Memory performance influences male reproductive success in a wild bird

Rachael C. Shaw[1,3,*], Regan D. MacKinlay[1], Nicola S. Clayton[2] & Kevin C. Burns[1]

[1]*School of Biological Sciences, Victoria University of Wellington, Wellington 6012, New Zealand*

[2]*Department of Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom*

[3]*Lead Contact*

*Correspondence: rachael.shaw@vuw.ac.nz

1   **SUMMARY**

2   Despite decades of comparative research, how selection shapes the evolution of

3   cognitive traits remains poorly understood [1–3]. Several lines of evidence suggest that

4   natural selection acts on spatial memory in food-caching species [3–6]. However, a link

5   between reproductive fitness and spatial memory ability has yet to be demonstrated in

6   any caching species [1,3,6]. Here we show that memory performance influences

7   reproductive success differentially for males and females in a caching songbird, the

8   New Zealand robin (*Petroica longipes*). Males' memory performance in a spatial task

9   during winter influenced their subsequent breeding success; individuals with more

10   accurate performance produced more fledglings and independent offspring per nesting

11   attempt. Males with superior memory performance also provided an increased

12   proportion of large prey items to chicks in the nest and spent more time flying while

13   foraging and provisioning. No such effects were found for females. Previous research

14   reveals that trade-offs may constrain selection and act to maintain variation in cognitive

15   traits [7]. The gender dimorphism in the reproductive benefits of robin memory

16   performance suggests an additional role for divergent selection between the sexes in

17   constraining runaway selection on male memory ability [8], ultimately maintaining

18   variation in this cognitive trait.

19   **KEYWORDS**

20   spatial memory; food caching; cognitive evolution; reproductive success; evolutionary

21   ecology

RESULTS AND DISCUSSION

Individual variation in cognitive performance is well documented across the animal kingdom [9], yet the ecological and evolutionary significance of cognitive variation is poorly understood [2,3]. Understanding the evolutionary ecology of cognitive traits requires examining whether and how selection acts on cognitive traits in the wild [1]. Pioneering studies of the fitness consequences of cognition have primarily examined the link between reproductive success and problem solving performance [7,10–12] or the 'general' cognitive performance captured by cognitive test batteries [13,14] (but see [15]). However, making clear predictions about how these broad measures of cognitive performance (which presumably underpin a suite of behaviours) should influence fitness remains challenging [13,16]. To further our understanding of cognitive evolution, we must therefore target specific cognitive traits [16] underpinning behaviours directly linked to survival and reproduction in the wild [2,3].

Food caching is a well-established study system for investigating the evolution of spatial memory, making caching species excellent candidates for studying the evolutionary ecology of a specific cognitive trait [1,3]. There is intraspecific evidence that spatial memory and its neural correlates can undergo fine scale selection in response to environmental variation [4,6]. For example, food storing chickadees (*Poecile* sp.) from harsher, more unpredictable, high altitude environments possess neurological adaptations and more accurate memory for cache locations compared to conspecifics from lower altitudes [17]. To date, however, the reproductive fitness consequences of intraspecific variation in spatial memory have not been examined in any caching species [1,3,6].

In the current study we investigate how memory performance in a spatial task influences reproductive success in a caching songbird, New Zealand's North Island

47   robin (*Petroica longipes*; hereafter referred to by their Māori name, toutouwai). The

48   toutouwai is a socially and genetically monogamous [18] forest dwelling species.

49   Toutouwai regularly consume some of the largest invertebrates on earth [19] and will

50   cache year round, with a peak in intensity in winter, particularly for males [20]. These

51   bold and curious birds are highly territorial; traits that make wild, free living individuals

52   amenable to cognitive testing across a range of domains [21–23], as well as easily

53   monitored during the breeding season [24]. During the austral winter (June 2016) we

54   measured individuals' (N_{male} = 36, N_{female} = 27) performance in an associative-learning

55   task where a spatial cue signalled the presence of a reward [25]. We hid a mealworm

56   inside one of eight compartments in a circular apparatus (Figure 1A). Each toutouwai

57   was tested once in the experiment: we gave the apparatus to a bird several times in a

58   single day, always at the same location within their territory, with the reward concealed

59   in the same compartment, such that it was in a consistent location relative to territory

60   landmarks across all trials (Figure 1A). Toutouwai cache their invertebrate prey over

61   short time intervals [19], we therefore used an ecologically relevant delay interval of one

62   hour between trials in our spatial memory experiment (the experiment consisted of an

63   initial probe and consolidation trial, followed by four separate test trials, ending with a

64   control trial).

