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

Camilla Soravia[1]; Benjamin J. Ashton[1,2,4]; Alex Thornton[3]; Amanda R. Ridley[1,4]

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

[2]School of Natural Sciences, Macquarie University, Sydney, NSW, Australia

[3]Centre for Ecology and Conservation, University of Exeter, Penryn, UK[4]

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

Camilla Soravia (corresponding author) ORCID: 0000-0002-9795-9677

Benjamin J. Ashton ORCID: 0000-0003-3357-0395

Alex Thornton ORCID: 0000-0002-1607-2047

Amanda R. Ridley ORCID: 0000-0001-5886-0992

ABSTRACT

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 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. General cognitive performance 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.

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

1. INTRODUCTION

1   The mental mechanisms through which animals acquire, process, store and act on information from

2   the environment represent animal cognition [1]. Animals use cognitive mechanisms to adjust their

3   behavioural responses to the environmental and social context, remember the location of resources,

4   and learn which environmental cues indicate presence of food, mates or predators [2]. Different

5   animal species rely on cognitive mechanisms to different extents to solve ecological problems [e.g. 3,

6   4], and even within a species, cognitive performance can vary significantly among individuals [e.g. 5,

7   6]. Such great inter-and intraspecific variation has led to the question: what selective pressures shape

8   the evolution of cognition?

9   Intraspecific studies of animal cognition have shown that in some cases cognitive performance is

10   heritable [7]. Additionally, cognitive performance has been linked to mate choice [8], reproductive

11   investment [6], reproductive success [5, 9], and survival [10]. The existence of differential fitness

12   arising from heritable variation in cognitive performance means cognitive traits can evolve [11].

13   However, better cognitive performance is not always associated with increased fitness, for example,

14   faster learning is associated with reduced longevity in fruit flies (*Drosophila melanogaster*) [12]. This

15   may occur because the energetic costs of enhanced cognitive function lead to a trade-off between

16   resource allocation to cognitive performance and somatic maintenance or reproduction [11, 13].

17   Therefore, we expect selection to favour cognitive performance only when the benefits outweigh the

18   costs [13].

19   To understand how selection acts on cognition, the link between cognitive variation and fitness

20   consequences needs to be identified, as well as the proximate causes of individual variation in

21   cognitive performance. Several factors have been associated with differences in individual cognitive

22   performance, including the physical and social environment [14, 15], and individual attributes such as

23   age [e.g. 16], rank [e.g. 17], and sex [e.g. 18]. Sex differences in cognitive performance often arise as

24   a consequence of mating strategies or sex-specific ecological constraints [19, 20]. For example, in the

25   brood-parasitic brown headed cowbird (*Molothrus ater*), females outperform males on a large-scale

26   spatial memory task, likely because the breeding strategy of this species relies on females finding

27   potential host nests to lay their eggs in [21].

28   Age differences in cognitive performance have been mostly found when comparing juveniles and

29   adults [22, 23]. However, cognitive performance can also change during adult life [24, 25]. The gradual

30   reduction of cognitive function with age is known as cognitive senescence [26]. Evidence for cognitive

31   senescence in non-human animals is largely limited to captive studies [24, 27]. For example, homing

32   pigeons (*Columba livia*) older than 10 years returned more often to feeders that they had just depleted

33   despite them being empty, showing impaired short-term memory [16]. To date, little is known about

34   cognitive senescence in the wild, due to logistical limitations including difficulties of estimating

35   individuals' age [28] and testing cognition in the wild [29].

36   During social interactions, individuals can differ in rank, where dominant individuals often monopolize

37   resources [30]. There is growing evidence that cognitive performance is related to rank, but the

38   direction of this relationship varies across studies [31, 32]. It has been suggested that cognitive

39   performance may not be related to social status *per se*, but to factors correlated to rank, such as

40   vigilance, neophobia or motivation to find alternative food sources [17, 32]. For example, in Arabian

41   babblers (*Argya squamiceps*) subordinates were the first to learn to remove black lids in a novel

42   foraging task, likely because they were more explorative, but dominants, which tend to be older in

43   this species, were better able to generalise the solution to white lids because of experience [33, 34].

44   Finally, individual cognitive performance may also be linked to social group size because living in larger

45   groups may require better cognitive performance in order to monitor the state and actions of group

46   members, remember their identity, and the outcome of past interactions [5, 14]. Despite the growing

47   number of studies investigating intraspecific differences in cognitive performance, these individual

48   and social attributes (rank, sex, age and group size) have rarely been examined simultaneously while

49 controlling for proxies of motivation, and evidence of their relative importance in driving cognitive

50 variation is scarce.

51 If we identify what selection pressures drive variation in cognition, the question of whether these act

52 on each cognitive trait separately, or jointly as part of general cognitive processes, remains. In humans,

53 it has been repeatedly demonstrated that individual performance correlates positively across different

54 cognitive tasks, and approximately 40% of the total variation in performance can be explained by a

55 single general cognitive factor g [35]. This factor, also referred to as general intelligence or intelligence

56 quotient (IQ), predicts important life outcomes, such as occupational attainment, health and longevity

57 [36]. Recently, several studies in non-human animals have also described something akin to a general

58 cognitive factor g explaining between 30%-60% of variation in cognitive performance across a battery

59 of cognitive tasks [reviewed in 37]. The evidence for g provided by animal studies however has

60 encountered criticism. First, generating reliable measures of g in non-human animals requires the use

61 of robust psychometric test batteries targeting well-studied cognitive traits [38]. It is also worth noting

62 that variation in the combination of cognitive tasks used in a test battery can lead to different

63 estimates of g [39]. Second, results indicative of g may also arise in the absence of a truly general

64 cognitive factor if performance on different tasks is underpinned by the same cognitive mechanism –

65 for instance, variation in associative learning performance could potentially impact performance

66 across a range of tasks [40]. Therefore, the single factor extracted from animal cognitive test batteries

67 does not necessarily equate to general intelligence or g as described in humans. Nonetheless, if

68 performance measured across a battery of cognitive tasks can be explained by a single factor,

69 hereafter referred to as "general cognitive performance (GCP)" [5, 41], and this factor predicts fitness

70 in the wild [5], then it may represent a measurable cognitive trait which may be under selection in

71 animal populations [35].

72 Here, we tested wild adult southern pied babblers (hereafter "babblers", *Turdoides bicolor*) on a

73 psychometric test battery containing three tasks designed to quantify (1) associative learning, (2)

74   reversal learning, (3) inhibitory control. These are well-studied cognitive traits that span different

75   domains [38, 42]. Additionally, they are likely to be ecologically relevant as they allow individuals to:

76   learn predictive contingencies between environmental cues (associative learning); learn a new

77   association when the previous one stops being rewarding (reversal learning); and control prepotent

78   motor responses when counterproductive (inhibitory control) [42, 43]. To achieve a comprehensive

79   understanding of the relationship between different cognitive traits, the factors underpinning

80   interindividual variation in cognition and the link between cognitive performance and fitness in a wild

81   animal population, we: (a) tested whether individual cognitive performance was positively correlated

82   across tasks and could be explained by a single factor (GCP); (b) measured proxies of motivation and

83   attributes of the individual and social group (age, sex, rank, group size) to identify determinants of

84   individual cognitive performance; and (c) related individual cognitive performance to multiple

85   measures of reproductive success.

86   **2. METHODS**

87   **2.1 Study site and species**

88   Data were collected at the Kuruman River Reserve (26°58' S, 21°49' E; South Africa, 33 km^2) between

89   September-March in 2018, 2019 and 2021. The reserve is situated within the semi-arid Kalahari region,

90   which is characterized by vegetated sand dunes [44]. Pied babblers are medium-sized (60-90 g),

91   sexually monomorphic passerines endemic to this region. They are cooperative breeders and live in

92   groups, which include a dominant breeding pair and subordinate helpers [45]. The dominant pair

93   produces approximately 95% of the offspring [46, 47]. Pair bond tenure varies greatly (from < 1 month

94   to > 5 years) [48]. All adult group members (> 1 year post-hatching) engage in care of young and

95   territory defence [45]. Each group defends a territory of 50-80 hectares year-round [49]. On average

96   only 4% of subordinates live in non-natal groups each year [46].

