Received: 5 May 2022 | Revised: 23 November 2022 | Accepted: 1 December 2022

DOI: 10.1111/eth.13351

RESEARCH ARTICLE

ethology WILEY

Ratcheting up tool innovation in Goffin's cockatoos (*Cacatua goffiniana*): The effect of contextually diverse prior experience

Paula Ibáñez de Aldecoa[1] | Alice M. I. Auersperg[2] | Andrea S. Griffin[3] | Sabine Tebbich[1]

[1]Department of Behavioural and Cognitive Biology, University of Vienna, Vienna, Austria

[2]Messerli Research Institute, University of Veterinary Medicine Vienna, University of Vienna, Medical University of Vienna, Vienna, Austria

[3]School of Psychology, University of Newcastle, Callaghan, New South Wales, Australia

Correspondence
Paula Ibáñez de Aldecoa, Department of Behavioural and Cognitive Biology, Biologiezentrum University of Vienna, Schlachthausgasse 43, 1030 Vienna, Austria.
Email: paula.i.aldecoa@univie.ac.at

Funding information
Austrian Science Fund (FWF), Grant/Award Number: Y01309 and W1262-B29; WWTF, Grant/Award Number: CS18 023

Editor: Christian Rutz

Abstract

The ability to gain information from one situation, acquire new skills and/or perfect existing ones, and subsequently apply them to a new situation is a key element in behavioural flexibility and a hallmark of innovation. A flexible agent is expected to store these skills and apply them to contexts different from that in which learning occurred. Goffin's cockatoos (*Cacatua goffiniana*) are highly innovative parrots renowned for their problem-solving and tool-using skills and are thus excellent candidates to study this phenomenon. We hypothesized that birds allowed to use a tool in a larger variety of contingencies would acquire a broader expertise in handling it, facilitating its transfer to new tasks. In our study, we compared the performance of two groups of captive Goffin's cockatoos ($N = 13$): A test group received more diverse learning and motor experiences on multiple applications of a hook-type tool, while a control group received intensive, total trial-matched, experience with a single application of the same tool. Then, both groups were tested on two novel tasks to determine whether experience with the tool in multiple contexts would facilitate performance during transfer. While both groups transferred to both novel tasks, group differences in performance were apparent, particularly in the second transfer task, where test birds achieved a higher success rate and reached criteria within fewer trials than control birds. These results provide support for the prediction that experiencing a diverse range of contingencies with a tool appears to allow birds to acquire generalizable knowledge and transferrable skills to tackle an untrained problem more efficiently. In contrast, intensive experience with the tool in a single context might have made control birds less flexible and more fixated on previously learned tool-dependent instances.

KEYWORDS
behavioural flexibility, parrot cognition, problem-solving, tool use transfer

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2022 The Authors. *Ethology* published by Wiley-VCH GmbH.

1 | INTRODUCTION

The ability to acquire information and/or skills from previous experiences and then apply them to new situations is a key element of behavioural flexibility and it is considered to be a hallmark of innovation. While innovating, animals sometimes use new techniques to tackle novel challenges, whereas in other cases they may apply a known or a similar approach to a new purpose (Reader & Laland, 2003). Across the different phases of the innovation process, from discovering an opportunity to establishing an interaction with it, agents can potentially acquire knowledge of varying nature about that opportunity. This knowledge can facilitate solutions to new problems by sorting and applying memory traces of similar experiences (Bobrowicz et al., 2020; Diamond & Bond, 1999; Dukas, 2004; Kamil, 1988; Reader & Laland, 2003; Tebbich et al., 2016). The more general this knowledge is (i.e. less directly connected to a specific opportunity), the larger its scope to be expanded to further unexploited opportunities (Dienes & Perner, 1999). Thus, transfer of knowledge and/or skills (by re-exploiting a learnt tactic) can be an efficient way to reduce the costs of innovating (e.g. energy and time invested in locating and inspecting a food source; Klump et al., 2015), eventually allowing innovations to extend beyond the initial context (Tebbich et al., 2016). In this study, we aimed to manipulate the diversity of experiences that subjects had prior to encountering a novel problem to test how the resulting variation in knowledge (specific or broad) may affect performance during transfer.

Tool use tasks seem particularly suited to this purpose, as the same tools can often be applied in multiple contexts: innovations that incorporate objects (and tools in particular) can potentially have a wider applicability (Tebbich et al., 2016). The concept of "tooling" was proposed by Fragaszy and Mangalam (2018) to delineate the different stages of tool use and the tool-object interactions they entail. While tooling, toolers (1) perceive the prospect of producing an effect upon a target with a grasped object, (2) transform their body into a body + object system and (3) create a mechanical interface with said target by establishing spatial relation(s) between the target and the tool. As a result, when handling a tool to solve a problem, animals can acquire knowledge about its functional features and affordances such as its shape, size and weight, by incorporating actions with objects and managing the tool–target relations established (Auersperg, Gajdon, & von Bayern, 2012; Call, 2012; Diamond & Bond, 1999; Fragaszy & Mangalam, 2018; Griffin & Guez, 2014). They could then potentially use this information to assist in future novel problems and innovative foraging that require the use of tools (Auersperg, 2015; Diamond & Bond, 1999; Lambert et al., 2017; Taylor, Hunt, et al., 2009; Taylor, Roberts, et al., 2009). Transfer of such acquired tool-related proficiency also requires behavioural adjustments, such as refining the motor output and combinations of different motor actions (Griffin et al., 2014; Griffin & Guez, 2014; Guez & Griffin, 2016; Tebbich et al., 2016). Therefore, transfer in tool-using contexts can be studied particularly well in species with a high capacity to innovate in technically challenging situations and who have a wide motor repertoire.

We tested Goffin's cockatoos (*Cacatua goffiniana*, henceforth Goffins), a species that deploys numerous foraging innovations and which is renowned for its extraordinary examples of flexible tool use in captivity (Auersperg et al., 2013; Auersperg, Gajdon, & von Bayern, 2012; Auersperg, Szabo, et al., 2014; Auersperg, van Horik, et al., 2014; Auersperg, von Bayern, et al., 2014) and, as we recently learned, also in the wild (O'Hara et al., 2019, 2021; Osuna-Mascaró & Auersperg, 2018). This trait seems to be facilitated by their beak–foot motor coordination to hold and control a tool, which, especially in parrots, permits a great diversity of manipulation alternatives (Mettke-Hofmann et al., 2002). Their feeding opportunism and extractive foraging behaviour in the wild may make it adaptive to innovate based on experience when faced with similar but slightly different problems. A recent study (O'Hara et al., 2021) showed that Goffins flexibly implement the use of a tool depending on the characteristics of the target encountered, a complex technique that requires careful consideration of two different concurrent spatial relations: holding the food with the foot and orienting the tool for insertion with the beak. New evidence also indicates that Goffins can selectively recombine relevant previous experiences to solve novel problems that partially match (either functionally or perceptually), by relying on flexible memory functions (Bobrowicz et al., 2021). This faculty had already been demonstrated in great apes: Ebel and Call (2018) found that gorillas (*Gorilla gorilla*), bonobos (*Pan paniscus*), chimpanzees (*Pan troglodytes*) and orangutans (*Pongo abelii*) exposed to an empty apparatus before the test (with a baited apparatus) gained information that was later used to solve a problem-solving task more efficiently, while Bobrowicz et al. (2020) showed that chimpanzees and orangutans attend to relevant aspects of a problem and draw on those past experiences to assist current problem-solving.

