Long-term repeatability and age-related plasticity of female behaviour in a free-living passerine

1   **Long-term repeatability and age-related plasticity of female behaviour in a free-living passerine**

2

3   Bert Thys[a*], Rianne Pinxten[a,b], Marcel Eens[a]

4

5   [a] Department of Biology, Behavioural Ecology and Ecophysiology Group, University of Antwerp, Antwerp,

6   Belgium

7   [b] Faculty of Social Sciences, Antwerp School of Education, University of Antwerp, Antwerp, Belgium

12   *Correspondence: B. Thys, Campus Drie Eiken, D.D. 123, Universiteitsplein 1, 2610 Wilrijk, Belgium.

13   E-mail address: bert.thys@uantwerpen.be (B. Thys).

14

15   Behaviour is often both repeatable among individuals (i.e. personality) and plastically adjusted within

16   individuals according to environmental conditions or age. Yet, little is known about the repeatability and age-

17   related plasticity in behavioural traits across the lifetime of free-living animals, which is, however, crucial for

18   understanding the development and evolutionary consequences of personality in the wild. Here, we explored

19   long-term (co)variation and age-dependent expression of two female behaviours, female–female aggression and

20   hissing behaviour, using longitudinal data of free-living great tits, *Parus major*. Both behaviours were repeatable

21   across years but did not covary among females into a behavioural syndrome. In contrast to hissing behaviour,

22   female aggression declined with age on the population level, which was due to within-individual plasticity in

23   aggression and not selective disappearance. Moreover, individual females differed in their level of plasticity,

24   with more aggressive females showing a steeper decline in aggression than less aggressive females, resulting in

25   a pattern where females became more similar in aggression with age. At the same time, the repeatability of

26   aggression increased with age, mostly driven by a decrease in within-individual variance across age. Large

27   between- and within-individual differences in same-sex aggression in early life and a higher repeatability in later

28   life may be functionally linked to age-dependent requirements and constraints. Overall, our findings suggest

29   that how female aggression will respond to selection might depend on when during the life cycle selection is

30   operating, which has important consequences for the understanding of the evolutionary dynamics of animal

31   personality in the wild.

32

33   Keywords: age-related plasticity, age-specific repeatability, behavioural reaction norms, female–female

34   aggression, hissing behaviour, *Parus major*, personality, senescence

35

36   Individuals typically express certain behaviours many times across their lifetime. This behavioural trait

37   expression is often highly plastic, where individuals adjust their behaviour in response to external environmental

38   variables, as well as to internal factors such as age (Dingemanse, Kazem, Réale, & Wright, 2010; Brommer &

39   Class, 2015). At the same time, individuals within populations often show repeatable between-individual

40   differences in behavioural traits (i.e. personality) and multiple behavioural traits often covary among individuals

41   into suites of traits (i.e. behavioural syndromes; Sih, Bell, Johnson, & Ziemba, 2004; Réale, Dingemanse, Kazem,

42   & Wright, 2010). Behavioural trait expression is hence multidimensional and can be both repeatable among

43   individuals and plastically adjusted within individuals. Investigating both patterns of individual behavioural

44   (co)variation in conjunction is essential in explaining how selection might act on, and therefore shape,

45   behavioural phenotypes (Dingemanse & Wolf, 2010; Dingemanse & Wolf, 2013).

46   Age-related changes in behaviour represent one specific aspect of behavioural plasticity (Dingemanse

47   et al., 2010; Brommer & Class, 2015). Age-related plasticity occurs when behavioural trait expression changes

48   within individuals with age, resulting in changes in population mean behaviour over age. Importantly,

49   progressive changes in the behavioural composition of a population, due to selective (dis)appearance of

50   individuals with certain behavioural types, can confound patterns of within-individual change (van de Pol &

51   Verhulst, 2006). For example, risk-taking behaviours (e.g. aggression, boldness, exploratory behaviour) might be

52 plastically adjusted within individuals with age (e.g. due to life history trade-offs or senescence; detailed below),

53 but at the same time, individuals that on average take more risks might have lower survival probabilities and

54 hence selectively disappear from the population (Smith & Blumstein, 2008; Moiron, Laskowski, & Niemelä,

55 2019). Studying patterns of age-related behavioural trait expression therefore requires approaches that allow

56 within- and among-individual age effects to be disentangled (van de Pol & Verhulst, 2006).

57   Whenever there are within-individual age-related changes in behaviour, individuals may be expected

58 to differ in the rate at which they adjust their behaviour with age (Nussey, Wilson, & Brommer, 2007; Brommer

59 & Class, 2015). The latter would be characterized by between-individual differences in age-related behavioural

60 plasticity (i.e. Individual×Age interaction; I×A), with some individuals showing a greater or smaller degree of

61 plasticity when ageing compared to others in the population. Importantly, the existence of I×A in behaviour

62 implies that the behaviour of individuals relative to one another changes over age and hence that repeatability

63 of behaviour can change with age (Brommer, 2013; Dingemanse & Wolf, 2013). Moreover, since I×A can have

64 underlying additive genetic variation (i.e. Genotype×Age interaction; G×A), the heritability of behavioural traits

65 and hence their response to selection might change across ages. In other words, in the presence of G×A,

66 evolutionary changes in behavioural traits will depend on when during the life cycle, hence at what age(s),

67 selection is operating (Lynch & Walsh, 1998; Dohm, 2002; Brommer & Class, 2015). Although I×A (and G×A) in

68 behaviour is expected to be common, empirical evidence in free-living species is still scarce (e.g. Carere, Drent,

69 Privitera, Koolhaas, & Groothuis, 2005, Fisher, David, Tregenza, & Rodríguez-Muñoz, 2015; Class & Brommer,

70 2016; Araya-Ajoy & Dingemanse, 2017; Class, Brommer, & van Oers, 2019). Such studies are, however, a crucial

71 first step in revealing the potential consequences of selection on, and therefore the evolutionary dynamics of,

72 behavioural phenotypes in wild populations.

73   At least two lines of theory can explain concurrent patterns of behavioural repeatability (i.e.

74 personality) and age-related plasticity. First, life history (asset protection) theory predicts that individuals differ

75 consistently in behaviour because they differ in assets (i.e. residual reproductive value; Wolf, Van Doorn, Leimar,

76 & Weissing, 2007). Specifically, individuals with low future fitness expectations (i.e. low assets) should

77 consistently take more risks, to favour current reproduction over survival, compared to individuals with higher

78 future fitness expectation (i.e. high assets; Wolf et al., 2007). At the same time, future fitness expectations (i.e.

79 assets to protect) are not stable over the course of an individual's life but typically decrease with age.

80 Consequently, an individual should take more risks with age. Second, senescence can explain age-related

81 plasticity in behaviour. Senescence is defined as a decline in organismal performance, and therefore fitness, with

82 age (Williams, 1957; Hamilton, 1966). Hence, fitness-related traits are expected to change towards values

83 associated with lower fitness. For risk-taking behaviours, which are generally assumed to increase current

84 reproductive success, senescence would result in a decline in risk-taking behaviours within individuals with age

85 (Brommer & Class, 2015). Asset protection and senescence can hence explain within-individual age-related

86 behavioural changes, yet they predict changes in opposite directions. Probably not surprisingly, this is reflected

87 in the current empirical evidence, where risk-taking behaviours have been found to increase (asset protection;

88 Dammhahn, 2012; Fisher et al., 2015; Hall et al., 2015) or decrease (senescence; Class & Brommer, 2016; Araya-

89 Ajoy & Dingemanse, 2017; Dingemanse, Moiron, Araya-Ajoy, Mochet, & Abbey-Lee, 2020) with age (see also

90 reviews in Brommer & Class, 2015; Kok et al., 2019).

91   Here, we evaluate the support for predictions within life history and senescence theory regarding age

92 effects on behaviour, using longitudinal data collected over 4 years (2016–2019) in a population of free-living

93 great tits, *Parus major*. We focused on female–female aggression during territorial intrusion (Slagsvold, 1993)

94 and female hissing behaviour when confronted with a predator inside the nest (Sibley, 1955), both of which have

95 been shown to be repeatable in the short term (i.e. within-year; Thys et al., 2017; Thys, Lambreghts, Pinxten, &

96 Eens, 2019). Since both behaviours are expressed in risky situations they allow for appropriate evaluation of

97 theory (Wolf et al., 2007; Brommer & Class, 2015). Notably, the functional consequences of age-related changes

98 might differ between aggression and hissing behaviour due to behaviour-specific hypothesized relationships

99 with fitness. That is, female–female aggression is predicted to increase reproductive success (Smith & Blumstein,

100 2008; Rosvall, 2011; see also Thys, Eens, Pinxten, & Iserbyt, 2020), since more aggressive females outcompete

101 less aggressive females for access to males that own a territory and are better at preventing nest site or mate

102 take-overs (Gosler, 1993; Slagsvold, 1993). Hence, within-individual changes towards higher aggression would

103 align with predictions of the asset protection principle, while changes in the opposite direction would align with

104 senescence. In addition, according to life history theory, more aggressive females should have lower survival

105 probabilities and selectively disappear from the population (Wolf et al., 2007). In contrast, hissing behaviour is

106 a form of antipredator nest defence, performed by some incubating and brooding females when confronted

107 with a predator inside the nest (Sibley 1955). As hissing calls deter predators (Zub, Czeszczewik, Ruczyński,

108   Kapusta, & Walankiewicz, 2017; Dutour et al., 2020), females that produce more hissing calls may have both

109   increased nest success and increased survival probabilities (Krams et al., 2014). However, recent evidence

110   indicates that more fiercely hissing behavioural types may pay a reproductive cost in terms of egg production

111   (Thys et al., 2019), fledging success (Tilgar & Koosa, 2019) or fledgling quality (Thys, Eens et al., 2020). In addition,

112   nonhissing females may decrease their mortality risk, but increase the predation risk of offspring, by hiding in

113   the nest or (partially) moving aside their clutch/brood upon predator confrontation (e.g. Fresneau, Kluen, &

114   Brommer, 2014). Studying patterns of repeatability and age-related plasticity, as well as selective

115   (dis)appearance, across the lifetime of females is therefore necessary to help elucidate the functional

116   significance of variation in hissing behaviour.

