

Research



Cite this article: Soravia C, Ashton BJ, Thornton A, Ridley AR. 2022 General cognitive performance declines with female age and is negatively related to fledging success in a wild bird. *Proc. R. Soc. B* **289**: 20221748. <https://doi.org/10.1098/rspb.2022.1748>

Received: 5 September 2022

Accepted: 24 November 2022

Subject Category:

Behaviour

Subject Areas:

cognition, evolution, ecology

Keywords:

cooperative breeding, southern pied babbler, general intelligence, cognitive senescence, sex differences, cognition

Author for correspondence:

Camilla Soravia

e-mail: camilla.soravia@research.uwa.edu.au

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6330234>.

General cognitive performance declines with female age and is negatively related to fledging success in a wild bird

Camilla Soravia¹, Benjamin J. Ashton^{1,2,4}, Alex Thornton³ and Amanda R. Ridley^{1,4}

¹Centre for Evolutionary Biology, School of Biological Sciences, University of Western Australia, Perth, WA, Australia

²School of Natural Sciences, Macquarie University, Sydney, New South Wales, Australia

³Centre for Ecology and Conservation, University of Exeter, Penryn, UK

⁴FitzPatrick Institute of African Ornithology, University of Cape Town, Cape Town, South Africa

id CS, 0000-0002-9795-9677; BJA, 0000-0003-3357-0395; AT, 0000-0002-1607-2047; ARR, 0000-0001-5886-0992

Identifying the causes and fitness consequences of intraspecific variation in cognitive performance is fundamental to understand how cognition evolves. Selection may act on different cognitive traits separately or jointly as part of the general cognitive performance (GCP) of the individual. To date, few studies have examined simultaneously whether individual cognitive performance covaries across different cognitive tasks, the relative importance of individual and social attributes in determining cognitive variation, and its fitness consequences in the wild. Here, we tested 38 wild southern pied babblers (*Turdoides bicolor*) on a cognitive test battery targeting associative learning, reversal learning and inhibitory control. We found that a single factor explained 59.5% of the variation in individual cognitive performance across tasks, suggestive of a general cognitive factor. GCP varied by age and sex; declining with age in females but not males. Older females also tended to produce a higher average number of fledglings per year compared to younger females. Analysing over 10 years of breeding data, we found that individuals with lower general cognitive performance produced more fledglings per year. Collectively, our findings support the existence of a trade-off between cognitive performance and reproductive success in a wild bird.

1. Introduction

The mental mechanisms through which animals acquire, process, store and act on information from the environment represent animal cognition [1]. Animals use cognitive mechanisms to adjust their behavioural responses to the environmental and social context, remember the location of resources, and learn which environmental cues indicate presence of food, mates or predators [2]. Different animal species rely on cognitive mechanisms to different extents to solve ecological problems (e.g. [3,4]), and even within a species, cognitive performance can vary significantly among individuals (e.g. [5,6]). Such great inter- and intraspecific variation has led to the question: what selective pressures shape the evolution of cognition?

Intraspecific studies of animal cognition have shown that in some cases cognitive performance is heritable [7]. Additionally, cognitive performance has been linked to mate choice [8], reproductive investment [6], reproductive success [5,9] and survival [10]. The existence of differential fitness arising from heritable variation in cognitive performance means cognitive traits can evolve [11]. However, better cognitive performance is not always associated with increased fitness, for example, faster learning is associated with reduced

longevity in fruit flies (*Drosophila melanogaster*) [12]. Indeed, the energetic costs of enhanced cognitive function may lead to a trade-off with somatic maintenance or reproduction [11,13]. At the interspecific level, this energetic constraint may actually result in a positive relationship between longevity and cognitive performance: larger brains require a longer development, and as a consequence large-brained species often mature and reproduce later in life [13]. Therefore, to understand how selection acts on cognitive performance we need to consider its costs along with its benefits [13].

To understand how selection acts on cognition, the link between cognitive variation and fitness consequences needs to be identified, as well as the factors associated with individual variation in cognition. Several factors may explain differences in individual cognitive performance, including the physical and social environment [14,15], and individual attributes such as age (e.g. [16]), rank (e.g. [17]) and sex (e.g. [18]). Sex differences in cognitive performance often arise as a consequence of mating strategies or sex-specific ecological constraints [19]. For example, in the brood-parasitic brown headed cowbird (*Molothrus ater*), females outperform males on a large-scale spatial memory task, likely because the breeding strategy of this species relies on females finding host nests to lay their eggs in [20].

Age differences in cognitive performance have been mostly found when comparing juveniles and adults [21,22]. However, cognitive performance can also change during adult life [23,24]. The gradual reduction of cognitive function with age is known as cognitive senescence [25]. For example, homing pigeons (*Columba livia*) older than 10 years returned more often to feeders that they had just depleted despite them being empty, showing impaired short-term memory [16]. To date, little is known about cognitive senescence in the wild, due to the difficulties of quantifying individuals' age and testing cognition in the wild.

Growing evidence suggests that cognitive performance is related to social rank, but the direction of this relationship varies across studies [26,27]. Rank is determined based on social interactions, where dominant individuals monopolize resources [28]. It has been suggested that cognitive performance may not be related to dominance *per se*, but to covarying factors, such as vigilance, experience or motivation to find alternative food sources [17,27]. For example, in Arabian babblers (*Argya squamiceps*) subordinates were the first to learn to remove black lids in a novel foraging task, likely because they were more explorative, but dominants, which tend to be older in this species, were better able to generalize the solution to white lids because of experience [29].

Finally, individual cognitive performance can be influenced by group size because living in larger groups may require better cognitive performance in order to monitor the state and actions of group members, remember their identity, and the outcome of past interactions [5,14]. Despite the growing number of studies investigating intraspecific differences in cognitive performance, these individual and social attributes (rank, sex, age and group size) have rarely been examined simultaneously while controlling for proxies of motivation, and evidence of their relative importance in driving cognitive variation is scarce.

If we identify what selection pressures drive variation in cognition, the question of whether these act on each cognitive trait separately, or jointly as part of general cognitive processes, remains. In humans, it has been repeatedly

demonstrated that individual performance correlates positively across different cognitive tasks, and approximately 40% of the total variation in performance can be explained by a single general cognitive factor g [30]. This factor, also referred to as general intelligence or intelligence quotient (IQ), predicts important life outcomes, such as occupational attainment, health and longevity [30]. Recently, several studies in non-human animals have also described something akin to a general cognitive factor g explaining between 30% and 60% of variation in performance across a battery of cognitive tasks [31]. The evidence for g provided by animal studies, however, has encountered criticism. First, generating reliable measures of g in non-human animals requires robust psychometric test batteries targeting well-studied cognitive traits [32]. It is also worth noting that variation in the combination of cognitive tasks used in a test battery can lead to different estimates of g [33]. Second, results indicative of g may arise in the absence of a truly general cognitive factor if performance on different tasks is underpinned by the same cognitive mechanism—for instance, variation in associative learning performance could potentially impact performance across a range of tasks [34]. Therefore, the single factor extracted from animal cognitive test batteries does not necessarily equate to general intelligence or g as described in humans. Nonetheless, if performance measured across a battery of cognitive tasks can be explained by a single factor, hereafter referred to as 'general cognitive performance (GCP)' [5,35], and this factor predicts fitness in the wild [5], then it represents a measurable cognitive trait which may be under selection in animal populations [30].

Here, we tested wild adult southern pied babblers (*Turdoides bicolor*, hereafter 'babblers') on a psychometric test battery containing three tasks designed to quantify (i) associative learning, (ii) reversal learning, (iii) inhibitory control. These are well-studied cognitive traits that span different domains [32,36]. Additionally, they are likely to be ecologically relevant as they allow individuals to: learn predictive contingencies between environmental cues (associative learning); learn a new association when the previous one stops being rewarding (reversal learning) and control prepotent motor responses when counterproductive (inhibitory control) [36,37]. To achieve a comprehensive understanding of the relationship between different cognitive traits, the factors underpinning interindividual variation in cognition and the link between cognitive performance and fitness in a wild animal, we (i) tested whether individual cognitive performance was positively correlated across tasks and could be explained by a single factor (GCP); (ii) measured proxies of motivation and attributes of the individual and social group (age, sex, rank, group size) to identify determinants of individual cognitive performance and (iii) related individual cognitive performance to multiple measures of reproductive success.