Here, we investigated whether previously acquired experiences about a tool and its possible applications can be transferred to solve a new, untrained problem. For this, we provided birds with a single tool (a hook) that could be used for solving different tasks by adapting a combination of learned motor patterns. In phase 1 of the experiment, a test group was given experience with three tasks where the tool was freely moving. Each task required a different pulling motion with the tool, however. Each one also required motor coordinating a different tool–target interaction: (1) inserting the hook through a handle from above and pulling *upwards* (basket task), (2) inserting the hook through a window from the front and pulling horizontally *towards the self* (skate task) and (3) inserting the hook into a vertically oriented ring from below and pulling *downwards* (trap task; see Figure 1a–c). In contrast, the control group received intensive, total trial-matched experience with the same tool in only one (basket/pull *up*) task. Using a similar methodological approach to ours (by controlling and varying pretesting experience but focusing on how groups differed in their knowledge of the task rather than in that of the tool), von Bayern et al. (2009) found that New Caledonian crows (*Corvus moneduloides*) can innovate after acquaintance with the functional properties of a task. Similarly, a study in great-tailed grackles (*Quiscalus mexicanus*) suggests that behavioural flexibility

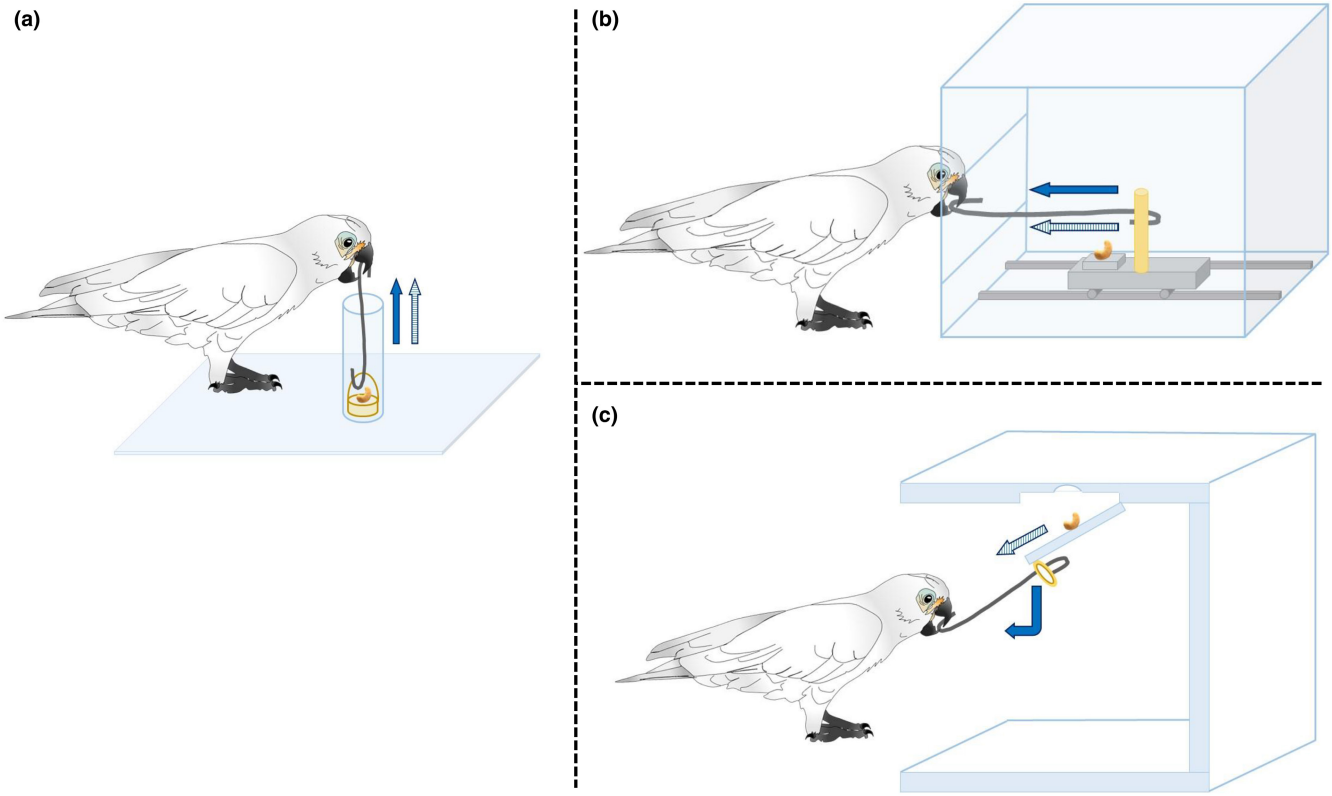

FIGURE 1 Apparatuses used for phase 1. (a) Basket: The reward is obtained by probing the hook through the transparent tube and inserting it through the basket's handle, then lifting it up. (b) Skate: The reward is obtained by passing the hook through the front window (left side of the drawing) and grabbing the protruding stick on the skate with the hook, then pulling it out. (c) Trap: The reward is obtained by inclining the hook upwards and grasping the ring with it, then pulling the trap down to open. Elements in yellow indicate target in the apparatus. Dark-coloured arrows denote the movement displayed by the bird and light-coloured arrows reflect the directionality of the reward following bird's movement.

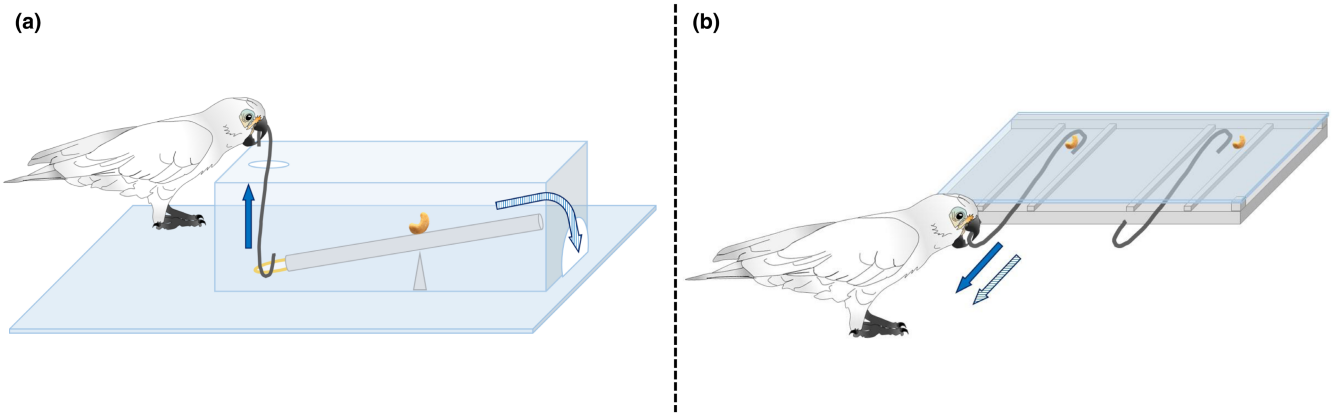

FIGURE 2 Apparatuses used for phase 2. (a) Seesaw task: The reward is obtained by actively inserting the hook through a hole located on the top of the box, grasping a lever (target element, coloured in yellow) at the end of the seesaw and pulling it upwards. (b) Cane task: The reward is obtained by pulling the pre-inserted functional hook towards the bird. Dark-coloured arrows denote the movement displayed by the bird and light-coloured arrows reflect the directionality of the reward following bird's movement.

can be shaped through experimental manipulation using reversal learning, and lead to an improvement of problem-solving in a new context, namely a multi-access box (Logan et al., 2022, preprint).

In phase 2 of the experiment, we measured the effect of the aforementioned differential training on performance in a new task, the seesaw (see Figure 2a). Two main features made this task novel: (1) the tool–target interaction differed from that of any apparatus presented in phase 1 (the hook was used to hoist a horizontally oriented lever located on the side of the seesaw), and (2) lifting the seesaw up caused the reward to slide away from, rather than towards, the self and be delivered at the opposite side of the apparatus. After the birds had completed testing with the seesaw

task (while data collection was still on going and prior to analysis), it become apparent that, despite the very distinct features mentioned above, inserting the tool from above and pulling it upwards were actions similar to those used in the basket task, with which both groups had had prior experience. Therefore, we decided to conduct a follow-up experiment by testing the birds on a second task that was perceptually even more different, so as to rule out the role of the motor actions. In our second transfer task (the cane; see Figure 2b), the contribution of the motor component was strongly reduced, as the tool was not freely manoeuvrable (i.e. it was already pre-inserted in the apparatus). In addition, in contrast to the seesaw, using the tool made direct contact with the reward (without mediation of a target). Hence, the birds had only to pay attention to the functional relationship between the tool and the reward rather than negotiate a tool to target contact to indirectly retrieve the reward. Furthermore, prior work with the cane task (Darwin's finches: Teschke et al., 2011) and other similar transfer tasks (Corvids: Bird & Emery, 2009; Emery & Clayton, 2009) allowed for performance of Goffins to be compared with other bird taxa.