65   Spatially based associative-learning tasks can provide a measure of spatial

66   memory performance [25]; individuals who remember the correct location and quickly

67   form an association between the spatial cues and reward will open the fewest lids

68   during test trials. Toutouwai learned the specific location of the food reward, as they

69   opened fewer compartment lids to retrieve it over the course of the trials (Figure 1B;

70   GLMM trial coefficient estimate, CE = -0.064, 95% confidence interval, CI = -0.103 to

71   -0.026; also see Table S1). There was moderate individual consistency in memory

72   performance over time, as there was repeatability in the number of lids that an

73   individual opened across their test trials (when statistically controlling for the influence

74   of trial sequence [26], $R_{adjusted}$ ± SE = 0.199 ± 0.062, P = 0.0001, 95% CI = 0.046 to

75   0.295). This moderate repeatability estimate for toutouwai memory performance is

76   consistent with repeatability estimates for cognitive performance in a range of non-

77   human species [27]. To examine the link between reproductive success and spatial

78   memory we quantified individual memory performance as the total number of lids

79   opened during the four test trials (following the methods of [28,29]). We investigated

80   whether this memory performance measure was influenced by possible confounds;

81   there was no effect of age, body condition or sex on individual memory performance

82   (GLM: body condition CE = 1.345, 95% CI = -0.671 to 3.397; Cohort CE = 0.032, 95%

83   CI = -0.016 to 0.082; Sex CE = -0.041, 95% CI = -0.240 to 0.161).

84   To examine whether memory performance influenced an individual's subsequent

85   reproductive success we monitored all test subjects that remained in the study area

86   during the following breeding season (September 2016 – March 2017; N_{males} = 31,

87   $N_{females}$ = 18) and used a multi-model averaging approach [30], controlling for life history

88   traits. Table 1 provides the factors that were included in these models; we ran all

89   possible models based on combinations of these predictors and calculated estimates

90   for model parameters by averaging across models (as none of the top candidate

91   models were clearly the best fit, i.e. AICw ≥ 0.9, see Table S2) [30]. Individual memory

92   performance did not affect the reproductive success of either sex in the earlier stages of

93   nesting (i.e. season start date, the number of clutches laid and hatching success; see

94   Table S3). However, spatial memory performance influenced male, but not female

95   reproductive success during the later stages of each breeding attempt (Figures 2A-D;

96   Table S3). Compared to males with poor spatial task performance, males with superior

97   memory performance fledged more chicks per clutch (N_{nests} = 54; mean β ± SE = -0.075

98   ± 0.035, 95% CI = -0.146 to -0.003; Figure 2A) and produced more independent

99   offspring per clutch (i.e. independently foraging and beginning to disperse; N_{nests} = 54;

100   mean β ± SE = -0.071 ± 0.032, 95% CI = -0.135 to -0.006; Figure 2C), but only tended

101   to produce more independent offspring over the whole season (N_{males} = 31; mean β ±

102   SE = -0.052 ± 0.30, 95% CI = -0.113 to 0.010). While our analyses controlled for

103   several potential determinants of toutouwai reproductive success (see Table 1), without

104   experimental manipulation we cannot completely exclude the possiblity that memory

105   performance covaried with unexamined ecological or life history factors [7].

106   Nonetheless, this evidence that memory performance in a spatial context is associated

107   with reproductive success in the wild supports the prediction that spatial memory is

108   likely to be under directional selection in food-caching species [2].

109   Cognition is only visible to selection via ecologically relevant behavioural

110   variation [3]. In a toutouwai pair the female builds the nest, incubates eggs and broods

111   chicks, while the male assists in provisioning her until the chicks have hatched, at which

112   point both sexes provision the young [31]. The brood is divided post fledging, with

113   males often caring for more young than females and frequently caring for fledglings for

114   longer periods of time [31]. Thus male provisioning behaviour is likely to make a larger

115   contribution to the overall success of the later stages of each breeding attempt (i.e. from

116   post-hatching through to fledgling independence). We observed a pair's foraging and

117   provisioning behaviour when chicks in the nest were 15-16 days old, fully feathered and

118   no longer reliant on brooding by the female. We estimated the total mass of food

119   delivered to each chick per hour (see methods for details). The total mass fed to chicks

120   increased as parents spent more time foraging (Table S4) and was lower for females

121   with superior memory performance (mean β ± SE = 0.015 ± 0.006, 95% CI = 0.001 to

122 0.030). By contrast, there was no effect of male memory performance on the total mass

123 of food delivered to chicks (Table S4). However, memory performance influenced a

124 male's provisioning strategy. Males with superior memory performance delivered a

125 larger proportion of large prey to the nest (Figure 3A; mean $\beta \pm$ SE = -0.120 ± 0.041,

126 95% CI = -0.207 to -0.033) and an increased proportion of large food items was

127 associated with a lower provisioning rate (i.e. the number of provisioning trips made per

128 minute spent foraging; mean $\beta \pm$ SE = -8.915 ± 3.565, 95% CI = -16.456 to -1.373). In

129 sparrow chicks (*Passer domesticus*), the delivery rate of large prey is also negatively

130 associated with overall provisioning rate, yet only the provisioning rate of the largest

131 food items is associated with increased fledging mass and future recruitment [32]. Thus

132 although male toutouwai memory performance did not influence the overall provisioning

133 rate, if toutouwai chick growth rates are similarly dependent on the provisioning of

134 larger prey items [32], selection may act on male memory via provisioning behaviour

135 during the breeding season, by favouring those males that are better able to provision

136 larger prey to offspring.