97   The study population has been monitored since 2003 and is habituated to human presence [45], which

98   allows researchers to observe the birds' natural behaviour from a close distance (< 5 m) and to present

6

99   them with cognitive tasks. Ringing and blood sampling for molecular sexing are performed on nestlings

100   11 days post-hatching [50]. Therefore, each bird in the study population is identifiable by a unique

101   ring combination, and sex and age are known for all adult birds. Adult immigrants are trapped with a

102   walk-in trap for ringing and blood sampling. We considered immigrants to be at least one year old at

103   the time they immigrated into our study population, and if they immigrated as dominants and bred

104   on the first year in which they immigrated we considered them to be at least two years old, as dispersal

105   and first breeding are rarely recorded before these ages respectively [48, 51]. On average, subordinate

106   individuals are younger than dominants [51]. Rank (dominant vs subordinate) is easily inferred from

107   aggressive displays by the dominant individuals towards subordinates [45], and distinctive affiliative

108   behaviours between dominants [48]; in addition, only the dominant female incubates the nest

109   overnight [45]. During the study years (2018-2021), the population comprised 14 groups ranging in

110   size from two to seven adults. We tested different individuals each year: 13 individuals from six groups

111   in 2018, 18 from 10 groups in 2019 and seven from four groups in 2021. Among the birds tested in

112   2021, four were unringed when tested because we found them as yearlings after a year's gap in data

113   collection (fieldwork in 2020 was suspended due to the COVID-19 outbreak). We were able to identify

114   these birds based on distinctive individual features (e.g. plumage or scarring), but their sex was

115   unknown. Sex was unknown also for two other ringed individuals tested in 2021 due to delays in the

116   analysis of blood samples caused by COVID-19.

2.2 Cognitive test battery

118   The cognitive test battery consisted of three tasks designed to quantify (1) associative learning, (2)

119   reversal learning, (3) inhibitory control. These cognitive tasks tapped into the natural terrestrial

120   foraging behaviour of babblers [45], as they required them to peck downwards at a lid or move around

121   a barrier on the ground to retrieve a food reward: a mealworm (*Tenebrio molitor* larva). The original

122   cognitive test battery included a spatial memory task, but this was later excluded because individuals'

123   behaviour when interacting with the task did not deviate from a random sampling strategy (see

124   Supplementary Material section 3).

125   Cognitive testing was conducted between 5 am and 7 pm, when babblers were active. Cognitive tasks

126   were always presented in the shade when the birds were not showing any heat dissipation behaviours

127   (i.e. panting and wingspreading) to avoid potential confounding effects of heat stress on cognitive

128   performance [52]. All trials in a cognitive test were performed when the focal individual was

129   temporarily out of sight of other group members. This was achievable because of the short trial

130   duration (< 1 min) and because babblers often forage over 10 m apart from each other [53]. The three

131   cognitive tests were carried out at least 24h apart and the order was randomised within individual,

132   except for the reversal learning, which was always carried out the day after the associative learning

133   test. Prior to quantifying learning performance, individuals were trained to peck the lids in a cognitive

134   task to find a food reward using unpainted lids (see Supplementary Material section 1). In all tasks, if

135   the focal bird did not interact with the task for 30 min, the test was paused and continued the

136   following day, and if the passing criterion was not reached by 120 trials, the test was stopped.

137   *2.2.1 Associative and reversal learning*

138   The task used to quantify associative and reversal learning consisted of a small wooden block (180 x

139   70 x 30 mm) with two equidistant circular wells (30 mm diameter, 20 mm depth) covered by painted

140   wooden lids. The lids were held in place by elastic bands; in this way, they fitted snugly into the wells,

141   preventing the bird from using visual cues to identify the rewarded well, but they could swivel when

142   pecked, making the food reward accessible to the bird (Figure 1A). The two lids were painted a dark

143   and light shade of the same colour rather than two different colours to avoid effects of past experience

144   or colour salience on learning performance [e.g. 29, 41; hereafter "colours" instead of "colour shades"

145   for brevity]. Each day before the start of cognitive testing, two mealworms were temporarily placed

146   in both wells of the cognitive task in order to prevent the bird from relying on olfactory cues to choose

147   the rewarded well during testing. The associative and reversal learning tests followed the protocols

8

148   used by Shaw et al. [41] and Ashton et al. [5]. One of the two colours was randomly assigned to be the

149   rewarded colour for each test bird. In each trial, the first peck of the individual when approaching the

150   task was counted as correct (1 = rewarded lid) or incorrect (0 = unrewarded lid). During the first trial,

151   the individual was allowed to search both wells to see that only one hid the reward. In the following

152   trials, if the individual chose correctly, it ate the mealworm and the task was removed to replace it

153   out of its sight. If the individual chose incorrectly the task was removed before the individual could

154   peck the other lid and gain the reward. The position of the rewarded well was pseudorandomised

155   between trials to ensure the individual associated the colour of the lid with the reward, and not the

156   position of the lid. Associative learning performance was quantified as the number of trials required

157   to reach the passing criterion, which was six correct choices in a row (a significant deviation from a

158   random binomial probability: binomial test p = 0.016; i.e. the individual has a probability of 1.6% of

159   achieving six correct choices in a row by random chance). If the bird passed the associative learning

160   task, the reversal learning task was carried out 24h after. Reversal learning performance was

161   quantified using exactly the same protocol and passing criterion used for associative learning, but

162   rewarding the opposite colour.

163   *2.2.3 Inhibitory control*

164   We quantified inhibitory control using a detour-reaching task, which consisted of a transparent barrier

165   (clear smooth PVC, 200 µm thick) fixed onto a wooden base (Figure 1B), with a mealworm positioned

166   ~2 cm behind the barrier on the wooden base. The task was presented to the individual straight on so

167   that the mealworm was visible behind the barrier, but not accessible from the direction that the

168   individual was approaching the task. In this task, the individual had to inhibit the prepotent response

169   of pecking the barrier when seeing the food reward and instead detour around the barrier to retrieve

170   it. A trial was marked as correct if the individual retrieved the mealworm from behind the barrier

171   without pecking it. The passing criterion was six correct trials in a row and the performance measure

172   was the number of trials to criterion.

173   *2.2.4 Task variants*

174   This study was conceived as part of a long-term project that involved repeatedly quantifying individual

175   cognitive performance (see Supplementary Material section 6). To control for the potentially

176   confounding effects of memory on cognitive performance, causally identical but visually distinct

177   variants of each task were used over the course of the project [54]. Different colour combinations

178   (dark vs light green and purple in 2018; dark vs light green, purple, blue, orange, and pink in 2019 and

179   2021) and shapes of the transparent barrier (cylinder and wall in 2018; cylinder, wall, arch, umbrella,

180   and corner in 2019 and 2021; Figure S1) were randomly assigned to each individual tested. The variant

181   used did not significantly affect the number of trials taken to pass the associative and reversal learning

182   tasks nor the inhibitory control task, respectively (see Supplementary Material section 2).

183   **2.3 Proxies of motivation**

184   Performance in cognitive tasks may be influenced by the motivation of the individual to interact with

185   the task, especially in the wild, where researchers have limited control over environmental and

186   individual condition [11]. When completing a cognitive task based on a food reward, individual

187   performance might vary depending on hunger level and amount of food available in the environment.

188   For this reason, we measured several proxies of motivation: foraging efficiency, body mass, latency to

189   approach the task, and inter-trial interval.