117

118   Overall, our aims were three fold. As a first step, using a bivariate analysis, we assessed both short-term

119   (within-year) and long-term (across-year) repeatability of aggression and hissing behaviour, as well as whether

120   these traits covaried among females into a long-term behavioural syndrome. Second, we investigated within-

121   individual age-related plasticity in these behavioural traits. On the one hand, aggression and hissing behaviour

122   might be up- and/or down-regulated in concert within females over age. On the other hand, both traits might

123   be independently adjusted within females with age. Indeed, previous cross-sectional analyses within a single

124   breeding season have revealed age effects on aggression, but not on hissing behaviour, with first-year breeding

125   females being on average more aggressive than older females (Thys et al., 2017; Thys et al., 2019). Yet, the

126   underlying causes of population level age effects are unknown and require longitudinal data (van de Pol &

127   Verhulst, 2006). Third and finally, for traits showing age-related plasticity, we investigated whether females

128   differed in the rate at which they adjusted their behaviour with age (i.e. I×A) and we describe patterns of among-

129   individual variance and repeatability of behaviour over age.

130   <H1>METHODS

131   *<H2>Field procedures and data collection*

132   The study was performed in a semiurban population of free-living great tits in the surroundings of

133   Antwerp, Belgium (51°09′44″N–4°24′15″E). This population has been monitored since 1997, with at present

134   approximately 150 nestboxes for great tits. Birds in the population are provided with a metal leg ring as nestlings

135   or upon first capture, and all adults receive a unique combination of three plastic colour rings, allowing

136   identification in the field. Reproductive activities of breeding pairs are monitored to determine lay date, clutch

137   size and onset of incubation.

138

139   For four years (2016–2019), behavioural tests were performed on females with first clutches ($N_{clutches}$ =

140   437; i.e. clutches initiated within 30 days after the first egg of the year was laid in the population; van Noordwijk,

141   McCleery, & Perrins, 1995). First, female–female aggression was assessed using simulated same-sex territorial

142   intrusion tests during the egg-laying period (henceforth 'aggression tests'; Thys et al., 2017). A stuffed female

143   great tit (decoy, one of five) was placed on top of the focal female's nestbox, at days 2 and 5 of the egg-laying

144   period (with day 1 the day the first egg was laid). After the focal female entered within a radius of 15 m around

145   the nestbox, or when she was already present at the start of the test, her behaviour was observed for 5 min.

146   From a distance of approximately 15 m, the observer (B.T. or an observer trained by B.T., 11 in total) scored the

147   following aggression parameters: the number of alarm calls produced, the minimum distance to the decoy

148   (approach distance, m), the time spent on the decoy (s) and the number of attacks towards the decoy. Second,

149   during the incubation period, female hissing behaviour was assessed using simulated predator intrusion tests

150   (henceforth 'hissing tests'; Grunst et al., 2018). The observer (one of 11) slowly and quietly approached the

151   nestbox and inserted the head of a stuffed specimen of the great spotted woodpecker, *Dendrocopos major* (one

152   of three) into the entrance hole of the focal female's nestbox, at days 2 and 5 of the incubation period, thereby

153   blocking the only entrance to the nestbox and preventing the incubating female from escaping. The woodpecker

154   was held in this position for 1 min, during which the observer counted the hissing calls produced (easily heard

155   from outside the nestbox when standing close to it), used as a measure of hissing behaviour (Supplementary

156   Figure S1; see Krams et al., 2014; Grunst et al., 2018; Thys et al., 2019).

157

158   The number of behavioural observations differed between females and behavioural traits, depending

159   on the number of years present in the population and whether or not focal females were observed inside their

160   territory (for aggression tests) or nestbox (for hissing tests) at the time of testing. Over the course of 2016–2019,

161   a total of 686 aggression tests and 866 hissing tests were successfully performed on 290 and 311 unique (i.e.

162   ringed) females, respectively (Table 1). For 289 of these females, both aggression and hissing tests were

163   successfully performed during the same breeding attempt, with four females repeatedly tested for both traits

164   in 4 years, 21 females in 3 years, 56 in 2 years and 208 within 1 year.

6

165

166   Age of birds (with age = 0 representing the year of birth) was determined using hatching records (local

167   recruits) or plumage characteristics upon first capture (immigrant birds; first-year versus older; Svensson, 1992).

168   Hence, absolute age was known for local recruits (N = 96 of 312 birds; 30.8%), as well as for immigrant birds that

169   were first captured as first-years (N = 192; 61.5%). Absolute age was unknown for immigrant birds with an adult

170   plumage upon first capture, implying they were 2 years old or older (N = 24; 7.7%). For this latter category of

171   immigrant birds, we assumed they recruited into the population as 2-year-olds (Bouwhuis, Sheldon, Verhulst, &

172   Charmantier, 2009; Class & Brommer, 2016; Dingemanse et al., 2020). This assumption leads to age being

173   underestimated in only a very small proportion of this category of immigrants, since for local recruits in the data

174   (N = 96) we know that 93.8% recruited as first-years, 5.2% as second-years and only 1% as being older than 2

175   years. Overall, breeding age ranged between 1 and 7 and average life span was 1.85 years.

176

177   Importantly, age at recruitment does not necessarily correspond to first age of behavioural testing since

178   some individuals in the population were already of a certain age when testing started in 2016 (see Table 1).

179   Similarly, the last age of behavioural testing does not necessarily correspond to the final age an individual was

180   found breeding in the population (i.e. some individuals were found breeding again after 2019; Table 1). Birds

181   not found breeding in the population for two consecutive breeding seasons can be considered to have died

182   (Bouwhuis et al., 2009). Consequently, in our data, individuals with complete life histories were restricted to

183   those recruiting in 2016 (or thereafter) and found breeding at the latest in 2018 (N = 119 for aggression; N = 130

184   for hissing behaviour).

185

186   *<H2>Statistical analyses*

187   Prior to all other analyses, a principal component analysis was performed on the aggression parameters

188   scored during aggression tests. This resulted in a single principal component (PC1) with eigenvalue > 1 (EV =

189   1.40) explaining 49% of the total variance (Table A1; see also Thys et al., 2017). High scores on PC1 reflected

190   closer approach distance, more time on the decoy and more attacks. In contrast, low scores on PC1 reflected

191   more alarm calls produced from a larger distance. This component was used in further analyses as a measure of

192   aggression (henceforth aggression).

193   Next, we conducted three general sets of analyses. First, we fitted a bivariate mixed model (model 1)

194   with aggression and hissing behaviour as response variables. The model included the following fixed effects: age,

195   year (2016–2019), lay date (relative to the first-egg date in the given year) and clutch size. Age was left-centred

196   (i.e. fitted as actual age – 1) to ensure that model intercepts corresponded to female behaviour at the earliest

197   age of first reproduction (Dingemanse et al., 2020). Clutch size was centred and standardized within individuals,

198   thereby partitioning effects of clutch size on behavioural traits into its among-individual (i.e. mean clutch size

199   per individual) and within-individual (clutch deviation; i.e. the deviation of each observation from an individual's

200   mean clutch size) components (van de Pol & Wright, 2009). Random intercepts were included for female identity

201   (ID, 289 levels) and the unique combination of ID and year (ID_Year, 399 levels), the latter denoting a period

202   (here breeding season) during which repeated observations were obtained for individuals. This random-effect

203   structure enabled us to quantify behavioural (co)variation on different hierarchical levels (Araya-Ajoy, Mathot,

204   & Dingemanse, 2015). Specifically, the among-individual level (random effect ID) reflects the long-term

205   (co)variation caused by permanent environment effects and genetic differences. The within-individual/among-

206   year level (random effect ID_Year) reflects short-term (co)variation caused by common environmental effects

207   (e.g. food availability, breeding density, temperature) that vary across years. Finally, the residual level reflects

208   within-individual plasticity (and/or measurement errors) caused by short-term effects that vary across

209   observations within years. From this model, we calculated both short- (R_{ST}) and long-term (R_{LT}) adjusted

210   repeatability (i.e. after correcting for fixed effects in the model) of aggression and hissing behaviour, by dividing

211   the variance of interest by the total variance (i.e. sum of among-individual, within-individual/among-year and

212   residual variance in the denominator). The distinction between the formulas for R_{ST} and R_{LT} is that the numerator

213   for R_{ST} includes the sum of the among-individual and within-individual/among-year variance, while the formula

214   for R_{LT} includes only the among-individual variance in the numerator (Araya-Ajoy et al., 2015). For ease of

215   interpretation, covariances between aggression and hissing behaviour on different hierarchical levels were

216   converted into correlations by dividing the respective covariance by the square root of the product of the

217   respective variances.

