# No reproductive benefits of dear enemy recognition in a territorial songbird

Michael S. Reichert[1,2,*], Jodie M.S. Crane[2,3], Gabrielle L. Davidson[2,4], Eileen Dillane[2], Ipek G. Kulahci[2], James O'Neill[2], Kees van Oers[5,6], Ciara Sexton[2], John L. Quinn[2]

1. Department of Integrative Biology, Oklahoma State University, USA
2. School of Biological Earth and Environmental Sciences, University College Cork, Ireland
3. Kākāpō Recovery Programme, Department of Conservation, Invercargill, New Zealand
4. Department of Psychology, University of Cambridge, Cambridge, UK
5. Behavioural Ecology Group, Wageningen University, Wageningen, The Netherlands
6. Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, The Netherlands

Correspondence:
Michael Reichert
501 Life Sciences West, Department of Integrative Biology, Oklahoma State University, Stillwater, OK 74078 USA
*email: michael.reichert@okstate.edu

ABSTRACT

Territorial animals often exhibit the dear enemy effect, in which individuals respond less aggressively to neighbours than to other individuals. The dear enemy effect is hypothesized to be adaptive by reducing unnecessary aggressive interactions with individuals that are not a threat to territory ownership. A key prediction of this hypothesis, that individual fitness will be affected by variation in the speed and extent to which individuals reduce their aggression towards neighbours relative to strangers, has never been tested. We used a series of song playbacks to measure the change in response of male great tits on their breeding territories to a simulated establishment of a neighbour on an adjacent territory. Males reduced their approach to the speaker and sang fewer songs on later repetitions of the playback trials, consistent with a dear enemy effect through habituation learning. However, not all males discriminated between the neighbour and stranger playbacks at the end of the series of trials, and there was evidence that individuals consistently differed from one another in performing this discrimination. We monitored nests and analysed offspring paternity to determine male reproductive success. Unexpectedly, individuals that exhibited dear enemy behaviour towards the simulated neighbour did not have higher reproductive success, and in fact one measure, total offspring biomass, was lower for individuals that showed the dear enemy effect. Although the general capability to recognize neighbours is most likely adaptive, it seems that individuals who decrease their responsiveness to familiar neighbours too quickly may gain no advantage or even be at a disadvantage.

Key words: learning, individual recognition, playback, great tit, territorial behaviour, cognition

## INTRODUCTION

Territorial animals live in a social environment in which they may benefit by responding differently to their territorial neighbours than they do to other individuals (McGregor, 1993). For instance, many species show 'dear enemy effects', in which individuals respond less aggressively towards neighbours than towards strangers (Stoddard, 1996; Temeles, 1994; Tumulty, 2018; Ydenberg et al., 1988). A leading hypothesis for the evolution of the dear enemy effect is that it is adaptive because it enables territory holders to avoid wasted time, energy and injury risk in aggressive interactions with individuals that are not a threat because they hold their own territory (the 'threat level' hypothesis (Getty, 1987; Temeles, 1994)). This hypothesis predicts that intraspecific variation in the speed and extent to which individuals reduce their aggression towards non-threatening neighbours will affect individual fitness. However, this key prediction has never been tested, and indeed the fitness consequences of neighbour recognition have mostly been studied indirectly, for instance by demonstrating that individuals holding territories with long-term neighbours have higher reproductive success than those with new neighbours (Beletsky & Orians, 1989; Grabowska-Zhang, Sheldon, et al., 2012; Grabowska-Zhang, Wilkin, et al., 2012; Siracusa et al., 2021). The dear enemy effect is far from ubiquitous among species. In some species, individuals respond more aggressively to neighbours than strangers (Bruintjes et al., 2016; Muller & Manser, 2007; Temeles, 1990), and in others they do not discriminate between them at all (Battiston et al., 2015; Bee, 2003a; Tumulty et al., 2018). Within species, variation in the expression of the dear enemy effect is often context dependent (Courvoisier et al., 2014; Hyman, 2005; Pratt & McLain, 2006), but these studies examined population-level responses, and individual variation and its fitness consequences have not been measured. Thus, while it is clear that neighbour recognition systems vary and have evolved, there is still a limited understanding of the selection pressures that have driven this evolution, especially at the within-species level.

The dear enemy effect is facilitated by mechanisms for neighbour recognition learning (Wiley, 2013). Thus, variation among individuals in the expression of the dear enemy effect (i.e. the extent and speed with which aggression is reduced towards neighbours relative to strangers) may arise because of variation in cognitive abilities associated with neighbour recognition. If this cognitive variation affects reproductive success, for instance if those individuals on breeding territories that quickly learn to recognize their neighbours can invest the time and energy saved by avoiding unnecessary agonistic encounters into their offspring, then individuals with superior cognitive ability for neighbour recognition will have higher fitness and be favoured by selection. However, non-cognitive factors also affect the expression of the dear enemy effect. For instance, individuals may differ in aggressiveness or other personality traits, and context-dependent factors such as territory size and the density of neighbouring territory-holders also likely affect responsiveness to neighbours (Amy et al., 2010; Jacobs et al., 2014; Pratt & McLain, 2006; Yoon et al., 2012). Regardless of the source, this individual variation in dear enemy expression has potentially significant, but unexplored, consequences for fitness. Furthermore, few studies have examined the trajectory over which the dear enemy effect develops upon establishment of a new neighbour (Bee & Gerhardt, 2001a), but this time period in which relationships are being established is likely when individual differences will be greatest, and have strongest effects on fitness.