We predicted that providing the test group with the opportunity to use the tool in a larger variety of contingencies in phase 1 (three different problem-solving situations as opposed to one) would allow them to explore the affordances of the hook and acquire more diverse expertise in handling it, increasing their tool-related knowledge and ultimately facilitating a more skilful solving competency in phase 2 (Griffin et al., 2013; Griffin & Guez, 2014; von Bayern et al., 2009).

2 | MATERIALS AND METHODS

2.1 | Subjects

Five female and eight male adult captive-reared Goffin's cockatoos (6–10 years of age) participated in the study. Subjects were assigned to a group at random by using the *sample* function in R, after establishing an equal sex distribution as a grouping criterion. The composition of the groups, once balancing the number of males and females in each, was: Control ($n = 6$): two females (Olympia, Moneypenny) and four males (Kiwi, Konrad, Muki, Zozo); and Test ($n = 7$): three females (Fini, Heidi, Mayday) and four males (Dolittle, Figaro, Muppet, Pipin). Although age was not considered when forming the groups, age differences (values at the time of testing, in 2019) were not significant (Test = 9.85 ± 1.78; Control = 8.83 ± 0.41; $t = 0.246$; $p = .813$). For additional subjects' information and details on housing conditions see Table S1. All subjects received a diet consisting of a source of carbohydrates mixed with fruits, bird food pellets, powdered vitamin supplements and seeds, as well as ad libitum drinking water. Participation was voluntary and birds were not food deprived prior to the experiment. Subjects were in semi-free housing conditions, with access to indoor and outdoor aviaries, without social isolation but visually isolated during the testing sessions, when the experimental subject was briefly separated from the group to avoid social learning. None of the toys provided as environmental enrichment resembled the experimental apparatuses or the tool. Data collection took place between March and September 2019 at the Goffin Lab Goldegg, located in Lower Austria.

2.1.1 | Experimental histories and STRANGEness of our test sample

To acknowledge the potential influence of subjects past experience in our study, following the recommendations made by the STRANGE framework (Webster & Rutz, 2020), we declare that all birds had previously participated in a number of experiments on tool use and other technical problem-solving tasks involving the use of acrylic glass apparatuses (for some examples see Auersperg, Szabo, et al., 2012; Habl & Auersperg, 2017; Laumer et al., 2016, 2017). However, although birds had a different hatching date (see Table S1), tool use experiments in the Goffin Lab only started in 2012. This means that age does not equate to experimental history, since all subjects started taking part in experiments at the same time; hence, older birds did not necessarily accumulate more testing experience. All tasks (with the exception of the basket task, which was previously used in Laumer et al., 2017) were novel to all individuals of both groups. In that study, which used a bendable hook-shaped tool (a pipe cleaner), only Fini and Moneypenny learned to consistently make and use this type of tool.

2.2 | Apparatuses and procedures

All tasks required the use of a hooked tool to gain access to the reward. At the beginning of the experiment, all birds had equal prior experience in using the tool, as it was used in the training stage of a previous study published by our group (Laumer et al., 2017). The tool was a 7-cm-long S-shaped hook made of rigid stainless steel, a material that ensured that no modifications could be made by the birds. The food reward (a 1/8th piece of a cashew nut, which birds only received when participating in the experiment) was visible yet unreachable without the tool, and it was located inside one of five apparatuses made of 0.5-mm thick Plexiglas and plywood. Apparatuses were always baited out of the bird's sight. At the beginning of the experiment, there was a 2-day habituation period for each task separately. First, habituation was undertaken in a group setting by placing the unbaited apparatus without the tool in the aviary until all birds had approached it voluntarily. This habituation phase ended when all subjects had contacted the apparatus or fed in its proximity at least once. Next, habituation was undertaken in an individual setting by presenting each subject with the unbaited apparatus without a tool and allowing the bird to inspect it. In both settings (collective and individual), subjects were encouraged to approach the apparatus by spreading sunflower seeds around it. Habituation to the hook was also conducted prior to testing: each bird had to pick up the tool

TABLE 1 Experiment timeline for the test group (top three rows) and control group (bottom row) in phase 1 (left column) and phase 2 (right column)

	Phase 1				Phase 2
Test group	**Order 1** (Heidi, Mayday)	*Hab–train - test* **Basket** →	*Hab–train - test* **Skate** →	*Hab–train - test* **Trap**	*Test (no training)* **Seesaw** → *Test (no training)* **Cane**
Test group	**Order 2** (Dolittle, Muppet)	*Hab–train - test* **Trap** →	*Hab–train - test* **Basket** →	*Hab–train - test* **Skate**	
Test group	**Order 3** (Figaro, Fini, Pipin)	*Hab–train - test* **Skate** →	*Hab–train - test* **Trap** →	*Hab–train - test* **Basket**	
Control group	**No order** (Kiwi, Konrad, Muki, Moneypenny, Olympia, Zozo)	*Train - test (+ extra time)* **Basket**			Control and Test group

Note: Phase 1: for the Test group, the three possible orders of presentation are shown (order 1, 2 and 3), indicating between brackets the names of the subjects tested on each order. For each of the three tasks (basket, skate and trap), the experiment sequence was habituation ("*hab*") > experience or training ("*train*") > testing ("*test*") > extra time (applicable only for some subjects of the control group). Phase 2: all subjects were equally tested on the seesaw task and the cane task; unlike in phase 1, in phase 2 testing commenced directly, without previous training.

from the table and exchange it for a seed with the experimenter on five consecutive trials during three sessions.

In the experimental sessions, the experimenter remained in the testing room, wearing mirrored sunglasses and looking down with hands off the table to avoid interfering with the bird's performance or providing any indirect cues. Testing consisted of two phases: in phase 1 a test group and a control group received different experiences and in phase 2 both groups received two new tasks. Throughout both phases, all birds were tested for two sessions a week, where the order of subject's testing was randomised within sessions to prevent the same individuals from being tested always at the same time of the day. All trials were video recorded (HD-Camcorder HC-V160, Panasonic), and latencies were measured using a stopwatch.

2.2.1 | Phase 1

Phase 1 differed between groups: while birds in the test group gained a broader diversity of experiences on three different tasks (namely the basket, the skate and the trap), control group received intensive experience on the basket task only. To solve the problem and obtain the reward, each task required birds to perform a motor sequence that also differed in the direction of pulling with respect to gravity: in the basket, birds had to pull the hook *upwards* (see Figure 1a), whereas in the skate they had to pull the hook *towards the self* (see Figure 1b). Finally, the trap required birds to pull the hook *downwards* (Figure 1c). Similarly, the tool–target interaction differed for each task: lift a handle upwards in the basket, drag a mobile miniwagon horizontally in the skate and pull a ring downwards in the trap. Training with these tasks in phase 1 could enhance flexibility for transfer in phase 2 because the motor action needed and the position of the bird's body with respect to the tool were different for each apparatus; rapid motor adjustments were required to decouple the learnt pattern from a specific problem and permit its flexible adaptation to a new task (Fragaszy & Mangalam, 2018; Griffin et al., 2014): from leaning the head down in the basket, to bending the upper body forward in the skate or stretching the neck and tilting the head slightly up in the trap (see Figure 1a–c). The order of presentation of these three tasks was randomized across subjects in the test group as shown in Table 1.

Phase 1 consisted of the aforementioned habituation followed by a training stage and a testing stage (12 trials per session). A detailed summary of the testing scheme is shown in Table 1. During training and testing, birds were given an unlimited amount of time per trial to solve the task, but if they did not interact with the tool or the apparatus for 15 min, the trial was terminated. The purpose of the training stage was for the birds to learn the tool functioning in a particular task. To aid slowly engaging individuals (e.g. those who did not manage to grab the tool and direct it towards the target at this stage), we adapted our experimental protocol by providing step-by-step shaping (i.e. the experimenter demonstrated how to handle the tool rightfully and how to insert it in the apparatus) until such individuals could solve the task autonomously in five consecutive trials. The success criterion for the training stage was solving correctly 80% of trials within a session. Once this criterion was met, birds moved on to the testing stage, where the goal was to put into practice the skills acquired in the training stage without receiving any help from the experimenter. In the testing stage, birds had to solve 100% of trials per session in two consecutive sessions to ensure that they had completely mastered the use of the tool in a given apparatus before proceeding to the next one. In the case of the control group, subjects continued being tested on the basket task even after reaching solving criterion on the testing stage, to ensure an equivalent overall number of trials as the test group and a comparable amount of experience time with the hook. The purpose of this extensive training was to try to guarantee that all birds achieved the same level of competence before moving on to phase 2. These additional sessions were labelled as "extra time." The number of sessions

required per bird in each task can be found in Table S5 (test group) and Table S6 (control group) of the SI.

