" (Özge, 2024). The birds did so by leaning the tool against the front of the apparatus, then grabbed the end resting on the table and slid it into the tool opening. Similar strategies have been observed in the previous Golf Club task while handling the stick tool (Osuna-Mascaro et al., 2022). Additionally, during a study investigating social transmission of tool use, different strategies in retrieving a reward with a tool were observed, after subjects had acquired the use of a stick-tool to retrieve a food reward via emulation learning (Auersperg et al., 2014). This indicates a very consistent pattern of strategies applied in various problem-solving tasks and presents an opportunity for further investigation on distinct and individual preferences of tool use and dexterity in the tool use of Goffin cockatoos.

Limitations

The shape and size of the tools might be key limitations of this investigation and play a role in the cockatoos' difficulties in combining them. There is a possibility that the boats higher insertion resistance, due to friction at the tool opening of the apparatus, has influenced the predominant choice for the cone during short- and medium-distance tasks, rather than the necessity for using the cone. Likewise, disadvantages of the particular shape and size of the cone cannot be ruled out. Compared to other stick-type tools, the cones have a more ununiform structure with a diameter of 1.7 cm only at the wider part, which can possibly impact the birds' ability to either properly grasp the tools or combine them into a compound tool.

Likewise, the pre-constructed compound tools also have potential limitations. Accidental disassembly of the pre-constructed compound tools occurred 39 times during 150 trials of the pre-combined test, due to rough handling by the birds. On less occasions, some individuals appeared to intentionally deconstruct the compound tools before using them on the apparatus.

Some of the birds, especially Mayday, had a tendency to attempt handling two tools simultaneously or touch one tool with another, indicating that they potentially recognized the need for a second tool during the medium-distance task. Possibly they may have succeeded eventually had the experiment continued for a longer duration, as Goffin cockatoos tend to engage in combinatory actions when they reach an impasse (Osuna-Mascaro & Auersperg, 2018; O'Hara, unpublished observation). Therefore, future investigations on the topic of compound tool construction in Goffin cockatoos can extend the amount of time they are confronted with this problem-solving task and increase the chance of successful object combinations. This claim is additionally supported by previous cases of associative tool use of Goffin cockatoos, such as composite tool use (Osuna-Mascaro et al., 2022), tool-set use (O'Hara & Mioduszewska et al., 2022; Osuna-Mascaro et al., 2023) and sequential tool use (Rössler et al., in prep.).

Conclusion

This study confirms that Goffin cockatoos attempt object-to-object combinations once they have reached an impasse with a problem-solving task, as well as during object play. Even though these attempts did not result in a successful construction of a compound tool during the experiment, the birds have engaged mainly in combination attempts that include the cone element necessary for the construction of a functional tool. Additionally, they touched two tools simultaneously on numerous instances during the medium-distance test, suggesting that they adjusted their tool use depending on the test condition to some degree.

Regarding future studies on compound tool use in Goffin cockatoos, some steps can be taken to potentially facilitate compound tool construction for the birds, without providing excessive assistance: 1) Replacing the cone with an angular element, rather than a rounded one, may help secure the tool in a fixed position with less risk of rolling or slipping; 2) To prevent accidental deconstruction of the compound tools – both CC and BB– the elements need to be more firmly connected, for example by allowing them to snap into place; 3) Replacing the cone element with a more uniform element could improve the effectiveness of tool manipulation; 4) Reducing the size of the tools could potentially facilitate tool handling and compound tool construction, and finally; 5) Using a tool holder, as was done with New Caledonian crows (Bayern et al., 2018), helps secure the first component while the birds insert the second. The latter option provides the most assistance to the construction of a compound tool.

Despite of their capability to solve ill-structured problems, such as hook bending (Laumer et al., 2017), it is possible that the Goffin cockatoos face certain cognitive or anatomical limitations in other ill-structured problems, such as compound tool use.

Goffin cockatoos have previously exhibited the ability to adjust their tool choice based on the functionality of a tool (Laumer, Bugnyar, & Auersperg, 2016). Their strong preference for the compound tool pre-assembled from two cone elements supports these earlier findings.

On several occasions, Goffin cockatoos have demonstrated that they are capable of associative tool use, involving the use of more than one tool. This includes not only the use (O'Hara et al., 2021) but also the transport of tool sets (Osuna-Mascaro et al., 2023), as well as sequential tool use (Rössler et al., in prep.) and composite tool use in the golf club task (Osuna-Mascaro et al., 2022). Compound tool use is the rarest form of tool use observed in animals, as it is cognitively very demanding due to the intermediate, combinatorial step (Price et al., 2009). Goffin cockatoos have demonstrated object-to-object combinations on during object play (Auersperg et al., 2015; Zewald et al., in prep.) and attempted several combinations during the experiment. Thus, they have repeatedly shown their capabilities of innovating tool use in response to the specific problem-solving tasks. Based on all previous findings, supporting the Goffin cockatoos' capacity for innovative associative tool use, as well as their specific object combination attempts during this experiment, it can be assumed that the ability to create compound tools may emerge in the future, if an ecologically suitable set up can be provided to the cockatoos.