Like many songbirds, great tits (*Parus major*) defend territories around their nest during the breeding season (Gosler, 1993). Male great tits defend territories by singing and approaching intruding individuals to engage in visual displays and occasionally, physical fighting (Blurton Jones, 1968; Krebs, 1977). Great tits nest in natural cavities or artificial nest boxes, and the primary threat from intruders is the potential for them to usurp the limited resource of a high-quality nesting location (Krebs, 1971, 1976, 1982),

although intruders also present other threats including potentially mating with the territory-holder's social mate (Hill et al., 2011) or foraging on the territory (Hinde, 1956). The dear enemy effect has been demonstrated previously in great tits: in playback experiments, individuals showed a reduced response to songs of their territorial neighbours compared to songs of strangers (Falls et al., 1982; Krebs, 1971; McGregor & Avery, 1986), and the effect was stronger when the neighbour's song was played from its actual territory (McGregor & Avery, 1986). Recognition of neighbours in great tits is not based on a simple discrimination between the categories of familiar and unfamiliar individuals (Wiley, 2013); instead they can learn to discriminate among the songs of different specific individuals (McGregor & Avery, 1986; Weary & Krebs, 1992). There is also among-individual variation in aggressive responses to playback resulting from personality differences (Amy et al., 2010). Great tits trade off the time invested in foraging and territory defence (Ydenberg, 1984; Ydenberg & Krebs, 1987); thus, individuals that rapidly reduce their aggression towards their neighbours may benefit directly by increasing their foraging intake, and this in turn could enhance their ability to provision for their offspring (Martin, 1987). However, vigilance must be maintained against potential territorial usurpers, because the longer a usurper is on the territory, the more effort is required to expel it (Krebs, 1982).

To determine whether individual variation in dear enemy behaviour affects reproductive success, we performed a series of acoustic playback experiments in which we monitored territorial males' responses to the simulated arrival of a new male on a neighbouring territory. We broadcast the playback stimulus multiple times across three days to determine if males' responses to the playback declined over time, and tested whether any decline was specific to the song used as the playback stimulus and therefore consistent with discrimination between neighbours and strangers under the dear enemy effect. We classified individuals as either expressing dear enemy behaviour or not after nine repetitions of the playback, and then monitored each male's nest to test the prediction that individuals that rapidly

reduced their response to neighbours would have higher reproductive success in terms of number of offspring fledged and/or nestling mass, which is positively correlated with subsequent survival (Tinbergen & Boerlijst, 1990). We also examined the paternity of nestlings to determine if the reduction in aggressive response towards neighbours affected males' abilities to defend against extrapair copulations. By using experiments in a wild population to test a key prediction for the adaptive function of the dear enemy effect, we provide novel insights into a widespread territorial defence behaviour.

## METHODS

Playbacks were performed during the spring breeding season (April-May in 2017 and 2018) in eight small forestry plots in County Cork, Ireland (Table S1). Each site contained an array of nestboxes, which are preferentially used by great tits for breeding (East & Perrins, 1988). We identified potential playback subjects by listening for males singing near nestboxes, and examining the progress of nest building in the box. Males chosen for the experiment were then subject to a series of song playbacks as described below.

The project received ethical approval from the Animal Welfare Body at University College Cork (HPRA license number AE19130-P017), and was in accordance with the ASAB (Association for the Study of Animal Behaviour) Guidelines for the Treatment of Animals in Behavioural Research and Teaching. All research was conducted under licenses from the National Parks and Wildlife Service of Ireland and the British Trust for Ornithology as part of ongoing research in these populations.

### *Playback procedure*

Male great tits sing on territories containing their nesting site. Males typically have a repertoire of approximately three to four distinctive song types (Gompertz, 1961; McGregor & Krebs, 1982b). However, during a song bout they usually repeat the same song type for several minutes before switching to a different song type (Krebs, 1976). The playbacks were designed to mimic this repetition of a single song type. Although simulating a larger song repertoire through playbacks may have captured additional aspects of the dear enemy phenomenon, this was not necessary to address our primary aims and would have reduced our power to detect dear enemy behaviour because: 1. Males habituate to the presentation of both a single song type and multiple song types, but habituation is slower to playback of multiple song types (Krebs, 1976), and 2. Males will respond to the presentation of a single song type in a manner consistent with the dear enemy effect: responding more strongly to a single song type of a stranger than to a single song type of an established neighbour (Krebs, 1971).

The playback stimuli consisted of recordings of natural male songs made using Wildlife Acoustics SM4 audio recorders (24 kHz sampling rate) placed at nestboxes in May 2016. We scanned audio files for exemplars of male songs with a high signal to noise ratio and containing no other bird songs in the background. We inserted each chosen song into a new audio file in Audacity software, bandpass filtered the song between 1.5 and 11.5 kHz, and manipulated the song to contain 6 phrases (the basic repeating unit of the song (McGregor & Krebs, 1982a)) by copying or deleting phrases as needed. The song exemplar was then copied so that it was played back at a rate of either five ("low rate treatment") or ten ("high rate treatment") songs per minute for five minutes. These two stimulus rates were used to test the prediction that the likelihood of recognition learning depended on the stimulus repetition rate (Thompson & Spencer, 1966). A total of 29 song exemplars were chosen, and each exemplar was recorded at a different nestbox; songs came from seven of the eight study sites. The playback stimulus selected for the subject male and the song rate treatment were chosen randomly, with the restriction

that the stimulus song was not recorded from the same site as the subject (all sites were at least 2.25 km apart from each other; Table S1). This ensured that subjects would not have been familiar with the playback song already. Stimuli were broadcast as .wav files from an EasyAcc X02s speaker mounted on a tripod at approximately 1 m height, at a sound-pressure level of 90 dB (A) measured at 1 m using an Extech 407730 sound level meter.