2. Methods

(a) Study site and species

Data were collected at the Kuruman River Reserve (26°58' S, 21°49' E; South Africa, 33 km²) between September–March in 2018, 2019 and 2021 (fieldwork in 2020 was suspended due to the COVID-19 outbreak). The reserve is situated within the semi-arid Kalahari region, which is characterized by vegetated sand dunes. Pied babblers are medium-sized (60–90 g), sexually monomorphic passerines endemic to this region. They are cooperative

breeders and live in groups, which include a dominant pair and subordinate helpers [38]. The dominant pair produces approximately 95% of the offspring [39,40] and pair bond tenure varies greatly (from less than 1 month to greater than 5 years) [41]. On average only 4% of subordinates live in non-natal groups each year [39]. Each group defends a territory of 50–80 hectares year-round [42].

The study population has been monitored since 2003 and is habituated to human presence [38], which allows researchers to present the birds with cognitive tasks. Ringing and blood sampling for molecular sexing are performed on nestlings 11 days post-hatching [43]. Therefore, each bird in the study population is identifiable by a unique ring combination, and sex and age are known for most adult birds. Adult immigrants are trapped with a walk-in trap for ringing and blood sampling. We considered immigrants to be at least 1 year old at the time they immigrated into our study population, and if they immigrated as dominants and bred on the first year in which they immigrated we considered them to be at least 2 years old, as dispersal and first breeding are rarely recorded before these ages, respectively [41,44]. On average, subordinate individuals are younger than dominants [44]. Rank (dominant versus subordinate) is easily inferred from aggressive displays by dominant individuals towards subordinates [38], affiliative behaviours between dominants [41], and overnight incubation by the dominant female [38]. During the study (2018–2021), the population comprised 14 groups (average group size 4 ± 1 s.d., range 2–7 adults). We tested different individuals each year: 13 individuals from six groups in 2018, 18 from 10 groups in 2019 and seven from four groups in 2021.

(b) Cognitive test battery

The cognitive test battery consisted of three tasks designed to quantify (i) associative learning, (ii) reversal learning, (iii) inhibitory control. These cognitive tasks tapped into the natural terrestrial foraging behaviour of babblers [38], as they required them to peck downwards at a lid or move around a barrier on the ground to retrieve a food reward (a mealworm, *Tenebrio molitor* larva). The original cognitive test battery included a spatial memory task, but this was later excluded because individuals' behaviour when interacting with the task did not deviate from a random sampling strategy (see electronic supplementary material, section S3).

Cognitive testing was conducted between 05.00 and 19.00, when babblers were active. Cognitive tasks were always presented in the shade when the birds were not showing any heat dissipation behaviours (i.e. panting and wingspreading) to avoid potential confounding effects of heat stress [45]. All trials in a cognitive test were performed when the focal individual was out of sight of other group members. This meant starting testing when the focal individual was foraging apart from other group members (babblers often forage over 10 m apart from each other [46]) and repositioning the task between trials (each lasting less than 1 min) if a group member approached. The three cognitive tests were carried out at least 24 h apart and the order was randomized within individual, except for the reversal learning, which was always carried out the day after the associative learning test. Prior to quantifying learning performance, individuals were trained to peck the lids in a cognitive task to find the reward using unpainted lids (electronic supplementary material, section S1). In all tasks, if the focal bird did not interact with the task for 30 min, the test was paused and continued the following day; when this occurred, the count of correct and incorrect trials was carried over from one session to the following. If the passing criterion was not reached by 120 trials, the test was stopped and individual performance was recorded as 120 trials.

(i) Associative and reversal learning

The task used to quantify associative and reversal learning consisted of a small wooden block (180 × 70 × 30 mm) with two circular wells (30 mm diameter, 20 mm depth) covered by painted wooden lids. The lids were held in place by elastic bands; in this way, they fitted snugly into the wells, preventing the bird from using visual cues to identify the rewarded well, but they could swivel when pecked, making the food reward accessible (figure 1a). The two lids were painted a dark and light shade of the same colour rather than two different colours to avoid confounding effects of individual colour preferences due to past experience on learning performance (e.g. [35,47]; hereafter 'colours' instead of 'colour shades' for brevity). Each day before the start of cognitive testing, two mealworms were temporarily placed in both wells of the cognitive task to prevent the bird from relying on olfactory cues to choose the rewarded well during testing. The associative and reversal learning tests followed the protocols used by Shaw *et al.* [35] and Ashton *et al.* [5]. One of the two colours was randomly assigned to be the rewarded colour for each test bird. In each trial, the first peck of the individual when approaching the task was counted as correct (1 = rewarded lid) or incorrect (0 = unrewarded lid). During the first trial, we waited for all individuals to search both wells to see that only one hid the food reward, in subsequent trials, if the individual chose incorrectly the task was removed before the individual could gain the reward. The position of the rewarded well was pseudorandomized between trials to ensure the individual associated the colour and not the position of the lid with the reward. Associative learning performance was quantified as the number of trials required to reach the passing criterion, which was six correct choices in a row (a significant deviation from a random binomial probability: binomial test $p = 0.016$). If the bird passed the associative learning task, the reversal learning task was carried out 24 h after. Reversal learning performance was quantified using the same protocol and passing criterion used for associative learning, but rewarding the opposite colour.

(ii) Inhibitory control

We quantified inhibitory control using a detour-reaching task, which consisted of a transparent barrier (clear smooth PVC, 200 μ m thick) fixed onto a wooden base, with a mealworm positioned approximately 2 cm behind the barrier on the wooden base (figure 1b). The task was presented to the individual straight on so that the mealworm was visible behind the barrier, but not accessible from the direction that the individual was approaching the task. In this task, the individual had to inhibit the prepotent response of pecking the barrier when seeing the food reward and instead detour around the barrier to retrieve it. A trial was marked as correct if the individual retrieved the mealworm without pecking the barrier. We did not include a training phase (see electronic supplementary material, section S1) but during the first trial we allowed the individual to eat the mealworm even if it pecked the barrier before detouring to make sure that the individual was aware of the presence of a barrier. The passing criterion was six correct trials in a row and the performance measure was the number of trials to criterion.

Detour-reaching tasks have recently been criticized because performance measures can be influenced by confounding factors, such as (i) past experience, (ii) body condition, (iii) neophobia or (iv) persistence [48]. However, our protocol allowed us to control for or test these factors: (i) none of the tested individuals had prior experience with a transparent barrier; (ii) we quantified individual body mass prior to cognitive testing; (iii) we quantified latency to approach the task, a proxy of neophobia, during testing (see S2c) and (iv) we measured performance as the number of trials taken to reach six correct detours around the