218   Second, we modelled within-individual age effects on behaviour. Given the absence of strong support

219   for covariation between aggression and hissing behaviour on any hierarchical level (see Results), we ran a

220   separate univariate mixed model (model 2) for each behavioural trait, thereby using the full data available for

8

each trait. Age was partitioned into its among- and within-individual effects following Class and Brommer (2016). That is, individuals' 'mean age' was calculated as (last observed age + 1)/2. This represents the among-individual age effect and since it is linked directly with an individual's life span it estimates selective (dis)appearance of certain behavioural phenotypes (van de Pol & Verhulst, 2006). 'Age deviation' was calculated as the deviation of each observation from an individual's mean age, representing the within-individual age effect. Both 'mean age' and 'age deviation' were included as fixed effects in the models (along with the fixed effects included in model 1). Higher-order polynomial age effects (e.g. 'age deviation' squared) were not considered as their estimation requires more repeated measurements per female than we were able to collect. Random intercepts were included for ID (290 levels for aggression; 311 levels for hissing behaviour) and ID_Year (403 levels for aggression; 436 levels for hissing behaviour). We did not include random intercepts for Year, Observer or Decoy ID in these models since preliminary analyses revealed they explained little to no variation in behaviour (Table A2; thereby also preventing overparameterization leading to singular fit and model convergence issues). Since only two females more than 5 years old were tested for their behaviour (Table 1), we reran the same models by either combining age categories 5, 6 and 7 into one age category (Dingemanse et al., 2020) or removing behavioural observations made when females were older than 5 years (Bouwhuis et al., 2009). Neither analysis changed our estimates (see Table A3), showing there were no indications that our results were driven by these few very old females. Additionally, we reran the same models using only the data of females with complete life histories (detailed above), thereby obtaining more biologically relevant estimates of selective (dis)appearance from the population (i.e. 'mean age' effects). This revealed that estimates for 'mean age' were similar in sign to, but slightly higher in magnitude than, estimates from the main analyses (see Table A4). Yet, since estimates of both analyses found similar support, only findings of the main analyses are discussed further.

Third and finally, for behavioural traits showing within-individual age effects, we further investigated whether individuals differed in age-related plasticity (i.e. I×A). We restricted this analysis to include only observations made up to when females were 4 years of age (i.e. 670 observations on 285 females; representing 98% of all data) since sample sizes in older age classes were too low (Table 1). Specifically, we fitted a random regression model (model 3) which included 'mean age' (i.e. among-individual age effect) and age (fitted as actual age – 1) as fixed effects, the former avoiding biases in estimates of within-individual age effects (van de Pol & Verhulst, 2006). Random intercept and random slopes were included for female ID, thereby estimating among-

249   individual variance in mean behaviour ($ID_{intercept}$), age-related plasticity (ID_{slope}) and their covariance ($Cov_{intercepts-slopes}$).

250   The intercept–slope covariance was converted into a correlation ($Cor_{intercepts-slopes}$) using the standard

251   formula for correlations (see above). Since residual variance is likely to vary across age groups, thereby

252   potentially influencing repeatability, we allowed for heterogeneous residuals across age classes (i.e. age-specific

253   residual variances for age classes 1, 2, 3 and 4; e.g. Cleasby & Nakagawa, 2011; Brommer, 2013; Class et al.,

254   2019). Age-specific repeatabilities were then calculated by dividing the respective age-specific among-individual

255   variance by the sum of the among-individual and age-specific residual variance.

256   All analyses were performed in R 3.6.1 (R Core Team, 2019). Prior to analyses, continuous response and

257   explanatory variables were centred and standardized to unit variance, unless stated otherwise. All models were

258   fitted assuming a (multivariate) Gaussian error distribution. The MCMCglmm package (Hadfield, 2010) was used

259   throughout to simulate values of the posterior distribution of model parameters (see Appendix for details on

260   prior specification for different models). Based on 2000 simulations, we extracted 95% credible intervals (CrI)

261   around the mean. Fixed-effect and covariance/correlation estimates were considered to find strong support if

262   95% CrI did not overlap with zero. Support for variance components was assessed using permutation tests

263   (Araya-Ajoy & Dingemanse, 2017; Thys et al., 2019). Specifically, female identity (i.e. ID) was randomly reshuffled

264   across observations and mixed modelling approaches as described above were then performed on the new data

265   set to obtain posterior mean estimates for the variance components of interest. This procedure was repeated

266   100 times to obtain a 'null' distribution of posterior means for each variance component. Using this 'null'

267   distribution, we calculated the probability (permutation.P) that the observed posterior mean variance estimate

268   was greater than any posterior mean value of the 'null' distribution.

269   *<H2>Ethical Note*

270   This study was approved by the ethical committee of the University of Antwerp (ID 2017-23 and 2017-

271   61), performed in accordance with Belgian and Flemish laws regarding animal welfare and adhered to the

272   ASAB/ABS guidelines for the use of animals in behavioural research and teaching. The Royal Belgian Institute of

273   Natural Sciences (KBIN) provided ringing licences for all authors and technicians. Nestbox monitoring, ringing

274   and behavioural experiments (which mimicked naturally occurring conspecific and predator confrontations)

275   created only very low levels of stress that did not cause nest abandonment or nestling mortality.

<H1>RESULTS

<H2>(Co)variance in aggression and hissing behaviour

We found strong support for the existence of among-individual and within-individual/among-year

variation in both aggression and hissing behaviour (all permutation.P < 0.01), with aggression and hissing

behaviour being repeatable both within (R_{ST}) and across (R_{LT}) years (Table 2). Moreover, aggression decreased

with age on the population level (Fig. 1). For hissing behaviour, we found strong support for effects of lay date

and mean clutch size. Specifically, females that produced more hissing calls started laying relatively earlier and

produced, on average, smaller clutches (Table 2). Although both aggression and hissing behaviour were

repeatable, we found no support for among-individual, within-individual/among-year or residual correlations

between both behaviours (Table 3).

<H2>Age effects on behaviour

The within-individual age effect on female aggression was strongly supported, with aggression

decreasing when females aged ('age deviation'; β [95% CrI]; -0.30 [-0.44;-0.15]). No strong support was found

for a within-individual age effect on hissing behaviour (-0.08 [-0.20;0.03]). Among-individual age effects (i.e.

'mean age') were not supported, either for female aggression (-0.13 [-0.32;0.06]) or for hissing behaviour (0.04

[-0.16;0.27]), indicating the absence of strong support for the selective (dis)appearance of certain behavioural

phenotypes (see also Table A4). Output of full models can be found in Table A3.

<H2>Age-related plasticity and age-specific repeatability

Given we found support for within-individual age-related plasticity in aggression, we focussed on this

trait in further analyses. The random regression model revealed strong support for between-individual

differences in both intercepts and slopes (both permutation.P ≤ 0.01; Table 4, Fig. 2), indicating that females

differed both in mean aggression and in age-related plasticity in aggression. Moreover, we found strong support

for a negative correlation between intercepts and slopes, indicating that females that were on average more

aggressive also showed a larger (i.e. more negative) decline in aggression with age.

302   Among-individual variance (V_{ID}) in aggression declined between ages 1 and 2 (Fig. 3a). At age 3, V_{ID} was

303   lower than at age 1 but slightly higher than at age 2. This was subsequently followed by an apparent increase in

304   V_{ID} at age 4, but uncertainty around the estimate was large (Fig. 3a; age-specific variances calculated from model

305   3 following Brommer, 2013). In addition, residual variance (V_R) was notably higher at age 1 than at older ages

306   (Table 4, Fig. 3a). Age-specific repeatabilities (R [95% CrI]) of aggression, calculated based on age-specific

307   estimates of V_{ID} and V_R, were 0.30 [0.16;0.44] for age 1, 0.33 [0.17;0.49] for age 2, 0.57 [0.42;0.76] for age 3 and

308   0.66 [0.45;0.86] for age 4, indicating an increase in repeatability with age (Fig. 3b).