Acknowledgements

I would like to express my sincere gratitude to Alice Auersperg for her profound and patient feedback on my data collection and written master‘s thesis, and to Thomas Bugnyar for his role as exceptional supervisor. I am also incredibly grateful to Theresa Rössler for her support with the statistical analysis and to Nicola Auersperg for the adjustment of the apparatus. Additionally, I would like to thank entire Goffin lab team for their cooperation and kindness. My heartfelt thanks go to my parents and friends for their emotional support, and to my partner, who supported me during the most stressful times. I am also deeply grateful to my fellow students and colleagues I got to know throughout this process, whose feedback and support have been invaluable. For their cooperation, I would like to give special thanks to the cockatoos who participated in my experiment; Dolittle, Figaro, Fini, Heidi, Kiwi, Konrad, Mayday, Moneypenny, Muki, Muppet, Olympia, Pipin and Zozo.

References

Adelmant, C., Osuna-Mascaró, A. J., Folkertsma, R., & Auersperg, A. M. I. (2024). Goffin's cockatoos use object mass but not balance cues when making object transport decisions. *Scientific Reports*, 14(1), 27265–10. https://doi.org/10.1038/s41598-024-76104-7

Auersperg, A. M. I., Borasinski, S., Laumer, I., & Kacelnik, A. (2016). Goffin's cockatoos make the same tool type from different materials. *Biology Letters* (2005), 12(11), 20160689–20160689. https://doi.org/10.1098/rsbl.2016.0689

Auersperg, A. M. I., Gajdon, G. K., & Huber, L. (2010). Kea, *Nestor notabilis*, produce dynamic relationships between objects in a second-order tool use task. *Animal Behaviour*, 80(5), 783–789. https://doi.org/10.1016/j.anbehav.2010.08.007

Auersperg, A. M. I., Huber, L., & Gajdon, G. K. (2011). Navigating a tool end in a specific direction: stick-tool use in kea (*Nestor notabilis*). *Biology Letters* (2005), 7(6), 825–828. https://doi.org/10.1098/rsbl.2011.0388

Auersperg, A. M. I., Köck, C., O'Hara, M., & Huber, L. (2018). Tool making cockatoos adjust the lengths but not the widths of their tools to function. *PloS One*, 13(11), e0205429–e0205429. https://doi.org/10.1371/journal.pone.0205429

Auersperg, A. M. I., Köck, C., Pledermann, A., O'Hara, M., & Huber, L. (2017). Safekeeping of tools in Goffin's cockatoos, *Cacatua goffiniana*. *Animal Behaviour*, 128, 125–133. https://doi.org/10.1016/j.anbehav.2017.04.010

Auersperg, A. M. I., Laumer, I. B., & Bugnyar, T. (2013). Goffin cockatoos wait for qualitative and quantitative gains but prefer 'better' to 'more.' *Biology Letters* (2005), 9(3), 20121092–20121092. https://doi.org/10.1098/rsbl.2012.1092

Auersperg, A. M. I., Szabo, B., von Bayern, A. M. P., & Kacelnik, A. (2012). Spontaneous innovation in tool manufacture and use in a Goffin's cockatoo. *Current Biology*, 22(21), R903–R904. https://doi.org/10.1016/j.cub.2012.09.002

Auersperg, A. M. I., van Horik, J. O., Bugnyar, T., Kacelnik, A., Emery, N. J., & von Bayern, A. M. P. (2015). Combinatory Actions During Object Play in Psittaciformes (*Diopsittaca nobilis, Pionites melanocephala, Cacatua goffini*) and Corvids (*Corvus corax, C. monedula, C. moneduloides*). *Journal of Comparative Psychology* (1983), 129(1), 62–71. https://doi.org/10.1037/a0038314

Auersperg, A. M. I., von Bayern, A. M. I., Weber, S., Szabadvari, A., Bugnyar, T., & Kacelnik, A. (2014). Social transmission of tool use and tool manufacture in Goffin cockatoos (*Cacatua goffini*). *Proceedings of the Royal Society. B, Biological Sciences*, 281(1793), 20140972–20140972. https://doi.org/10.1098/rspb.2014.0972

Auersperg, A. M. I., von Bayern, A. M. P., 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–e20231. https://doi.org/10.1371/journal.pone.0020231

Bania, A. E., Harris, S., Kinsley, H. R. & Boysen, S. T. (2009). Constructive and Deconstructive Tool Modification by Chimpanzees (*Pan troglodytes*). *Animal cognition*, 12(1), 85-95.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software,* 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., & Jagan, M. (2025). Matrix: Sparse and Dense Matrix Classes and Methods [Computer software]. https://CRAN.R-project.org/package=Matrix

Baum, M., Rössler, T., Osuna-Mascaró, A. J., Auersperg, A., & Brock, O. (2024). Mechanical Problem Solving in Goffin's Cockatoos—Towards Modeling Complex Behavior. *Adaptive Behavior*, 32(6), 551–562. https://doi.org/10.1177/10597123241270764

Beck, S. R., Apperly, I. A., Chappell, J., Guthrie, C., & Cutting, N. (2011). Making tools isn't child's play. *Cognition*, 119(2), 301–306. https://doi.org/10.1016/j.cognition.2011.01.003

Bluff, L. A., Troscianko, J., Weir, A. A. S., Kacelnik, A., & Rutz, C. (2010). Tool use by wild New Caledonian crows *Corvus moneduloides* at natural foraging sites. *Proceedings of the Royal Society B: Biological Sciences*, 277(1686), 1377–1385. https://doi.org/10.1098/rspb.2009.1953