2.2.2 | Phase 2

In phase 2, both groups were tested on two novel, untrained tasks: the seesaw and the cane. In contrast to phase 1, birds commenced phase 2 testing immediately (i.e. without previous training). In the seesaw task, birds had to first insert the hook through a hole located in the top of the covering box, grasp a lever at the end of a seesaw and pull it upwards. This motor action lifted the seesaw, making the reward slide away from the self, down to a location where it could be collected by the bird (Figure 2a). Hence, the seesaw task differed from the other tasks in the immediate effect following the tool's movement: while in the rest of the tasks the hook always moved the reward closer to the bird, in the seesaw task its effect was to push the reward away before it became accessible to the bird elsewhere. Another crucial characteristic was that the target element in this apparatus (lever) had a horizontal orientation, in contrast to the vertical orientation of the target element in the basket (handle). Birds were administered a maximum of three 12-trial sessions for this task, and success criterion was solving 80% of the trials in one session.

Once testing with the seesaw had concluded, each bird was tested on a second task that lacked any common features with previously presented tasks. The cane task was a two-choice task in which two identical-looking baited hooks were laid out parallel to each other on a wooden platform covered with a transparent Plexiglas sheet (Figure 2b). Each hook was fitted into a groove to prevent it from flipping and to ensure that it could only be pulled straight. A similar version of this task was used in a previous study with Darwin's finches (Teschke et al., 2011). In each trial only one of the hooks was functional (i.e. it led to obtain food when pulled). Which hook was made functional was counterbalanced across trials. If a subject picked the functional hook, it could access the reward; if the hook picked was the non-functional, the apparatus was removed from the bird's reach, ending the trial. Most importantly, the main aspect distinguishing this task from the other tasks used in our study was that the hooks were already pre-inserted in the apparatus, hence not requiring birds to apply any complex motor action other than picking a hook and slightly drag it towards themselves. To make the present study results comparable to those from the Darwin's finches' study (where birds received a maximum of 140 trials and had to make 15 or more correct choices within two consecutive blocks of 10 trials), our Goffins received a maximum of ten 12-trial sessions in the cane task. Given that this was a two-choice task, we also applied a more stringent solving criterion to be more confident that success was not attributable to chance. The solving criterion was set at 80% of correct trials per session in two consecutive sessions. Testing in the cane task ended either when criterion was reached or once the bird had received 10 sessions.

In both tasks, a trial started as soon as the subject touched the tool and lasted either for 5 min or until the bird solved the task. In contrast to phase 1, phase 2 involved no shaping or feedback to the birds. The variables measured (values between brackets) were *session number* (1–3 in the seesaw task; 1–10 in the cane task), *trial number* (1–12 in both tasks), *solved* (0 = not solved; 1 = solved) and *time to solve* each trial (0–300 s).

2.3 | Statistical analyses

2.3.1 | Success probability in the seesaw task (model 1) and the cane task (model 2)

All statistical analyses were performed in RStudio (version 4.0.3, R Core Team, 2020). To analyse the effect of *sex*, *group* and *session number* on the *probability to solve a trial* in phase 2, we fitted two (one model per task) generalized linear mixed models (*glmer* function) with binomial error structure and *logit link* function of the package lme4 (version 1.1–25, McCullagh & Nelder, 1989). The response variable *probability to solve a trial* was a matrix with two columns: the first one grouping number of solved trials per individual and session, the second one with the unsolved trials (i.e. trial performance was aggregated at the session level). We included an interaction between *group* and *session number*, and we further included *individual* as random effect. We also z-transformed *session number* to a mean of zero and a standard deviation of one to achieve easier interpretable estimates while avoiding potential convergence issues (Schielzeth, 2010). We used Wald's *z*-approximation (Field, 2005) to infer about the significance of individual effects.

We divided the original data set into two subsets, namely the seesaw (model 1) and the cane (model 2) to fit two separate models, one for each of the tasks in phase 2. The number of observations in these subsets was 300 for model 1 and 1116 for model 2. In model 1, we fitted the full model with a control structure and added the optimizing algorithm *bobyqa* (Bound Optimization BY Quadratic Approximation; Powell, 2009), aiming to maximize convergence probability. In model 2, the full model suggested that the correlation among random slope of intercept for session number (a parameter which is recommended to routinely include; see Barr et al., 2013) was estimated to be essentially 0 (i.e. it was not appropriately identifiable due to the complexity of the model and the small sample size; Matuschek et al., 2017). Therefore, we fitted a full model without the correlation parameter, which showed an almost identical log-likelihood and no substantial differences with respect to the fixed effects, and we used this model for further inferences. The distribution of individual-specific deviations from the common intercept and slopes was symmetrical. The full model did not present any collinearity issues, and it was stable. The null model, lacking *group*, was fitted separately for model 1 and 2. In both cases, full-null model comparisons were conducted

using a χ^2 test. To infer the significance of the terms involved in the interaction, we conducted likelihood ratio tests by dropping one term at a time from the full model and comparing it to the null model, whereby the *p*-values obtained are considered significant only if the comparison reveals significance (Mundry, 2014). We determined the effect size for the full as compared to the null model with the function *r.squaredGLMM* of the package *MuMIn* (Barton, 2009), and confidence intervals of fitted values at 95% using the function *confint* of the package *base*. To rule out collinearity issues, we calculated Variance Inflation Factors (VIF; Field, 2005), with the function *vif* of the package *car* (version 3.0–10, Fox & Weisberg, 2011), applied to a model lacking the interaction and the random effects.

2.3.2 | Time to solve in the seesaw task (model 3) and the cane task (model 4)

Furthermore, we checked whether group's performance would differ throughout testing in both tasks, that is that groups would differ in time to solve a trial across sessions and across trials within session. To address this question, we performed two mixed effects Cox regressions (also known as survival analysis), one for each task separately, using the previously generated subsets (*seesaw* for model 3 and *cane* for model 4). The number of observations was 468 for model 3 and 1560 for model 4. In these models, the response variable time to event was measured as "time" in seconds needed to solve each trial in phase 2, and *solved* >0 was used as the argument for "event." We fitted the models using the function *coxme* of the package *coxme* (Therneau, 2020) including interaction terms between *group*, *trial ID*, and *session number*, a fixed effect of *sex*, and a random effect of *individual*. The 3-way interaction, formulated as *group*trial ID*session number*, included all the 2-way interactions and the main effects it encompassed. We created the survival object with the function *surv* of the package *survminer* (Kassambara et al., 2020). Prior to fitting the models, we *z*-transformed *session number* and *trial ID* to a mean of zero and a standard deviation of one. To control for cryptic multiple testing (Forstmeier & Schielzeth, 2011), we performed full-null model comparisons between the full models (comprising *group* and all the interactions this term was involved in) and the null models (identical to the full models with respect to the random effects structure but lacking *group* as a main effect but also without all the 3- and 2-way interactions that *group* was involved in); comparisons were conducted by using a likelihood ratio test (Dobson, 2002). We aimed to assess model stability (which reveals how much the estimated coefficients change when individual cases are excluded) by dropping datapoints, one at a time, then fitting the full model to each of these subsets and comparing the range of model estimates obtained with those obtained for the full data set. Model stability could not be determined due to the complexity of the models. However, although this might slightly influence the confidence that we can have on the conclusions drawn from our models, we did not expect their validity to be compromised, as we took all the necessary steps prior to fitting the models to ensure their reliability.

3 | RESULTS

3.1 | Qualitative description of performance in the seesaw task and the cane task

In the seesaw task, groups did not differ in the number of sessions required to reach criterion (mean ± SD = test: 1.85 ± 0.89 vs. control: 2 ± 1.1; range = 1–3; $t_{(11)} = -1.869$; $p = .092$). Although the same number of test subjects as control subjects reached criterion in session 1 (test group ($n = 3$): Figaro, Fini, Muppet; control group ($n = 3$): Kiwi, Konrad, Zozo), across all 3 sessions the test group contained a higher proportion of successful individuals (100% vs. 66%). This was due to two individuals from the control group (Muki and Olympia) not reaching criterion for this task. No group differences were detected in the average number of trials solved per session (test group: 11 ± 1.41; control group: 11 ± 1; $t_{(11)} = -0.582$; $p = .577$): the best-performing test group birds (Figaro and Muppet) solved as many trials per session (12 trials) as the best-performing control group bird (Kiwi). The average time to solve a trial in session 1 in the control group was almost double than that of the test group (125 s vs. 67 s), with Figaro (test group) being the fastest (27.4 s) and Olympia (control group) the slowest bird overall (260.6 s). For detailed information on learning criterion and solving accuracy, the reader is referred to Tables S3, S7b,d, S9b and S10b of the SI.