309   # DISCUSSION

310   We found that female–female aggression and female hissing behaviour were repeatable within and

311   across years, but that they did not form a behavioural syndrome. In contrast to hissing behaviour, aggression

312   was found to decline on the population level. The latter was due to within-individual age-related plasticity in

313   aggression and not selective disappearance of more aggressive females from the population. Moreover, females

314   differed in the rate at which aggression changed with age. Specifically, more aggressive behavioural types

315   showed a steeper decline in aggression, resulting in females becoming more similar in their aggression when

316   they age. At the same time, female aggression became more repeatable with age, mostly due a decrease in

317   within-individual variance across age.

318

319   ## (Co)variation in female behaviours

320   In line with our previous findings, females were found to differ consistently in aggression and hissing

321   behaviour within a given year (Thys et al., 2017; Thys et al., 2019). Here, we extend these findings by showing

322   that both behavioural traits were also repeatable across years (i.e. long-term repeatable). Since repeatability is

323   generally thought to set an upper limit to heritability (Boake, 1989; Dochtermann, Schwab, & Sih, 2015; but see

324   also Dohm, 2002), both traits may have the potential to evolve in response to selection. Although both

325   behaviours were repeatable, aggression and hissing behaviour did not covary among females in the long term

326   (i.e. no behavioural syndrome; see also Thys, Eens et al., 2020). Long-term among-individual correlations can

327   arise from permanent environmental effects and genetic correlations. Hence, the absence of a strong among-

328   individual correlation, assumed to be at least partially underpinned by a genetic correlation (Dochtermann,

12

329 2011), suggests that female aggression and hissing behaviour may be able to evolve independently in response

330 to selection. Moreover, female aggression and hissing behaviour also did not covary in the short term (i.e. the

331 within-individual/among-year level), suggesting the absence of strong common environmental effects

332 influencing both behavioural traits in conjunction across years. Hence, we have demonstrated that aggression

333 and hissing behaviour are two independent aspects of a female's behavioural phenotype, which are potentially

334 influenced by different underlying genetic mechanisms, intrinsic factors (e.g. age; discussed below) and/or

335 environmental conditions.

336

337 *<H2>Population level age effects on behaviour*

338 An individual's age is predicted to influence the level of behavioural trait expression (Dingemanse et al.,

339 2010; Brommer & Class, 2015). Here, we have shown that female great tits adjusted their level of same-sex

340 aggression in relation to age, resulting in a decline on the population level. Importantly, this population level

341 decrease in aggression over age was due to plastic changes of aggression within individuals and not due to

342 selective disappearance of more aggressive females over age. Hence, we have provided scarce empirical

343 evidence for age-related behavioural plasticity in a personality trait in a free-living species, while simultaneously

344 showing that female aggressive behavioural types do not appear to have different survival probabilities.

345 Female aggression in great tits is used in same-sex competition for access to males with a territory

346 (Gosler, 1993) and more aggressive females are thought to be better at preventing intruding females from taking

347 over the nest site or mate (Slagsvold, 1993). Moreover, we have previously shown that female aggression is

348 linked with parental investment and short-term (annual) reproductive success, although fitness consequences

349 associated with different aggressive behavioural types varied across years (see Thys, Eens et al., 2020).

350 Nevertheless, female aggression probably reflects investment in reproduction and, following the asset

351 protection principle, we would have expected an increase in aggression as females aged. In contrast, we found

352 a within-individual decline in aggression with age, contradicting predictions of asset protection but aligning with

353 predictions within senescence. Interestingly, the pattern of within-individual age-related plasticity in female

354 aggression we observed generally follows previously described patterns of reproductive senescence in female

355 great tits, where reproductive performance is highest in early life and decreases afterwards (Bouwhuis et al.,

356   2009; Dingemanse et al., 2020). Formally linking reproductive senescence with patterns of age-related plasticity

357   in female aggression will hence be a fruitful future endeavour.

358   Alternatively, or in addition to senescence, differences in the cost–benefit ratio associated with female

359   aggression across age may have contributed to within-individual age-related plasticity. Since great tits commonly

360   reoccupy their territory of the previous breeding season (i.e. high nest site fidelity; Kluijver, 1951; Harvey,

361   Greenwoord, & Perrins, 1979), new females have to either outcompete already established (i.e. older) females

362   for access to males or settle with other vacant territory-owning males. Consequently, selection might favour

363   higher aggression in early life, while selective benefits (relative to costs) of high aggression progressively

364   decrease later in life. In line with this, social and breeding experience (hence prior knowledge about the breeding

365   area) increase with age, which potentially leads to more accurate assessment of the threat associated with same-

366   sex territorial intrusion, in terms of both nest site and partner loss (see also Araya-Ajoy & Dingemanse, 2017).

367   More accurate threat assessment, combined with varying cost–benefit ratios, may also coincide with the

368   adoption of different strategies to settle same-sex contests across age (Johnstone & Norris, 1993; Kokko, 1997).

369   For example, when ageing, female great tits may increasingly rely on status signalling, instead of overt

370   aggression, to settle same-sex contests (see e.g. Thys, Pinxten, & Eens, 2020). Hence, future studies should aim

371   at identifying whether the age-related plasticity in female same-sex aggression we observed is caused by

372   senescence, age-related changes in cost–benefit ratios and/or breeding experience.

373   In contrast, female hissing behaviour was not plastically adjusted within females with age, neither did

374   certain hissing behavioural types selectively (dis)appear from the population. Consequently, and contrasting

375   with predictions from either asset protection or senescence, we found no strong support for either population

376   or individual level plasticity in hissing behaviour across age. The absence of age-related plasticity, combined with

377   the high short- and long-term repeatability, suggests that there may be strong permanent environmental and/or

378   genetic effects underlying variation in hissing behaviour, which awaits confirmation (but see Timm, Koosa, &

379   Tilgar, 2019). Moreover, variation in hissing behaviour has been argued to be maintained by life history trade-

380   offs, with hissing females prioritizing survival over investment in current reproduction, while nonhissing females

381   do the opposite (Krams et al., 2014; Thys et al., 2019; Tilgar & Koosa, 2019). In line with this, we indeed showed

382   that females that produced more hissing calls started breeding relatively earlier and on average had smaller

383   clutches. Hence, more fiercely hissing behavioural types seemed to pay a reproductive cost in terms of the

384　average number of eggs they produced (Thys et al., 2019). However, this apparent cost was not outweighed by

385　increased survival, given the absence of selective disappearance of certain hissing behavioural types, arguing

386　against hissing behaviour being involved in the trade-off between reproduction and survival. The latter might be

387　partially due to our use of artificial nestboxes, which are less prone to predation than natural cavities (e.g.

388　Mitrus, 2003). Nevertheless, hissing behaviour has been found to differentially affect female survival in another

389　nestbox population of great tits (Krams et al., 2014). Assessing lifetime reproductive success associated with

390　different hissing behavioural types may help to reveal how long-term among-individual variation in hissing

391　behaviour is maintained within our population.

392

393　<H2>*Age-related plasticity and age-specific repeatability*

394　For female aggression, which declined on the population level, we additionally revealed that females

395　differed in the rate at which they plastically adjusted their aggression with age (i.e. between-individual

396　differences in age-related plasticity; I×A). Moreover, more aggressive behavioural types decreased their

397　aggression to a larger extent than less aggressive types, leading to a decrease in among-individual variance

398　across most of the life span of female great tits. As a result, aggression followed a pattern where females in our

399　population became more similar as they aged. Our findings largely corroborate previously observed patterns of

400　age-related plasticity in aggressive behaviour in response to handling in blue tits, *Cyanistes caeruleus* (Class &

401　Brommer, 2016), but are in contrast with, for example, findings in field crickets, *Gryllus campestris*, where

402　boldness and activity increased with age and no evidence was found for individual differences in plasticity (i.e.

403　no I×A; Fisher et al., 2015; see also reviews in Brommer & Class, 2015; Kok et al., 2019).

404　Among-individual variation provides the raw material for selection to act on and the existence of I×A

405　indicates that measurements at one point in time do not necessarily reflect true among-individual variance (and

406　repeatability; Brommer, 2013). Consequently, the outcome of selection may differ depending on when during

407　the life cycle selection is operating and age classes are hence not necessarily equivalent with respect to their

408　evolutionary potential and/or response (Lynch & Walsh, 1998; Nussey et al., 2007). Here, we have shown that

409　among-individual variance in aggression follows a roughly U-shaped pattern over the life span of females, first

410　decreasing and then increasing in later life. Note that variances at the extremes are probably overestimated,

411 especially when sample sizes are low (e.g. Promislow, Tatar, Khazaeli, & Curtsinger, 1996). Hence, our limited

412 sample size in age class 4 prevents us from drawing strong conclusions regarding the increase in among-

413 individual variance late in life, as also reflected by the large uncertainty around the estimate. Consequently,

414 determining whether among-individual variance indeed follows a U-shaped pattern or rather decreases

415 monotonically with age and/or potentially flattens off in later life would require a larger sample size in old age

416 classes. Interestingly, because within-individual (i.e. residual) variance was also found to vary across age classes,

417 repeatability did not follow the same age-related pattern as the among-individual variance. That is, we found

418 that residual variance in female aggression was particularly high at 1 year old compared to later in life.