190   Weekly 20-min behavioural focal observations were carried out for all the individuals tested. Focal

191   observations were conducted by continuously recording the behaviours of the individual (to the

192   nearest second) using a customised programme created in the free software Cybertracker. Foraging

193   efficiency [grams of biomass consumed per foraging minute; following 55] was calculated from focal

194   observations comprising at least five minutes of foraging. Food items that were provisioned to young

195   were excluded from the calculation to better approximate individual hunger level. As a previous study

196   found babblers forage more efficiently in the early morning [56], we paired the timing of the focal

197   observation and cognitive testing by performing both either in the early morning (before 9 am) or

198   later in the day (after 9 am). As an additional proxy of hunger level, we measured the body mass of

199   each individual (accuracy 0.1 g) within the four hours prior to each cognitive test by enticing the

200   individual to jump on a top-pan scale to retrieve a mealworm [57]. Finally, we measured the latency

201   to approach the task as the time elapsed between the focal individual being within 5 m of the task and

202   first making contact with the task [5] and the average inter-trial interval (see Supplementary Material

203   section 4).

204   **2.4 Measures of reproductive success**

205   Since 2003, each year during the breeding season (September-March) researchers perform weekly

206   visits to the babbler groups during which the number and identity of individuals (adults, fledglings,

207   juveniles) and any breeding activity are noted [44]. Nests are located by observing nest building, and

208   accurate hatch and fledge dates are recorded by checking the nests every two-three days once they

209   have been located. If fledglings are missing after two consecutive visits to the group, they are

210   considered dead. We assumed only dominant individuals bred [46], therefore the offspring produced

211   in each breeding attempt were attributed to the dominant male and female in the group at the time

212   the breeding attempt was recorded. The extensive life history database allowed us to determine the

213   number of fledglings produced per year, the number of fledglings surviving to independence [i.e. 90

214   days post-hatching, when offspring receive < 1 feed/hour; 57] per year, and the number of fledglings

215   recruited into the adult population [i.e. surviving to one year post-hatching; 58] per year for each

216   dominant individual.

217   **2.5 Statistical analyses**

218   All analyses were performed with R statistical software version 4.2.0 [59]. To investigate the causes

219   and fitness consequences of variation in cognitive performance, we fitted different sets of Generalized

220   Linear Mixed Models (GLMMs) using the *lmerTest* package [60] and tested the relative importance of

221   different candidate explanatory terms by ranking them by Akaike information criterion score

222   corrected for small sample sizes (AICc). Models within 2 ΔAICc of the best model and with predictors

223   whose 95% confidence intervals did not intersect zero were included in the top model set, and were

224   considered to explain variation in the dependent variable better than other candidate models [61].

225   Continuous predictors were scaled by centring on the mean and dividing by one standard deviation.

226   Normality of residuals, presence of outliers and dispersion were checked using the *DHaRMa* package

227   [62].

228   *2.5.1 Relationships between individual cognitive performances across tasks*

229   First, we tested whether cognitive performance was correlated across tasks by performing Spearman's

230   rank correlations on the scores (i.e. number of trials to pass) of each pair of tasks. Note that a lower

231   score in this case indicates fewer trials to pass the task, and hence, better cognitive performance. To

232   determine whether individual performance in different tasks could be explained by a single factor

233   (GCP), we then performed an unrotated principal component analysis (PCA) on the scores of the

234   associative learning, reversal learning and inhibitory control tasks, using the *FactoMineR* package [63].

235   Following Shaw et al. [41], to test whether the mean and standard deviation of the loadings onto the

236   first principal component (PC1) deviated from what is expected by chance, we performed 10000 PCA

237   simulations using the function *randomizeMatrix* in the *picante* package [64]. For each simulation, the

238   cognitive scores within each task were randomised among individuals and a PCA was performed. We

239   then compared the real mean and standard deviation of the loadings onto PC1 to the 95% confidence

240   intervals (CI) of the simulated means and standard deviations of the loadings onto PC1.

241   *2.5.2 Factors explaining interindividual variation in cognitive performance*

242   To determine whether individual and group attributes or proxies of motivation explained inter-

243   individual variation in cognitive performance, we fitted LMMs containing group identity as a random

244   term and GCP as dependent variable, where GCP was the individual coordinate along PC1 but with the

245   opposite sign so that higher values corresponded to higher general cognitive performance. We used

246   GCP as a measure of individual cognitive performance because performance on all tasks loaded

247   strongly and positively onto PC1 (section 2.5.1). The individual and group attributes considered as

248   candidate explanatory terms were age, sex, rank, and group size. The proxies of motivation tested

249   were: average latency to approach (s), inter-trial interval (min), body mass (g) and foraging efficiency

250   (g/min), all of which were averaged across the three tasks used to compute GCP. We also included

251   testing order within a group to test for any potential effect of social learning (*sensu* Ashton et al.

252   2018a). If social learning was occurring, we predicted that individuals tested later within a group would

253   perform better than those tested earlier. Additionally, to test for a potential effect of different activity

254   levels throughout the day, we included the explanatory term "time of day", which was calculated as

255   follows: each task was assigned a 1 if the test started before 9 am or a 0 if the test started later in the

256   day, then this value was summed for the three tasks, obtaining a value between 0 (all tests started

257   after 9 am) and 3 (all tests started before 9 am). Finally, we included study year (2018; 2019 or 2021)

258   as a predictor to check for any differences in overall conditions across years that might have affected

259   cognitive performance. We also tested all additive models and pairwise interactions among sex, age,

260   rank, group size, body mass and study year. Individuals of unknown sex (N = 6) were excluded from

261   this analysis.

262   *2.5.3 The relationship between cognitive performance and reproductive success*

263   When analysing reproductive success, we considered only dominant individuals because subordinates

264   do not have the opportunity to breed [46]. We included two individuals that were subordinates in the

265   early years of testing but were retested once they gained dominance, for a total of N = 19 dominant

266   individuals. First, we checked whether the individual attributes that determine variation in cognitive

267   performance, i.e. age and sex (based on the results of section 2.5.2), were also associated with

268   variation in the average number of fledglings produced per year since year two of age, which is the

269   earliest age at which individuals in our dataset bred (see Supplementary Material section 8). Hence,

270   we determined if individual cognitive performance was directly related to reproductive success. We

271   considered three measures of reproductive success: number of fledglings produced per year, number

272   of fledglings that survived to independence per year and number of fledglings that survived to

13

273   recruitment per year. When there were multiple breeding attempts within a year we used cumulative

274   numbers, and we assigned a 0 for years in which dominant individuals did not successfully breed. The

275   average number of years with breeding data per dominant individual tested was 4.7 (range 1-11

276   years), where a year encompasses the austral breeding season (from September of one year to August

277   of the next year). For each of the three measures of reproductive success we fitted a set of GLMMs

278   with a Poisson error distribution and year and individual ID as random terms. Group ID was not

279   included in these models as a random term because it resulted in overfitting (singular fit). The

280   candidate explanatory terms tested were GCP, age, sex, group size, and drought (1= drought vs 0 = no

281   drought occurring during the breeding season). Group size and drought [defined as rainfall ≤ 137 mm;

282   see 65] were included among the explanatory terms because a recent study found they predicted the

283   number of offspring surviving to independence in babblers [65]. We also tested the interaction

284   between GCP and sex to determine if the relationship between cognition and reproduction differed

285   in males and females. To identify the minimum determinable effect of two-way interactions given our

286   sample sizes [66, 67], we conducted a power analysis with the *pwr* package [68].

287   **3. RESULTS**

288   **3.1 Relationships between individual cognitive performance across tasks**

289   The 38 tested babblers completed the associative learning, reversal learning and inhibitory control

290   tasks in a mean of 39.26 trials (range 6-120), 63.18 trials (range 6-120) and 34.58 trials (range 6-105)

291   respectively; the range indicating great variation in cognitive performance. We found positive

292   correlations in cognitive performance for all pairwise comparisons across tasks, but only the

293   correlation between associative and reversal learning performance was significant (associative and

294   reversal learning: Spearman's rho = 0.65, p < 0.001; reversal learning and inhibitory control:

295   Spearman's rho = 0.29, p = 0.08; associative learning and inhibitory control: Spearman's rho = 0.14, p

296   = 0.42). The consistent positive direction of pairwise correlations between tasks aligns with the output

297   of the PCA, which showed that all cognitive scores loaded positively onto PC1 extracted with an

298   eigenvalue over one (Table 1). PC1 explained 59.5 % of the total variation in cognitive performance

299   across tasks (Table 1).