Bugnyar, T. (2019). Tool use: New Caledonian crows engage in mental planning. *Current Biology*, *29*(6), R200-R202. https://doi.org/10.1016/j.cub.2019.01.059

Cespedes-Gonzalez, A., Osuna-Mascaro, A. J., O'Hara, M., Roessler, T., Hanon, L., & Auersperg, A. M. I. (2024). The Effect of Four Different Object Properties on Latency to Approach in Goffin's Cockatoos (*Cacatua goffiniana*). *Journal of Experimental Psychology. Animal Learning and Cognition*, 50(2), 131–143. https://doi.org/10.1037/xan0000373

Chappell, J., & Kacelnik, A. (2002). Tool selectivity in a non-primate, the New Caledonian crow (*Corvus moneduloides*). *Animal Cognition*, 5(2), 71–78. https://doi.org/10.1007/s10071-002-0130-2

Colbourne, J. A. D., Auersperg, A. M. I., Lambert, M. L., Huber, L., & Völter, C. J. (2021). Extending the Reach of Tooling Theory: A Neurocognitive and Phylogenetic Perspective. *Topics in Cognitive Science*, 13(4), 548–572. https://doi.org/10.1111/tops.12554

Cutting, N., Apperly, I. A., Chappell, J., & Beck, S. R. (2014). The puzzling difficulty of tool innovation: Why can't children piece their knowledge together? *Journal of Experimental Child Psychology*, 125(1), 110–117. https://doi.org/10.1016/j.jecp.2013.11.010

de Beaune, S. (2004). The Invention of Technology: Prehistory and Cognition. *Current Anthropology*, 45(2), 139–162. https://doi.org/10.1086/381045

Diamond, J., & Bond, A. (2003). A Comparative Analysis of Social Play in Birds. *Behaviour*, 140(8–9), 1091–1115. https://doi.org/10.1163/156853903322589650

Emery, N. J. (2006). Cognitive ornithology: the evolution of avian intelligence. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences*, 361(1465), 23–43. https://doi.org/10.1098/rstb.2005.1736

Emery, N. J. (2013). Insight, imagination and invention: Tool understanding in a non-tool-using corvid. In *Tool Use in Animals* (pp. 67–88). Cambridge University Press. https://doi.org/10.1017/CBO9780511894800.006

Fox, E., & Bin'Muhammad. (2002). Brief communication: New tool use by wild Sumatran orangutans (*Pongo pygmaeus abelii*). *American Journal of Physical Anthropology, 119*(2), 186–188. https://doi.org/10.1002/ajpa.10105

Fox, J., & Weisberg, S. (2019). An R companion to applied regression (3rd ed.). Sage.

Gajdon, G. K., Lichtnegger, M., & Huber, L. (2014). What a parrot's mind adds to play: the urge to produce novelty fosters tool use acquisition in kea. *Open Journal of Animal Sciences*, 4(2), 51-58. https://doi.org/10.4236/ojas.2014.42008

Goodman, M., Hayward, T., & Hunt, G. R. (2018). Habitual tool use innovated by free-living New Zealand kea. *Scientific Reports*, 8(1), 13935–12. https://doi.org/10.1038/s41598-018-32363-9

Grahl, W. de. (1985). Papageien: Lebensweise, Arten, Zucht (Völlig neubearb. u. neugestaltete 8. Aufl.). *Ulmer*.

Gruber, R., Schiestl, M., Boeckle, M., Frohnwieser, A., Miller, R., Gray, R. D., Clayton, N. S., & Taylor, A. H. (2019). New Caledonian Crows Use Mental Representations to Solve Metatool Problems. *Current Biology*, 29(4), 686-692.e3. https://doi.org/10.1016/j.cub.2019.01.008

Habl, C., & Auersperg, A. M. I. (2017). The keybox: Shape-frame fitting during tool use in Goffin's cockatoos (*Cacatua goffiniana*). *PLoS One*, 12(11), e0186859. https://doi.org/10.1371/journal.pone.0186859

Hansell, M., & Ruxton, G. D. (2008). Setting tool use within the context of animal construction behaviour. *Trends in Ecology & Evolution (Amsterdam)*, 23(2), 73–78. https://doi.org/10.1016/j.tree.2007.10.006

Hartig, F. (2024). DHARMa: Residual diagnostics for hierarchical (multi-level / mixed) regression models [Computer software]. https://CRAN.R-project.org/package=DHARMa

Hihara, S., Obayashi, S., Tanaka, M., & Iriki, A. (2003). Rapid learning of sequential tool use by macaque monkeys. *Physiology & behavior*, 78(3), 427-434. https://doi.org/10.1016/S0031-9384(02)01006-5

Hunt, G. R. (1996). Manufacture and use of hook-tools by New Caledonian crows. *Nature (London)*, 379(6562), 249–251. https://doi.org/10.1038/379249a0

Jacobs, I. F., Osvath, M., Osvath, H., Mioduszewska, B., von Bayern, A. M. P., & Kacelnik, A. (2014). Object caching in corvids: Incidence and significance. *Behavioural Processes*, 102, 25–32. https://doi.org/10.1016/j.beproc.2013.12.003