419 Consequently, although among-individual variance in aggression declined with age (from 1 year old to 2–3 years

420 old), the relatively larger age-related decline in residual variance ultimately resulted in an increase in

421 repeatability. Hence, female aggression became more consistent (i.e. repeatable) with age, which appeared to

422 be mainly driven by a reduction in within-individual variance. Age-related increases in the repeatability of

423 behavioural traits has been found in a number of species, including humans, but only a few studies have aimed

424 at identifying underlying processes (i.e. changes in among-individual and/or within-individual variance; reviews

425 in Brommer & Class, 2015; Kok et al., 2019). In the particular case of female great tit aggressive behaviour,

426 relatively large among- and within-individual (i.e. residual) variance when 1 year old could be, at least partially,

427 linked to high and variable same-sex competition faced by females at this stage (i.e. first potential age of

428 breeding). Later in life, female aggression seems to become less plastic (as suggested by reduced within-

429 individual variance) and more stabilized at, on average, a lower level of expression (see also Araya-Ajoy &

430 Dingemanse, 2017), which could be linked with high nest site fidelity, reduced competition and/or females

431 becoming more accurate in threat assessment (see also above). Hence, age-dependent requirements and

432 constraints may result in age-related changes in the repeatability (and underlying variance components) of

433 female same-sex aggression.

434 From an ultimate perspective, the existence of I×A in female aggression could be caused by differences

435 in reproductive senescence rates between aggressive behavioural types. This remains to be determined, but

436 recent evidence suggests this might be the case. That is, a recent study in female great tits found that more

437 exploratory behavioural types did not selectively disappear from the population but instead showed higher

438 reproductive senescence than less exploratory types (Dingemanse et al., 2020). The existence of I×A and

439   apparent absence of survival selection on female aggression that we observed, combined with previously

440   reported positive covariation between female aggression and exploratory behaviour (Thys et al., 2017), raises

441   the possibility that female aggression follows similar behaviour–senescence patterns as recently reported for

442   exploratory behaviour (Dingemanse et al., 2020). If so, more aggressive females would pay the cost of higher

443   early investment in reproduction by an earlier onset of senescence, instead of decreased survival. Notably, we

444   found that variation in aggression was not linked with either mean clutch size (see also Thys, Eens, et al., 2020)

445   or deviations from mean clutch size across years, suggesting the absence of a link between aggression and

446   reproductive investment in terms of egg production. Whether this is also the case for other reproductive traits

447   awaits confirmation.

448   Importantly, our findings do not preclude fluctuating (across-year) selection. That is, fluctuating

449   environmental conditions (i.e. extrinsic factors; such as density, predation and resource abundance) may not

450   only affect fitness consequences of variation in female aggression (Thys, Eens et al., 2020) but also influence

451   rates of reproductive senescence (e.g. by speeding up or slowing down senescence rates; see review in Lemaître

452   & Gaillard, 2017). Along these lines, evolutionary theories of senescence predict that I×A in fitness-related traits

453   has underlying additive genetic variance (i.e. G×A), which can lead to age-dependent heritability and age-

454   dependent rates of evolutionary change in response to selection (Lynch & Walsh, 1998; Dohm, 2002; Nussey et

455   al., 2007; Brommer & Class, 2015). Empirical evidence for G×A in behaviour is scarce (Class & Brommer, 2016;

456   Class et al., 2019; review in Brommer & Class, 2015) and whether this is true for female great tit aggression

457   deserves further investigation, as it will help reveal whether the I×A we observed represents potentially evolved

458   differences in age-related plasticity.

459

460   *<H2>Conclusion*

461   Same-sex aggression and hissing behaviour were repeatable in the long term but represented

462   independent aspects of a female great tit's behavioural phenotype. In line with this, female aggression, but not

463   hissing behaviour, declined on the population level due to within-individual age-related plasticity. Moreover,

464   females differed in the rate at which they adjusted their aggression with age, resulting in a decline in among-

465   individual variance across most of the life span of female great tits. At the same time, repeatability of aggression

466   increased with age, with potentially important implications for the consequences of selection at different ages.

467   Our findings illustrate the importance of considering individual differences in age-related behavioural plasticity

468   to improve our ultimate understanding of animal personality in free-living species.

469   ACKNOWLEDGMENTS

474

475

476   REFERENCES

477   Araya-Ajoy, Y. G., & Dingemanse, N. J. (2017). Repeatability, heritability, and age-dependence of seasonal

478   plasticity in aggressiveness in a wild passerine bird. *Journal of Animal Ecology*, *86*(2), 227–238.

479   https://doi.org/10.1111/1365-2656.12621

480   Araya-Ajoy, Y. G., Mathot, K. J., & Dingemanse, N. J. (2015). An approach to estimate short-term, long-term and

481   reaction norm repeatability. *Methods in Ecology and Evolution*, *6*(12), 1462–1473.

482   https://doi.org/10.1111/2041-210X.12430

483   Boake, C. R. B. (1989). Repeatability: Its role in evolutionary studies of mating behavior. *Evolutionary Ecology*,

484   *3*(2), 173–182. https://doi.org/10.1007/BF02270919

485   Bouwhuis, S., Sheldon, B. C., Verhulst, S., & Charmantier, A. (2009). Great tits growing old: Selective

486   disappearance and the partitioning of senescence to stages within the breeding cycle. *Proceedings of the

487   Royal Society B: Biological Sciences*, *276*(1668), 2769–2777. https://doi.org/10.1098/rspb.2009.0457

488   Brommer, J. E. (2013). Variation in plasticity of personality traits implies that the ranking of personality measures

489   changes between environmental contexts: Calculating the cross-environmental correlation. *Behavioral*

490   *Ecology and Sociobiology*, *67*(10), 1709–1718. https://doi.org/10.1007/s00265-013-1603-9

491   Brommer, J. E., & Class, B. (2015). The importance of genotype-by-age interactions for the development of

492   repeatable behavior and correlated behaviors over lifetime. *Frontiers in Zoology*, *12*(1), S2.

493   https://doi.org/10.1186/1742-9994-12-S1-S2

494   Carere, C., Drent, P. J., Privitera, L., Koolhaas, J. M., & Groothuis, T. G. G. (2005). Personalities in great tits, *Parus*

495   *major*: stability and consistency. *Animal Behaviour, 70*, 795-805. doi:10.1016/j.anbehav.2005.01.003

496   Class, B., & Brommer, J. E. (2016). Senescence of personality in a wild bird. *Behavioral Ecology and Sociobiology*,

497   *70*(5), 733–744. https://doi.org/10.1007/s00265-016-2096-0

498   Class, B., Brommer, J. E., & van Oers, K. (2019). Exploratory behavior undergoes genotype–age interactions in a

499   wild bird. *Ecology and Evolution*, *9*(16), 8987–8994. https://doi.org/10.1002/ece3.5430

500   Cleasby, I.R., Nakagawa, S. (2011). Neglected biological patterns in the residuals. *Behavioral Ecology and*

501   *Sociobiology, 65*, 2361-2372. doi:10.1007/s00265-011-1254-7.

502   Dammhahn, M. (2012). Are personality differences in a small iteroparous mammal maintained by a life-history

503   trade-off? *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2645–2651.

504   https://doi.org/10.1098/rspb.2012.0212

505   Dingemanse, N. J., Kazem, A. J. N., Réale, D., & Wright, J. (2010). Behavioural reaction norms: animal personality

506   meets individual plasticity. *Trends in Ecology and Evolution*, *25*(2), 81–89.

507   https://doi.org/10.1016/j.tree.2009.07.013

508   Dingemanse, N. J., Moiron, M., Araya-Ajoy, Y. G., Mouchet, A., & Abbey-Lee, R. N. (2020). Individual variation in

509   age-dependent reproduction: Fast explorers live fast but senesce young? *Journal of Animal Ecology*, *89*(2),

510   601–613. https://doi.org/10.1111/1365-2656.13122

511   Dingemanse, N. J., & Wolf, M. (2010). Recent models for adaptive personality differences: A review.

512   *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1560), 3947–3958.

513   https://doi.org/10.1098/rstb.2010.0221

514   Dingemanse, N. J., & Wolf, M. (2013). Between-individual differences in behavioural plasticity within

515   populations: Causes and consequences. *Animal Behaviour*, *85*(5), 1031–1039.

516   https://doi.org/10.1016/j.anbehav.2012.12.032

517   Dochtermann, N. A. (2011). Testing Cheverud's conjecture for behavioral correlations and behavioral

518   syndromes. *Evolution*, *65*(6), 1814–1820. https://doi.org/10.1111/j.1558-5646.2011.01264.x

519   Dochtermann, N. A., Schwab, T., & Sih, A. (2015). The contribution of additive genetic variation to personality

520   variation: heritability of personality. *Proceedings of the Royal Society B: Biological Sciences*, *282*, 20142201.