300   The PCA results were highly unlikely to occur by chance because the real mean loading onto PC1 was

301   higher than the 95% CI of the randomly simulated mean loadings (95% CI of simulated means for PC1

302   = 0.01-0.67, real mean = 0.76; Figure S3), and while the real SD was within the 95% CI of the simulated

303   SD, it was at the lower end of the distribution (95% CI of simulated SD for PC1 = 0.08-0.83, real SD =

304   0.16; Figure S3). In other words, of the 10000 random simulations, only 0.03% had a larger mean

305   loading on PC1 and only 8.07% had a smaller SD. Additionally, when we examined the cognitive scores

306   of 18 individuals that were tested twice on the cognitive test battery during the study years (2018-

307   2021), individual scores from the second replicate of the cognitive test battery also loaded positively

308   onto PC1, which explained 46.3% of the total variance in cognitive performance, providing further

309   evidence for general cognitive performance (GCP); importantly, GCP was significantly repeatable (R =

310   0.50; SE = 0.18; 95% CI = 0.09; 0.78; p = 0.015) (see Supplementary Material section 6).

311   **3.2 Factors explaining interindividual variation in cognitive performance**

312   The factors that best explained variation in GCP were age and sex (Table 2). GCP declined with age in

313   females but not males (females: coefficient ± SE = -0.77 ± 0.20, 95% CI = -1.18; -0.37, males: coefficient

314   ± SE = 0.13 ± 0.24, 95% CI = -0.36; 0.62; N = 32, of which 16 females and 16 males; see Figure 2). Group

315   size was not a significant predictor of general cognitive performance (see Supplementary Material

316   section 9 for a discussion of this result). Importantly, study year and the proxies of motivation

317   examined (i.e. latency to approach the task, inter-trial interval, body mass, foraging efficiency, time of

318   day) did not significantly explain variation in GCP (Supplementary Material Table S3).

319   **3.3 The relationship between general cognitive performance and reproductive success**

320   The average number of fledglings produced per year since age two tended to increase with age in

321   females but not in males (Supplementary Material section 8). Hence, in females, the relationship

15

322   between reproductive success and age followed an opposite trend compared to the relationship

323   between GCP and age: older females tended to produce more fledglings per year on average but

324   showed lower general cognitive performance. In line with this result, we found that individual general

325   cognitive performance was negatively related with the number of fledglings produced per year (Table

326   3A, Figure 3). We did not find evidence that this relationship differed in males and females (non-

327   significant interaction GCP × sex), but we only had power to detect very large effects of two-way

328   interactions (Cohen's $f^2 = 0.51$).

329   The main predictor of the number of fledglings surviving to independence was the occurrence of

330   droughts, with more fledglings reaching nutritional independence in non-drought years (Table 3B).

331   None of the explanatory terms tested were a significant predictor of the number of fledglings surviving

332   to recruitment (Table 3C). However, we had to exclude the two dominant females who showed the

333   highest GCP from the latter analysis due to missing data on the number of fledglings surviving to

334   recruitment, therefore the lack of an effect of GCP on the number of offspring recruited per year

335   should be interpreted with caution.

336   **4. DISCUSSION**

337   We quantified individual cognitive performance in a wild bird population with the aim to answer three

338   central questions in cognitive ecology: (a) does performance co-vary across cognitive tasks, (b) what

339   drives these individual differences, and (c) is individual cognitive performance related to reproductive

340   success. We found that most of the variation in individual cognitive performance across tasks could

341   be explained by a single factor (GCP or general cognitive performance). Individual differences in GCP

342   depended on age and sex. Older females (but not males) showed lower GCP and tended to produce

343   more fledglings per year on average. Accordingly, we found that GCP was negatively related to the

344   number of fledglings produced per year. These findings support the existence of general cognitive

345   processes in wild babblers and suggest that individuals might trade off the resources invested in

346   reproduction against cognitive performance.

16

347 **4.1 Is there evidence for a general cognitive factor in wild babblers?**

348 Individuals that learnt an association faster, were also faster at reversing the learnt association and

349 showed better inhibitory control, as indicated by positive (albeit not always significant) correlations in

350 cognitive performance across tasks. Indeed, approximately 60% of the variance in individual cognitive

351 performance across tasks could be explained by a single factor: GCP. Additionally, GCP was

352 significantly repeatable (R = 0.50), indicating that our measure of general cognitive performance

353 captured consistent inter-individual differences in cognition. While we cannot completely exclude that

354 motivation to interact with the cognitive tasks affected cognitive performance, we are confident that

355 its effect on our measure of GCP was minimal because none of the measured proxies of motivation,

356 such as average latency to approach the task or inter-trial interval, significantly explained variation in

357 GCP. It is also worth noting that all the tested birds interacted with the tasks and always ate the food,

358 further indicating that the birds were motivated to interact with the tasks. Therefore, our findings are

359 consistent with the existence of a general cognitive factor underpinning performance across different

360 cognitive domains in babblers. However, our test battery included only three cognitive tasks, which is

361 the minimum number required to test for a general cognitive factor [35]. Therefore, future studies

362 should consider expanding the test battery by including, for example, spatial memory tasks redesigned

363 so that there is scope to quantify spatial memory [e.g. adding a presentation at 72h or changing the

364 scale of the spatial task; 41], and tasks assessing social cognition, the ability to make inferences, and

365 reaction time [37, 41].

366 Alternative explanations for the single factor GCP underpinning individual cognitive performance

367 across tasks are also possible. First, the different tasks used may tap into the same cognitive process.

368 For example, it has been suggested that associative learning may underlie variation in performance in

369 animal test batteries [40]. Second, positive correlations between tasks could be the consequence of

370 underlying variation in individual phenotypic or genetic quality; for example, a single genetically-

371 determined component of the nervous system may determine differences in neuronal function that

372 affect all cognitive domains simultaneously [69]. Therefore, whether statistical evidence for GCP

373 indicates a truly general cognitive ability underlying performance across different cognitive domains

374 remains to be determined.

375 **4.2 Age-related cognitive decline and individual reproductive success**

376 In babblers, individual variation in general cognitive performance was predicted by an interaction

377 between age and sex, with cognitive performance declining with age in females but not in males.

378 Faster cognitive ageing in females has been previously reported in humans [70], nematodes

379 (*Caenorhabditis remanei*) [71], mice (*Mus musculus*) [72], and captive marmosets (*Callithrix jacchus*)

380 [25]. However, the only study testing for cognitive senescence in the wild found no decline in spatial

381 memory performance in mountain chickadees (*Poecile gambeli*) from one to six years of age [26].

382 Hence, to our knowledge, our finding represents the first evidence of sex differences in age-related

383 cognitive decline in a wild animal.

384 Senescence has been explained by two main evolutionary theories [reviewed in 28]. The "selection

385 shadow" theory states that selection strength decreases with age after sexual maturity [73]. Our data

386 do not support this theory because babblers were still breeding up to 13 years of age, leaving ample

387 opportunity for selection to act on cognitive traits among older individuals. A second theory is the life

388 history theory of ageing, which encompasses two convergent theories: the first states that due to the

389 limited resources available to organisms, these must be traded-off between reproduction and somatic

390 maintenance ("disposable soma theory") [74]; the second states that alleles with beneficial effects

391 early in life but detrimental effects later in life can be favoured by selection ("antagonistic pleiotropy")

392 [75]. Based on these theories we would expect that in babblers (1) cognitive performance in older

393 females is traded-off against increased reproductive output, and/or that (2) females have been

394 selected for higher cognitive performance early in life even at the expenses of reduced cognitive

395 performance later in life.