Jacobs, I. F., von Bayern, A. M. P. & Osvath, M. (2016). A novel tool-use mode in animals: New Caledonian crows insert tools to transport objects. *Animal cognition*, 19, 1249-1252. doi:10.1007/s10071-016-1016-z

Jones, T. B., & Kamil, A. C. (1973). Tool-Making and Tool-Using in the Northern Blue Jay. *Science (American Association for the Advancement of Science)*, 180(4090), 1076–1078. https://doi.org/10.1126/science.180.4090.1076

Kenward, B., Rutz, C., Weir, A. A. S., & Kacelnik, A. (2006). Development of tool use in New Caledonian crows: inherited action patterns and social influences. *Animal Behaviour*, 72(6), 1329–1343. https://doi.org/10.1016/j.anbehav.2006.04.007

Kenward, B., Weir, A. A. S., Rutz, C., & Kacelnik, A. (2005). Behavioural ecology: Tool manufacture by naive juvenile crows. *Nature (London)*, 433(7022), 121–121. https://doi.org/10.1038/433121a

Kleiber, C., & Zeileis, A. (2008). Applied econometrics with R. *Springer*.

Kohler, W. (2018). The mentality of apes. *Routledge*. https://doi.org/10.4324/9781351294966

Lambert, M. L., Jacobs, I., Osvath, M., & von Bayern, A. M. P. (2019). Birds of a feather? Parrot and corvid cognition compared. *Behaviour*, 156(5–8), 505–594. https://doi.org/10.1163/1568539X-00003527

Laumer, I. B., Bugnyar, T., & Auersperg, A. M. I. (2016). Flexible decision-making relative to reward quality and tool functionality in Goffin cockatoos (*Cacatua goffiniana*). *Scientific Reports*, 6(1), 28380–28380. https://doi.org/10.1038/srep28380

Laumer, I. B., Bugnyar, T., Reber, S. A., & Auersperg, A. M. I. (2017). Can hook-bending be let off the hook? Bending/unbending of pliant tools by cockatoos. *Proceedings of the Royal Society. B, Biological Sciences*, 284(1862), 20171026–20171026. https://doi.org/10.1098/rspb.2017.1026

Lenth, R. V. (2025). emmeans: Estimated marginal means, aka least-squares means [Computer software]. https://CRAN.R-project.org/package=emmeans

Lethmate, J. (1982). Tool-using skills of orang-utans. *Journal of Human Evolution*, 11(1), 49-64. https://doi.org/10.1016/S0047-2484(82)80031-6

Mangalam, M., & Fragaszy, D. M. (2016). Transforming the body-only system into the body-plus-tool system. *Animal Behaviour*, 117, 115–122. https://doi.org/10.1016/j.anbehav.2016.04.016

Mannu, M. & Ottoni, E. B. (2009). The Enhanced Tool-Kit of Two Groups of Wild Bearded Capuchin Monkeys in the Caatinga: Tool Making, Associative Use, and Secondary Tools. *American Journal of Primatology*, 71, 242–251. https://doi.org/10.1002/ajp.20642

Martin-Ordas, G., Schumacher, L., & Call, J. (2012). Sequential tool use in great apes. *PLoS One*, 7(12), e52074. https://doi.org/10.1371/journal.pone.0052074

McGillycuddy, E., Brooks, M. E., & Kristensen, K. (2025). glmmTMB: Generalized linear mixed models using template model builder [Computer software]. https://CRAN.R-project.org/package=glmmTMB

Mioduszewska, B., Auersperg, A. M. I., & O'Hara, M. (2022). Treasure islands: foraging ecology and the emergence of tool use in wild Goffin's cockatoos. *Current Opinion in Behavioral Sciences*, 45, 101118. https://doi.org/10.1016/j.cobeha.2022.101118

Mundry, R., Nunn, C. L., Michael C. Whitlock, & Emília P. Martins. (2009). Stepwise Model Fitting and Statistical Inference: Turning Noise into Signal Pollution. *The American Naturalist*, 173(1), 119–123. https://doi.org/10.1086/593303

Myowa-Yamakoshi, M., & Matsuzawa, T. (2000). Imitation of Intentional Manipulatory Actions in Chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology (1983)*, 114(4), 381–391. https://doi.org/10.1037/0735-7036.114.4.381

Nasa, Ö. (2024). Handling of a Disbalanced Stick Tool in the Goffin's Cockatoo (Master's Thesis). *University of Vienna*. https://doi.org/10.25365/thesis.77645

O'Hara, M., Mioduszewska, B., Mundry, R., Yohanna, Haryoko, T., Rachmatika, R., Prawiradilaga, D. M., Huber, L., & Auersperg, A. M. I. (2021). Wild Goffin's cockatoos flexibly manufacture and use tool sets. *Current Biology*, 31(20), 4512-4520.e6. https://doi.org/10.1016/j.cub.2021.08.009

Osuna-Mascaró, A. J., & Auersperg, A. M. I. (2018). On the brink of tool use? Could object combinations during foraging in a feral Goffin's cockatoo (*Cacatua goffiniana*) result in tool innovations. *Animal Behavior and Cognition*, 5(2), 229-234. https://doi.org/10.26451/abc.05.02.05.2018

Osuna-Mascaro, A. J., Mundry, R., Tebbich, S., Beck, S. R., & Auersperg, A. M. I. (2022). Innovative composite tool use by Goffin's cockatoos (*Cacatua goffiniana*). *Scientific reports*, 12, 1510. https://doi.org/10.1038/s41598-022-05529-9