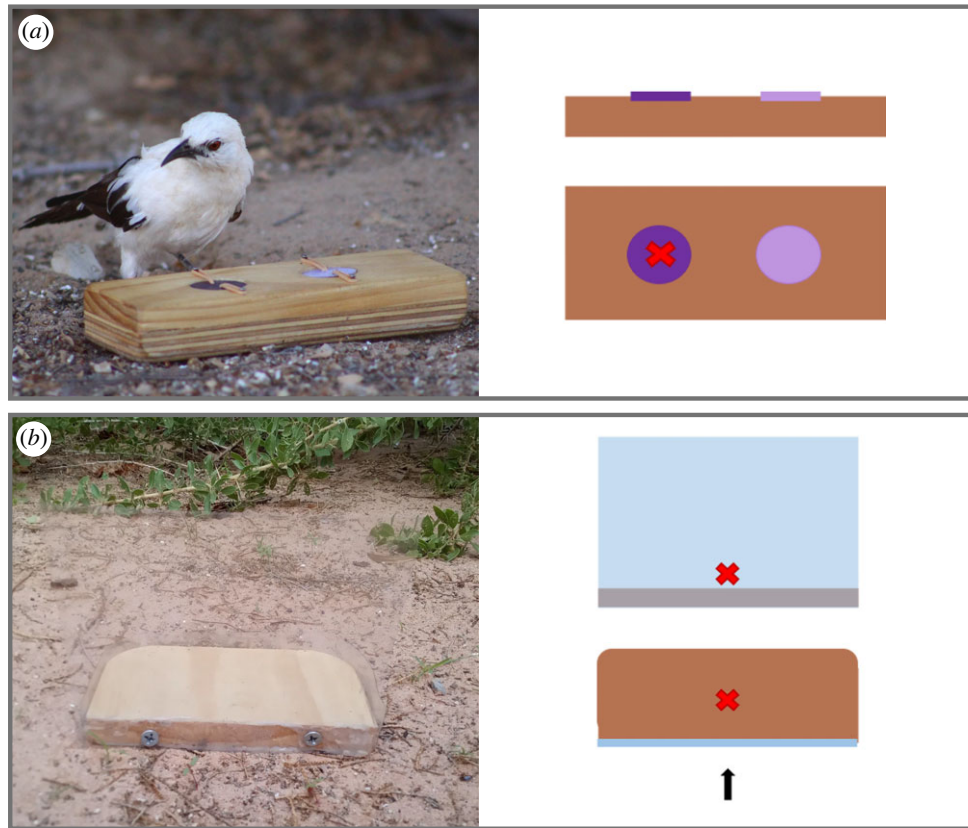


Figure 1. (a) Wild pied babbler interacting with the cognitive task used to quantify associative and reversal learning (left); and schematic of the associative and reversal learning task (right) showing both the overhead and front view. (b) Example of task used to quantify inhibitory control (left) and schematic of the inhibitory control task (right), which includes a frontal view and an overhead view with a black arrow indicating the direction of approach. The red X represents the location of the food reward: behind the barrier in the inhibitory control task and in the well of the rewarded colour shade (assigned randomly for each test subject) in the associative or reversal learning task. Photo credits: Nicholas Pattinson. (Online version in colour.)

barrier instead of the number of pecks to the barrier (the latter may be confounded by motivation and/or persistence).

(iii) Task variants

This study was conceived as part of a long-term project that involved repeatedly quantifying individual cognitive performance (see electronic supplementary material, section S7). To control for the potentially confounding effects of memory on cognitive performance, causally identical but visually distinct variants of each task (i.e. different colour combinations or shapes of the transparent barrier) were used for different replicates of the cognitive test battery over the course of the project [49]. The variant of the cognitive task used in a given replicate was assigned randomly to each individual. Task variant did not significantly affect the number of trials taken to pass the associative and reversal learning tasks nor the inhibitory control task (see electronic supplementary material, section S2). Therefore, in this study, we used individual cognitive performance measures obtained from the first replicate of the test battery, which comprised different task variants (e.g. associative learning might have been measured using dark/light green lids for one individual and dark/light purple lids for another).

(c) Proxies of motivation

Performance in cognitive tasks may be influenced by the motivation of the individual to interact with the task [11]. When completing a cognitive task based on a food reward, individual performance might vary depending on hunger level and amount of food available in the environment. For this reason, we measured several proxies of motivation: foraging efficiency, body mass, latency to approach the task and inter-trial interval.

Weekly 20-min behavioural focal observations were carried out for all the individuals tested. Focal observations were conducted by continuously recording the behaviours of the individual (to the nearest second) using a customized program created in the free software Cybertracker. Foraging efficiency (grams of biomass consumed per foraging minute; following [50]) was calculated from focal observations comprising at least 5 min of foraging. As a previous study found babblers forage more efficiently in the early morning [51], we paired the timing of the focal observation and cognitive testing by performing both either in the early morning (before 09.00) or later in the day (after 09.00). Additionally, we measured the body mass of each individual (accuracy 0.1 g) within the 4 h prior to each cognitive test by enticing the individual to jump on a top-pan scale to retrieve a mealworm [52]. Finally, we measured the latency to approach the task as the time elapsed between the focal individual being within 5 m of the task and first making contact with the task [5], and the average inter-trial interval, which quantifies how quickly the individual approached the task again after a trial (electronic supplementary material, section S4).

(d) Measures of reproductive success

Since 2003, each year during the breeding season (September–March) researchers perform weekly visits to the babbler groups during which the number and identity of individuals (adults, fledglings, juveniles) and any breeding activity are noted. Nests are located by observing nest building, and accurate hatch and fledge dates are recorded by checking the nests every 2–3 days. If fledglings are missing after two consecutive visits to the group, they are considered dead. We assumed only dominant individuals bred [39]; therefore, the offspring produced in each

breeding attempt were attributed to the dominant male and female in the group at the time the attempt was recorded. The extensive life-history database allowed us to determine the number of fledglings produced, the number of fledglings surviving to independence (i.e. 90 days post-hatching, when offspring receive less than one feed per hour [52]), and the number of fledglings recruited (i.e. surviving to 1 year post-hatching [53]) per year for each dominant individual.

(e) Statistical analyses

All analyses were performed with R statistical software v.4.2.0 [54]. We fitted different sets of generalized linear mixed models (GLMMs) using the *lmerTest* package [55] and tested the relative importance of different candidate explanatory terms by ranking them by Akaike information criterion corrected for small sample sizes (AICc). Models within two Δ AICc of the best model and with predictors whose 95% confidence intervals (CIs) did not intersect zero were included in the top model set, and were considered to explain variation in the dependent variable better than other candidate models [56]. Continuous predictors were scaled by centring on the mean and dividing by 1 s.d. Normality of residuals, presence of outliers and dispersion were checked using the *DHARMA* package [57]. To identify the minimum determinable effect of two-way interactions given our sample sizes [58], we conducted a power analysis with the *power* package [59].

(i) Relationships between individual cognitive performances across tasks

First, we tested whether cognitive performance was correlated across tasks by performing Spearman's rank correlations on the scores (i.e. number of trials to pass) of each pair of tasks. Note that a lower score in this case indicates fewer trials to pass the task, and hence, better cognitive performance. To determine whether individual performance in different tasks could be explained by a single factor (GCP), we then performed an unrotated principal component analysis (PCA) on the scores of the associative learning, reversal learning and inhibitory control tasks, using the *FactoMineR* package [60]. Following Shaw *et al.* [35], to test whether the mean and standard deviation of the loadings onto the first principal component (PC1) deviated from what is expected by chance, we performed 10 000 PCA simulations using the function *randomizeMatrix* in the *picante* package [61]. For each simulation, the cognitive scores within each task were randomized among individuals and a PCA was performed. We then compared the real mean and standard deviation of the loadings onto PC1 to the 95% CIs of the simulated means and standard deviations of the loadings onto PC1.

(ii) Factors explaining interindividual variation in cognitive performance

To determine whether individual and group attributes or proxies of motivation explained interindividual variation in cognitive performance, we fitted LMMs containing group identity as a random term and GCP as dependent variable, where GCP was the individual coordinate along PC1 but with the opposite sign so that higher values corresponded to higher GCP. We used GCP as a measure of individual cognitive performance because performance on all tasks loaded strongly and positively onto PC1 (§2e(i)). We considered as candidate explanatory terms: (i) individual and group attributes, i.e. age, sex, rank and group size; (ii) proxies of motivation, i.e. average latency to approach (s), inter-trial interval (min), body mass (g) and foraging efficiency ($g \text{ min}^{-1}$), all of which were averaged across the three tasks used to compute GCP; (iii) time of day; (iv) study year (2018, 2019 or 2021). We also included 'testing order' as a candidate

explanatory term to confirm that individuals tested later within a group did not outperform those tested earlier (i.e. social learning was not occurring, *sensu* [5]). Overall, the model set included separate LMMs for each of the candidate explanatory terms listed above (i.e. age, sex, rank, group size, latency, inter-trial interval, body mass, foraging efficiency, testing order, time of day and study year) and for all possible pairwise additive models and pairwise interactions among sex, age, rank, group size, body mass and study year. Six individuals were excluded from this analysis due to unknown sex (molecular sexing not completed).