521   https://doi.org/10.1098/rspb.2014.2201

522   Dohm, M. R. (2002). Repeatability estimates do not always set an upper limit to heritability. *Functional Ecology*,

523   *16*(2), 273–280. https://doi.org/10.1046/j.1365-2435.2002.00621.x

524   Dutour, M., Lévy, L., Lengagne, T., Holveck, M.-J., Crochet, P.-A., Perret, P., Doutrelant, C., & Grégoire, A. (2020).

525   Hissing like a snake: bird hisses are similar to snake hisses and prompt similar anxiety behavior in a

526   mammalian model. *Behavioral Ecology and Sociobiology*, *74*(1), 1. https://doi.org/10.1007/s00265-019-

527   2778-5

528   Fisher, D. N., David, M., Tregenza, T., & Rodríguez-Muñoz, R. (2015). Dynamics of among-individual behavioural

529   variation over adult lifespan in a wild insect. *Behavioral Ecology*, *26*(4), 975–985.

530   https://doi.org/10.1093/beheco/arv048

531   Fresneau, N., Kluen, E., Brommer, J. E. (2014) A sex-specific behavioral syndrome in a wild passerine. *Behavioral*

532   *Ecology, 25*(2), 359-367. http://doi.org/10.1093/beheco/aru008

533   Gosler, A. (1993). *The great tit*. London, U.K.: Hamlyn.

534 Grunst, A. S., Grunst, M. L., Thys, B., Raap, T., Daem, N., Pinxten, R., & Eens, M. (2018). Variation in personality

535 traits across a metal pollution gradient in a free-living songbird. *Science of the Total Environment*, *630*,

536 668–678. https://doi.org/10.1016/j.scitotenv.2018.02.191

537 Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R

538 package. *Journal of Statistical Software*, *33*(2), 1–22. https://doi.org/10.18637/jss.v033.i02

539 Hall, M. L., van Asten, T., Katsis, A. C., Dingemanse, N. J., Magrath, M. J. L., & Mulder, R. A. (2015). Animal

540 personality and pace-of-life syndromes: Do fast-exploring fairy-wrens die young? *Frontiers in Ecology and*

541 *Evolution*, *3*,28. https://doi.org/10.3389/fevo.2015.00028

542 Hamilton, W. D. (1966). The moulding of senescence by natural selection. *Journal of Theoretical Biology*, *12*(1),

543 12–45. https://doi.org/10.1016/0022-5193(66)90184-6

544 Harvey, P. H., Greenwood, P. J., & Perrins, C. M. (1979). Breeding Area Fidelity of Great Tits (*Parus major*). *The*

545 *Journal of Animal Ecology*, *48*(1), 305–313. https://doi.org/10.2307/4115

546 Johnstone, R. A., & Norris, K. (1993). Badges of status and the cost of aggression. *Behavioral Ecology and*

547 *Sociobiology*, *32*(2), 127–134. https://doi.org/10.1007/BF00164045

548 Kluijver, H. N. (1951). The population ecology of the great tit (*Parus m. major L.*). *Ardea,* 39, 1-135.

549 Kok, E. M. A., Burant, J. B., Dekinga, A., Manche, P., Saintonge, D., Piersma, T., & Mathot, K. J. (2019). Within-

550 individual canalization contributes to age-related increases in trait repeatability: A longitudinal experiment

551 in red knots. *The American Naturalist*, *194*(4), 455–469. https://doi.org/10.1086/704593

552 Kokko, H. (1997). Evolutionarily stable strategies of age-dependent sexual advertisement. *Behavioral Ecology*

553 *and Sociobiology*, *41*(2), 99–107. https://doi.org/10.1007/s002650050369

554 Krams, I., Vrublevska, J., Koosa, K., Krama, T., Mierauskas, P., Rantala, M. J., & Tilgar, V. (2014). Hissing calls

555 improve survival in incubating female great tits (*Parus major*). *Acta Ethologica*, *17*(2), 83–88.

556 https://doi.org/10.1007/s10211-013-0163-3

557   Lemaître, J. F., & Gaillard, J. M. (2017). Reproductive senescence: new perspectives in the wild. *Biological*

558   *Reviews*, *92*(4), 2182–2199. https://doi.org/10.1111/brv.12328

559   Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer.

560   Mitrus, C. (2003). A comparison of the breeding ecology of Collared Flycatchers nesting in boxes and natural

561   cavities. *Journal of Field Ornithology*, *74*(3), 293–299. https://doi.org/10.1648/0273-8570-74.3.293

562   Moiron, M., Laskowski, K. L., & Niemelä, P. T. (2020). Individual differences in behaviour explain variation in

563   survival: a meta-analysis. *Ecology Letters*, *23*(2), 399–408. https://doi.org/10.1111/ele.13438

564   Nussey, D. H., Wilson, A. J., & Brommer, J. E. (2007). The evolutionary ecology of individual phenotypic plasticity

565   in wild populations. *Journal of Evolutionary Biology*, *20*(3), 831–844. https://doi.org/10.1111/j.1420-

566   9101.2007.01300.x

567   Promislow, D. E. L., Tatar, M., Khazaeli, A. A., & Curtsinger, J. W. (1996). Age-specific patterns of genetic variance

568   in *Drosophila melanogaster*. I. Mortality. *Genetics*, *143*(2), 839–848.

569   R Core Team (2019). *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for

570   Statistical Computing. http://www.r-project.org.

571   Réale, D., Dingemanse, N. J., Kazem, A. J. N., & Wright, J. (2010). Evolutionary and ecological approaches to the

572   study of personality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1560),

573   3937–3946. https://doi.org/10.1098/rstb.2010.0222

574   Rosvall, K. A. (2011). Intrasexual competition in females: Evidence for sexual selection? *Behavioral Ecology*,

575   *22*(6), 1131–1140. https://doi.org/10.1093/beheco/arr106

576   Sibley, C. G. (1955). Behavioral mimicry in the titmice (*Paridae*) and certain other birds. *The Wilson Bulletin*,

577   *67*(2), 128–132.

578   Sih, A., Bell, A. M., Johnson, J. C., & Ziemba, R. E. (2004). Behavioral Syndromes: An Integrative Overview. *The*

579   *Quarterly Review of Biology*, *79*(3), 241–277. https://doi.org/10.1086/422893

580   Slagsvold, T. (1993). Female-female aggression and monogamy in great tits *Parus major*. *Ornis*

581   *Scandinavica*, *24*(2), 155-158. https://doi.org/10.2307/3676366

582   Smith, B. R., & Blumstein, D. T. (2008). Fitness consequences of personality: A meta-analysis. *Behavioral Ecology*,

583   *19*(2), 448–455. https://doi.org/10.1093/beheco/arm144

584   Svensson, L. (1992). *Identification guide to European passerines*. Stockholm, Sweden: Published by the author.

585   Thys, B., Lambreghts, Y., Pinxten, R., & Eens, M. (2019). Nest defence behavioural reaction norms: Testing life-

586   history and parental investment theory predictions. *Royal Society Open Science*, *6*(4).

587   https://doi.org/10.1098/rsos.182180

588   Thys, B., Pinxten, R., & Eens, M. (2020). Does the tie fit the female? Melanin-based colouration, aggressive

589   personality and reproductive investment in female great tits (*Parus major*). *Behavioral Ecology and*

590   *Sociobiology*, *74*(4), 1–11. https://doi.org/10.1007/s00265-020-2828-z

591   Thys, B., Eens, M., Pinxten, R., & Iserbyt, A. (2020). Pathways linking female personality with reproductive

592   success are trait- and year-specific. *Behavioral Ecology*, in press. https://doi.org/10.1093/beheco/araa110

593   Thys, B., Pinxten, R., Raap, T., De Meester, G., Rivera-Gutierrez, H. F., & Eens, M. (2017). The Female Perspective

594   of Personality in a Wild Songbird: Repeatable Aggressiveness Relates to Exploration Behaviour. *Scientific*

595   *Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-08001-1

596   Tilgar, V., & Koosa, K. (2019). Hissing females of great tits (*Parus major*) have lower breeding success than non-

597   hissing individuals. *Ethology*, *125*(12), 949–956. https://doi.org/10.1111/eth.12952

598   Timm, K., Koosa, K., & Tilgar, V. (2019). The serotonin transporter gene could play a role in anti-predator

599   behaviour in a forest passerine. *Journal of Ethology, 37*(2), 221-227. doi:10.1007/s10164-019-00593-7.

600   van de Pol, M., & Verhulst, S. (2006). Age-dependent traits: A new statistical model to separate within- and

601   between-individual effects. *The American Naturalist*, *167*(5), 766–773. https://doi.org/10.1086/503331

602   van de Pol, M., & Wright, J. (2009). A simple method for distinguishing within- versus between-subject effects

603   using mixed models. *Animal Behaviour*, *77*(3), 753–758. https://doi.org/10.1016/j.anbehav.2008.11.006

604   van Noordwijk, A. J., McCleery, R. H., & Perrins, C. M. (1995). Selection for the Timing of Great Tit Breeding in

605   Relation to Caterpillar Growth and Temperature. *The Journal of Animal Ecology*, *64*(4), 451.

606   https://doi.org/10.2307/5648

607   Williams, G. C. (1957). Pleiotropy, Natural Selection, and the Evolution of Senescence. *Evolution*, *11*(4), 398–411.