On each of three consecutive days we exposed males to three playbacks, performed at 1 hour intervals, of one of the playback stimuli (Figure 1). Each subject was therefore exposed to a total of nine playbacks of the same stimulus (henceforth the "neighbour" stimulus), which simulated a newly arrived male on a neighbouring territory. Playbacks were performed between 0800h and 1540h; each male was tested at approximately the same time on each of the three days of the experiment. The playbacks were always performed from the same location, and at a distance of 25 m from the nestbox. This location was chosen because territoriality is likely strongest near the nestbox (Giraldeau & Ydenberg, 1987), 50 m is a typical nearest-neighbour distance (Krebs, 1971), and neighbour-stranger discrimination is typically strongest near the territorial border (Falls & Brooks, 1975; Stoddard et al., 1991). We therefore standardized the absolute distance of the speaker from the territory's centre in order to standardize sound amplitude at the nestbox for all males. There is an inevitable trade-off between the advantages of this approach and those of estimating males' actual territorial boundaries and placing the speaker at each male's estimated territorial edge. Both for logistical reasons and because absolute stimulus intensity is likely to play a role in the speed of dear enemy learning (Bee, 2001), we used a standardized playback distance. The specific location of the playback was chosen in a randomized direction, with the constraint that playbacks were only performed from areas the experimenter could access, and that did not overlap with the territory of another male great tit.