(iii) The relationship between cognitive performance and reproductive success

When analysing reproductive success, we considered only dominant individuals [39]. We included two individuals that were subordinates in the early years of testing but were retested once they gained dominance (total $N = 19$ dominant individuals). First, we checked whether the individual attributes that determine variation in cognitive performance, i.e. age and sex (based on the results of §2e(ii)), were also associated with variation in the average number of fledglings produced per year since 2 years of age, which is the earliest age at which individuals in our dataset bred (see electronic supplementary material, section S9). Hence, we determined if individual cognitive performance was directly related to reproductive success. We considered three measures of reproductive success: number of fledglings per year, number of independent offspring per year and number of offspring recruited per year. When there were multiple breeding attempts within a year we used cumulative numbers, and we assigned a 0 for years in which dominant individuals did not successfully breed. The average number of years with breeding data per dominant individual tested was 4.7 (range 1–11 years), where a year encompasses the austral breeding season (from September of one year to August of the next year). For each of the three measures of reproductive success we fitted a set of GLMMs with a Poisson error distribution and year and individual ID as random terms. Group ID was not included in these models as a random term because it resulted in overfitting (singular fit), but a parallel analysis including group ID instead of individual ID confirmed the results presented here. Each model set included separate GLMMs for GCP, age, sex, group size and drought (1 = drought versus 0 = no drought during the breeding season, where drought is defined as rainfall less than or equal to 137 mm [53]). We also included the interaction between GCP and sex to test if the relationship between cognition and reproduction differed in males and females.

3. Results

(a) Relationships between individual cognitive performance across tasks

The 38 tested babblers completed the associative learning, reversal learning and inhibitory control tasks in a mean of 39.26 trials (range 6–120), 63.18 trials (range 6–120) and 34.58 trials (range 6–105), respectively. The range in the number of trials taken to pass the tasks indicates considerable intraspecific variation in cognitive performance (see electronic supplementary material, section S5 for a comparison with studies using similar test batteries on wild birds). We found positive correlations in cognitive performance for all pairwise comparisons across tasks, but only the correlation between associative and reversal learning performance was significant (associative and reversal learning: Spearman's $\rho = 0.65$, $p < 0.001$; reversal learning and inhibitory control: Spearman's $\rho = 0.29$, $p = 0.08$; associative learning and inhibitory control: Spearman's $\rho =$

Table 1. Output of the principal component analysis on the scores (i.e. number of trials to pass) obtained by 38 pied babblers on three cognitive tasks quantifying associative learning, reversal learning and inhibitory control.

cognitive task	PC1
associative learning	0.82
reversal learning	0.88
inhibitory control	0.58
eigenvalue	1.79
% variance explained	59.54

0.14, $p = 0.42$). The consistent positive direction of pairwise correlations between tasks aligns with the output of the PCA, which showed that all cognitive scores loaded positively onto PC1 extracted with an eigenvalue over 1 (table 1). PC1 explained 59.5% of the total variation in cognitive performance across tasks (table 1).

The PCA results were highly unlikely to occur by chance because the real mean loading onto PC1 was higher than the 95% CI of the randomly simulated mean loadings (95% CI of simulated means for PC1 = 0.01–0.67, real mean = 0.76; electronic supplementary material, figure S3), and while the real s.d. was within the 95% CI of the simulated s.d., it was at the lower end of the distribution (95% CI of simulated s.d. for PC1 = 0.08–0.83, real s.d. = 0.16; electronic supplementary material, figure S3). Additionally, when we examined the cognitive scores of 18 individuals that were tested twice on the cognitive test battery during 2018–2021, individual scores from the second replicate of the cognitive test battery also loaded positively to PC1, which explained 46.3% of the total variance in cognitive performance, providing further evidence for GCP; importantly, GCP was significantly repeatable ($R = 0.50$; s.e. = 0.18; 95% CI = 0.09; 0.78; $p = 0.015$) (see electronic supplementary material, section S7).

(b) Factors explaining interindividual variation in cognitive performance

The factors that best explained variation in GCP were age and sex (table 2). GCP declined with age in females but not males (females: coefficient \pm s.e. = -0.77 ± 0.20 , 95% CI = -1.18 ; -0.37 , males: coefficient \pm s.e. = 0.13 ± 0.24 , 95% CI = -0.36 ; 0.62 ; $N = 32$, of which 16 females and 16 males; figure 2). Group size was not a significant predictor of GCP. Importantly, study year and the proxies of motivation examined (i.e. latency to approach, inter-trial interval, body mass, foraging efficiency, time of day) did not significantly explain variation in GCP (electronic supplementary material, table S3). The power analysis confirmed that we had enough power to detect moderate effects of main terms (Cohen's $f^2 = 0.26$) and large effects of two-way interactions (Cohen's $f^2 = 0.39$) given $N = 32$ individuals.

(c) The relationship between general cognitive performance and reproductive success

The average number of fledglings produced per year since age 2 tended to increase with age in females but not in males (electronic supplementary material, section S9). Hence, in females,

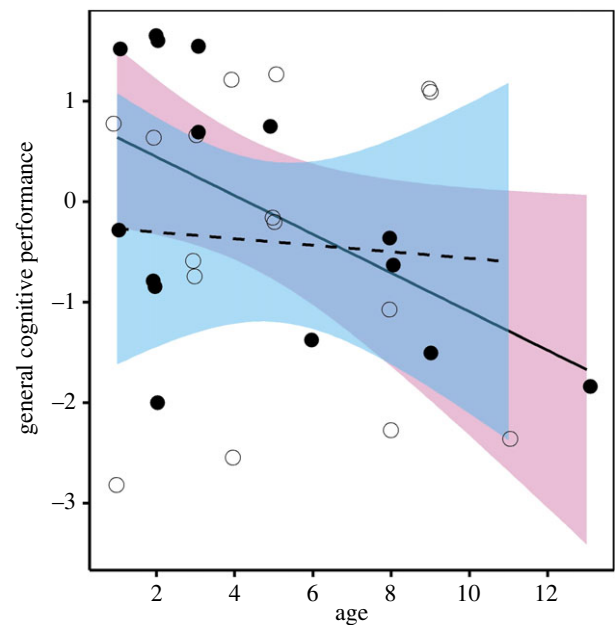


Figure 2. Variation in pied babblers' general cognitive performance by age and sex (females: pink colour, solid line, filled dots, $N = 16$; males: blue colour, dashed line, empty dots, $N = 16$). Points are raw data; fitted lines and 95% confidence bands are generated from the output of the model presented in table 2. (Online version in colour.)

Table 2. Top model set of candidate terms affecting general cognitive performance in pied babblers. All models included group ID as a random term. AICc and Δ AICc are provided for models within 2 Δ AICc of the top model and with predictors whose 95% CI do not intersect zero. Coefficient estimates \pm s.e. and 95% CI are given below the top model set. $N = 32$ individuals from 11 groups. See electronic supplementary material, table S3 for full model selection outputs.

top model set	AICc	Δ AICc
age \times sex	103.53	0.00
null model (intercept only)	109.46	5.93
effect size of explanatory terms	estimate \pm s.e.	95% CI
age	-0.76 ± 0.19	-1.14 ; -0.37
sex (male)	-0.48 ± 0.28	-1.05 ; 0.09
age \times sex (male)	0.90 ± 0.29	0.29 ; 1.49

the relationship between reproductive success and age followed an opposite trend compared to the relationship between GCP and age: older females tended to produce more fledglings per year on average but showed lower GCP. In line with this result, we found that individual GCP was negatively related with the number of fledglings produced per year (table 3a, figure 3). We did not find evidence that this relationship differed in males and females (non-significant interaction GCP \times sex), but we only had power to detect very large effects of two-way interactions (Cohen's $f^2 = 0.75$) given $N = 19$ individuals.