608   https://doi.org/10.2307/2406060

609   Wolf, M., Van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs favour the evolution of

610   animal personalities. *Nature*, *447*(7144), 581–584. https://doi.org/10.1038/nature05835

611   Zub, K., Czeszczewik, D., Ruczyński, I., Kapusta, A., & Walankiewicz, W. (2017). Silence is not golden: the hissing

612   calls of tits affect the behaviour of a nest predator. *Behavioral Ecology and Sociobiology*, *71*(5).

613   https://doi.org/10.1007/s00265-017-2313-5

Table 1: Sample sizes for aggression and hissing tests. (A) Number of individuals with regard to the number of repeated observations (between 1 and 8). (B) Number of individuals tested at each age, as well as the first and last age of observation for all individuals (age ranged from age 1 to 7, averaging 1.85 years).

(A)	Repeated observations								
	1	2	3	4	5	6	7	8	Total
Aggression	69	145	22	30	8	13	1	2	290
Hissing	3	218	2	57	1	26	0	4	311
(B)	Age								
	1	2	3	4	5	6	7		
Aggression	215	100	53	26	7	1	1		
First age	215	43	15	12	4	0	1		
Last age	156	65	38	23	6	1	1		
Hissing	235	110	52	27	10	1	1		
First age	235	43	14	12	6	0	1		
Last age	167	72	38	23	9	1	1		

618

620 **Table 2**: Results of the bivariate mixed model (model 1) predicting female aggression and hissing
621 behaviour in a population of great tits (*N* = 289) between 2016 and 2019.

	Aggression	Hissing
Fixed effects	β (CrI)	β (CrI)
Intercept	0.09 (-0.45;0.61)	0.19 (-0.38;0.80)
Age[a]	**-0.23 (-0.32;-0.15)**	-0.04 (-0.12;0.04)
Year 2017	0.13 (-0.10;0.38)	**0.40 (0.21;0.60)**
Year 2018	**0.29 (0.07;0.55)**	**0.40 (0.19;0.62)**
Year 2019	-0.15 (-0.37;0.10)	**0.55 (0.31;0.76)**
Lay date	-0.03 (-0.12;0.06)	**-0.13 (-0.21;-0.05)**
Mean clutch	0.01 (-0.06;0.07)	**-0.07 (-0.14;0.00)**
Clutch deviation	-0.13 (-0.29;0.01)	-0.01 (-0.12;0.09)
Random effects	σ² (CrI)	σ² (CrI)
ID	0.18 (0.07;0.29)	0.61 (0.48;0.74)
ID_Year	0.21 (0.09;0.33)	0.14 (0.08;0.21)
Residual	0.57 (0.49;0.66)	0.21 (0.18;0.23)
Repeatability (R)[b]	R (CrI)	R (CrI)
R_{ST}	0.42 (0.32;0.50)	0.79 (0.74;0.82)
R_{LT}	0.19 (0.07;0.29)	0.64 (0.55;0.71)

Mean estimates for fixed and random effects, as well as short-term (R_{ST}) and long-term (R_{LT})
repeatability, are given with 95% credible intervals in parentheses. Fixed effects with strong
support are in bold. Support for variances was assessed using permutation tests (see text for
details). Covariances from the model are given in Table 3.
[a] Linear covariate, fitted as age - 1.
[b] Adjusted repeatability, i.e. corrected for all fixed effects in the model.

621

622 **Table 3**: Covariances/correlations between female aggression and hissing behaviour on different
623 hierarchical levels.

Level	Covariance	Correlation
ID	0.002 (-0.093;0.090)	0.01 (-0.29;0.30)
ID_Year	0.05 (-0.01;0.12)	0.29 (-0.11;0.63)
Residual	0.02 (-0.02;0.06)	0.05 (-0.05;0.16)

Covariances/correlations, estimated from the bivariate mixed model (model 1) presented in Table
2, are given with 95% credible intervals in parentheses.

624

625

626 **Table 4**: Results from random regression (model 3) modelling aggression as a function of age.

	Estimate
Fixed effects	β (CrI)
Intercept	-0.09 (-0.42;0.22)
Mean age[a]	0.21 (-0.06;0.42)
Age[b]	**-0.37 (-0.51;-0.24)**
Random effects	σ^2 (CrI)
$ID_{intercept}$	0.40 (0.19;0.64)
ID_{slope}	0.17 (0.06;0.28)
$Cov_{intercepts-slopes}$	-0.21 (-0.36;-0.03)
$Cor_{intercepts-slopes}$	-0.80 (-0.95;-0.46)
Residual[c]	
Age 1	0.91 (0.71;1.11)
Age 2	0.37 (0.25;0.50)
Age 3	0.17 (0.11;0.25)
Age 4	0.36 (0.15;0.63)

Mean estimates for fixed and random effects are given with
95% credible intervals in parentheses. Fixed effects with
strong support are in bold. Support for variances was
assessed using permutation tests (see text for details).
[a] Calculated as (last age + 1)/2.
[b] Linear covariate, fitted as age - 1.
[c] Residual variances were allowed to differ across age
classes.

627

628
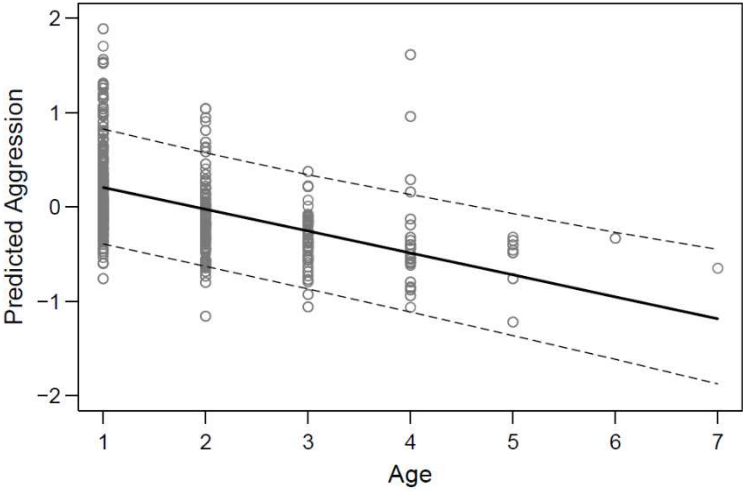
Figure 1: Female aggression with respect to age on the population level. The average population regression line (solid line), as predicted by model 1, is given with associated 95% credible intervals (dashed lines). Points depict predicted values for aggression of individuals at each age. Estimates from model 1 with left-centred age (presented in Table 2) were back-transformed to actual age.
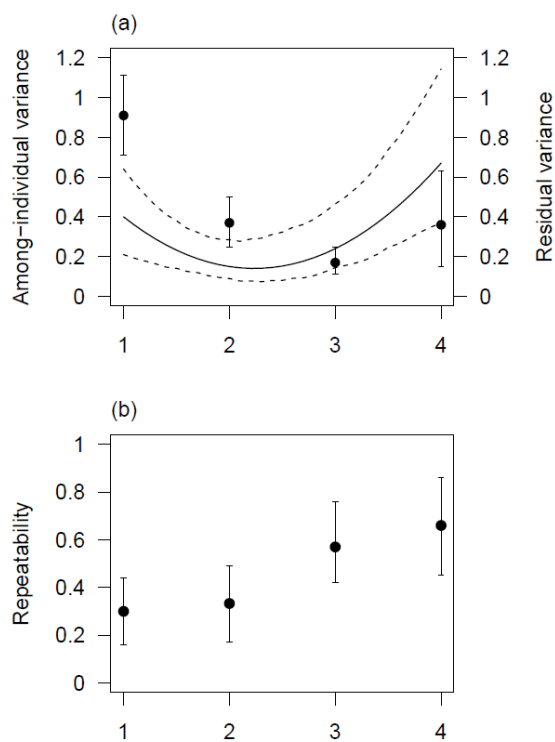
633

634
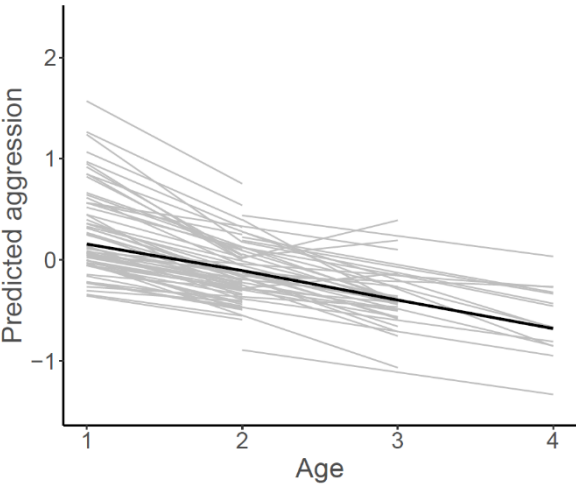
Figure 2: Individual (grey lines) and population (black line) reaction norms of female aggression with respect to age. Estimates from model 3 with left-centred age (presented in Table 4) were back-transformed to actual age.