137   We suggest that our memory task may provide a measure of an individual's

138 ability to accurately form associations between food rewards and spatial cues and/or

139 landmarks. In the context of provisioning, this type of recall may allow birds to efficiently

140 locate large prey, as some of the large invertebrate species in the toutouwai's diet have

141 a clumped diurnal spatial distribution (e.g. Wellington tree wētā, *Hemideina crassidens*

142 [33]). Moreover, due to their size, such prey must be broken into smaller pieces before

143 they can be consumed [34]. This activity is usually carried out in a secluded, ground-

144 level location on the territory, with pieces then carried to the nest in multiple trips [34].

145 This behaviour also occurs in the context of caching, as large prey are broken into

146 smaller pieces before being moved to individual cache sites in the canopy [19,34]. Thus

147   in both contexts, memory for spatial or landmark cues may enable individuals to

148   efficiently return to the processing site to collect any remaining pieces after feeding

149   chicks or caching. Experimental evidence from caching corvid and parid species

150   suggests that accurate cache retrieval is underpinned by associations formed between

151   food caches and spatial cues or landmarks [35–37]. The use of landmarks and spatial

152   cues during cache retrieval has yet to be investigated in toutouwai; however, our

153   behavioural observations do provide some putative evidence that male memory

154   performance influences provisioning behaviour. Males with superior memory

155   performance spent more time flying per hour (Figure 3B; mean $\beta \pm SE$ = -0.102 ± 0.027,

156   95% CI = -0.159 to -0.045), while this was not the case for females (Table S4). For

157   males, flight time was also negatively associated with the proportion of large prey

158   delivered to the nest (mean $\beta \pm SE$ = -1.692 ± 0.671, 95% CI = -3.131 to -0.253), but

159   not with overall provisioning rate (see Table S4). These links cannot be attributed to a

160   correlation between territory quality and memory performance, as we found no

161   associations between memory performance and foraging rates (Table S4) or breeding

162   territory size (median size = 3640 m^2, range = 1300 – 8340 m^2; correlation for males: R_s

163   = 0.05, N = 32, P = 0.78; correlation for females R_s = -0.28, N = 19, P = 0.24). Instead,

164   these patterns suggest that males with better memory performance may be able to

165   forage across a larger area of their territory when provisioning chicks. To further

166   investigate how selection may act on memory performance in the context of both

167   provisioning and caching, future research should aim to test the links between

168   toutouwai memory performance, natural prey search and processing efficiency and the

169   spatial cues used by toutouwai during cache retrieval.

170   Males and females did not differ in their performance in our spatially based

171   associative-learning task, yet individual variation in cognitive performance influenced

172   provisioning behaviour and reproductive success differentially for each sex in our study

173   population. This difference may arise both because female toutouwai typically

174   contribute less to the overall provisioning of offspring (see above discussion) and

175   because they are less reliant on their own caches of food at all times of the year,

176   particularly when they must compete for food with their more dominant mates [20].

177   When the sexes have different selective optima for shared phenotypic traits, divergent

178   selection may act to maintain variation within a population [8,38]. Our results raise the

179   possibility that male and female toutouwai differ in terms of their selective optimum for

180   memory performance. Ultimately, this difference may constrain runaway selection on

181   male memory ability and contribute to the maintenance of interindividual variation in

182   spatial memory within the toutouwai population [38]. In addition, as there was only a

183   tendency for males with more accurate memory performance to produce more

184   independent young across the whole season, it also remains possible that undetected

185   costs or life history trade-offs constrain selection on male memory ability in this

186   population [7]. Our data represents selective processes shaping cognitive variation

187   within a single season for this relatively long-lived passerine. Nonetheless, it points

188   towards mechanisms that may maintain individual variation in spatial memory in food-

189   caching species.

190   **SUPPLEMENTAL INFORMATION**

191   Supplemental Information includes 4 tables and Data S1-S3.

201   **AUTHOR CONTRIBUTIONS**

202   R.C.S. conceived the study, designed the experiment, collected cognition and breeding

203   data, analysed data and wrote the manuscript. R.D.M collected behavioural observation

204   and breeding data. All authors discussed the results and commented on the

205   manuscript.