The main predictor of the number of fledglings surviving to independence was the occurrence of droughts, with more fledglings reaching nutritional independence in non-drought years (table 3b). None of the explanatory terms tested were a significant predictor of the number of fledglings surviving to recruitment (table 3c). However, we had to exclude the two

Table 3. Model set of candidate terms affecting three measures of reproductive success in pied babblers. The models included year and individual ID as random terms. Coefficient estimates \pm s.e. and 95% CI are given below the model sets for models within 2 Δ AICc of the top model and with predictors whose 95% CI do not intersect zero. The measures of reproductive success examined over 14 years were: (a) number of fledglings produced per year, $N = 90$ observations for 19 dominant individuals; (b) number of fledglings reaching independence per year, $N = 81$ observations for 18 dominant individuals; (c) number of fledglings recruited per year, $N = 79$ observations for 14 dominant individuals.

model selection	AICc	Δ AICc
(a) number of fledglings per year		
GCP	356.08	0.00
GCP \times sex ^a	357.38	1.30
sex ^a	357.46	1.38
age	358.76	2.68
null model (intercept only)	359.01	2.93
drought	359.75	3.67
group size	361.17	5.09
(b) number of independent offspring per year		
drought	258.70	0.00
null model (intercept only)	264.59	5.89
sex	264.80	6.10
age	266.09	7.39
GCP	266.26	7.56
group size	266.56	7.86
GCP \times sex	269.21	10.51
(c) number of offspring recruited per year		
drought ^a	204.48	0.00
null model (intercept only)	205.73	1.25
age	205.77	1.29
sex	206.05	1.57
GCP	206.74	2.26
group size	207.80	3.32
GCP \times sex	209.27	4.79
effect size of explanatory terms	estimate \pm s.e.	95% CI
(a) number of fledglings per year		
GCP	-0.18 ± 0.07	$-0.34; -0.03$
(b) number of independent offspring per year		
drought	-0.99 ± 0.34	$-1.82; -0.36$

^aNot included in the top model set because 95% CI intersect zero.

dominant females who showed the highest GCP from the latter analysis due to missing data on the number of fledglings surviving to recruitment; therefore, the lack of an effect of GCP on the number of offspring recruited per year should be interpreted with caution.

4. Discussion

We quantified individual cognitive performance in a wild bird population to answer three central questions in cognitive

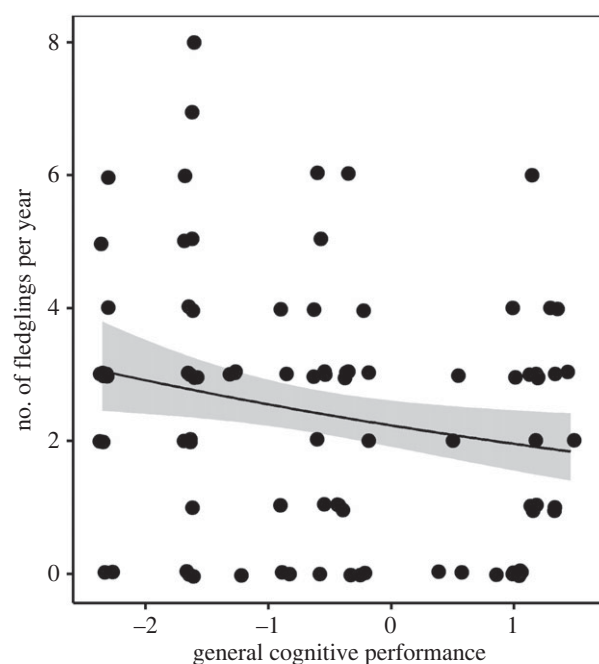


Figure 3. The relationship between the number of fledglings produced per year and general cognitive performance in dominant pied babblers ($N = 90$ observations for 19 dominant individuals over 14 years). Points are raw data; the fitted line and 95% confidence band are generated from the output of the model presented in table 3.

ecology: (i) does performance co-vary across cognitive tasks, (ii) what drives these individual differences and (iii) is individual cognitive performance related to reproductive success.

Individuals differed greatly in their associative learning, reversal learning and inhibitory control performance. Those that learnt an association faster, were also faster at reversing the learnt association and showed better inhibitory control, as shown by positive (albeit not always significant) correlations in cognitive performance across tasks. Indeed, approximately 60% of the variance in individual cognitive performance across tasks could be explained by a single factor: GCP. Additionally, GCP was significantly repeatable ($R = 0.50$), indicating that our measure of GCP captured consistent inter-individual differences in cognition. While we cannot completely exclude that motivation affected cognitive performance, we are confident that its effect on our measure of GCP was minimal because all the birds that were presented with the task interacted with it successfully and always ate the reward; more importantly, none of the measured proxies of motivation significantly explained variation in GCP. Therefore, our findings are consistent with the existence of a general cognitive factor underpinning performance across different cognitive domains in babblers. However, our test battery included only three cognitive tasks, which is the minimum number required to test for a general cognitive factor [30]. Therefore, future studies should consider expanding the test battery by including, for example, tasks assessing social cognition, the ability to make inferences and reaction time [31,35].

Alternative explanations for GCP are also possible. For example, the different tasks used may tap into the same cognitive process [34], so that they are effectively measuring the same cognitive trait instead of a common cognitive factor underpinning performance across domains. Therefore, whether statistical evidence for GCP indicates a truly general

cognitive ability underlying performance across different cognitive domains remains to be determined.

We found no evidence that living in larger groups is linked to cognitive performance in pied babblers. This is in contrast to recent findings on Western Australian magpies (*Gymnorhina tibicen dorsalis*) [5]. The difference between the two species could depend on the degree to which their societies involve individualized relationships [62]. Babblers live in groups with high within-group relatedness and high reproductive skew [39,40]. By contrast, in Western Australian magpies, multiple individuals breed within the group, offspring care by helpers is facultative [63] and within-group relatedness is low [64]. These conditions give more scope for conflicts of interest and hence for a society where individuals may benefit from tracking the outcome of past interactions with other group members, negotiating access to resources and breeding opportunities, and engaging in strategic decision-making [65]. Therefore, rather than social living *per se*, we must consider the extent to which social systems present individuals with information-processing challenges, such as the need to navigate multiple differentiated social relationships [66]. To test this, future cognition studies should quantify the amount and diversity of social interactions among group members through social network analysis.

Individual variation in GCP was predicted by an interaction between age and sex, with cognitive performance declining with age in females but not in males. Faster cognitive ageing in females has been previously reported in humans [67], nematodes (*Caenorhabditis remanei*) [68], mice (*Mus musculus*) [69] and marmosets (*Callithrix jacchus*) [24]. However, the only study testing for cognitive senescence in the wild found no decline in spatial memory performance in mountain chickadees (*Poecile gambeli*) [25]. Hence, to our knowledge, our finding represents the first evidence of sex differences in age-related cognitive decline in a wild animal.

Senescence has been explained by two main evolutionary theories [70]. The 'selection shadow' theory states that selection strength decreases with age after sexual maturity [71]. Our data do not support this theory because babblers were still breeding up to 13 years of age, leaving ample opportunity for selection to act on cognitive traits among older individuals. A second theory is the life-history theory of ageing, which encompasses two convergent theories: the first states that due to the limited resources available to organisms, these must be traded-off between reproduction and somatic maintenance (disposable soma theory) [72]; the second states that alleles with beneficial effects early in life but detrimental effects later in life can be favoured by selection (antagonistic pleiotropy) [73]. Previous studies on babbler life history [44,74] provide some support for both these theories.

First, female babblers (but not males) engage in costly breeding competition [74,75]. Subordinate females compete both indirectly, by courting and nest building with unrelated dominant males, and directly, by destroying the eggs of the dominant female [74]. This competition forces dominant (older) females to engage in frequent aggressive displays towards subordinate (younger) females and repeatedly re-lay clutches [74], which entails an additional energetic cost [76]. Hence, in older (dominant) females the cost of maintaining a high reproductive output, even in the presence of competitors, might be traded-off against the maintenance of the energetically costly nervous system [13]. For example, experiments in the fruit fly and the cabbage white butterfly

(*Pieris rapae*) have revealed a trade-off between learning performance and competitive ability [77] or female fecundity [78], respectively. In line with this explanation, when analysing long-term reproductive success in dominant babblers, we found that higher cognitive performance was associated with a lower number of fledglings produced per year. However, a larger sample size will be necessary to test whether this negative relationship between cognition and reproduction differs between males and females.