638

639

640 **Figure 3**: (a) Variances in female aggression as a function of age: among-individual variance (solid
641 line) with associated 95% credible intervals (CrI; dashed lines) across ages and age-specific residual
642 variances (dots) with associated 95% CrI (whiskers) for each age class. (b) Repeatability of aggression
643 (dots) with associated 95% CrI (whiskers) for each age class.

644 **Appendix**

645   All models ran for 2.2×10^6 iterations, discarding the first 2×10^6 iterations (burn-in phase) and using a

646 thinning interval of 1000. This resulted in a sample of 2000 values for each estimate. To assess prior sensitivity,

647 we ran each model with different prior specifications (e.g. inverse Wishart, inverse gamma). Using different

648 relatively uninformative priors gave qualitatively similar results. Results presented in this article are from models

649 with the following prior specifications

650 (1) Bivariate mixed model

651   list(R=list(V=diag(2)/3,nu=2), G=list(G1=list(V=diag(2)/3, nu=2), G2=list(V=diag(2)/3, nu=2)))

652 (2) Univariate mixed models

653   list(R=list(V=1/3,nu=1), G=list(G1=list(V=1/3, nu=1), G2=list(V=1/3, nu=1)))

654 (3) Random regression model

655 list(R=list(V=diag(4)/4,nu=0.01), G=list(G1=list(V=diag(2)/3, nu=0.01)))

656 **Table A1:** PCA loadings for aggression parameters scored during territorial intrusion experiments

	PC1
Eigenvalue	1.40
Proportion of total variance	0.49
No. of alarm calls	-0.317
Approach distance[a]	0.332
Time on decoy	0.646
No. of attacks	0.610

[a] Approach distance was multiplied by -1 prior to analysis.

657

658 **Table A2:** Random effect estimates in explaining variance in aggression and hissing behaviour

	Aggression	Hissing
Fixed effects (β)		
Intercept	-0.03 (-0.32;0.26)	-0.08 (-0.39;0.25)
Random effects (σ^2)		
ID	0.30 (0.18;0.43)	0.71 (0.59;0.84)
Decoy[a]	0.01 (0.00;0.04)	0.01 (0.00;0.02)
Observer[b]	0.04 (0.00;0.12)	0.02 (0.00;0.09)
Year[c]	0.02 (0.00;0.18)	0.03 (0.00;0.28)
Residual	0.67 (0.59;0.79)	0.24 (0.21;0.27)

Mean estimates for fixed and random effects are given with 95% credible intervals in parentheses.

[a] Refers to same-sex decoy (five random levels) for aggression and to woodpecker (three random levels) for hissing behaviour.

[b] 11 random levels associated with observer identity.

[c] Four random levels associated with year (2016–2019).

659

Table A3: Output of univariate mixed models (model 2) using different (sub)sets of data

	Full data		Age 5, 6 and 7 pooled		Age > 5 removed	
	Aggression	Hissing	Aggression	Hissing	Aggression	Hissing
Fixed effects (β)						
Intercept	0.29 (-0.39;0.92)	0.32 (-0.71;1.18)	0.32 (-0.37;0.99)	0.31 (-0.36;0.98)	0.34 (-0.31;1.03)	0.33 (-0.33;1.00)
Mean age	-0.13 (-0.32;0.06)	0.04 (-0.16;0.27)	-0.16 (-0.34;0.05)	0.03 (-0.20;0.25)	-0.17 (-0.36;0.02)	0.02 (-0.19;0.24)
Age deviation	**-0.30 (-0.44;-0.15)**	-0.08 (-0.20;0.03)	**-0.30 (-0.44;-0.15)**	-0.08 (-0.20;0.03)	**-0.30 (-0.44;-0.15)**	-0.08 (-0.20;0.04)
Year 2017	0.14 (-0.10;0.36)	**0.49 (0.30;0.68)**	0.15 (-0.09;0.40)	**0.48 (0.30;0.67)**	0.16 (-0.09;0.38)	**0.48 (0.29;0.68)**
Year 2018	**0.32 (0.10;0.56)**	**0.43 (0.20;0.63)**	**0.32 (0.07;0.56)**	**0.41 (0.22;0.65)**	**0.32 (0.09;0.56)**	**0.40 (0.19;0.64)**
Year 2019	-0.09 (-0.34;0.16)	**0.62 (0.34;0.91)**	-0.10 (-0.37;0.18)	**0.61 (0.35;0.89)**	-0.09 (-0.36;0.17)	**0.60 (0.32;0.87)**
Lay date	-0.01 (-0.02;0.01)	**-0.02 (-0.03;-0.01)**	-0.01 (-0.02;0.01)	**-0.02 (-0.04;-0.01)**	-0.01 (-0.03;0.01)	**-0.02 (-0.03;-0.01)**
Mean clutch	-0.10 (-0.24;0.04)	-0.06 (-0.16;0.05)	-0.10 (-0.25;0.02)	**-0.06 (-0.13;0.00)**	-0.10 (-0.24;0.03)	**-0.06 (-0.12;0.00)**
Clutch deviation	0.00 (-0.06;0.06)	0.01 (-0.13;0.13)	0.00 (-0.06;0.06)	0.01 (-0.09;0.08)	0.00 (-0.06;0.06)	0.00 (-0.08;0.09)
Random effects (σ²)						
ID	0.19 (0.06;0.32)	0.62 (0.49;0.75)	0.18 (0.06;0.31)	0.63 (0.50;0.78)	0.18 (0.06;0.31)	0.63 (0.51;0.77)
ID_Year	0.19 (0.06;0.31)	0.13 (0.07;0.18)	0.19 (0.07;0.33)	0.12 (0.07;0.18)	0.19 (0.07;0.34)	0.13 (0.07;0.19)
Residual	0.57 (0.49;0.67)	0.20 (0.17;0.23)	0.57 (0.48;0.66)	0.20 (0.17;0.23)	0.57 (0.48;0.66)	0.20 (0.18;0.23)
N (ID;obs)	290;686	311;866	290;686	311;866	289;683	310;862

Mean estimates for fixed and random effects are given with 95% credible intervals in parentheses. Fixed effects with strong support are in bold. *N* (ID;obs): number of individuals/number of observations per analysis.

1   **Table A4**: Output of univariate mixed models using only data of females with full life histories

	Aggression	Hissing
Fixed effects (β)		
Intercept	0.63 (-0.49;1.81)	-0.09 (-1.25;1.03)
Mean age	-0.29 (-0.67;0.08)	0.19 (-0.26;0.63)
Age deviation	**-0.54 (-0.86;-0.24)**	-0.21 (-0.51;0.06)
Year 2017	0.19 (-0.19;0.52)	**0.56 (0.24;0.85)**
Year 2018	0.37 (-0.06;0.79)	**0.48 (0.02;0.89)**
Lay date	-0.01 (-0.03;0.02)	-0.01 (-0.04;0.01)
Mean clutch	-0.04 (-0.15;0.08)	-0.04 (-0.15;0.07)
Clutch deviation	0.06 (-0.14;0.27)	0.09 (-0.14;0.27)
Random effects (σ²)		
ID	0.29 (0.12;0.49)	0.66 (0.44;0.88)
ID_Year	0.00 (0.00;0.00)	0.09 (0.02;0.18)
Residual	0.68 (0.53;0.85)	0.21 (0.16;0.25)
N (ID;obs)	119;255	130;329

Mean estimates for fixed and random effects are given with 95% credible intervals in parentheses. Fixed effects with strong support are in bold. *N* (ID;obs): number of individuals/number of observations per analysis.

2

3
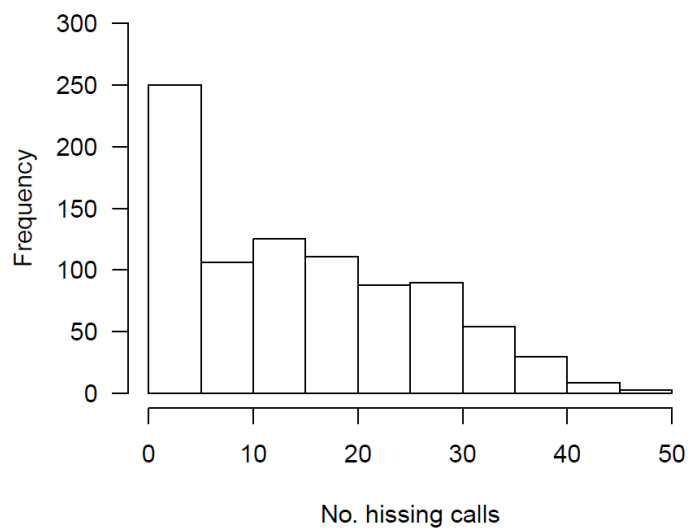
4

5   **Figure A1**: Distribution of the number of hissing calls produced.