206   **DECLARATION OF INTERESTS**

207   The authors declare no competing interests.

208

209 **Figure 1. The experimental set-up (A) and the number of lids opened by**

210 **toutouwai during the memory test (B)**. The dashed line in (B) represents the number

211 of lids opened to find the reward if birds search at random, calculated following [39]. By

212 the second test trial (2 h after the probe and consolidation trials) the birds' search

213 efficiency was better than the random expectation and remained so until the end of the

214 experiment (see also Table S1, Data S1). Error bars give the 95% confidence interval,

215 adjusted for within subjects repeated measures.

216 **Figure 2. The association between memory test performance and the number of**

217 **fledglings and independent young produced per nest.** The mean memory

218 performance (measured as the number of lids opened during test trials) associated with

219 each level of nest success is shown for males in blue (A, C) and females in red (B, D).

220 A smaller value (on the right of each x-axis) represents a more accurate performance in

221 the memory test (i.e. fewer lids opened). Error bars give the standard error. The number

222 of nests associated with each mean is shown above each point. As very few nests had

223 3 fledglings/independent young, these nests are grouped together with nests that had 2

224 fledglings/independent young. See also Tables S2, Table S3 and Data S2.

225 **Figure 3. The relationships between male memory performance and provisioning**

226 **behaviour.** Males with superior memory performance fed their chick(s) a larger

227 proportion of food that was equivalent to or larger than a mealworm in size (i.e. ~1.5 cm

228 in length and 0.1 g in weight; A). The amount of time males spent flying per hour also

229 increased for males with superior spatial memory performance (B). See also Table S4

230 and Data S3.

231 **Table 1. The factors included in models of our measures of reproductive success**

232 **in the 2016 – 2017 breeding season.** Bird ID was specified as a random factor in all

233 GLMMs where the response was at the level of the nest (these models were also

234 weighted by clutch size). Models were run separately for males and females, as the

235 sample size was insufficient to analyse at the level of pairs.

Response	Variables included in the full model
Start date for the season* *Quasipoisson GLM*	parent cohort, paired last season, memory performance
Total clutches produced in season *Poisson GLM*	parent cohort, start date, parent survived entire season (Y/N), nest predation in season ('1' if at least one nest predation, '0' for no confirmed nest predation), memory performance
Chicks hatched per nest *Poisson GLMM*	parent cohort, start date, parent survived nesting† ('1' if survived until fledgling independence, '0' if not), sequence of the clutch within the season (e.g. 1, 2, 3; hereafter 'clutch number'), memory performance
Total fledglings per nest *Poisson GLMM*	parent cohort, start date, parent survived nesting†, clutch number, memory performance
Independent young per nest *Poisson GLMM*	parent cohort, start date, parent survived nesting†, clutch number, memory performance
Total independent young per season *Poisson GLM*	parent cohort, start date, parent survived entire season, nest predation in season, memory performance

236 * We excluded pairs where the partner had already attempted to breed with another
237 mate.

238 † Only included in the full models for males, as all females survived all nesting
239 attempts.

STAR+METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Rachael Shaw (rachael.shaw@vuw.ac.nz).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The North Island robin (*Petroica longipes*; here we use their Māori name, toutouwai) study population is located within a 25 hectare area at Zealandia Wildlife Sanctuary in Wellington, New Zealand (the site is described in more detail elsewhere [21]). Since 2014 all birds holding territories or hatched within the study site have been banded with a unique combination of three leg band colours for individual identification. We have monitored resident adult birds at our study site since 2014. In the 2016/2017 season the median age of breeding birds was 3 years, with a range of 1-10 years (N_{males} = 40, $N_{females}$ = 32). Toutouwai may successfully nest up to three times per season, with the first clutch typically containing 2 eggs and later clutches 2 – 3 eggs [24,31]. Chicks fledge at around 21 days old and both parents provision offspring from hatching through to independence (defined as juveniles surviving until at least 4.5 weeks post-fledge, which is when young independently forage and may begin dispersing [31]). This research was approved by Victoria University of Wellington's Animal Ethics Committee and carried out under permit from New Zealand's Department of Conservation (Authorisation number: 38497-FAU).

METHOD DETAILS

Memory test

262 In the winter (6th June 2016 – 29th June 2016) we gave a memory test to 63 (N_{male} =

263 36, N_{female} = 27) toutouwai. The test apparatus was a grey plastic ring (outer diameter =

264 40 cm, inner diameter = 30 cm, height = 1.5 cm) with eight wells (width = 1.5. cm, depth

265 = 1 cm) evenly spaced around the circumference. Each well was covered with a white

266 plastic lid that was held in place with a screw and could be swivelled open (Figure 1A).

267 All individuals had previously learned how to open these lids [22]. We chose a test

268 location on a bird's winter territory that was at least 5 m from the nearest neighbour

269 (territory boundaries were identified by observing boundary displays between

270 neighbours). Trials were only conducted when no conspecific intruders were present

271 (including mates, as toutouwai pairs typically maintain separate territories in winter

272 [40]). The apparatus was placed on a cleared, flat area that was ca. 10cm from a single

273 tree (minimum diameter of 6 cm). The tree was used as a landmark to ensure that the

274 orientation and location of the apparatus, as well as the position of the mealworm

275 reward, were consistent across all trials. In every trial the eight well lids were fully

276 closed. The location of the mealworm (relative to the landmark tree) was randomised

277 between birds, but was consistent across all trials for each individual.