Secondly, babblers show sex differences in dispersal strategies. Females are more likely than males to gain a breeding position by overthrowing a dominant female in a non-natal group [44]. Accordingly, juvenile females are more aggressive than males, and higher female aggressiveness is associated with younger age at dispersal [79]. On the contrary, males are more sedentary and disperse only when search costs are low [80]. It is possible that these sex differences lead to selection on females for higher cognitive performance early in life, even at the expenses of faster cognitive senescence. Indeed, cognitive performance in young females might be crucial to gain access to breeding positions by enabling them to navigate across territories, identify the sex and rank of conspecifics in non-natal groups, and decide when to engage in aggressive displays [81]. Additionally, in babblers the number of immigrant competitors decreases with pair bond tenure, while reproductive success increases [41]. This suggests that on average the risk of losing a breeding position and thus, potentially, the need to maintain high cognitive performance might decrease with female age. Overall, our findings paired with evidence from previous research in babblers suggest that females may be under selection for higher cognitive performance earlier in life despite faster cognitive senescence and/or cognitive senescence may be accelerated by investment in reproduction and breeding competition. However, longitudinal studies are ultimately needed to describe cognitive ageing trajectories in males and females and clarify how individual rank interacts with age to determine cognitive decline.

Since we used a cross-sectional design, we cannot determine whether cognitive performance declined throughout life in females, or whether only females with lower cognitive performance survived until old ages. Therefore, a third potential explanation for the observed sex differences in age-related cognitive decline is that cognitive performance is negatively linked to survival, at least in females. For example in pheasants (*Phasianus colchicus*), survival in the wild was negatively related to reversal learning performance [82]. As most of the birds tested in the present study are still alive to date, we could not perform a survival analysis to examine the relationship between cognitive performance and survival, but this will be necessary to confirm whether the present findings are due to cognitive senescence or reduced survival of smarter females.

We found that individuals with better GCP produced fewer fledglings per year, which is consistent with a trade-off between cognition and reproduction. GCP did not predict the number of fledglings surviving to nutritional independence, which depended instead on the occurrence of droughts during the breeding season, in line with a previous study [83]. It is possible that parental traits influence offspring survival in the nestling stage but not in the post-fledgling stage, where survival may be more strongly influenced by environmental conditions [43]. Therefore, the extent to which

cognitive performance may be under negative selection remains to be determined. Cognitive performance might also be simultaneously associated with life-history traits linked to fitness in different directions [84]. For example, while cognitive performance is negatively related to the number of fledglings produced per year by dominant individuals, it might be positively related to the age at which dominance is acquired in the first place. Future comparisons of the age at acquisition of dominance among individuals whose cognition was tested as subordinates will allow us to address this hypothesis.

Overall, we found that individual cognitive performance covaried across tasks, which is consistent with a general cognitive factor, though alternative explanations cannot be excluded. We considered the effect of individual and social attributes and several proxies of motivation on cognitive performance. We found that GCP depended on sex and age, declining with age in females but not males. Older females also tended to fledge more nestlings per year. By analysing over 10 years of breeding data, we show that individuals with lower GCP produced more fledglings per year. Our findings suggest that cognitive performance is traded-off against reproduction, demonstrating that in order to understand how selection acts on cognition we need to consider not only its benefits but also its costs.

Ethics. This research was approved by the Animal Ethics Committee, University of Western Australia (RA/3/100/1619 and 2021/ET000414).

Data accessibility. Data are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.d51c5b06v> [85].

Additional information is provided in the electronic supplementary material [86].

Authors' contributions. C.S.: conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing—original draft, writing—review and editing; B.J.A.: conceptualization, formal analysis, funding acquisition, methodology, supervision, writing—review and editing; A.T.: conceptualization, funding acquisition, writing—review and editing; A.R.R.: conceptualization, formal analysis, funding acquisition, methodology, project administration, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was funded by the Australian Government Research Training Program through a scholarship awarded to C.S. at the University of Western Australia. The Kuruman River Reserve (KRR), the study site where this research was based, was financed by the Universities of Cambridge and Zurich, the MAVA Foundation and the European Research Council (grant no. 294494 to Tim Clutton-Brock) and received logistical support from the Mammal Research Institute of the University of Pretoria.

Acknowledgements. We thank the managers at the Kuruman River Reserve (KRR) and the families at the surrounding farms for making this work possible. We thank Amanda Bourne, Lina Peña-Ramirez, Grace Blackburn, Amy Hunter and Samantha Wagstaff for assistance during data collection. We are grateful to the Hot Birds Research Project for access to the KRR rainfall data and to Amanda Bourne for calculation of drought/non-drought years. We thank the FitzPatrick Institute of African Ornithology for ongoing logistical support. We also thank Prof. Leigh Simmons for the advice given to C.S. as part of her supervisory team. We thank two anonymous reviewers for their insightful comments that greatly improved this manuscript.

References

- Shettleworth SJ. 2010 *Cognition, evolution, and behavior*. Oxford, UK: Oxford University Press.
- Szabo B, Valencia-Aguilar A, Damas-Moreira I, Ringler E. 2022 Wild cognition—linking form and function of cognitive abilities within a natural context. *Curr. Opin. Behav. Sci.* **44**, 101115. (doi:10.1016/j.cobeha.2022.101115)
- Bond AB, Kamil AC, Balda RP. 2007 Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus*, *Nucifraga columbiana*, *Aphelocoma californica*). *J. Comp. Psychol.* **121**, 372–379. (doi:10.1037/0735-7036.121.4.372)
- MacLean EL *et al.* 2014 The evolution of self-control. *Proc. Natl Acad. Sci. USA* **111**, E2140–E2148. (doi:10.1073/pnas.1318476111)
- Ashton BJ, Ridley AR, Edwards EK, Thornton A. 2018 Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature* **554**, 364–367. (doi:10.1038/nature25503)
- Branch CL, Pitera AM, Kozlovsky DY, Bridge ES, Pravosudov VV. 2019 Smart is the new sexy: female mountain chickadees increase reproductive investment when mated to males with better spatial cognition. *Ecol. Lett.* **22**, 897–903. (doi:10.1111/ele.13249)
- Croston R, Branch CL, Kozlovsky DY, Dukas R, Pravosudov VV. 2015 Heritability and the evolution of cognitive traits. *Behav. Ecol.* **26**, 1447–1459. (doi:10.1093/beheco/arv088)
- Chen J, Zou Y, Sun Y-H, ten Cate C. 2019 Problem-solving males become more attractive to female budgerigars. *Science* **363**, 166–167. (doi:10.1126/science.aau8181)
- Shaw RC, MacKinlay RD, Clayton NS, Burns KC. 2019 Memory performance influences male reproductive success in a wild bird. *Curr. Biol.* **29**, 1498–1502. (doi:10.1016/j.cub.2019.03.027)
- Sonnenberg BR, Branch CL, Pitera AM, Bridge E, Pravosudov VV. 2019 Natural selection and spatial cognition in wild food-caching mountain chickadees. *Curr. Biol.* **29**, 670–676.e673. (doi:10.1016/j.cub.2019.01.006)
- Boogert NJ, Madden JR, Morand-Ferron J, Thornton A. 2018 Measuring and understanding individual differences in cognition. *Phil. Trans. R. Soc. B* **373**, 20170280. (doi:10.1098/rstb.2017.0280)
- Burger JM, Kolss M, Pont J, Kawecki TJ. 2008 Learning ability and longevity: a symmetrical evolutionary trade-off in *Drosophila*. *Evolution* **62**, 1294–1304. (doi:10.1111/j.1558-5646.2008.00376.x)
- Heldstab SA, Isler K, Graber SM, Schuppli C, van Schaik CP. 2022 The economics of brain size evolution in vertebrates. *Curr. Biol.* **32**, R697–R708. (doi:10.1016/j.cub.2022.04.096)
- Ashton BJ, Thornton A, Ridley AR. 2018 An intraspecific appraisal of the social intelligence hypothesis. *Phil. Trans. R. Soc. B* **373**, 20170288. (doi:10.1098/rstb.2017.0288)
- Pravosudov VV, Roth II TC, LaDage LD, Freas CA. 2015 Environmental influences on spatial memory and the hippocampus in food-caching chickadees. *Comp. Cogn. Behav. Rev.* **10**, 25–43. (doi:10.3819/cabr.2015.100002)
- Coppola VJ, Hough G, Bingman VP. 2014 Age-related spatial working memory deficits in homing pigeons (*Columba livia*). *Behav. Neurosci.* **128**, 666–675. (doi:10.1037/bne0000013)
- Boogert NJ, Reader SM, Laland KN. 2006 The relation between social rank, neophobia and individual learning in starlings. *Anim. Behav.* **72**, 1229–1239. (doi:10.1016/j.anbehav.2006.02.021)
- Lucon-Xiccato T, Bisazza A. 2017 Sex differences in spatial abilities and cognitive flexibility in the guppy. *Anim. Behav.* **123**, 53–60. (doi:10.1016/j.anbehav.2016.10.026)
- Jones CM, Braithwaite VA, Healy SD. 2003 The evolution of sex differences in spatial ability. *Behav. Neurosci.* **117**, 403–411. (doi:10.1037/0735-7044.117.3.403)
- Guigueno MF, Snow DA, MacDougall-Shackleton SA, Sherry DF. 2014 Female cowbirds have more accurate spatial memory than males. *Biol. Lett.* **10**, 20140026. (doi:10.1098/rsbl.2014.0026)
- Franks VR, Thorogood R. 2018 Older and wiser? Age differences in foraging and learning by an endangered passerine. *Behav. Process.* **148**, 1–9. (doi:10.1016/j.beproc.2017.12.009)