In the cane task, groups differed in the number of sessions needed: individuals in the test group required between 4 and 7 sessions to reach criterion (mean ± SD = 5.57 ± 1.27), while the control group required between 7 and 10 sessions (9 ± 1.26; $t_{(11)} = -4.227$; $p = .001$). The fastest subjects in the test group (number of sessions in brackets) were Dolittle and Fini (4), followed by Pipin (5), Muppet and Figaro (6), and Mayday and Heidi (7). The fastest bird in the control group (Kiwi) needed as many sessions (7) as the slowest birds in the test group (Mayday and Heidi). Kiwi was followed closely by Zozo (8) and Konrad (9) in the control group. The same three individuals from the control group (Muki, Olympia and Moneypenny) that did not reach criterion for the seesaw task also did not reach it for the cane task. Groups also differed in their solving accuracy in this task, with the test group having a higher average number of solved trials per session (8.42 ± 0.45) compared with the control group (7.76 ± 0.75; $t_{(11)} = 2.645$; $p = .024$). Muppet (test group) was the bird with the highest average number of solved trials (9.2) whereas Olympia (control group) had the lowest (6.7). There were no differences in the average time to solve a trial in session 1 between the test and the control group (6.26 vs. 10.78 s; $t_{(11)} = -0.992$; $p = .351$), with Fini (test group) being the fastest (2.4 s) and Muki (control group) the slowest (69.7 s) of birds overall. For detailed information on learning criterion and solving accuracy, the reader is referred to Tables S2, S7a,c, S9a and S10a of the SI.

3.2 | Success in the seesaw task (model 1) and in the cane task (model 2)

In the seesaw task, the full-null model (model 1) comparison indicated that no group differences were found in the probability to solve a trial across sessions ($\chi^2_{(2)} = 1.545$; $p = .219$; see Table S4 of the SI). However, the fact that both groups' performance was above 50% of accuracy already from session 1 (test group: 78.3%; control group: 62.5%) indicates that they both indeed transferred immediately (see Table S9b of the SI).

In the cane task, the full model (model 2) detected a significant interaction between *group* and *session number* ($F_{(1, 2)} = 5.631$; $p = .018$; see Table 2), indicating that groups differed with respect to their probability of solving a trial across sessions, being this higher for the test group. The full-null model comparison revealed an effect of *group* on the number of solved trials per *session number* ($\chi^2_{(2)} = 19.631$; $p < .001$; $R^2_{\text{fixed+random effects}} = .475$). Hence, while transfer occurred in both groups, it was not immediate (i.e. it did not occur from session 1), and the test group solved the task faster (as reflected by a steeper learning curve; see Figure 3).

3.3 | Time to solve a trial in the seesaw task (model 3) and the cane task (model 4)

In the seesaw task, we did not find a significant effect of *group* on *time to solve a trial across sessions* ($\chi^2_{(4)} = 2.925$; $p = .571$; see Table S8 of the SI), suggesting that both groups' performance remained equally fast throughout testing. In the cane task, on the contrary, the full-null model comparison revealed a significant effect of *group* on *time to solve a trial across sessions* ($\chi^2_{(4)} = 29.104$; $p < .001$.; see Table 3; see Figure 4), indicating that individuals from the test group solved a trial significantly faster than individuals from the control group across sessions. This meant that, not only did the test group learn the task faster but they also showed more improvement in their solving accuracy along testing.

A summary of performance of both groups in both tasks is shown in Table 4.

4 | DISCUSSION

We studied tool transfer in Goffin's cockatoo by investigating whether the capacity to apply previously acquired information about a tool and its use from one context to a new one changes as a function of contextually diverse prior experience with the same tool. Our findings indicate that Goffin's cockatoos can transfer knowledge and derived skilfulness about a hooked tool to new contexts in which they received no training and different from those in which learning occurred and that individuals who received broader experience perform better in transfer.

Transfer was originally planned and tested on one task only, the seesaw. We initially considered this task to be sufficiently new to all birds for two major reasons: First, the application of the tool onto the target element in the seesaw (lever) did not allow to retrieve the reward immediately, as it was delivered to the opposite side of the apparatus once the seesaw was lifted up; and second, the orientation of said lever was horizontal, as opposed to the vertically oriented target in the basket (handle). However, it became apparent when most birds from both groups transferred almost immediately that despite the design modifications, the seesaw task shared a similarity in the sequence of movements required to find a solution with the basket (which was common for both groups in phase 1). This resemblance affected in particular the motor operating, as solving the seesaw involved some steps in which both groups had gained experience (e.g. inserting the hook vertically and pulling upwards). This could therefore explain why performance of both groups did not differ in the seesaw, as we did not manage to significantly change the task's appearance, a requisite that transfer tasks must fulfil (Seed et al., 2006).

To address this concern, we presented the birds with a second transfer task. To amend the limitation of the seesaw, we incorporated two major distinctions to make this task unique and completely novel for all birds: (1) the tool–reward interaction was direct (i.e. the tool was applied directly onto the reward without mediation of a target), and (2) the tool was already pre-inserted in the apparatus (i.e. not requiring the individual to apply any motor action other than simply picking a hook). We found that birds who had received in phase 1 more diverse learning and motor experiences on multiple

TABLE 2 Coefficients for fixed effects of the glmer (cane task, model 2, full model)

	Estimate	SE	*p*-value	Lower CI	Upper CI
Intercept	0.569	0.108	**<.001***	0.347	0.804
Group	0.917	0.176	**<.001***	0.563	1.277
Session nr.	0.435	0.082	**<.001***	0.276	0.597
Sex	−0.191	0.148	.196	−0.500	0.118
Group:Session nr.	0.438	0.185	**.018***	0.080	0.807

Note: Reference values for *group* and *sex* are "control" and "male," respectively; *Session number* was *z*-transformed; original mean ± SD = 3.805 ± 2.512.

Abbreviations: nr., number; SE:,standard error; CI, 95% confidence intervals; (:), interaction term.

*Significant *p*-values at $\alpha = 0.05$ are shown in bold.

usages of the hooked tool (test group) had a higher solving accuracy (i.e. probability to solve a trial), needed less time to solve (i.e. in terms of latency to solve a trial and lower number of sessions to reach criterion altogether) and had a higher proportion of successful individuals in the cane task. These results provide support in favour of the prediction that the test birds would be better able to apply the hooking capacity of the tool to a task that did not resemble any other known task and that required a new solving principle with which none of the groups had experience. The fact that there were no group differences in the average latency to solve a trial in session 1 seems to indicate that there was no strong involvement of the motor component to find a solution. Hence, this two-choice task allowed us to investigate subject's sensitivity to the functional spatial relationship between a reward and the tool used to obtain it (Teschke et al., 2011). Transfer in this context, at the very least, suggests that by having experienced a broader range of contingencies with the tool, the test group acquired a more general knowledge about the tool's affordances and function/s (including, but not limited to, a set of transferable skills), which varied in form and content from that acquired by the control group (Kamil, 1988; Tebbich et al., 2016). This may, in turn, have enabled the test group to grasp the new tool–reward interaction faster and to improve their problem-solving competency. Therefore, the test group would have benefited from a broader knowledge with the potential to influence behaviour in various situations (i.e. be transferred and lead to the discovery and exploitation of adjacent opportunities) and to consequently increase behavioural flexibility (Anderson, 1976; Dienes & Perner, 1999; Tebbich et al., 2016).