Osuna-Mascaró, A. J., O'Hara, M., Folkertsma, R., Tebbich, S., Beck, S. R., & Auersperg, A. M. I. (2023). Flexible tool set transport in Goffin's cockatoos. *Current Biology*, 33(5), 849-857.e4. https://doi.org/10.1016/j.cub.2023.01.023

Pepperberg, I. M., & Shive, H. R. (2001). Simultaneous Development of Vocal and Physical Object Combinations by a Grey Parrot (*Psittacus erithacus*): Bottle Caps, Lids, and Labels. *Journal of Comparative Psychology* (1983), 115(4), 376–384. https://doi.org/10.1037/0735-7036.115.4.376

Price, E. E., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2009). A potent effect of observational learning on chimpanzee tool construction. *Proceedings of the Royal Society B: Biological Sciences*, 276(1671), 3377–3383. https://doi.org/10.1098/rspb.2009.0640

R Core Team. (2024). R: A language and environment for statistical computing (Version 4.4.2) [Computer software]. *R Foundation for Statistical Computing*. https://www.R-project.org/

Reindl, E., Tennie, C., Apperly, I. A., Lugosi, Z. & Beck, S. R. (2022). Young children spontaneously invent three different types of associative tool use behaviour. *Evolutionary Human Sciences*, 4, e5. doi:10.1017/ehs.2022.4

Rutz, C., & St Clair, J. J. (2012). The evolutionary origins and ecological context of tool use in New Caledonian crows. *Behavioural Processes*, 89(2), 153-165. https://doi.org/10.1016/j.beproc.2011.11.005

Rutz, C., & St Clair, J. J. H. (2012). The evolutionary origins and ecological context of tool use in New Caledonian crows. *Behavioural Processes*, 89(2), 153–165. https://doi.org/10.1016/j.beproc.2011.11.005

Schick, K. D., Toth, N., Garufi, G., Savage-Rumbaugh, E. S., Rumbaugh, D., & Sevcik, R. (1999). Continuing Investigations into the Stone Tool-making and Tool-using Capabilities of a Bonobo (*Pan paniscus*). *Journal of Archaeological Science*, 26(7), 821–832. https://doi.org/10.1006/jasc.1998.0350

Seed, A., Emery, N., & Clayton, N. (2009). Intelligence in Corvids and Apes: A Case of Convergent Evolution? *Ethology*, 115(5), 401–420. https://doi.org/10.1111/j.1439-0310.2009.01644.x

Szabo, B., Bugnyar, T., & Auersperg, A. M. I. (2017). Within-group relationships and lack of social enhancement during object manipulation in captive Goffin's cockatoos (*Cacatua goffiniana*). *Learning & Behavior*, 45(1), 7–19. https://doi.org/10.3758/s13420-016-0235-0

Taylor, A. H., Hunt, G. R., Holzhaider, J. C., & Gray, R. D. (2007). Spontaneous Metatool Use by New Caledonian Crows. *Current Biology*, 17(17), 1504–1507. https://doi.org/10.1016/j.cub.2007.07.057

Tebbich, S., Taborsky, M., Fessl, B., & Dvorak, M. (2002). The ecology of tool-use in the woodpecker finch (*Cactospiza pallida*). *Ecology Letters*, 5(5), 656–664. https://doi.org/10.1046/j.1461-0248.2002.00370.x

Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics with S (4th ed.). *Springer*.

Von Bayern, A. M. P., Danel, S., Auersperg, A. M. I., Mioduszewsja, B. & Kacelnik, A. (2018). Compound tool construction by New Caledonian crows. *Scientific reports,* 8(*1), 15676.* https://doi.org/10.1038/s41598-018-33458-z

Wadley, L. (2010). Compound-Adhesive Manufacture as a Behavioral Proxy for Complex Cognition in the Middle Stone Age. *Current Anthropology*, 51(S1), S111–S119. https://doi.org/10.1086/649836

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. *Springer*.

Wickham, H., & Bryan, J. (2023). readxl: Read Excel files [Computer software]. https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., & Müller, K. (2023). dplyr: A grammar of data manipulation [Computer software]. https://CRAN.R-project.org/package=dplyr

Wickham, H., Vaughan, D., & Girlich, M. (2024). tidyr: Tidy messy data [Computer software]. https://CRAN.R-project.org/package=tidyr

Wimpenny, J. H., Weir, A. A. S., Clayton, L., Rutz, C., & Kacelnik, A. (2009). Cognitive Processes Associated with Sequential Tool Use in New Caledonian Crows. *PloS One*, 4(8), e6471–e6471. https://doi.org/10.1371/journal.pone.0006471

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. https://doi.org/10.1371/journal.pone.0006471

Wood, P. K. (1983). Inquiring systems and problem structure: Implications for cognitive development. Human development, 26(5), 249-265. https://doi.org/10.1159/000272887

Yellen, J. E., Brooks, A. S., Cornelissen, E., Mehlman, M. J., & Stewart, K. (1995). A Middle Stone Age worked bone industry from Katanda, upper Semliki Valley, Zaire. *Science (American Association for the Advancement of Science)*, 268(5210), 553–556. https://doi.org/10.1126/science.7725100

Appendix

Participants

Table 3: Participating subjects and their sex are listed here. Olympia and Muki had to be excluded prior to the medium-distance test. Konrad was excluded in the pre-combined test. During the last phase of the experiment, the pre-combined test, only ten individuals remained.