278   We conducted trials between 8:30 and 14:30. We began with a probe trial in

279 which we allowed a bird to open all lids to search for the mealworm. At the start of the

280 probe trial we placed a single mealworm on the forest floor in the centre of the

281 apparatus (to ensure that every bird would begin a trial from the centre). After a bird

282 took this mealworm we gave it 6 min to open all lids. If the bird did not open all lids

283 within 6 min, we opened the remaining closed lids slightly (1 mm) and gave the bird an

284 additional 3 min to open these lids, before removing the apparatus. All birds completed

285 the probe trial. We began a consolidation trial 2 min after the probe trial had ended. At

286 the start of the consolidation trial we threw a small (< 2cm) stick into the centre of the

287   apparatus; toutouwai typically followed the stick and thus began the trial in the centre of

288   the apparatus. We gave birds up to 3 min to find the mealworm. The apparatus was

289   removed once the toutouwai had retrieved the mealworm. Following the consolidation

290   trial, we gave a bird five more trials that day, with trials spaced 1 h apart. Each trial

291   followed the same procedures as the consolidation trial. The final trial (5 h after the

292   consolidation trial) served as a control. In this trial no mealworm was present in the

293   apparatus to control for the possibility that birds relied on non-spatial cues (e.g.

294   olfactory cues, or subtle visual cues indicating the presence of the mealworm) to locate

295   the hidden mealworm. To check motivation and weight, all toutouwai hopped on

296   electronic scales to retrieve a mealworm before the first probe trial and after the final

297   control trial. All birds took both worms. Individual performance data in the memory test

298   is available in Data S1.

299   **Breeding season monitoring and foraging observations**

300   We monitored the breeding success of resident pairs at our study site following

301   established protocols [24]. In the 2016/2017 breeding season we monitored 32 of the

302   males and 19 of the females that had participated in the memory test. One pair never

303   nested and so could not be included in analyses. Four male and eight female subjects

304   disappeared prior to the breeding season (and were assumed to have either died or

305   dispersed). Females that remained in the study area during the breeding season did not

306   differ in terms of memory performance compared to those that either died or dispersed

307   prior to breeding (Mann-Whitney U test: N_{stayed} = 19; $N_{dispersed}$ = 8, W = 96.5, P = 0.29).

308   We also mapped pairs' territories by using a Garmin® GPSMAP 62c to mark

309   boundaries. Boundaries were found by observing territorial disputes, noting individuals'

310   refusal to approach and take mealworms from the ground, or by the presence of

311   geographic boundaries (large bodies of water or open spaces). We recorded at least

312   eight GPS points on the territory boundary for each pair. ArcMap® was used to create

313   territory polygons from these points and obtain area estimates. Breeding success data

314   and territory data is available in Data S2.

315   In the 2016/2017 breeding season we conducted behavioural observations for 25

316   pairs when they were caring for the first clutch of the season with chick(s) that survived

317   to 15 – 16 days post hatching. This observation timing was chosen to ensure that

318   chicks were capable of thermoregulation, so that females were not overly constrained in

319   terms of the time they could spend away from the nest, and that the breeding stage and

320   testing conditions were as comparable as possible across pairs. Within these pairs, 18

321   males and 14 females had participated in the spatial memory experiment. Observations

322   occurred on days with wind speed <40kph and without heavy rain and were preceded

323   by 15 min habituation time after the observer (R.D.M.) arrived on the territory. Each

324   parent was observed for 1 h; the order (male or female first) was determined by a coin

325   toss. The first observation began by 0830 and the second by 1000. The observer was

326   positioned with a view of the nest and as much of the territory as possible and

327   minimised their movement during the observation. If a bird moved out of visual range it

328   was followed only to the point that the nest still remained in view. All prey acquisition

329   and food sharing events were recorded. Prey size was recorded as smaller, equivalent

330   to, or larger than a mealworm (which are typically ca. 1.5 cm in length and 0.1 g in

331   weight). We used these size categories to estimate the mass of food delivered to each

332   chick during the observation. We estimated that provisioning bouts equivalent in size to

333   a mealworm provided 0.1g of food; sharing bouts that were larger were conservatively

334   estimated to contain 0.15 g of food, and smaller bouts 0.05 g of food. We divided the

335   summed mass of all sharing bouts by the number of chicks in the nest to estimate the

336   total prey mass delivered to each chick per hour. The observer had extensive prior

337   experience monitoring and feeding mealworms to robins, but was naïve to individual's

338   memory performance scores. Data were scored in the field using the app ATracker