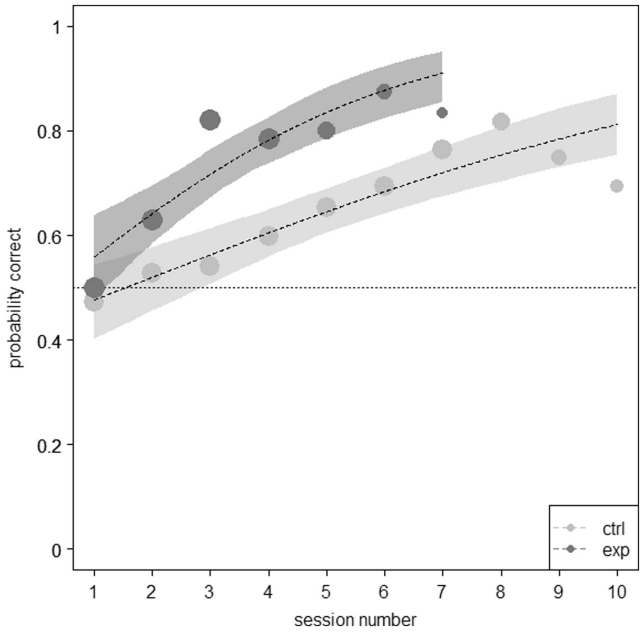

FIGURE 3 Probability of a trial being solved in the cane task as a function of session number. Performance of control group (ctrl) is shown in light grey, and test group (exp) is depicted in dark grey. Point area is proportional to the number of subjects that participated on each session (i.e. areas have different sizes because of some individuals reaching solving criterion earlier than others). Shaded line corresponds to 95% confidence intervals. Horizontal dotted line reflects performance at chance level. The plot shows the model with *sex* centred and 1000 bootstraps

A crucial difference between the test and control groups in phase 1 was that the test group was given the opportunity to practice how to use the tool in three different ways (i.e. to reach a target in three differently operated apparatuses). This, in turn, may have (a) strengthened the test birds' understanding of the complexity of the tool–target interactions and about how to correctly orientate the tool in relation to their own body in each apparatus (Fragaszy & Mangalam, 2018), and (b) assisted the development of a less rigid sensorimotor competence (i.e. not bound to a single tool-use instance). However, even if the latter was the case, the test group would not have had a notorious advantage in the cane task, where the contribution of the motor component is considered practically neutralized, as birds did not have to manipulate the tool. Hence, no skilful manoeuvring or handling of the tool was required except for slightly pulling it, making both groups equal with respect to the influence of motor dexterity. For this reason, we do not think that enhanced sensorimotor competence played a significant driving role in the test group's performance on the cane task. We rather suggest that group differences might be attributable to differences in the knowledge acquired or potentiated during phase 1. Whether the cognitive mechanisms underlying these observed differences are based on associative learning and/or causal understanding is

TABLE 3 Coefficients for fixed effects of the Cox regression (cane task, model 4)

	Coef.	Exp.	SE	*p*-value	Lower CI	Upper CI
Group	1.801	6.057	0.691	**.009***	0.446	3.156
Trial nr.	0.319	1.377	0.055	**<.001***	0.211	0.428
Session nr.	1.135	3.110	0.122		0.894	1.375
Sex	0.161	1.175	0.700	.83	−1.211	1.534
Group:Trial nr.:Session nr.	−0.124	0.884	0.086	**<.001***	−0.292	0.044

Note: Reference values for *group* and *sex* are "control" and "male," respectively; *Session number* and *trial number* were z-transformed; original mean ± standard deviation = 5.5 ± 2.87 and 6.5 ± 3.45, respectively.

Abbreviations: nr., number; SE, standard error; CI, 95% confidence intervals; Coef., coefficients; Exp., exponentiated coefficients; :, interaction term.

*Significant *p*-values at $\alpha = 0.05$ are shown in bold.

something that our data do not allow to determine, and a dichotomous vision that we do not endorse. Instead, we agree with previous research suggesting that they rather result from a non-exclusive and mutually influencing combination of both processes (Tebbich et al., 2016; Teschke et al., 2013).

In contrast to the performance of the test group, the control group—which had an intensive experience with the tool in a single context—appeared less flexible to switch its use when faced with the cane task. Although they did transfer, they showed a lower probability of solving and a slower speed in doing so. At the proximate level, we speculate that this pattern could be due to the control group having had access to a narrower information of the tool's multi-versatility and its potential applicability, which was intentionally limited during phase 1. This could have resulted in (a) a lack of knowledge about the relations between the problem elements, or (b) that such knowledge was coupled to a specific problem (Martin-Ordas et al., 2008), namely the basket task, and hence less generalizable (Tebbich et al., 2016). Furthermore, studies in macaques indicate that tool use creates mental representations (Iriki et al., 1996; Maravita & Iriki, 2004) which tend to be rigidly used in relation to the behaviour in which they were originated (Macellini et al., 2012), leaving little space for adjustments and eventually leading to inflexibility. Our results suggest that control birds' acquisition of a less plastic substrate to fine-tune performance might have led them to form a single representation of the tool (i.e. as a mean to hoist an out-of-reach reward in the vertical axis) by which they got fixated, as opposed to the three-way representation formed by the test group. Eventually, the nature of our training might have driven subjects in the control group towards becoming more prone to suffer from functional fixedness, which needed to be overcome in order to approach the new problem efficiently (Fujita et al., 2011): by being intensively trained in the basket task (and repeating the corresponding behaviour in the seesaw task, where the hook was operated in a similar way), they might have been hindered to discover a new solution for the cane task different from the one that worked out for the basket and the seesaw (Duncker, 1945; Ebel et al., 2021; Harrison & Whiten, 2018). Despite the cane being perceptually and functionally dissimilar from the basket, this pattern of results could have emerged from the (over)training itself inducing proclivity to fixation in the control group.

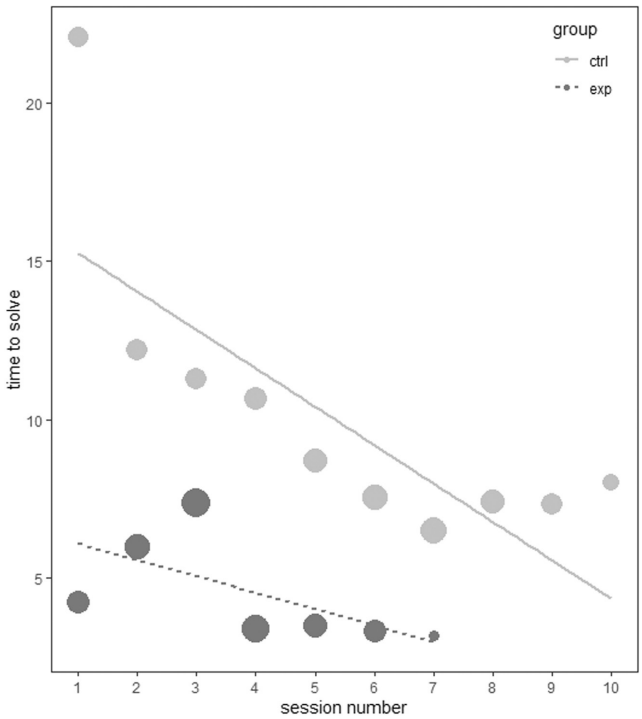

FIGURE 4 Time to solve a trial (in seconds) in the cane task as a function of session number. Performance of control group (ctrl) is shown in light grey with a solid line, and test group (exp) is depicted in dark grey with a dashed line. Point area corresponds to the number of solved trials per session across all subjects of the same group; lines represent the prediction made by the fitted model. Values on the *y*-axis are in seconds.

Goffin's cockatoos' ability to gain different degrees of information about a tool and its purpose depending on the tool-use training provided, and to use this information to tackle new problems more efficiently is a faculty that has been already observed in other bird taxa. Teschke et al. (2011), using a version of the cane task that involved several experimental conditions, found that woodpecker finches (WPF; *Cactospiza pallida*) and small tree finches (STF; *Camarhynchus parvulus*) were equally fast at solving the transfer task and showed similar improvement over the sequence of trials, with no difference between tool-using and non-tool-using WPF, although STF outperformed non-tool-using WPF in success probability. In rooks (*Corvus frugilegus*; a non-tool-using species), Seed et al. (2006) tested eight subjects in a series of experiments with the modified

TABLE 4 Summary of group differences in the seesaw task and the cane task

Variable	Seesaw task	Cane task
Number of sessions to reach criterion	Test = Control	Test < Control
Proportion of successful individuals	Test > Control	Test > Control
Number of trials solved per session	Test = Control	Test > Control
Time to solve a trial in session 1	Test < Control	Test < Control
Probability to solve a trial across sessions	Test = Control	Test > Control
Time to solve a trial across sessions	Test = Control	Test < Control

Abbreviations: =, groups did not differ in their performance on the corresponding variable; <, first group (test) had a lower score in the corresponding variable; >, first group (test) had a higher score in the corresponding variable.

two-trap, a transfer task similar to our cane task in the sense that the tool was already placed inside the tube. In Experiment 2, birds that had solved configuration A of the tube in Experiment 1 were tested on configuration B, and vice versa, to assess whether they could transfer (which authors defined as "responding significantly correctly within the first 20 trials"). All rooks that had learnt the initial two-trap task configuration (7 out of 8 subjects) were able to transfer in the new configuration. In Experiment 3, where subjects could not solve the task by using arbitrary cues or procedural rules based on the task configuration, all birds but one failed to transfer (which was defined as ≥9/10 correct trials or ≥15/20 in two 10-trial blocks). Hence, our results provide the first evidence that parrots can, at the very least, parallel corvids' performance in transfer tasks.