Subject	**Sex**
Dolittle	male
Figaro	male
Fini	female
Heidi	female
Kiwi	male
Konrad	male
Mayday	female
Moneypenny	female
Muki	male
Muppet	male
Olympia	female
Pipin	male
Zozo	male

Model outputs

1A. Preference test: Tool choice – Duration:

Family: tweedie (log)
Formula:
Touch_duration ~ Tool_type + Session + Gender + (1 | Individual)
Zero inflation: ~1

AIC	BIC	logLik	2*log(L)	df.resid
432,4	450,3	-208,2	416,4	62

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	3.676e-09	6.063e-05

Number of obs: 70, groups: Individual, 13
Dispersion parameter for tweedie family (): 3.89

	Estimate	Std. Error	z-value	p-value
Intercept	1,289	0,684	1,884	0,060
Cone	0,717	0,360	1,993	0,046 *
Session	0,437	0,216	2,029	0,042 *
Sex_m	1,232	0,397	3,102	0,002 *

Zero-inflation model:

Estimate	Std. Error	z-value	p-value
-0,389	0,424	-0,191	0,358

1B. Preference test: Tool choice – Frequency:

Family: nbinom2 (log)
Formula: Total_count ~ Tool_type + Touched_first + Session + Gender + (1|Subject)

AIC	BIC	logLik	2*log(L)	df.resid
335,7	349,6	-160,8	321,7	47

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	0,252	0,502

Number of obs: 54, groups: Subject, 13
Dispersion parameter for nbinom2 family (): 1.86

	Estimate	Std. Error	z-value	p-value
Intercept	1,935	0,438	4,417	0,00001 *
Cone	-0,207	0,232	-0,892	0,372
TouchedFirst_Cone	-0,658	0,309	-2,131	0,033 *
Session	0,087	0,139	0,630	0,529
Sex_m	0,031	0,381	0,080	0,931

2A. Short-distance test: Tool choice – Duration:

Family: tweedie (log)
Formula:
Touch_duration ~ Tool_type + Session + Gender + (1 | Individual)

AIC	BIC	logLik	2*log(L)	df.resid
687,3	703,2	-336,6	673,3	65

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	0,3533	0,5944

Number of obs: 72, groups: Individual, 13
Dispersion parameter for tweedie family (): 7,76

	Estimate	Std. Error	z-value	p-value
Intercept	2,609	0,514	5,079	3.8e-07 *
Cone	1,027	0,291	3,529	0.0004 *
Session	0,310	0,151	2,049	0.040 *
Sex_m	0,371	0,449	0,828	0.408 *

2B. Short-distance test: Tool choice – Frequency:

Family: nbinom1 (log)
Formula: Total_count ~ Tool_type + Touched_first + Session + Gender + (1|Subject)

AIC	BIC	logLik	2*log(L)	df.resid
333,4	346,5	-159,7	319,4	41

Random effects:

Groups	Name	Variance	Std.Dev.
Subject	Intercept	0,07813	0,2795

Number of obs: 48, groups: Subject, 13
Dispersion parameter for nbinom1 family (): 4,2

	Estimate	Std. Error	z-value	p-value
Intercept	2,380	0,356	6,683	2.35e-11 *
Cone	0,788	0,200	3,945	7.97e-05 *
TouchedFirst_Cone	-0,119	0,226	-0,526	0,599
Session	-0,061	0,055	-1,123	0,261
Sex_m	-0,253	0,242	-1,043	0,297

3A. Medium-distance test: Tool choice – Duration:

Family: tweedie (log)
Formula: Touch_duration ~ Tool_type + Session + Gender + (1 | Individual)

AIC	BIC	logLik	2*log(L)	df.resid
2192,6	2214,7	-1088,8	2177,6	195

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	0,04048	0,2012

Number of obs: 202, groups: Individual, 11
Dispersion parameter for tweedie family (): 4,28

	Estimate	Std. Error	z-value	p-value
Intercept	4,258	0,194	21,944	< 2e-16 *
Cone	0,806	0,126	6,379	1.79e-10 *
Session	-0,042	0,021	-2,022	0,0432 *
Sex_m	-0,016	0,178	-0,089	0,929

3B. Medium-distance test: Tool choice – Frequency:

Family: nbinom1 (log)
Formula: Total_count ~ Tool_type + Touched_first + Session + Gender + (1|Subject)

AIC	BIC	logLik	2*log(L)	df.resid
1447,5	1470,7	-716,7	1433,5	197

Random effects:

Groups	Name	Variance	Std.Dev.
Subject	Intercept	0,07831	0,2798

Number of obs: 204, groups: Subject, 11
Dispersion parameter for nbinom1 family (): 5,87

	Estimate	Std. Error	z-value	p-value
Intercept	2,563	0,197	13,011	< 2e-16 *
Cone	0,565	0,096	5,868	4.41e-09 *
TouchedFirst_Cone	-0,114	0,104	-1,101	0,271
Session	-0,022	0,016	-1,319	0,187
Sex_m	-0,140	0,201	-0,700	0,484