339   Pro®. Behavioural observation data is available in Data S3.

340   **QUANTIFICATION AND STATISTICAL ANALYSIS**

341   All analyses were conducted in R (v. 3.1.1). We first investigated the birds' performance

342   as a group in the spatial memory experiment. We calculated search efficiency as the

343   number of lids an individual opened to find the mealworm (e.g. if a bird found the food

344   under the third lid it opened, its search efficiency was 3 for that trial). To examine

345   whether search efficiency improved over the spatial memory trials (excluding the

346   consolidation and control trials), we ran a generalized linear mixed model (GLMM, using

347   the *lme4* package in R) with a Poisson error structure and log link, including trial

348   number as a fixed factor and individual as a random factor. We also compared the

349   birds' actual search efficiency to a 'sampling without replacement' random search

350   strategy (i.e. once a lid was open, a bird could not re-open it; calculated using equation

351   8 in [39]). We used a two-tailed, one-sample Wilcoxon sign ranks test to evaluate

352   whether the actual search efficiency was better than the random search expectation in

353   each trial (reported in Table S1). In addition, we examined whether an individual's

354   performance was repeatable across the four test trials (i.e., trials on hours 1-4,

355   excluding the initial probe, consolidation and control trial). We used the package 'rptR'

356   to implement a mixed-effect model approach with a Poisson error structure to estimate

357   the repeatability, adjusted for trial order [26]. Finally, we tested whether non-cognitive

358   factors affected birds' spatial memory performance. Following previous studies of avian

359   spatial learning [28,29], performance was the summed search efficiency, measured as

360   lids opened to find the food reward, of the test trials (i.e. trials on hours 1-4). We used a

361   Generalised Linear Model (GLM) with a Quasipoisson error distribution (as a Poisson

362   error structure with a logarithmic link yielded overdispersion) to examine whether

363   performance was affected by body condition (calculated as body mass divided by

364   tarsus length [22]), sex and banding cohort (used as a proxy for age, as not all

365   individuals were of known age). We calculated an individual's average body mass from

366   each reliable scale reading. We used profile likelihood to calculate the 95% confidence

367   intervals (CI) for the fixed factors included in these models.

368   For the reproductive success measures we examined 'start date', which was

369   defined as the date that incubation began for a pair (the first pair to breed in the study

370   population had a start date of 0, the start date for subsequent pairs was the number of

371   calendar days that had elapsed relative to this baseline, i.e. a pair that began incubation

372   two weeks later had a start date of '14'), the total number of clutches produced in a

373   season, the number of eggs hatched per clutch, the number of chicks fledged per nest,

374   the number of independent young produced per nest and the total independent young

375   produced in a season. We used a multi-model inference approach [30] to assess the

376   relationship between winter memory performance in 2016 and subsequent reproductive

377   success in the 2016/2017 breeding season, while controlling for other life history

378   variables. We modelled the factors influencing our reproductive success measures

379   using GLMs and GLMMs with a Poisson distribution and log link (with the exception of

380   'start date', which was overdispersed and thus modelled with a Quasipoisson

381   distribution, see Table 1). For each response variable we ran all possible combinations

382   of the predictors outlined in Table 1. For each model in the resulting set we calculated

383   the Akaike information criterion corrected for small samples sizes AICc (for Poisson

384   models), or quasi-AICc (QAICc, for Quasipoisson models), as well as the change in AIC

385   relative to the best model in a set (ΔAICc/ ΔQAICc) and the Akiake weight (AICw),

386   which gives the conditional probability of the model [30]. In Table S2 we report these

387   measures. We obtained averages of model parameters by averaging across the full

388   model set (as is recommended best practice [30]). We used profile likelihood to

389   calculate the 95% CI for the averaged predictor variables. The average models for each

390   reproductive measure are reported in Table S3. For all our analyses we examined the

391   sexes separately, as sample sizes were insufficient to analyse at the level of pairs. We

392   also tested the correlation between 2016/2017 breeding territory size (measured in m^2)

393   and memory performance using Spearman rank correlations.

394   Finally, we examined parental foraging and provisioning behaviour. We used a

395   multi-model inference approach (as described above) to investigate the factors affecting

396   four measures of foraging and provisioning behaviour: foraging rate (items acquired per

397   minute spent foraging), provisioning quantity (estimated mass of food delivered to a

398   chick per hour), the proportion of large items shared (the proportion of sharing bouts in

399   which the items shared were equivalent to or larger than a mealworm) and the amount

400   of time (in minutes) spent flying during the observation. We used linear models (LM) to

401   examine the influence of spatial memory performance and the number of chicks in the

402   nest on the foraging rate (log transformed). We used LMs to examine the influence of

403   foraging duration (min spent foraging in the hour), memory performance, the amount

404   eaten by the parent and the flying duration (min) on provisioning quantity. We examined

405   the proportion of large items shared using a binomial GLM, with the number of

406   mealworm sized or larger items out of the total items shared as the response, and

407   memory performance, the number of chicks in the nest and the provisioning rate (no. of

408   sharing trips made per min spent foraging) as predictors. Finally, we used a LM to

409   examine how memory performance, provisioning rate and the proportion of large items

410   shared influenced flying duration. We ran separate models for the two sexes and used

411   profile likelihood (LM) and bootstrapping (binomial GLM) to calculate 95% CI for the

averaged model parameters. The model averaging results for these models and their

parameters are reported in Table S4.