Altogether, our results highlight that the ability to transfer poses a clear advantage in offsetting the downsides of innovating (Auersperg et al., 2017; Klump et al., 2015), as it can enable individuals to reach or create novel opportunities while reducing the cognitive and ecological resources required (e.g. time investment for manufacturing a new tool for a new purpose and mastering its use). To unveil *how* exactly transfer operates in Goffin's cockatoos, we encourage upcoming studies to use tasks based on the principle of triangulation (Heyes, 1993), such as sequential transfer tasks, where arbitrary features are removed successively from the original problem (see Taylor, Hunt, et al., 2009) for a reliable identification of the contribution of each component in the transfer process (e.g. determining the precise role of the motor component). Future studies should also aim to include transfer tasks where the type of tool is unattached and freely manipulable and the type of action applied involves a dynamic interaction, orientable by the agent (Fragaszy & Mangalam, 2018), which was the case in our seesaw task but not in the cane task. A final consideration concerns the fact that age was not accounted for in our study. Although ideally the eldest birds should not have all been part of the test group, the average age did not differ across groups. We hence consider that this factor might have had a minimal impact on our results, as Goffin's age does not necessarily translate into how much experimental experience they have (since they all began participating in tool use experiments in the same year), and none of the subjects were considered juvenile or senile. However, we would like to make a remark on the role that experience (defined by the STRANGE framework as "the opportunities for individual learning, such as participation in earlier experiments"; Webster & Rutz, 2020) can play on behaviour, as long-lived subjects often accumulate experimental histories that could influence their test performance. To amend this and other potential sampling biases that may limit the generalisability of our findings (e.g. social background, genetic make-up or acclimation), ideally a study of similar characteristics (using a different type of tool) should be conducted, where the subjects formerly allocated to the control group are the test subjects, and vice versa. This becomes however challenging in the case of the Goffins, given the reduced number of captive individuals available, and it (a) warrants cautious interpretation of our findings due to our relatively low sample size and (b) begs consideration for replicating the study in wild populations to assess the ecological relevance of this behaviour. Nevertheless, we believe that our study adds up to the growing catalogue of aptitudes of this species and expands our understanding of parrot cognition, opening interesting paths for future research in this field.

AUTHOR CONTRIBUTIONS

Paula Ibáñez de Aldecoa: Conceptualization; investigation; writing – original draft; methodology; validation; visualization; formal analysis; software; project administration; data curation. **Alice M. I. Auersperg:** Conceptualization; investigation; funding acquisition; methodology; validation; project administration; supervision; resources. **Andrea S. Griffin:** Conceptualization; investigation; methodology; validation; project administration; supervision. **Sabine Tebbich:** Conceptualization; investigation; funding acquisition; methodology; validation; project administration; supervision; resources.

ACKNOWLEDGMENTS

We are grateful to Roger Mundry for statistical advice and for kindly providing R functions, Isabelle Laumer for helpful input on experimental design, and Nicola Pachmann and Antonio Osuna-Mascaró for assistance on building the experimental apparatuses. This project was funded by the Austrian Science Fund (FWF) W1262-B29 to S.T. The Goffin Lab is presently funded by WWTF Project CS18 023 and FWF START Project Y01309 to A.M.I.A. The funding source had no involvement in study design, data collection, statistical analysis or writing of the manuscript.

CONFLICT OF INTEREST

The authors declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

DATA AVAILABILITY STATEMENT

The data sets generated and/or analysed during the current study are available at https://phaidra.univie.ac.at/view/o:1620423

CODE AVAILABILITY

The code for the statistical analysis is available from the corresponding author upon request.

ORCID

Paula Ibáñez de Aldecoa https://orcid.org/0000-0003-3858-5107
Alice M. I. Auersperg https://orcid.org/0000-0001-7405-9791
Andrea S. Griffin https://orcid.org/0000-0003-4624-9904
Sabine Tebbich https://orcid.org/0000-0003-3971-2139

REFERENCES

Anderson, J. R. (1976). *Language, memory, and thought*. Erlbaum.
Auersperg, A. M. I. (2015). Exploration technique and technical innovations in corvids and parrots. In A. B. Kaufman & J. C. Kaufman (Eds.), *Animal creativity and innovation* (pp. 45–63). Academic Press.
Auersperg, A. M. I., Gajdon, G. K., & von Bayern, A. M. P. (2012). A new approach to comparing problem solving, flexibility and innovation.

Communicative & Integrative Biology, *5*, 140–145. https://doi.org/10.4161/cib.18787

Auersperg, A. M. I., Kacelnik, A., & von Bayern, A. M. P. (2013). Explorative learning and functional inferences on a five-step means-means-end problem in Goffin's cockatoos (*Cacatua goffini*). *PLoS One*, *8*, e68979. https://doi.org/10.1371/journal.pone.0068979

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., Szabo, B., von Bayern, A. M. P., & Bugnyar, T. (2014). Object permanence in the Goffin cockatoo (*Cacatua goffini*). *Journal of Comparative Psychology*, *128*(1), 88–98. https://doi.org/10.1037/a0033272

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. (2014). 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*, *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. https://doi.org/10.1098/rspb.2014.0972

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

Barton, K. (2009). Mu-MIn: Multi-model inference. R Package Version 0.12.2/r18. http://R-Forge.R-project.org/projects/mumin/

Bird, C. D., & Emery, N. J. (2009). Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10370–10375. https://doi.org/10.1073/pnas.0901008106

Bobrowicz, K., Johansson, M., & Osvath, M. (2020). Great apes selectively retrieve relevant memories to guide action. *Scientific Reports*, *10*, 12603. https://doi.org/10.1038/s41598-020-69607-6

Bobrowicz, K., O'Hara, M., Carminito, C., Auersperg, A. M. I., & Osvath, M. (2021). Goffin's cockatoos (*Cacatua goffiniana*) can solve a novel problem after conflicting past experiences. *Frontiers in Psychology*, *12*, 694719. https://doi.org/10.3389/fpsyg.2021.694719

Call, J. (2012). Three ingredients for becoming a creative tool user. In C. M. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals: Cognition and ecology* (p. 326). Cambridge University Press.

Diamond, J., & Bond, A. B. (1999). *Kea, bird of paradox: The evolution and behavior of a New Zealand parrot*. University of California Press.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735–808. https://doi.org/10.1017/S0140525X99002186

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

Dukas, R. (2004). Evolutionary biology of animal cognition. *Annual Review of Ecology, Evolution and Systematics*, *35*, 347–374. https://doi.org/10.1146/annurev.ecolsys.35.112202.130152

Duncker, K. (1945). On problem-solving (L. S. Lees, Trans.). *Psychological Monographs*, *58*, i-113. https://doi.org/10.1037/h0093599

Ebel, S. J., & Call, J. (2018). The interplay of prior experience and motivation in great ape problem-solving (*Gorilla gorilla*, *Pan paniscus*, *Pan troglodytes*, and *Pongo abelii*). *Journal of Comparative Psychology*, *132*(3), 294–305. https://doi.org/10.1037/com0000117

Ebel, S. J., Völter, C. J., & Call, J. (2021). Prior experience mediates the usage of food items as tools in great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, and *Pongo abelii*). *Journal of Comparative Psychology*, *135*(1), 64–73. https://doi.org/10.1037/com0000236

Emery, N. J., & Clayton, N. S. (2009). Tool use and physical cognition in birds and mammals. *Current Opinion in Neurobiology*, *19*(1), 27–33. https://doi.org/10.1016/j.conb.2009.02.003

Field, A. (2005). *Discovering statistics using SPSS*. Sage Publications.