18

396 Previous studies on babbler life history [51, 76] provide some support for both explanation (1) and (2),

397 which are not mutually exclusive. First, female babblers (but not males) engage in costly breeding

398 competition [76, 77]. Subordinate females compete both indirectly, by courting and nest-building with

399 unrelated dominant males, and directly, by destroying the eggs of the dominant female [76]. This

400 competition forces dominant (and older) females to engage in frequent aggressive displays towards

401 subordinate (and younger) females and repeatedly abandon breeding attempts and re-lay clutches

402 [76], which entails an additional energetic cost [78]. Hence, in older (dominant) females the cost of

403 maintaining a high reproductive output, even in the presence of competitors, might be traded-off

404 against the maintenance of the energetically costly nervous system [13]. For example, previous

405 experiments in the fruit fly and the cabbage white butterfly (*Pieris rapae*) have revealed a trade-off

406 between learning performance and competitive ability [79] or female fecundity [80], respectively. In

407 line with this explanation, when analysing long-term reproductive success in dominant babblers, we

408 found that higher cognitive performance was associated with a lower number of fledglings produced

409 per year. However, a larger sample size will be necessary to test whether this negative relationship

410 between cognition and reproduction differs between males and females.

411 The second explanation (i.e. selection for higher female cognitive performance earlier in life) is partly

412 supported by sex differences in babbler dispersal strategies. Females are more likely than males to

413 gain a breeding position by overthrowing a dominant female in a non-natal group [51, 81].

414 Accordingly, juvenile females are more aggressive than males, and higher female aggressiveness is

415 associated with younger age at dispersal [82]. On the contrary, males are more sedentary [83] and

416 disperse only when search costs are low [84]. It is possible that these sex differences lead to selection

417 on females for higher cognitive performance early in life, even at the expenses of faster cognitive

418 senescence. Indeed, cognitive performance in young females might be crucial to gain access to

419 breeding positions by enabling them to navigate across territories, identify the sex and rank of

420 conspecifics in non-natal groups, and decide when to engage in aggressive displays towards a

421 dominant female [85]. Additionally, in babblers the number of immigrant competitors decreases with

19

422   pair bond tenure, while reproductive success increases [48]. This suggests that on average the risk of

423   losing a breeding position and thus, potentially, the need to maintain high cognitive performance

424   might decrease with female age. Overall, our findings paired with evidence from previous research in

425   babblers suggest that females may be under selection for higher cognitive performance earlier in life

426   despite faster cognitive senescence and/or cognitive senescence may be accelerated by investment

427   in reproduction and breeding competition. However, longitudinal studies are ultimately needed to

428   describe cognitive ageing trajectories and test these hypotheses.

429   Since we used a cross-sectional design instead of a longitudinal design, we cannot determine whether

430   cognitive performance declined throughout life in females, or whether only females with lower

431   cognitive performance survived until old ages. Therefore, a third potential explanation for the

432   observed sex differences in age-related cognitive decline is that cognitive performance is negatively

433   linked to survival, at least in females. For example in pheasants (*Phasianus colchicus*), survival in the

434   wild was negatively related to reversal learning performance [86]. As most of the birds tested in the

435   present study are still alive to date, we could not perform a survival analysis to examine potential

436   effects of individual cognitive performance on survival, but this will be a necessary next step to confirm

437   whether the findings from the present study are due to cognitive senescence or reduced survival of

438   smarter females.

439   **4.3 A negative relationship between cognitive performance and reproductive success**

440   We found that individuals with better general cognitive performance produced fewer fledglings per

441   year, which is consistent with a trade-off between cognition and reproduction. However, GCP did not

442   predict the number of fledglings surviving to nutritional independence, which depended instead on

443   the occurrence of droughts during the breeding season, in line with a previous study [65]. It is possible

444   that parental traits influence offspring survival in the nestling stage but not in the post-fledgling stage,

445   where survival may be more strongly influenced by environmental conditions [87]. Therefore, the

446   extent to which cognitive performance may be under negative selection remains to be determined.

447   Cognitive performance might also be simultaneously associated with life-history traits linked to fitness

448   in different directions [88, 89]. For example, while cognitive performance is negatively related to the

449   number of fledglings produced per year by dominant individuals, it might be positively related to the

450   age at which dominance is acquired in the first place. Future comparisons of the age at acquisition of

451   dominance among individuals whose cognition was tested as subordinates will allow us to address

452   this hypothesis.

453   **4.4 Conclusion**

454   We found that individual cognitive performance covaried across tasks, which is consistent with a

455   general cognitive factor, though alternative explanations cannot be excluded. We considered the

456   effect of individual and social attributes and several proxies of motivation on cognitive performance.

457   We found that general cognitive performance depended on sex and age, declining with age in females

458   but not males. Older females also tended to fledge more nestlings per year. By analysing over 10 years

459   of breeding data, we show that individuals with lower general cognitive performance produced more

460   fledglings per year. Our findings suggest that cognitive performance is traded-off against

461   reproduction, demonstrating that in order to understand how selection acts on cognition we need to

462   consider not only its benefits but also its costs.

479 **Ethics statement**

480 This research was approved by the Animal Ethics Committee, University of Western Australia

481 (RA/3/100/1619 and 2021/ET000414).

482 **Tables**

483 **Table 1.** Output of the principal component analysis on the scores (i.e. number of trials to pass)

484 obtained by 38 pied babblers on three cognitive tasks quantifying associative learning, reversal

485 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

486

487   **Table 2.** Top model set of candidate terms affecting general cognitive performance in pied babblers.

488   All models included group ID as a random term. Corrected Akaike information criterion (AICc) and

489   ΔAICc are provided for models within 2 ΔAICc of the top model and with predictors whose 95%

490   confidence intervals (CI) do not intersect zero. Coefficient estimates ± standard errors (SE) and 95% CI

491   are given below the top model set. N = 32 individuals from 11 groups. See Supplementary Material,

492   Table S3 for full model selection outputs.

Top model set	AICc	ΔAICc
Age × Sex	103.53	0.00
Basic	*109.46*	*5.93*
Effect size of explanatory terms	**Estimate ± SE**	**95% CI**
Age	-0.76 ± 0.19	-1.14; -0.37
Sex (male)	-0.48 ± 0.28	-1.05; 0.09
Age × Sex (male)	0.90 ± 0.29	0.29; 1.49

493

24

494   **Table 3.** Model set of candidate terms affecting three measures of reproductive success in pied

495   babblers. The models included year and individual ID as random terms. Corrected Akaike information

496   criterion (AICc) and ΔAICc are provided for candidate explanatory terms. Coefficient estimates ±

497   standard errors (SE) and 95% confidence intervals (CI) are given below the model sets for models

498   within 2 ΔAICc of the top model and with predictors whose 95% confidence intervals (CI) do not

499   intersect zero. The measures of reproductive success examined were (A) number of fledglings

500   produced per year, N = 90 observations for 19 dominant individuals over 14 years; (B) number of

501   fledglings reaching independence per year, N = 81 observations for 18 dominant individuals over 14

502   years; (C) number of fledglings recruited per year, N = 79 observations for 14 dominant individuals

503   over 13 years.

| Model selection	AICc	ΔAICc
A) Number of fledglings per year		
General cognitive performance	356.08	0.00
General cognitive performance × Sex*	357.38	1.30
Sex*	357.46	1.38
Age	358.76	2.68
Basic	*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
Basic	*264.59*	*5.89*
Sex	264.80	6.10
Age	266.09	7.39
General cognitive performance	266.26	7.56
Group size	266.56	7.86
General cognitive performance × Sex	269.21	10.51
C) Number of offspring recruited per year		
Drought	204.48	0.00
Basic	*205.73*	*1.25*
Age	205.77	1.29
Sex	206.05	1.57
General cognitive performance	206.74	2.26
Group size	207.80	3.32
General cognitive performance × Sex	209.27	4.79
Effect size of explanatory terms	Estimate ± SE	95% CI
A) Number of fledglings per year		
General cognitive performance	-0.18 ± 0.07	-0.34; -0.03
B) Number of independent offspring per year		
Drought	-0.99 ± 0.34	-1.82; -0.36

504   * Not included in the top model set because 95% CI intersect zero

505   **Figures**

506

Figure 1. Wild pied babbler interacting with the cognitive task used to quantify associative and reversal learning (A); and example of task used to quantify inhibitory control (B). In A) the individual has to learn the association between a colour cue (dark versus light purple lids in the picture) and a food reward (mealworm inside the well). In B) a mealworm is placed behind the transparent barrier and the individual has to inhibit the prepotent instinct of pecking the barrier when seeing the food reward behind it and instead detour around it. Photo credits: Nicholas Pattinson.