Data availability. The data that support the findings of this study are included as Data

S1, Data S2 and Data S3.

Data S1. Individual performance in the memory test. Related to Figure 1, STAR

Methods.

Data S2. Reproductive success for the 2016/2017 breeding season. Related to

Figure 2, STAR Methods.

Data S3. Provisioning and foraging behavioural observation data. Related to

Figure 3, STAR Methods.

REFERENCES

1. Morand-Ferron, J., and Quinn, J.L. (2015). The evolution of cognition in natural populations. Trends Cogn. Sci. *19*, 235–237.

2. Morand-Ferron, J., Cole, E.F., and 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.

3. Cauchoix, M., and Chaine, A.S. (2016). How can we study the evolution of animal minds? Front. Psychol. *7*, 358.

4. Roth, T.C., LaDage, L.D., Freas, C.A., and Pravosudov, V. V. (2012). Variation in memory and the hippocampus across populations from different climates: a common garden approach. Proc. R. Soc. B Biol. Sci. *279*, 402–410.

5. Krebs, J.R., Sherry, D.F., Healy, S.D., Perry, V.H., and Vaccarino, A.L. (1989). Hippocampal specialization of food-storing birds. Proc. Natl. Acad. Sci. U. S. A. *86*, 1388–1392.

6. Sonnenberg, B.R., Branch, C.L., Pitera, A.M., Bridge, E., and Pravosudov, V. V (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. Curr. Biol. *29*, 670–676.e3.

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

8. Foerster, K., Coulson, T., Sheldon, B.C., Pemberton, J.M., Clutton-Brock, T.H., and Kruuk, L.E.B. (2007). Sexually antagonistic genetic variation for fitness in red

444   deer. Nature *447*, 1107–1110.

445   9.   Thornton, A., and Lukas, D. (2012). Individual variation in cognitive performance:
446   developmental and evolutionary perspectives. Philos. Trans. R. Soc. Lond. B.
447   Biol. Sci. *367*, 2773–2783.

448   10.   Cauchard, L., Boogert, N.J., Lefebvre, L., Dubois, F., and Doligez, B. (2013).
449   Problem-solving performance is correlated with reproductive success in a wild
450   bird population. Anim. Behav. *85*, 19–26.

451   11.   Wetzel, D.P. (2017). Problem-solving skills are linked to parental care and
452   offspring survival in wild house sparrows. Ethology *123*, 475–483.

453   12.   Keagy, J., Savard, J.-F., and Borgia, G. (2009). Male satin bowerbird problem-
454   solving ability predicts mating success. Anim. Behav. *78*, 809–817.

455   13.   Shaw, R.C., and Schmelz, M. (2017). Cognitive test batteries in animal cognition
456   research: evaluating the past, present and future of comparative psychometrics.
457   Anim. Cogn. *20*, 1003–1018.

458   14.   Ashton, B.J., Ridley, A.R., Edwards, E.K., and Thornton, A. (2018). Cognitive
459   performance is linked to group size and affects fitness in Australian magpies.
460   Nature *554*, 364–367.

461   15.   Raine, N.E., and Chittka, L. (2008). The correlation of learning speed and natural
462   foraging success in bumble-bees. Proc. R. Soc. B Biol. Sci. *275*, 803–8.

463   16.   Rowe, C., and Healy, S.D. (2014). Measuring variation in cognition. Behav. Ecol.
464   *25*, 1287–1292.

465   17.   Freas, C.A., LaDage, L.D., Roth, T.C., and Pravosudov, V. V. (2012). Elevation-
466   related differences in memory and the hippocampus in mountain chickadees,
467   *Poecile gambeli*. Anim. Behav. *84*, 121–127.

468   18.   Ardern, S.L., Ma, W., Ewen, J.G., Armstrong, D.P., and Lambert, D.M. (1997).
469   Social and sexual monogamy in translocated New Zealand robin populations
470   detected using minisatellite DNA. Auk *114*, 120–126.

471   19.   Clark, L., and Shaw, R.C. (2018). The ontogeny of food-caching behaviour in wild
472   New Zealand robins (*Petroica longipes*). Behav. Processes *151*, 27–33.