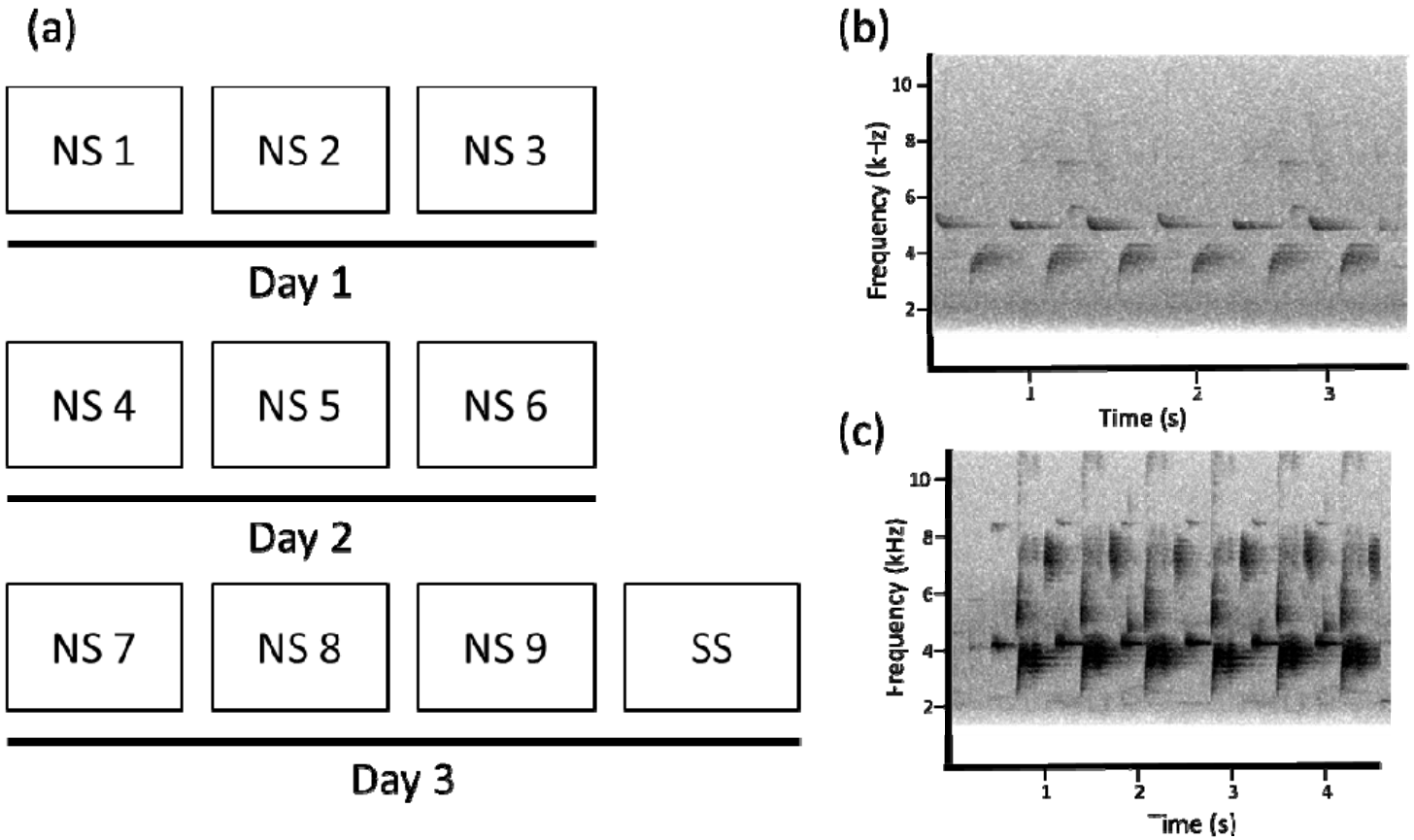


Figure 1. (a) Timeline of the playback experimental design. We broadcast the neighbour stimulus ('NS') for five minutes per trial (number in each box corresponds to the trial number), for a total of three trials per day with approximately one hour between trials. This procedure was repeated for three consecutive days. On the third day, immediately following the final neighbour playback (NS 9), we broadcast a different song from a different location ('SS': stranger stimulus, a test of stimulus specificity). In 2017, the entire procedure was repeated with different stimuli

192 after a period of ten days to test the repeatability of dear enemy behaviour. (b) Exemplar playback stimulus. The spectrogram shows a single
193 great tit song that was used to build one of the playback stimuli. (c) a second exemplar stimulus, taken from a different bird, illustrating the
194 acoustic variation in great tit songs that allows for dear enemy recognition. An example playback experiment could use the song in (b) for the
195 neighbour stimulus and that in (c) for the stranger stimulus. Note that the spectrograms do not illustrate the equalization of playback amplitude
196 that was used for the trials.

During each playback we monitored males' behaviour and made audio recordings of any songs using Marantz PMD 660 or PMD 661 audio recorders with Sennheiser ME67 directional microphones (16 bit, 44.1 kHz sampling rate). We noted the closest approach of the male to the playback speaker, in categories of 5 m (Nelson & Soha, 2004). Subjects were not always located at the nestbox at the start of the playback, so approach distances greater than 25 m were possible. Individuals that did not respond at all were given a value of "None" for approach distance, which in the ordinal analyses described below was considered the greatest distance. From the audio recordings we counted the number of songs produced by the male during the playback. We considered males to have responded to the playback if they sang at any point during the stimulus broadcast, or made any movement towards the playback speaker, and to have not responded if they did neither of those. The final dataset included 51 individuals; data from the first trial recording for one bird included in this total were lost because of a faulty microphone cable.

On the third day, immediately after the ninth exposure of the subject to the neighbour stimulus, we broadcast a different song stimulus from a different location (still 25 m away from the nestbox). The aim was to compare the response to this new stimulus, simulating a new stranger individual (henceforth, the "stranger" stimulus), to the original playback stimulus to which we hypothesized the subject would have developed some familiarity towards. This is an essential step for demonstrating individual recognition in territorial systems, in which individuals reduce their response to familiar neighbours, but maintain a heightened response to an unfamiliar stranger (Bee et al., 2016).

In 2017 only, we repeated this three-day procedure with 20 individuals ten days after their final trial from the first round of playbacks. The aim of this repetition was to investigate whether the expression of dear enemy behaviour was repeatable. Therefore, we performed playbacks as above, but with a

221 different song stimulus from a different location, simulating a different new neighbour establishing a
222 territory. A different stranger stimulus was also used in these repeated trials, and the male was given
223 the opposite rate treatment of what it had been exposed to during the first round of playbacks.
224
225 *Criterion for dear enemy recognition*
226
227 The classical method for demonstrating the dear enemy effect is to, in a single trial, present individuals
228 with the signals of an actual established neighbour, followed by a single trial presentation of the signals
229 of an individual that the subject would not have interacted with (Battiston et al., 2015; Brunton et al.,
230 2008; Wei et al., 2011). The dear enemy effect is inferred when individuals have a reduced response to
231 the neighbour stimulus compared to the stranger stimulus. In our protocol we simulated the
232 establishment of a new neighbour through a series of playbacks, which essentially served as training
233 sessions to give the subject the opportunity to learn to recognize its neighbour. Therefore, to test
234 whether or not the subject was indeed exhibiting dear enemy behaviour at the end of the necessarily
235 fixed number of trials, we used the classical criterion for testing the dear enemy effect: comparison of
236 the response to a final neighbour playback and a stranger playback. Our criterion (hereafter referred to
237 as the 'standard criterion') that the subject was expressing dear enemy behaviour was that it did not
238 respond in the final (ninth) playback of the simulated neighbour, but then did respond to the
239 subsequent playback of a simulated stranger.
240
241 Although we argue that our criterion effectively identifies individuals that recognized their neighbour
242 and were expressing dear enemy behaviour towards it, we recognize that this binary criterion has
243 limitations as a metric of dear enemy expression. Therefore, we explored additional measures of an
244 individual's change in response towards its neighbour over time. First, in some cases individuals