22. Mirville MO, Kelley JL, Ridley AR. 2016 Group size and associative learning in the Australian magpie (*Cracticus tibicen dorsalis*). *Behav. Ecol. Sociobiol.* **70**, 417–427. (doi:10.1007/s00265-016-2062-x)
23. Yu L, Tucci V, Kishi S, Zhdanova IV. 2006 Cognitive aging in zebrafish. *PLoS ONE* **1**, e14. (doi:10.1371/journal.pone.0000014)
24. Rothwell ES, Workman KP, Wang D, Lacreuse A. 2022 Sex differences in cognitive aging: a 4-year longitudinal study in marmosets. *Neurobiol. Aging* **109**, 88–99. (doi:10.1016/j.neurobiolaging.2021.09.015)
25. Heinen VK, Pitera AM, Sonnenberg BR, Benedict LM, Branch CL, Bridge ES, Pravosudov VV. 2021 Specialized spatial cognition is associated with reduced cognitive senescence in a food-caching bird. *Proc. R. Soc. B* **288**, 20203180. (doi:10.1098/rspb.2020.3180)
26. Barnard C, Luo N. 2002 Acquisition of dominance status affects maze learning in mice. *Behav. Process.* **60**, 53–59. (doi:10.1016/S0376-6357(02)00121-3)
27. Thornton A, Samson J. 2012 Innovative problem solving in wild meerkats. *Anim. Behav.* **83**, 1459–1468. (doi:10.1016/j.anbehav.2012.03.018)
28. Dehnen T, Arbon JJ, Farine DR, Boogert NJ. 2022 How feedback and feed-forward mechanisms link determinants of social dominance. *Biol. Rev.* **97**, 1210–1230. (doi:10.1111/brv.12838)
29. Keynan O, Ridley AR, Lotem A. 2016 Task-dependent differences in learning by subordinate and dominant wild Arabian babblers. *Ethology* **122**, 399–410. (doi:10.1111/eth.12488)
30. Burkart JM, Schubiger MN, van Schaik CP. 2017 The evolution of general intelligence. *Behav. Brain Res.* **40**, e195.
31. Flaim M, Blaisdell AP. 2020 The comparative analysis of intelligence. *Psychol. Bull.* **146**, 1174. (doi:10.1037/bul0000306)
32. Shaw RC, Schmelz M. 2017 Cognitive test batteries in animal cognition research: evaluating the past, present and future of comparative psychometrics. *Anim. Cogn.* **20**, 1003–1018. (doi:10.1007/s10071-017-1135-1)
33. van Horik JO, Langley EJ, Whiteside MA, Laker PR, Madden JR. 2018 Intra-individual variation in performance on novel variants of similar tasks influences single factor explanations of general cognitive processes. *R. Soc. Open Sci.* **5**, 171919. (doi:10.1098/rsos.171919)
34. van Horik JO, Lea SEG. 2017 Disentangling learning from knowing: does associative learning ability underlie performances on cognitive test batteries? *Behav. Brain Res.* **40**, e220.
35. Shaw RC, Boogert NJ, Clayton NS, Burns KC. 2015 Wild psychometrics: evidence for ‘general’ cognitive performance in wild New Zealand robins, *Petroica longipes*. *Anim. Behav.* **109**, 101–111. (doi:10.1016/j.anbehav.2015.08.001)
36. Morand-Ferron J, Cole EF, Quinn JL. 2016 Studying the evolutionary ecology of cognition in the wild: a review of practical and conceptual challenges. *Biol. Rev.* **91**, 367–389. (doi:10.1111/brv.12174)
37. Kabadayi C, Bobrowicz K, Osvath M. 2018 The detour paradigm in animal cognition. *Anim. Cogn.* **21**, 21–35. (doi:10.1007/s10071-017-1152-0)
38. Ridley AR. 2016 Southern pied babblers: the dynamics of conflict and cooperation in a group-living society. In *Cooperative breeding in vertebrates: studies of ecology, evolution and behavior*, pp. 115–132. Cambridge, UK: Cambridge University Press.
39. Nelson-Flower MJ, Hockey PA, O’Ryan C, Raihani NJ, Du Plessis MA, Ridley AR. 2011 Monogamous dominant pairs monopolize reproduction in the cooperatively breeding pied babbler. *Behav. Ecol.* **22**, 559–565. (doi:10.1093/beheco/arr018)
40. Nelson-Flower MJ, Flower TP, Ridley AR. 2018 Sex differences in the drivers of reproductive skew in a cooperative breeder. *Mol. Ecol.* **27**, 2435–2446. (doi:10.1111/mec.14587)
41. Wiley EM, Ridley AR. 2018 The benefits of pair bond tenure in the cooperatively breeding pied babbler (*Turdoides bicolor*). *Ecol. Evol.* **8**, 7178–7185. (doi:10.1002/ece3.4243)
42. Humphries DJ, Nelson-Flower MJ, Bell MB, Finch FM, Ridley AR. 2021 Kinship, dear enemies, and costly combat: the effects of relatedness on territorial overlap and aggression in a cooperative breeder. *Ecol. Evol.* **11**, 17 031–17 042. (doi:10.1002/ece3.8342)
43. Ridley AR. 2007 Factors affecting offspring survival and development in a cooperative bird: social, maternal and environmental effects. *J. Anim. Ecol.* **76**, 750–760. (doi:10.1111/j.1365-2656.2007.01248.x)
44. Raihani NJ, Nelson-Flower MJ, Golabek KA, Ridley AR. 2010 Routes to breeding in cooperatively breeding pied babblers *Turdoides bicolor*. *J. Avian Biol.* **41**, 681–686. (doi:10.1111/j.1600-048X.2010.05211.x)
45. Soravia C, Ashton BJ, Thornton A, Ridley AR. 2021 The impacts of heat stress on animal cognition: implications for adaptation to a changing climate. *Wiley Interdiscip. Rev.: Clim. Change* **12**, e713. (doi:10.1002/wcc.713)
46. Hollén L, Bell M, Wade H, Rose R, Russell A, Niven F, Ridley A, Radford A. 2011 Ecological conditions influence sentinel decisions. *Anim. Behav.* **82**, 1435–1441. (doi:10.1016/j.anbehav.2011.09.028)
47. Rowe C, Healy SD. 2014 Measuring variation in cognition. *Behav. Ecol.* **25**, 1287–1292. (doi:10.1093/beheco/aru090)
48. van Horik JO, Langley EJG, Whiteside MA, Laker PR, Beardsworth CE, Madden JR. 2018 Do detour tasks provide accurate assays of inhibitory control? *Proc. R. Soc. B* **285**, 20180150. (doi:10.1098/rspb.2018.0150)
49. Ashton BJ, Thornton A, Cauchoix M, Ridley AR. 2022 Long-term repeatability of cognitive performance. *R. Soc. Open Sci.* **9**, 220069. (doi:10.1098/rsos.220069)
50. Raihani NJ, Ridley AR. 2007 Adult vocalizations during provisioning: offspring response and postfledging benefits in wild pied babblers. *Anim. Behav.* **74**, 1303–1309. (doi:10.1016/j.anbehav.2007.02.025)
51. du Plessis KL, Martin RO, Hockey PA, Cunningham SJ, Ridley AR. 2012 The costs of keeping cool in a warming world: implications of high temperatures for foraging, thermoregulation and body condition of an arid-zone bird. *Glob. Change Biol.* **18**, 3063–3070. (doi:10.1111/j.1365-2486.2012.02778.x)
52. Ridley AR, Raihani NJ. 2007 Variable postfledging care in a cooperative bird: causes and consequences. *Behav. Ecol.* **18**, 994–1000. (doi:10.1093/beheco/arm074)
53. Bourne AR, Cunningham SJ, Spottiswoode CN, Ridley AR. 2020 Hot droughts compromise interannual survival across all group sizes in a cooperatively breeding bird. *Ecol. Lett.* **23**, 1776–1788. (doi:10.1111/ele.13604)
54. R Core Team. 2022 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <https://www.R-project.org/>.
55. Kuznetsova A, Brockhoff PB, Christensen RH. 2017 lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26. (doi:10.18637/jss.v082.i13)
56. Symonds MR, Moussalli A. 2011 A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike’s information criterion. *Behav. Ecol. Sociobiol.* **65**, 13–21. (doi:10.1007/s00265-010-1037-6)
57. Hartig F. 2021 DHARMA: residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.4.4. See <https://CRAN.R-project.org/package=DHARMA>.
58. Cohen J. 1992 Statistical power analysis. *Curr. Dir. Psychol. Sci.* **1**, 98–101. (doi:10.1111/1467-8721.ep10768783)
59. Champely S. 2020 Package ‘pwr’. R package version 1.3.0. See <https://CRAN.R-project.org/package=pwr>.
60. Lê S, Josse J, Husson F. 2008 FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18.
61. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010 Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464. (doi:10.1093/bioinformatics/btq166)
62. Thornton A, McAuliffe K. 2015 Cognitive consequences of cooperative breeding? A critical appraisal. *J. Zool.* **295**, 12–22. (doi:10.1111/jzo.12198)
63. Pike KN, Ashton BJ, Morgan KV, Ridley AR. 2019 Social and individual factors influence variation in offspring care in the cooperatively breeding Western Australian magpie. *Front. Ecol. Evol.* **7**, 92. (doi:10.3389/fevo.2019.00092)
64. Hughes JM, Mather PB, Toon A, Ma J, Rowley I, Russell E. 2003 High levels of extra-group paternity in a population of Australian magpies *Gymnorhina tibicen*: evidence from microsatellite analysis. *Mol.*