473   20.   Steer, J., and Burns, K.C. (2008). Seasonal variation in male-female competition,
474   cooperation and selfish hoarding in a monogamous songbird. Behav. Ecol.
475   Sociobiol. *62*, 1175–1183.

476   21.   Shaw, R.C., Boogert, N.J., Clayton, N.S., and Burns, K.C. (2015). Wild
477   psychometrics: evidence for 'general' cognitive performance in wild New Zealand
478   robins, *Petroica longipes*. Anim. Behav. *109*, 101–111.

479   22.   Shaw, R.C. (2017). Testing cognition in the wild: factors affecting performance
480   and individual consistency in two measures of avian cognition. Behav. Processes
481   *134*, 31–36.

482   23.   Shaw, R.C., MacKinlay, R.D., Clayton, N.S., and Burns, K.C. (2017). Male New
483   Zealand robins (*Petroica longipes*) cater to their mate's desire when sharing food
484   in the wild. Sci. Rep. *7*, 896.

485   24.   Powlesland, R.G. (1997). Protocols for monitoring New Zealand robins. Dep.
486   Conserv. Tech. Ser. *13*, 1–14.

487   25.   Roth, T.C., LaDage, L.D., and Pravosudov, V. V. (2012). Evidence for long-term

488   spatial memory in a parid. Anim. Cogn. *15*, 149–154.

489 26. Nakagawa, S., and Schielzeth, H. (2010). Repeatability for Gaussian and non-
490   Gaussian data: a practical guide for biologists. Biol. Rev. Camb. Philos. Soc. *85*,
491   935–56.

492 27. Cauchoix, M., Chow, P.K.Y., van Horik, J.O., Atance, C.M., Barbeau, E.J., Bize,
493   P., Boussard, A., Buechel, S.D., Dalesman, S., Devaud, J.M., *et al.* (2018). The
494   repeatability of cognitive performance: a meta-analysis. Philos. Trans. R. Soc. B
495   Biol. Sci. *373*, 20170281.

496 28. Pravosudov, V. V., Lavenex, P., and Omanska, A. (2005). Nutritional deficits
497   during early development affect hippocampal structure and spatial memory later
498   in life. Behav. Neurosci. *119*, 1368–1374.

499 29. Sewall, K.B., Soha, J.A., Peters, S., and Nowicki, S. (2013). Potential trade-off
500   between vocal ornamentation and spatial ability in a songbird. Biol. Lett. *9*,
501   20130344.

502 30. Burnham, K.P., and Anderson, D.R. (2002). Model Selection and Multimodel
503   Inference: A Practical Information-Theoretic Approach 2nd Editio. (New York:
504   Springer-Verlag).

505 31. Armstrong, D.P., Ewen, J.G., Dimond, W.J., Lovegrove, T.G., Bergstrom, A.,
506   Walter, B., Bergström, A., and Walter, B. (2000). Breeding biology of North Island
507   robins (*Petroica australis longipes*) on Tiritiri Matangi Island, Hauraki Gulf, New
508   Zealand. Notornis *47*, 106–118.

509 32. Schwagmeyer, P.L., and Mock, D.W. (2008). Parental provisioning and offspring
510   fitness: size matters. Anim. Behav. *75*, 291–298.

511 33. Kelly, C.D. (2006). Movement patterns and gallery use by the sexually dimorphic
512   Wellington tree weta. N. Z. J. Ecol. *30*, 273–278.

513 34. Higgins, P.J., and Peter, J.M. (2002). *Petroica australis* New Zealand Robin. In
514   Handbook of Australian, New Zealand and Antarctic Birds. 6. Pardalotes to
515   Shrike-thrushes (Oxford University Press), pp. 706–725.

516 35. Kelly, D.M., Kamil, A.C., and Cheng, K. (2010). Landmark use by Clark's
517   nutcrackers (*Nucifraga columbiana*): influence of disorientation and cue rotation
518   on distance and direction estimates. Anim. Cogn. *13*, 175–188.

519 36. Watanabe, S. (2005). Strategies of spatial learning for food storing in scrub jays.
520   J. Ethol. *23*, 181–187.

521 37. Clayton, N.S., and Krebs, J.R. (1994). Memory for spatial and object-specific
522   cues in food-storing and non-storing birds. J. Comp. Physiol. A *174*, 371–379.

523 38. Cox, R.M., and Calsbeek, R. (2009). Sexually antagonistic selection, sexual
524   dimorphism, and the resolution of intralocus sexual conflict. Am. Nat. *173*, 176–
525   187.

526 39. Tillé, Y., Newman, J. a., and Healy, S.D. (1996). New tests for departures from
527   random behavior in spatial memory experiments. Anim. Learn. Behav. *24*, 327–
528   340.

529 40. Menzies, I.J., and Burns, K.C. (2010). Temporal shifts in the pair-bond dynamics
530   of New Zealand robins (*Petroica australis*). N. Z. J. Ecol. *34*, 265–268.

531