245 responded to neither the final neighbour playback nor the stranger playback. These results are difficult
246 to interpret because it may indicate a general loss of motivation to defend the territory rather than a
247 failure to discriminate neighbours and strangers. We therefore repeated all analyses, defining dear
248 enemy behaviour as above, but only including those birds that responded to the stranger stimulus
249 (referred to as the 'standard criterion without nonresponders'; no individual failed to respond to the
250 stranger playback after responding to the final neighbour playback). Second, for each individual, we
251 extracted slope parameters from a logistic regression of the binary response variable on trial number
252 (not including the stranger playback), and a linear regression of the number of songs on trial number.
253 The aim was to obtain a more quantitative estimate of the change in response across trials that may
254 reveal more variation than our binary criterion. If individuals have developed dear enemy recognition of
255 the newly-established neighbour, these values would be expected to be negative, indicating a decline in
256 response across trials. However, a decline in response to a neighbour is not sufficient to demonstrate
257 the dear enemy effect, because individuals must continue to respond to strangers. Therefore, we only
258 included individuals that responded to the stranger playback for these analyses.
259
260 *Breeding data*
261 Breeding data were obtained as part of standard monitoring protocols for the project (O'Shea et al.,
262 2018). We recorded the date when the first egg was laid, the total clutch size, date of hatching and
263 number of fledglings. The identity of the subject male sometimes could be determined by identification
264 of unique colour rings, if the male had already been captured prior to the experiment. Some males were
265 also identified by the use of RFID-equipped nestbox entrance doors, which could read the unique
266 passive integrated transponder tag placed on the leg of previously captured males. Males could also be
267 identified when breeding adults were caught at the nest for ringing and measurements, 10-12 days after
268 the eggs hatched. Fifteen subjects were not identified because the male was not tagged, the nest was

269 abandoned prior to trapping, or the male could not be caught. However, our analyses do not rely on
270 knowing the specific identity of the subject, and we can safely assume that no male was recorded in the
271 study more than once based on the timing and distribution of boxes at which playbacks were carried out
272 across the eight sites. Breeding densities are low at our sites and in cases where males could not be
273 identified by color rings we nevertheless consider it highly unlikely that more than one individual
274 responded to the playback on different trials. Chicks were weighed at day 15, and we determined the
275 number of fledglings by inspecting the nest for any dead chicks after the breeding attempt was
276 complete.
277
278 *Paternity analysis*
279
280 Estimates of reproductive fitness in birds depend on levels of extra-pair paternity (EPP) (Webster et al.,
281 1995). Although rates of EPP are relatively low in great tits (Patrick et al., 2012; van Oers et al., 2008),
282 they could have altered the relationship between dear enemy behaviour and reproductive success,
283 particularly as paternity loss is one of the potential costs of territorial intrusions, and neighbours and
284 strangers may differ in the threat they pose to paternity (Schlicht et al., 2015). EPP levels may also
285 influence selection on males to cooperate via dear enemy effects (Eliassen & Jørgensen, 2014). We
286 therefore analyzed males' reproductive success using metrics that excluded any offspring that were
287 identified in a paternity analysis as being extra-pair offspring.
288
289 DNA was obtained from feathers taken from breeding pairs and offspring on their respective dates of
290 capture and ringing. DNA extraction was performed using the protocol of the E.Z.N.A. Tissue DNA Kit
291 (Omega Bio-Tek, Norcross, GA, USA), with the exception of the use of 80 µL elution buffer on a single
292 elution step. Samples were genotyped at eight microsatellite loci, selected based on previously observed

293 variability and utility, as well as potential for multiplexing in a single reaction (Pma69u (k=7 alleles)
294 (Kawano, 2003); PmaD22 (k=20), PmaCan1 (k=15), PmaGAn30 (k=5), PmaC25 (k=19), PmaTAGAn86
295 (k=19), PmaTGAn33 (k=17) and PmaTGAn45 (k=11) (Saladin et al., 2003)). Multiplex PCR was performed
296 in a 3.5 µl total volume, including 1 µl of DNA extract, and 1.75 µl of 2x Top-Bio™ Plain Combi PP
297 Mastermix, with a concentration of 0.03 µM for the Pma69U, PmaCan1, PmaGAn30 and PmaTAGAn86
298 primers and 0.06 µM for the PmaC25, PmaD22, PmaTGAn33 and PmaTGAn45 primers. The PCR was
299 programmed with an initial denaturation at 95°C (15 min) followed by five cycles of 94°C (30 s), 55°C (90
300 s), 72°C (60 s), then 27 cycles at 94°C (30 s), 57°C (90 s), 72°C (60 s), followed by an elongation step at
301 60°C for 30 min. The PCR products were diluted in 14 µL nuclease free water; and run on an Applied
302 Biosystems ABI3500xl DNA analyser using POP-7 polymer with GeneScan™ 600 LIZ™ Dye Size Standard
303 v2.0 (ThermoFisher Scientific). We used GeneMarker version 2.7.0 software (SoftGenetics, Pennsylvania,
304 USA) to determine allele sizes.
305
306 Paternity was assigned using CERVUS version 3.0.7 software (Kalinowski et al., 2007) using 10,000 cycles,
307 94 candidate fathers, a 0.02% error rate, two candidate parents and 93% of loci typed as simulation
308 parameters. Only individuals that were successfully genotyped at five or more loci were included in
309 paternity analyses. Individuals were determined to be within-pair offspring if all loci matched those of
310 the social father and social mother combination, or if there was a mismatch at only one of the loci but
311 the social father was identified as the most likely father using critical trio LOD scores returned by the
312 program. Offspring that did not meet these criteria were categorized as extra-pair offspring. We
313 assumed that offspring whose paternity was not determined (9 of 96 fledged offspring and 13 of 111
314 weighed offspring were of unknown paternity; two offspring fledged but were not weighed or analysed
315 for paternity because they fledged on the day of weighing) were within-pair offspring. The combined
316 exclusion probability for all eight microsatellites was >99.99%. Two of our loci significantly deviated from