513
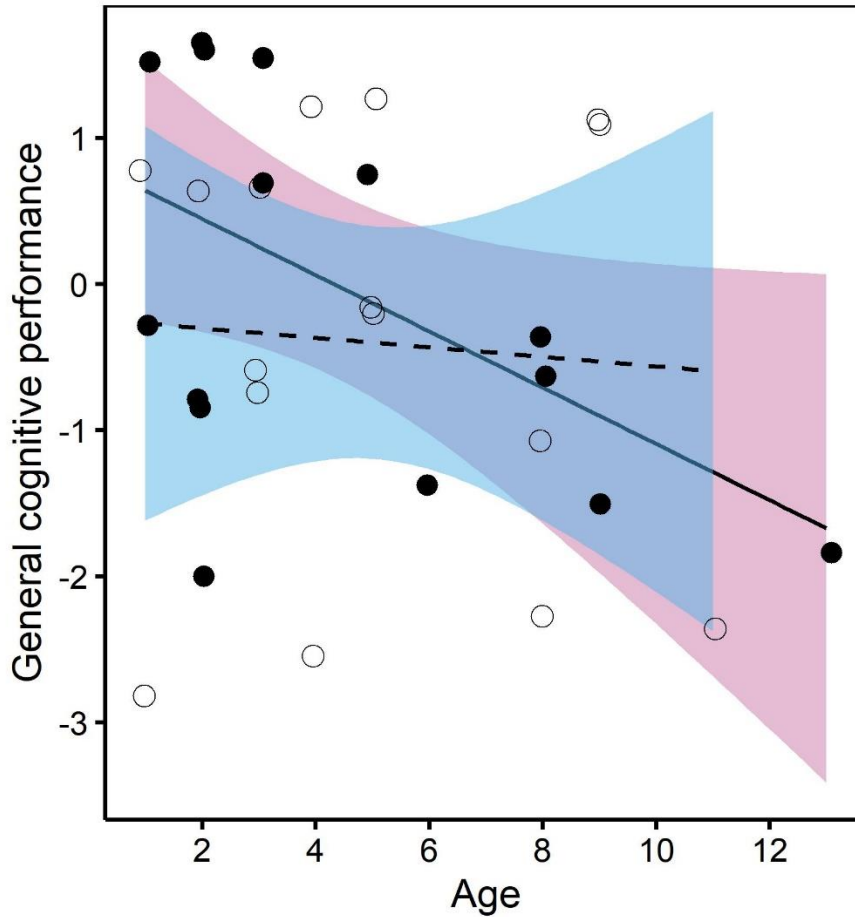
Figure 2. Variation in pied babblers' general cognitive performance by age and sex (females: pink colour, solid line, filled dots; males: blue colour, dashed line, empty dots). General cognitive performance declined with age in females but not in males (N = 16 females and 16 males). Points are raw data; fitted lines and 95% confidence interval bands are generated from the output of the model presented in Table 2.

519
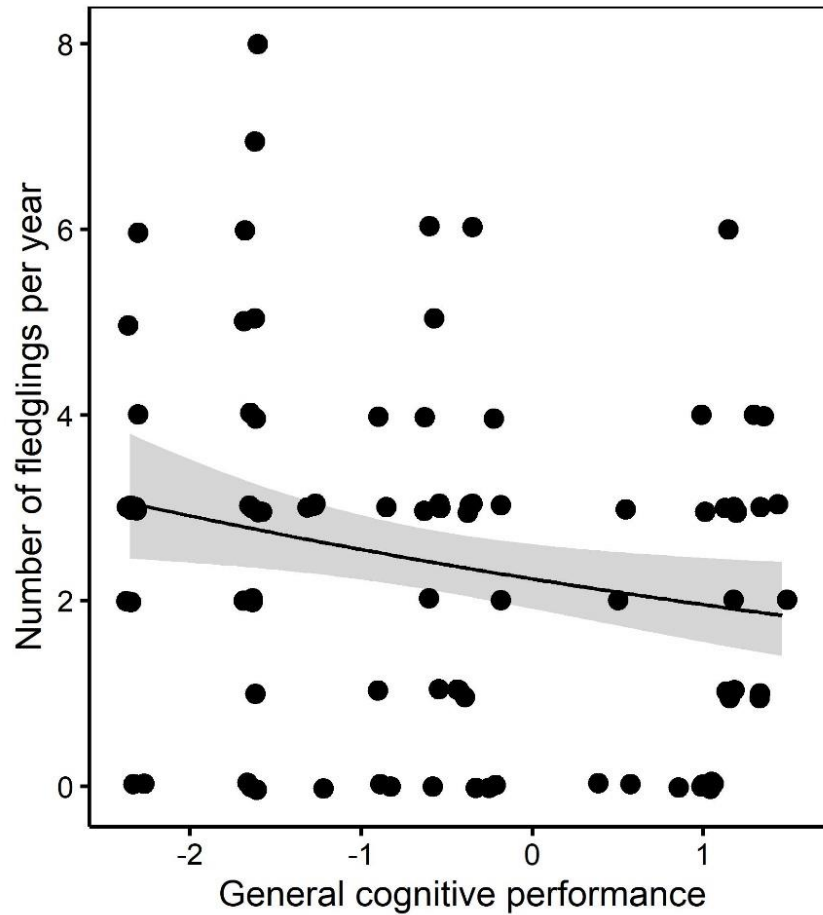
Figure 3. The relationship between the number of fledglings produced per year and general cognitive performance in dominant pied babblers (N = 90 observations for 19 dominant individuals over 14 years). Individuals showing higher general cognitive performance produced fewer fledglings per year. Points are raw data; the fitted line and 95% confidence interval band are generated from the output of the model presented in Table 3.

References

1 Shettleworth, S. J. 2010 *Cognition, evolution, and behavior*. Oxford: Oxford University Press.

2 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.

3 Bond, A. B., Kamil, A. C., Balda, R. P. 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.

4 MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., Aureli, F., Baker, J. M., Bania, A. E., Barnard, A. M. 2014 The evolution of self-control. *PNAS*. **111**, E2140-E2148.

5 Ashton, B. J., Ridley, A. R., Edwards, E. K., Thornton, A. 2018 Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*. **554**, 364-367.

6 Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., Pravosudov, V. V. 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.

7 Croston, R., Branch, C. L., Kozlovsky, D. Y., Dukas, R., Pravosudov, V. V. 2015 Heritability and the evolution of cognitive traits. *Behav. Ecol.* **26**, 1447-1459.

8 Chen, J., Zou, Y., Sun, Y.-H., ten Cate, C. 2019 Problem-solving males become more attractive to female budgerigars. *Science*. **363**, 166-167.

9 Shaw, R. C., MacKinlay, R. D., Clayton, N. S., Burns, K. C. 2019 Memory performance influences male reproductive success in a wild bird. *Curr. Biol.* **29**, 1498-1502

10 Sonnenberg, B. R., Branch, C. L., Pitera, A. M., Bridge, E., Pravosudov, V. V. 2019 Natural selection and spatial cognition in wild food-caching mountain chickadees. *Curr. Biol.* **29**, 670-676. e673.

11 Boogert, N. J., Madden, J. R., Morand-Ferron, J., Thornton, A. 2018 Measuring and understanding individual differences in cognition. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **373**, 20170280.