- Ecol.* **12**, 3441–3450. (doi:10.1046/j.1365-294X.2003.01997.x)
65. Lukas D, Clutton-Brock T. 2018 Social complexity and kinship in animal societies. *Ecol. Lett.* **21**, 1129–1134. (doi:10.1111/ele.13079)
 66. Kappeler PM. 2019 A framework for studying social complexity. *Behav. Ecol. Sociobiol.* **73**, 1–14. (doi:10.1007/s00265-018-2618-z)
 67. Levine DA *et al.* 2021 Sex differences in cognitive decline among US adults. *AMA Netw. Open.* **4**, e210169. (doi:10.1001/jamanetworkopen.2021.0169)
 68. Zwoinska MK, Kolm N, Maklakov AA. 2013 Sex differences in cognitive ageing: testing predictions derived from life-history theory in a dioecious nematode. *Exp. Gerontol.* **48**, 1469–1472. (doi:10.1016/j.exger.2013.09.008)
 69. Benice T, Rizk A, Kohama S, Pfankuch T, Raber J. 2006 Sex-differences in age-related cognitive decline in C57BL/6J mice associated with increased brain microtubule-associated protein 2 and synaptophysin immunoreactivity. *Neuroscience* **137**, 413–423. (doi:10.1016/j.neuroscience.2005.08.029)
 70. Nussey DH, Froy H, Lemaitre J-F, Gaillard J-M, Austad SN. 2013 Senescence in natural populations of animals: widespread evidence and its implications for bio-gerontology. *Ageing Res. Rev.* **12**, 214–225. (doi:10.1016/j.arr.2012.07.004)
 71. Medawar P. 1952 *An unsolved problem in biology*. London, UK: HK Lewis.
 72. Kirkwood TB, Rose MR. 1991 Evolution of senescence: late survival sacrificed for reproduction. *Phil. Trans. R. Soc. Lond. B* **332**, 15–24. (doi:10.1098/rstb.1991.0028)
 73. Williams GC. 1957 Pleiotropy, natural selection and the evolution of senescence. *Evolution* **11**, 393–411. (doi:10.1111/j.1558-5646.1957.tb02911.x)
 74. Nelson-Flower MJ, Hockey PA, O’Ryan C, English S, Thompson AM, Bradley K, Rose R, Ridley AR. 2013 Costly reproductive competition between females in a monogamous cooperatively breeding bird. *Proc. R. Soc. B* **280**, 20130728. (doi:10.1098/rspb.2013.0728)
 75. Nelson-Flower MJ, Ridley AR. 2015 Male–male competition is not costly to dominant males in a cooperatively breeding bird. *Behav. Ecol. Sociobiol.* **69**, 1997–2004. (doi:10.1007/s00265-015-2011-0)
 76. Williams TD. 2005 Mechanisms underlying the costs of egg production. *Bioscience* **55**, 39–48. (doi:10.1641/0006-3568(2005)055[0039:MUTCOE]2.0.CO;2)
 77. Mery F, Kawecki TJ. 2003 A fitness cost of learning ability in *Drosophila melanogaster*. *Proc. R. Soc. B* **270**, 2465–2469. (doi:10.1098/rspb.2003.2548)
 78. Snell-Rood EC, Davidowitz G, Papaj DR. 2011 Reproductive tradeoffs of learning in a butterfly. *Behav. Ecol.* **22**, 291–302. (doi:10.1093/beheco/arq169)
 79. Raihani NJ, Ridley AR, Browning LE, Nelson-Flower MJ, Knowles S. 2008 Juvenile female aggression in cooperatively breeding pied babblers: causes and contexts. *Ethology* **114**, 452–458. (doi:10.1111/j.1439-0310.2008.01482.x)
 80. Nelson-Flower MJ, Wiley EM, Flower TP, Ridley AR. 2018 Individual dispersal delays in a cooperative breeder: ecological constraints, the benefits of philopatry and the social queue for dominance. *J. Anim. Ecol.* **87**, 1227–1238. (doi:10.1111/1365-2656.12814)
 81. Ashton BJ, Kennedy P, Radford AN. 2020 Interactions with conspecific outsiders as drivers of cognitive evolution. *Nat. Commun.* **11**, 1–9. (doi:10.1038/s41467-019-13993-7)
 82. Madden JR, Langley EJJ, Whiteside MA, Beardsworth CE, van Horik JO. 2018 The quick are the dead: pheasants that are slow to reverse a learned association survive for longer in the wild. *Phil. Trans. R. Soc. B* **373**, 20170297. (doi:10.1098/rstb.2017.0297)
 83. Bourne AR, Cunningham SJ, Spottiswoode CN, Ridley AR. 2020 Compensatory breeding in years following drought in a desert-dwelling cooperative breeder. *Front. Ecol. Evol.* **8**, 190. (doi:10.3389/fevo.2020.00190)
 84. Cole EF, Morand-Ferron J, Hinks AE, Quinn JL. 2012 Cognitive ability influences reproductive life history variation in the wild. *Curr. Biol.* **22**, 1808–1812. (doi:10.1016/j.cub.2012.07.051)
 85. Soravia C, Ashton BJ, Thornton A, Ridley AR. 2022 Data from: General cognitive performance declines with female age and is negatively related to fledging success in a wild bird. Dryad Digital Repository. (doi:10.5061/dryad.d51c5b06v)
 86. Soravia C, Ashton BJ, Thornton A, Ridley AR. 2022 General cognitive performance declines with female age and is negatively related to fledging success in a wild bird. Figshare. (doi:10.6084/m9.figshare.c.6330234)