3C. Combination attempts: Duration

Family: tweedie (log)
Formula: Touch_duration ~ Combination + Session + Gender + (1 | Individual)
Zero inflation: ~1

AIC	BIC	logLik	2*log(L)	df.resid
1522,2	1562,3	-751,1	1502,2	398

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	1,999	1,414

Number of obs: 408, groups: Individual, 11
Dispersion parameter for tweedie family (): 1,75

	Estimate	Std. Error	z-value	p-value
Intercept	-0,459	0,733	-0,627	0,531
B-to-B	-0,476	0,218	-2,180	0,0293 *
B-to-C	1,876	0,167	11,208	< 2e-16 *
C-to-B	1,189	0,178	6,662	2.69e-11 *
Session	-0,031	0,019	-1,674	0,100
Sex_m	-0,218	0,911	-0,240	0,811

Zero-inflation model:

Estimate	Std. Error	z-value	p-value
-20,330	3030,640	-0,007	0,995

3C.1. Tukey-adjusted pairwise comparisons of estimated marginal means

Com. attempt	emmean	SE	df	asymp. L CL	asymp. U CL
C-to-C	-0,739	0,474	Inf	-1,668	0,191
B-to-B	-1,215	0,483	Inf	-2,162	-0,268
B-to-C	1,137	0,462	Inf	0,231	2,043
C-to-B	0,45	0,467	Inf	-0,465	1,365

Results are averaged over the levels of: Gender
Results are given on the log (not the response) scale.
Confidence level used: 0.95

contrast	estimate	SE	df	z-ratio	p-value
C-to-C - B-to-B	0,476	0,218	Inf	2,18	0,129
C-to-C - B-to-C	-1,876	0,167	Inf	-11,208	< 0,001
C-to-C - C-to-B	-1,189	0,178	Inf	-6,662	< 0,001
B-to-B - B-to-C	-2,351	0,186	Inf	-12,625	< 0,001
B-to-B - C-to-B	-1,665	0,19	Inf	-8,768	< 0,001
B-to-B - C-to-B	0,687	0,137	Inf	5,015	< 0,001

Results are averaged over the levels of: Gender
Results are given on the log (not the response) scale.

3D. Combination attempts: Freqeuncy

Family: nbinom2 (log)
Formula: Total_occurences ~ Behavior + Gender + Session + (1 | Subject)

AIC	BIC	logLik	2*log(L)	df.resid
804,9	836,9	-394,4	788,9	396

Random effects:

Groups	Name	Variance	Std.Dev.
Subject	Intercept	0,3377	0,5811

Number of obs: 404, groups: Subject, 11
Dispersion parameter for nbinom2 family (): 0,745

	Estimate	Std. Error	z-value	p-value
Intercept	-0,135	0,390	-0,345	0,730
B-to-B	-1,764	0,322	-5,484	4.16e-08 *
B-to-C	-0,132	0,232	-0,568	0,570
C-to-B	-0,755	0,253	-2,979	0,0029 *
Session	0,004	0,032	0,115	0,909
Sex_m	-0,163	0,420	-0,389	0,697

3E. Two tools simultaneously: Duration

Family: tweedie (log)
Formula: Two_tools_duration ~ Session + Gender + (1 |Individual)
Zero inflation: ~1

AIC	BIC	logLik	2*log(L)	df.resid
485,2	503,5	-235,6	471,2	94

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	0,499	0,7064

Number of obs: 101, groups: Individual, 11
Dispersion parameter for tweedie family (): 2,74

	Estimate	Std. Error	z-value	p-value
Intercept	1,805	0,493	3,659	0,000253 *
Session	-0,090	0,050	-1,812	0.070048
Sex_m	-0,795	0,535	-1,485	0,138

Zero-inflation model:

Estimate	Std. Error	z-value	p-value
-19,510	6707,400	-0,003	0,998

3F. Two tools simultaneously: Frequency

Family: nbinom2 (log)
Formula: Total_occurences ~ Gender + Session + (1 | Subject)

AIC	BIC	logLik	2*log(L)	df.resid
348,2	361,3	-169,1	338,2	96

Random effects:

Groups	Name	Variance	Std.Dev.
Subject	Intercept	0,1558	0,3947

Number of obs: 101, groups: Subject, 11
Dispersion parameter for nbinom2 family (): 3,07

	Estimate	Std. Error	z-value	p-value
Intercept	0,718	0,313	2,296	0,022 *
Session	-0,021	0,034	-0,622	0,534
Sex_m	-0,383	0,325	-1,181	0,238

4A. Pre-combined test: Tool choice – Duration

Family: tweedie (log)
Formula: Touch_duration ~ Tool_type + Session + Gender + (1 | Individual)

AIC	BIC	logLik	2*log(L)	df.resid
1665,7	1691,6	-825,9	1651,7	291

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	0,2785	0,5278

Number of obs: 298, groups: Individual, 10
Dispersion parameter for tweedie family (): 4,74

	Estimate	Std. Error	z-value	p-value
Intercept	0,367	0,351	1,044	0,297
CC	3,700	0,205	18,050	< 2e-16 *
Session	-0,060	0,015	-4,075	4.61e-05 *
Sex_m	-0,379	0,365	-1,037	0,300

4B. Pre-combined test: Tool choice – Frequency

Family: nbinom2 (log)
Formula: Total_count ~ Tool_type + Touched_first + Session + Gender + (1 + Session || Subject)