12 Burger, J. M., Kolss, M., Pont, J., Kawecki, T. J. 2008 Learning ability and longevity: a symmetrical evolutionary trade-off in *Drosophila*. *Evolution*. **62**, 1294-1304.

13 Heldstab, S. A., Isler, K., Graber, S. M., Schuppli, C., van Schaik, C. P. 2022 The economics of brain size evolution in vertebrates. *Curr. Biol.* **32**, R697-R708.

14 Ashton, B. J., Thornton, A., Ridley, A. R. 2018 An intraspecific appraisal of the social intelligence hypothesis. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **373**, 20170288.

15 Pravosudov, V. V., Roth II, T. C., LaDage, L. D., Freas, C. A. 2015 Environmental influences on spatial memory and the hippocampus in food-caching chickadees. *Comp. Cogn. Behav.* **10**,

16 Coppola, V. J., Hough, G., Bingman, V. P. 2014 Age-related spatial working memory deficits in homing pigeons (*Columba livia*). *Behav. Neurosci.* **128**, 666.

17 Boogert, N. J., Reader, S. M., Laland, K. N. 2006 The relation between social rank, neophobia and individual learning in starlings. *Anim. Behav.* **72**, 1229-1239.

18 Lucon-Xiccato, T., Bisazza, A. 2017 Sex differences in spatial abilities and cognitive flexibility in the guppy. *Anim. Behav.* **123**, 53-60.

19 Jones, C. M., Braithwaite, V. A., Healy, S. D. 2003 The evolution of sex differences in spatial ability. *Behav. Neurosci.* **117**, 403-411.

20 Keagy, J., Minter, R., Tinghitella, R. M. 2019 Sex differences in cognition and their relationship to male mate choice. *Curr. Zool.* **65**, 285-293.

21 Guigueno, M. F., Snow, D. A., MacDougall-Shackleton, S. A., Sherry, D. F. 2014 Female cowbirds have more accurate spatial memory than males. *Biol. Lett.* **10**, 20140026. (doi:10.1098/rsbl.2014.0026)

22 Franks, V. R., Thorogood, R. 2018 Older and wiser? Age differences in foraging and learning by an endangered passerine. *Behav. Processes*. **148**, 1-9.

23 Mirville, M. O., Kelley, J. L., Ridley, A. R. 2016 Group size and associative learning in the Australian magpie (*Cracticus tibicen dorsalis*). *Behav. Ecol. Sociobiol.* **70**, 417-427.

24 Yu, L., Tucci, V., Kishi, S., Zhdanova, I. V. 2006 Cognitive aging in zebrafish. *PloS one*. **1**, e14.

25 Rothwell, E. S., Workman, K. P., Wang, D., Lacreuse, A. 2022 Sex differences in cognitive aging: a 4-year longitudinal study in marmosets. *Neurobiol. Aging.* **109**, 88-99.

26 Heinen, V. K., Pitera, A. M., Sonnenberg, B. R., Benedict, L. M., Branch, C. L., Bridge, E. S., Pravosudov, V. V. 2021 Specialized spatial cognition is associated with reduced cognitive senescence in a food-caching bird. *Proc. R. Soc. B: Biol. Sci.* **288**, 20203180.

27 Forster, M. J., Dubey, A., Dawson, K. M., Stutts, W. A., Lal, H., Sohal, R. S. 1996 Age-related losses of cognitive function and motor skills in mice are associated with oxidative protein damage in the brain. *PNAS.* **93**, 4765-4769.

28 Nussey, D. H., Froy, H., Lemaitre, J.-F., Gaillard, J.-M., Austad, S. N. 2013 Senescence in natural populations of animals: widespread evidence and its implications for bio-gerontology. *Ageing Res. Rev.* **12**, 214-225.

29 Rowe, C., Healy, S. D. 2014 Measuring variation in cognition. *Behav. Ecol.* **25**, 1287-1292.

30 Dehnen, T., Arbon, J. J., Farine, D. R., Boogert, N. J. 2022 How feedback and feed-forward mechanisms link determinants of social dominance. *Biol. Rev.* **97**, 1210-1230.

31 Barnard, C., Luo, N. 2002 Acquisition of dominance status affects maze learning in mice. *Behav. Processes.* **60**, 53-59.

32 Thornton, A., Samson, J. 2012 Innovative problem solving in wild meerkats. *Anim. Behav.* **83**, 1459-1468.

33 Keynan, O., Ridley, A. R., Lotem, A. 2016 Task-dependent differences in learning by subordinate and dominant wild Arabian babblers. *Ethology.* **122**, 399-410.

34 Keynan, O., Ridley, A. R., Lotem, A. 2014 Social foraging strategies and acquisition of novel foraging skills in cooperatively breeding Arabian babblers. *Behav. Ecol.* **26**, 207-214.

35 Burkart, J. M., Schubiger, M. N., van Schaik, C. P. 2017 The evolution of general intelligence. *Behav. Brain Res.* **40**, e195.

36 Deary, I. J., Penke, L., Johnson, W. 2010 The neuroscience of human intelligence differences. *Nature reviews neuroscience.* **11**, 201-211.

37 Flaim, M., Blaisdell, A. P. 2020 The comparative analysis of intelligence. *Psychol. Bull.* **146**, 1174.

38 Shaw, R. C., Schmelz, M. 2017 Cognitive test batteries in animal cognition research: evaluating the past, present and future of comparative psychometrics. *Anim. Cogn.* **20**, 1003-1018.

39 van Horik, J. O., Langley, E. J., Whiteside, M. A., Laker, P. R., Madden, J. R. 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.

40 van Horik, J. O., Lea, S. E. G. 2017 Disentangling learning from knowing: Does associative learning ability underlie performances on cognitive test batteries? *Behav. Brain Res.* **40**, e220.

41 Shaw, R. C., Boogert, N. J., Clayton, N. S., Burns, K. C. 2015 Wild psychometrics: evidence for 'general'cognitive performance in wild New Zealand robins, *Petroica longipes*. *Anim. Behav.* **109**, 101-111.

42 Morand-Ferron, J., Cole, E. F., Quinn, J. L. 2016 Studying the evolutionary ecology of cognition in the wild: a review of practical and conceptual challenges. *Biol. Rev.* **91**, 367-389.

43 Kabadayi, C., Bobrowicz, K., Osvath, M. 2018 The detour paradigm in animal cognition. *Anim. Cogn.* **21**, 21-35.

44 Ridley, A., Van den Heuvel, I. 2012 Is there a difference in reproductive performance between cooperative and non-cooperative species? A southern African comparison. *Behaviour.* **149**, 821-848.

45 Ridley, A. R. 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*. (ed.^eds. pp. 115-132. Cambridge: Cambridge University Press.

46 Nelson-Flower, M. J., Hockey, P. A., O'Ryan, C., Raihani, N. J., Du Plessis, M. A., Ridley, A. R. 2011 Monogamous dominant pairs monopolize reproduction in the cooperatively breeding pied babbler. *Behav. Ecol.* **22**, 559-565.