317 Hardy-Weinberg equilibrium when the genotypes of all individuals in the analysis were included
318 (PmaC25: $\chi^2$ = 40.98, df = 15, p < 0.001, PmaD22: $\chi^2$ = 33.87, df = 15, p = 0.004). This was likely due to
319 the family structure of the data.
320
321 *Data analysis*
322 Except for the analyses of repeatability (see below), all analyses were performed only on the first
323 repetition of the playback trials (i.e., we excluded data from the repetition of the procedure that was
324 performed in 2017 only). We address the following questions in our analyses:
325
326 1. Does the response to the neighbour playback decrease over time? We tested whether males reduced
327 their aggressiveness as they became more exposed to the songs of other males using three different
328 variables, all of which have been demonstrated to be related to the aggressive response in great tits in
329 previous studies (Amy et al., 2010; Doutrelant et al., 2000; McGregor & Avery, 1986; Snijders et al.,
330 2016): "response", i.e. whether the male responded at all by singing or approaching (binary), approach
331 distance (ordinal categorical), and number of songs during individual trials. In all cases the trial number
332 (1-9) and stimulus rate treatment (low or high) were entered as a fixed effects and individual identity as
333 a random effect. Response was modelled using a binomial generalized linear mixed model using the
334 glmer function in the lme4 version 1.1-23 package (Bates et al., 2015) in R 4.0.2 software (R
335 Development Core Team, 2021). Approach category was modelled using a cumulative link mixed model
336 to account for the ordinal nature of the dependent variable using the clmm function in the ordinal
337 2019.12-10 package (Christensen, 2019). Number of songs was modelled as a poisson variable using the
338 glmmTMB function of the glmmTMB 1.0.2.1 package (Brooks et al., 2017), which accounts for zero
339 inflation in the dataset.
340

2. Is there a difference in the response to the stranger playback compared to the final (ninth) familiar neighbour playback? If there was a decrement in aggression in the analyses above, the next step to demonstrate the dear enemy effect is to show that this decrement is stimulus specific. Therefore, individuals were predicted to respond more strongly to the stranger playback than to the final neighbour playback. We tested the same variables as above, using the same analyses but with the trial variable a factor with two levels: final neighbour or stranger playback trial.

3. Is there a difference in reproductive success between individuals that did or did not express dear enemy behaviour? We ran separate models for each of five different metrics of reproductive success, and each of the four measures of dear enemy behaviour (see *Criterion for dear enemy recognition*, above). For these analyses, we excluded data from three males from one of the sites (Dunderrow, see Table S1) in 2018 because of widespread nest predation. Although predation is certainly a component of reproductive success, in this site almost every nest was completely predated by stoats, and therefore we consider that there could be no relationship between the male phenotypic characteristics under study and the survival of offspring in this site with unusually high predation, rendering reproductive measures of these individuals meaningless in the context of our hypotheses.

In all reproductive success models we included whether or not the individual exhibited dear enemy behaviour, as defined above, as a fixed factor. Initial models also included the date the first egg was laid in the clutch, with the first of March as day 1, as well as the date of the first playback trial relative to the first egg date, but neither of these variables ever explained variation in reproductive success and so were excluded from the final models (the lack of an effect of first egg date was unexpected, but see (O'Shea et al., 2018) for an explanation). 1. Clutch size was modelled as a poisson variable in a generalized linear model. For one subject, no eggs were laid at the nestbox, and it was given a clutch

size value of zero, as well as zeroes for the other measures of reproductive success described next. 2. Number of within-pair fledglings was a poisson variable, but there were a large number of zero values. We therefore modelled it accounting for zero-inflation as above. 3. The sum total of the mass of all within-pair offspring in the nest on day 15. There were a large number of zero values because in many nests all chicks died, in which case the nest was assigned a total biomass of zero, so this continuous variable was modelled with a tobit regression using the censReg function in the censReg 0.5-30 package (Henningsen, 2019). One nest was removed from this analysis because two offspring were observed fledging on the day of weighing, and therefore were not weighed, so an accurate measure for this variable could not be obtained 4. Average mass of within-pair offspring on day 15. Although this measure may be more straightforward than the total mass analysis described above, it has the disadvantage of excluding the large number of subjects (N=18) whose nests failed entirely before day 15. This reduces sample size and also excludes individuals facing the most severe fitness consequences. We entered average offspring mass as the dependent variable in a linear model with an additional factor of brood size (including both within- and extra-pair offspring). 5. The actual mass of each individual within-pair offspring. This analysis also excludes failed nests, but may reveal important variation in parental investment among those nests that did survive. Individual mass was entered as a Gaussian variable in a linear mixed model, with brood size as an additional factor.

Initial models with site included as a random effect could not be run because of singularity issues, because there were few samples from most sites (see Table S1). For the same reason, we did not perform a formal analysis of reproductive success by site, but there were no obvious qualitative differences in these variables between the four sites with the largest numbers of individuals tested (Table S1).