AIC	BIC	logLik	2*log(L)	df.resid
870,7	900,2	-427,3	854,7	290

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
Subject	Intercept	1.730e-01	0.415975	
	Session	3.072e-05	0.005543	0

Number of obs: 298, groups: Subject, 11
Dispersion parameter for nbinom2 family (): 2,05

	Estimate	Std. Error	z-value	p-value
Intercept	0,347	0,358	0,970	0,332
CC	2,290	0,180	12,673	< 2e-16 *
TouchedFirst_CC	-1,007	0,286	-3,522	0,0004 *
Session	-0,061	0,017	-3,609	0,0003 *
Sex_m	-0,380	0,305	-1,246	0,213

4C. Pre-combined test: Unsuccessful trials

Family: binomial (logit)

Formula: Nr_Unsuccessful_trial ~ BB_touched_first + Session + Gender + (1 | Individual)

AIC	BIC	logLik	2*log(L)	df.resid
84,9	99,9	-37,4	74,9	144

Scaled residuals:

Min	1Q	Median	3Q	Max
-1,2038	-0,2551	-0,1867	-0,1304	5,5883

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	1,501	1,225

Number of obs: 149, groups: Individual, 10

Fixed effects:

	Estimate	Std. Error	z-value	p-value
Intercept	-3,077	1,160	-2,653	0,008 *
BB_touched_first	0,093	1,027	0,091	0,928
Session	-0,112	0,083	-1,338	0,181
Sex_m	1,381	1,213	1,139	0,255

Correlation of Fixed Effects:

	Intercept	BB_first	Session
BB_first	-0,173		
Session	-0,423	0,325	
Sex_m	-0,683	-0,149	-0,084

4D. Pre-combined test: Successful trials

Family: binomial (logit)
Formula: PreCom_success ~ CC_touched_first + Session + Gender + (1 | Individual)

AIC	BIC	logLik	2*log(L)	df.resid
84,9	99,9	-37,4	74,9	144

Scaled residuals:

Min	1Q	Median	3Q	Max
-5,5883	0,1304	0,1867	0,2551	1,2038

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	1,501	1,225

Number of obs: 149, groups: Individual, 10

Fixed effects:

	Estimate	Std. Error	z-value	p-value
Intercept	2,984	1,410	2,116	0,034 *
CC_touched_first	0,093	1,027	0,091	0,928
Session	0,112	0,083	1,338	0,181
Sex_m	-1,381	1,213	-1,139	0,255

Correlation of Fixed Effects:

	Intercept	CC_first	Session
CC_first	-0,586		
Session	-0,111	-0,325	
Sex_m	-0,671	0,149	-0,084

4E. Pre-combined test: Time to success

Linear mixed model fit by REML ['lmerMod']
Formula: time_to_success_log ~ Session + Gender.x + (1 | Individual)
REML criterion at convergence: 322,9

Scaled residuals:

Min	1Q	Median	3Q	Max
-2,4331	-0,6353	-0,0626	0,4876	2,5231

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	0,1252	0,3539
Residual		0,5821	0,7629

Number of obs: 136, groups: Individual, 10

Fixed effects:

	Estimate	Std. Error	t-value
Intercept	4,562	0,31	14,723
Session2	-0,979	0,378	-2,589
Session3	-0,666	0,351	-1,896
Session4	-0,719	0,351	-2,047
Session5	-0,469	0,351	-1,335
Session6	-1,296	0,341	-3,799
Session7	-1,504	0,351	-4,279
Session8	-1,614	0,351	-4,593
Session9	-1,332	0,351	-3,794
Session10	-0,933	0,351	-2,656
Session11	-1,216	0,363	-3,346
Session12	-1,474	0,351	-4,194
Session13	-1,118	0,351	-3,184
Session14	-1,084	0,341	-3,177
Session15	-1,194	0,341	-3,498
Sex_m	-0,295	0,265	-1,113

5A. Cone preference: Duration

Family: beta (logit)
Formula: cone_ratio_time_new ~ Test + Gender + (1 | Subject)

AIC	BIC	logLik	2*log(L)	df.resid
-222,4	-204,2	117,2	-234,4	147

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	0,4832	0,6951

Number of obs: 153, groups: Subject, 13
Dispersion parameter for beta family (): 1.44

	Estimate	Std. Error	z-value	p-value
Intercept	0,720	0,430	1,677	0,094
Medium-distance	0,622	0,298	2,089	0,037 *
Short-distance	0,148	0,337	0,438	0,661
Sex_m	-0,916	0,452	-2,026	0,043 *

5B. Cone preference: Frequency

Family: beta (logit)
Formula:
cone_ratio_new ~ Test + Gender + (1 | Subject)

AIC	BIC	logLik	2*log(L)	df.resid
-106,5	-88,4	59,3	-118,5	147

Random effects:

Groups	Name	Variance	Std.Dev.
Individual	Intercept	0,3178	0,5637

Number of obs: 153, groups: Subject, 13
Dispersion parameter for beta family (): 1.76

	Estimate	Std. Error	z-value	p-value
Intercept	0,200	0,371	0,539	0,590
Medium-distance	0,997	0,282	3,541	0,0004 *
Short-distance	0,359	0,324	1,110	0,267
Sex_m	-0,769	0,382	-2,014	0,044 *