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*, 47–55. https://doi.org/10.1007/s00265-010-1038-5

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). SAGE.

Fragaszy, D. M., & Mangalam, M. (2018). Tooling. In M. Naguib, L. Barrett, S. D. Healy, J. Podos, L. W. Simmons, & M. Zuk (Eds.), *Advances in the Study of Behavior* (pp. 177–241). Academic Press. https://doi.org/10.1016/bs.asb.2018.01.001

Fujita, K., Sato, Y., & Kuroshima, H. (2011). Learning and generalization of tool use by tufted capuchin monkeys (*Cebus apella*) in tasks involving three factors: Reward, tool, and hindrance. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(1), 10–19. https://doi.org/10.1037/a0020274

Griffin, A. S., Diquelou, M., & Perea, M. (2014). Innovative problem solving in birds: A key role of motor diversity. *Animal Behavior*, *92*, 221–227. https://doi.org/10.1016/j.anbehav.2014.04.009

Griffin, A. S., & Guez, D. (2014). Innovation and problem solving: A review of common mechanisms. *Behavioural Processes*, *109*, 121–134. https://doi.org/10.1016/j.beproc.2014.08.027

Griffin, A. S., Guez, D., Lermite, F., & Patience, M. (2013). Tracking changing environments: Innovators are fast, but not flexible learners. *PLoS One*, *8*, e84907. https://doi.org/10.1371/journal.pone.0084907

Guez, D., & Griffin, A. S. (2016). Unraveling the key to innovative problem solving: A test of learning versus persistence. *Behavioral Ecology*, *27*(5), 1449–1460. https://doi.org/10.1093/beheco/arw055

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

Harrison, R. A., & Whiten, A. (2018). Chimpanzees (*Pan troglodytes*) display limited behavioral flexibility when faced with a changing foraging task requiring tool use. *Peer J*, *6*, e4366. https://doi.org/10.7717/peerj.4366

Heyes, C. M. (1993). Anecdotes, training, trapping, and triangulating: Do animals attribute mental states? *Animal Behavior*, *46*, 177–188. https://doi.org/10.1006/anbe.1993.1173

Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurons. *Neuroreport*, *14*, 2325–2330. https://doi.org/10.1097/00001756-199610020-00010

Kamil, A. C. (1988). A synthetic approach to the study of animal intelligence. In D. W. Leger (Ed.), *Comparative perspectives in modern psychology* (pp. 257–308). University of Nebraska Press.

Kassambara, A., Kosinski, M., & Biecek, P. (2020). survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.4.8. https://CRAN.R-project.org/package=survminer

Klump, B. C., van der Wal, J. E. M., St Clair, J. J. H., & Rutz, C. (2015). Context-dependent 'safekeeping' of foraging tools in new Caledonian crows. *Proceedings of the Royal Society B: Biological Sciences*, *282*, 20150278. https://doi.org/10.1098/rspb.2015.0278

Lambert, M. L., Schiestl, M., Schwing, R., Taylor, A. H., Gajdon, G. K., Slocombe, K. E., & Seed, A. M. (2017). Function and flexibility of object exploration in kea and new Caledonian crows. *Royal Society Open Science*, *4*, 170652. https://doi.org/10.1098/rsos.170652

Laumer, I. B., Bugnyar, T., & Auersperg, A. M. I. (2016). Flexible decision-making relative to reward quality and tool functionality in Goffin

cockatoos (*Cacatua goffiniana*). *Scientific Reports*, *6*, 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*, 20171026. https://doi.org/10.1098/rspb.2017.1026

Logan, C. J., Blaisdell, A. P., Johnson-Ulrich, Z., Lukas, D., MacPherson, M., Seitz, B. M., Sevchik, A., & McCune, K. B. (2022). Behavioral flexibility is manipulatable and it improves flexibility and problem solving in a new context. *EcoEvoRxiv*. https://doi.org/10.32942/osf.io/5z8xs

Macellini, S., Maranesi, M., Bonini, L., Simone, L., Rozzi, S., Ferrari, P. F., & Fogassi, L. (2012). Individual and social learning processes involved in the acquisition and generalization of tool use in macaques. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1585), 24–36. https://doi.org/10.1098/rstb.2011.0125

Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, *8*, 79–86. https://doi.org/10.1016/j.tics.2003.12.008

Martin-Ordas, G., Call, J., & Colmenares, F. (2008). Tubes, tables and traps: Great apes solve two functionally equivalent trap tasks but show no evidence of transfer across tasks. *Animal Cognition*, *11*, 423–430. https://doi.org/10.1007/s10071-007-0132-1

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

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

Mettke-Hofmann, C., Winkler, H., & Leisler, B. (2002). The significance of ecological factors for exploration and neophobia in parrots. *Ethology*, *108*(3), 249–272. https://doi.org/10.1046/j.1439-0310.2002.00773.x

Mundry, R. (2014). Chapter 6: Statistical issues and assumptions of phylogenetic generalized least squares. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology*. Springer. https://doi.org/10.1007/978-3-662-43550-2_6

O'Hara, M., Mioduszewska, B., Haryoko, T., Prawiradilaga, D. M., Huber, L., & Auersperg, A. (2019). Extraction without tooling around — The first comprehensive description of the foraging- and socio-ecology of wild Goffin's cockatoos (*Cacatua goffiniana*). *Behaviour*, *156*(5-8), 661–690. https://doi.org/10.1163/1568539x-00003523

O'Hara, M., Mioduszewska, B., Mundry, R., 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

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. Department of Applied Mathematics and Theoretical Physics, Cambridge England, technical report NA2009/06

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reader, S. M., & Laland, K. N. (2003). *Animal innovation*. Oxford University Press.

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, *1*, 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Seed, A. M., Tebbich, S., Emery, N. J., & Clayton, N. S. (2006). Investigating physical cognition in rooks, *Corvus frugilegus*. *Current Biology*, *16*, 697–701. https://doi.org/10.1016/j.cub.2006.02.066

Taylor, A. H., Hunt, G. R., Medina, F. S., & Gray, R. D. (2009). Do new Caledonian crows solve physical problems through causal reasoning? *Proceedings of the Royal Society B: Biological Sciences*, *276*, 247–254. https://doi.org/10.1098/rspb.2008.1107

Taylor, A. H., Roberts, R., Hunt, G. R., & Gray, R. D. (2009). Causal reasoning in new Caledonian crows. *Communicative & Integrative Biology*, *2*(4), 311–312. https://doi.org/10.4161/cib.2.4.8224

Tebbich, S., Griffin, A. S., Peschl, M. F., & Sterelny, K. (2016). From mechanisms to function: An integrated framework of animal innovation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*, 20150195. https://doi.org/10.1098/rstb.2015.0195

Teschke, I., Cartmill, E. A., Stankewitz, S., & Tebbich, S. (2011). Sometimes tool use is not the key: No evidence for cognitive adaptive specializations in tool-using woodpecker finches. *Animal Behaviour*, *82*, 945–956. https://doi.org/10.1016/j.anbehav.2011.07.032

Teschke, I., Wascher, C. A. F., Scriba, M. F., von Bayern, A. M. P., Huml, V., Siemers, B., & Tebbich, S. (2013). Did tool-use evolve with enhanced physical cognitive abilities? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1630), 20120418. https://doi.org/10.1098/rstb.2012.0418

Therneau, T. M. (2020). coxme: Mixed Effects Cox Models. R package version 2.2-16. https://CRAN.R-project.org/package=coxme

Von Bayern, A. M. P., Heathcote, R. J. P., Rutz, C., & Kacelnik, A. (2009). The role of experience in problem solving and innovative tool use in crows. *Current Biology*, *19*, 1965–1968. https://doi.org/10.1016/j.cub.2009.10.03

Webster, M. M., & Rutz, C. (2020). How STRANGE are your study animals? *Nature*, *582*, 337–340.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ibáñez de Aldecoa, P., Auersperg, A. M. I., Griffin, A. S., & Tebbich, S. (2023). Ratcheting up tool innovation in Goffin's cockatoos (*Cacatua goffiniana*): The effect of contextually diverse prior experience. *Ethology*, *129*, 133–145. https://doi.org/10.1111/eth.13351