We estimated the repeatability of dear enemy recognition behaviour using the rptR package (Nakagawa & Schielzeth, 2010) with the repetition (first or second) and stimulus rate treatment as fixed factors, individual as a random factor and whether the individual did or did not exhibit the dear enemy effect (standard criterion) as a binary dependent variable. One individual was removed from this analysis because it did not respond during any trial in the second repetition. Repeatabilities of the response and song slope measures were analysed similarly, but with the dependent variable modelled as Gaussian. Finally, we used a binomial generalized linear model to test whether there were differences between individuals that did or did not exhibit the dear enemy effect (standard criterion) in the proportion of fledglings that were extra-pair.

## RESULTS

*Change in response to the neighbour playback*

The likelihood of a response (either or both of singing or moving towards the speaker) significantly decreased across the neighbour playback trials (Table 1; Figure 2A). The pattern of responses shows an initial slight increase in the likelihood of responding over the first few trials, and then larger decreases especially in the final trials of the day. Approach distance was significantly greater on later trials, such that individuals were more likely to make a close approach to the speaker on earlier trials, and approached but stayed further away or did not approach at all on later trials (Table 1). Likewise, the number of songs produced by males during the playback decreased across trials (Table 1; Figure 2B). There was no effect of stimulus rate on response or approach distance, but there was a non-significant trend for an effect on the number of songs, with males giving more songs in response to stimuli presented at the higher rate (Table 1).

Table 1. Tests of change in response to neighbour playback across trials.

| Variable | Factor | Estimate (SE) | z | P |
|---|---|---|---|---|
| Response | Intercept | 0.53 (0.27) | | |
| | Trial number | −0.10 (0.04) | −2.52 | 0.01 |
| | Rate | 0.04 (0.28) | 0.13 | 0.89 |
| Approach distance | Trial number | 0.11 (0.04) | 3.10 | 0.002 |
| | Rate | −0.01 (0.30) | −0.03 | 0.97 |
| Number of songs | Intercept | 2.91 (0.09) | | |
| | Trial Number | −0.04 (0.01) | −5.88 | <0.001 |
| | Rate | 0.23 (0.12) | 1.89 | 0.058 |

Output is from a generalized linear mixed model for Response (binary), a cumulative link mixed model for approach distance (distances were placed into ordinal categories), and a zero-inflated glmm for number of songs (Poisson; the coefficients are for the conditional model output). The reference category for stimulus rate is the low rate, so positive estimates indicate more responses and a greater number of songs, and a greater distance from the speaker, in response to the high rate stimulus. Trial number was entered as a numerical variable. $N$ = 51 individuals.