624 47 Nelson-Flower, M. J., Flower, T. P., Ridley, A. R. 2018 Sex differences in the drivers of reproductive
625 skew in a cooperative breeder. *Mol. Ecol.* **27**, 2435-2446.
626 48 Wiley, E. M., Ridley, A. R. 2018 The benefits of pair bond tenure in the cooperatively breeding pied
627 babbler (*Turdoides bicolor*). *Ecol. Evol.* **8**, 7178-7185.
628 49 Humphries, D. J., Nelson-Flower, M. J., Bell, M. B., Finch, F. M., Ridley, A. R. 2021 Kinship, dear
629 enemies, and costly combat: The effects of relatedness on territorial overlap and aggression in a
630 cooperative breeder. *Ecol. Evol.* **11**, 17031-17042.
631 50 Bourne, A. R., Cunningham, S. J., Nupen, L. J., McKechnie, A. E., Ridley, A. R. 2021 No sex-specific
632 differences in the influence of high air temperatures during early development on nestling mass and
633 fledgling survival in the Southern Pied Babbler (*Turdoides bicolor*). *Ibis*. **164** 304-312.
634 51 Raihani, N. J., Nelson-Flower, M. J., Golabek, K. A., Ridley, A. R. 2010 Routes to breeding in
635 cooperatively breeding pied babblers *Turdoides bicolor*. *Journal of Avian Biology*. **41**, 681-686.
636 52 Soravia, C., Ashton, B. J., Thornton, A., Ridley, A. R. 2021 The impacts of heat stress on animal
637 cognition: implications for adaptation to a changing climate. *Wiley Interdisciplinary Reviews: Climate
638 Change*. **12**, e713.
639 53 Hollén, L., Bell, M., Wade, H., Rose, R., Russell, A., Niven, F., Ridley, A., Radford, A. 2011 Ecological
640 conditions influence sentinel decisions. *Anim. Behav.* **82**, 1435-1441.
641 54 Ashton, B. J., Thornton, A., Cauchoix, M., Ridley, A. R. 2022 Long-term repeatability of cognitive
642 performance. *R. Soc. Open Sci.* **9**, 220069.
643 55 Raihani, N. J., Ridley, A. R. 2007 Adult vocalizations during provisioning: offspring response and
644 postfledging benefits in wild pied babblers. *Anim. Behav.* **74**, 1303-1309.
645 56 du Plessis, K. L., Martin, R. O., Hockey, P. A., Cunningham, S. J., Ridley, A. R. 2012 The costs of
646 keeping cool in a warming world: implications of high temperatures for foraging, thermoregulation
647 and body condition of an arid-zone bird. *Glob. Change Biol.* **18**, 3063-3070.
648 57 Ridley, A. R., Raihani, N. J. 2007 Variable postfledging care in a cooperative bird: causes and
649 consequences. *Behavioral Ecology*. **18**, 994-1000.
650 58 Bourne, A. R., Cunningham, S. J., Spottiswoode, C. N., Ridley, A. R. 2020 Hot droughts compromise
651 interannual survival across all group sizes in a cooperatively breeding bird. *Ecol. Lett.* **23**, 1776-1788.
652 59 R Core Team. R: A Language and Environment for Statistical Computing. 2021 [cited; Available
653 from: https://www.R-project.org/
654 60 Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. 2017 lmerTest package: tests in linear mixed
655 effects models. *J. Stat. Softw.* **82**, 1-26.
656 61 Symonds, M. R., Moussalli, A. 2011 A brief guide to model selection, multimodel inference and
657 model averaging in behavioural ecology using Akaike's information criterion. *Behav. Ecol. Sociobiol.*
658 **65**, 13-21.
659 62 Hartig, F. 2021 DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression
660 models. R package version 0.4.4., https://CRAN.R-project.org/package=DHARMa.
661 63 Lê, S., Josse, J., Husson, F. 2008 FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.*
662 **25**, 1-18.
663 64 Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg,
664 S. P., Webb, C. O. 2010 Picante: R tools for integrating phylogenies and ecology. *Bioinform.* **26**, 1463-
665 1464.
666 65 Bourne, A. R., Cunningham, S. J., Spottiswoode, C. N., Ridley, A. R. 2020 Compensatory breeding in
667 years following drought in a desert-dwelling cooperative breeder. *Front. Ecol. Evol.* **8**, 190.
668 66 Cohen, J. 1992 Statistical Power Analysis. *Curr. Dir. Psychol. Sci.* **1**, 98-101.
669 67 Leon, A. C., Heo, M. 2009 Sample sizes required to detect interactions between two binary fixed-
670 effects in a mixed-effects linear regression model. *Comput. Stat. Data Anal.* **53**, 603-608.
671 68 Champely, S. Package 'pwr'. R package version 1.3.0 2020 [cited; Available from: https://CRAN.R-
672 project.org/package=pwr

69 Arden, R., Zietsch, B. P. 2017 An all-positive correlation matrix is not evidence of domain-general intelligence. *Behav. Brain Res.* **40**, e197.

70 Levine, D. A., Gross, A. L., Briceño, E. M., Tilton, N., Giordani, B. J., Sussman, J. B., Hayward, R. A., Burke, J. F., Hingtgen, S., Elkind, M. S. 2021 Sex differences in cognitive decline among US adults. *AMA Netw. Open*. **4**, e210169-e210169.

71 Zwoinska, M. K., Kolm, N., Maklakov, A. A. 2013 Sex differences in cognitive ageing: testing predictions derived from life-history theory in a dioecious nematode. *Exp. Gerontol.* **48**, 1469-1472.

72 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.

73 Medawar, P. 1952 *An unsolved problem in biology* London: HK Lewis.

74 Kirkwood, T. B., Rose, M. R. 1991 Evolution of senescence: late survival sacrificed for reproduction. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **332**, 15-24.

75 Williams, G. C. 1957 Pleiotrophy, natural selection and the evolution of senescence. *Evolution*. **11**, 393-411.

76 Nelson-Flower, M. J., Hockey, P. A., O'Ryan, C., English, S., Thompson, A. M., Bradley, K., Rose, R., Ridley, A. R. 2013 Costly reproductive competition between females in a monogamous cooperatively breeding bird. *Proc. R. Soc. B: Biol. Sci.* **280**, 20130728.

77 Nelson-Flower, M. J., Ridley, A. R. 2015 Male-male competition is not costly to dominant males in a cooperatively breeding bird. *Behav. Ecol. Sociobiol.* **69**, 1997-2004.

78 Williams, T. D. 2005 Mechanisms underlying the costs of egg production. *Bioscience*. **55**, 39-48.

79 Mery, F., Kawecki, T. J. 2003 A fitness cost of learning ability in *Drosophila melanogaster*. *Proc. R. Soc. B: Biol. Sci.* **270**, 2465-2469.

80 Snell-Rood, E. C., Davidowitz, G., Papaj, D. R. 2011 Reproductive tradeoffs of learning in a butterfly. *Behav. Ecol.* **22**, 291-302.

81 Nelson-Flower, M. J., Hockey, P. A., O'Ryan, C., Ridley, A. R. 2012 Inbreeding avoidance mechanisms: dispersal dynamics in cooperatively breeding southern pied babblers. *J. Anim. Ecol.* **81**, 876-883.

82 Raihani, N. J., Ridley, A. R., Browning, L. E., Nelson-Flower, M. J., Knowles, S. 2008 Juvenile female aggression in cooperatively breeding pied babblers: causes and contexts. *Ethology*. **114**, 452-458.

83 Nelson-Flower, M. J., Ridley, A. R. 2016 Nepotism and subordinate tenure in a cooperative breeder. *Biol. Lett.* **12**, 20160365.

84 Nelson-Flower, M. J., Wiley, E. M., Flower, T. P., Ridley, A. R. 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.

85 Ashton, B. J., Kennedy, P., Radford, A. N. 2020 Interactions with conspecific outsiders as drivers of cognitive evolution. *Nat. Commun.* **11**, 1-9.

86 Madden, J. R., Langley, E. J. G., Whiteside, M. A., Beardsworth, C. E., van Horik, J. O. 2018 The quick are the dead: pheasants that are slow to reverse a learned association survive for longer in the wild. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **373**, 20170297.

87 Ridley, A. R. 2007 Factors affecting offspring survival and development in a cooperative bird: social, maternal and environmental effects. *J. Anim. Ecol.*, 750-760.

88 Cole, E. F., Morand-Ferron, J., Hinks, A. E., Quinn, J. L. 2012 Cognitive ability influences reproductive life history variation in the wild. *Curr. Biol.* **22**, 1808-1812.

89 Johnson-Ulrich, L., Benson-Amram, S., Holekamp, K. E. 2019 Fitness consequences of innovation in spotted hyenas. *Front. Ecol. Evol.* **7**, 443.