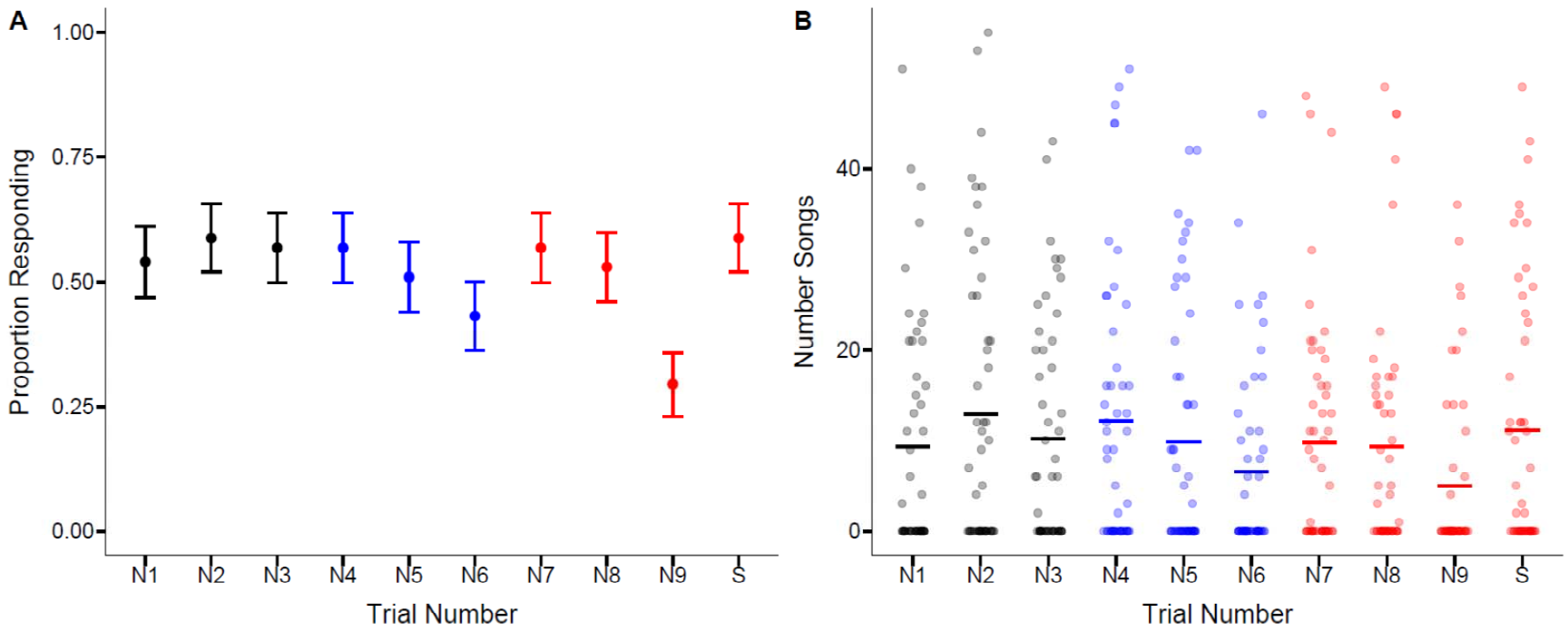


Figure 2. Response across trials. (A) The proportion (±SE) of individuals (*N* = 51 for all trials except *N* = 50 for N1; see Methods) that responded by either singing or approaching the playback speaker in each of the nine neighbour playback trials (labelled N1-N9) and on the stranger playback trial. Colours correspond to the day that the playback was performed (black = day 1, blue = day 2, red = day 3). (B) The number of songs produced by males during the playback trial. Dots represent an individual's response to that trial (points have been jittered along the x-axis and rendered partially transparent for ease of interpretation). Horizontal lines represent mean values.

*Stimulus specificity of response to neighbour playback*

Although twice as many individuals responded to the stranger playback compared to the immediately preceding final neighbour playback, this difference was not significant (Figure 2A; Table 2). However, there was a significant difference in both the approach distance and number of songs between the final neighbour playback and the stranger playback: individuals approached closer and sang more songs in response to the stranger playback (Figure 2B; Table 2). Stimulus rate did not affect the responses when only these two trials were considered (Table 2).

Table 2. Tests of stimulus specificity of response to neighbour playback.

| Variable | Factor | Estimate (SE) | z | P |
|---|---|---|---|---|
| Response | Intercept | −2.54 (2.51) | | |
| | Trial | 3.59 (3.33) | 1.08 | 0.28 |
| | Rate | −0.12 (1.42) | −0.08 | 0.94 |
| Approach distance | Trial | −2.09 (0.56) | −3.73 | <0.001 |
| | Rate | 0.13 (0.79) | 0.17 | 0.87 |
| Number of songs | Intercept | −0.04 (0.74) | | |
| | Trial | 0.30 (0.08) | 3.56 | <0.001 |
| | Rate | 0.04 (0.92) | 0.04 | 0.97 |

Output is from models as in Table 1, but now only including the response to the final neighbour playback and the stranger playback. Trial was therefore coded as a binary variable. The reference category for trial is the final neighbour playback, so positive estimates would indicate more responses and a greater number of songs, and a negative estimate indicates closer approach, during the stranger playback. The reference category for stimulus rate is the low rate. $N = 51$ individuals.

*Dear enemy behaviour and reproductive success*

Based on our standard criterion of no response to the final neighbour playback followed by a response to the stranger playback, 15 of 51 individuals exhibited dear enemy behaviour towards their simulated neighbour after the nine playbacks (15 of 30 when non-responders were removed i.e. standard criterion without nonresponders, see Methods; these results refer only to the first set of playbacks, and do not include the second set of playbacks that were performed in 2017). Whether an individual met the standard criterion did not depend on the date of testing, the date relative to the day on which the first egg was laid, the year or the stimulus rate (Table S2). These variables also did not affect the slope of the response or number of songs across trials (Table S2). There was also no effect of age (first year versus older) on meeting the standard criterion for dear enemy behaviour (Estimate of effect being older than first year ± SE = −1.29 ± 0.88, z = −1.46, P = 0.15; N = 33 individuals of known age). However, there was an effect of age on the slope of number of songs across trials, with older birds having a more positive slope than first year birds (Estimate = 2.96 ± 0.83, $t$ = 3.57, P = 0.003; N = 17 individuals of known age), and also on the slope of responses across trials (Estimate = 0.07 ± 0.03, $t$ = 2.33, P = 0.03; N = 17). There was a trend for significant repeatability in whether an individual exhibited dear enemy behaviour, with a moderate repeatability coefficient ($N$ = 19 individuals tested twice; R = 0.44, P = 0.074). Five of six individuals that showed dear enemy behaviour in the first set of playbacks also did so in the second, and six of 13 individuals that did not show dear enemy behaviour in the first set of playbacks also did not show it in the second set. There were insufficient numbers of individuals tested twice and responding to the stranger playback to allow for estimating the repeatability of the slope for response or number of songs across trials.

Sixteen offspring were identified as extra-pair young and were excluded from the analyses of number of offspring fledged and offspring mass. There was no difference between individuals that did or did not show dear enemy behaviour in the proportion of fledglings that were extrapair (Estimate = 0.20 ± 0.71, P = 0.78, *N* = 29 nests). There was no effect of whether an individual exhibited dear enemy behaviour (standard criterion) on the number of eggs laid in its nest (Figure 3A; Table S3), or on the number of within-pair offspring that fledged (Figure 3B; Table S3). However, the total mass of all within-pair offspring weighed at day 15 was significantly lower for individuals that showed dear enemy behaviour towards the simulated neighbour (Figure 3C; Estimate = −38.2 ± 15.4, P = 0.01, *N* = 47 nests). For the limited set of individuals that fledged young, there was no effect of whether it had met the standard criterion for dear enemy behaviour on the average mass of its within-pair offspring (Figure 3D; Table S3) or on the individual mass of within-pair offspring (Table S3). Qualitatively similar results were obtained when the standard criterion without nonresponders was used as the response variable (Table S3). When the slope of responses or number of songs across trial was used as the response variable, there were no significant effects on any of the reproductive success measures (Table S3).

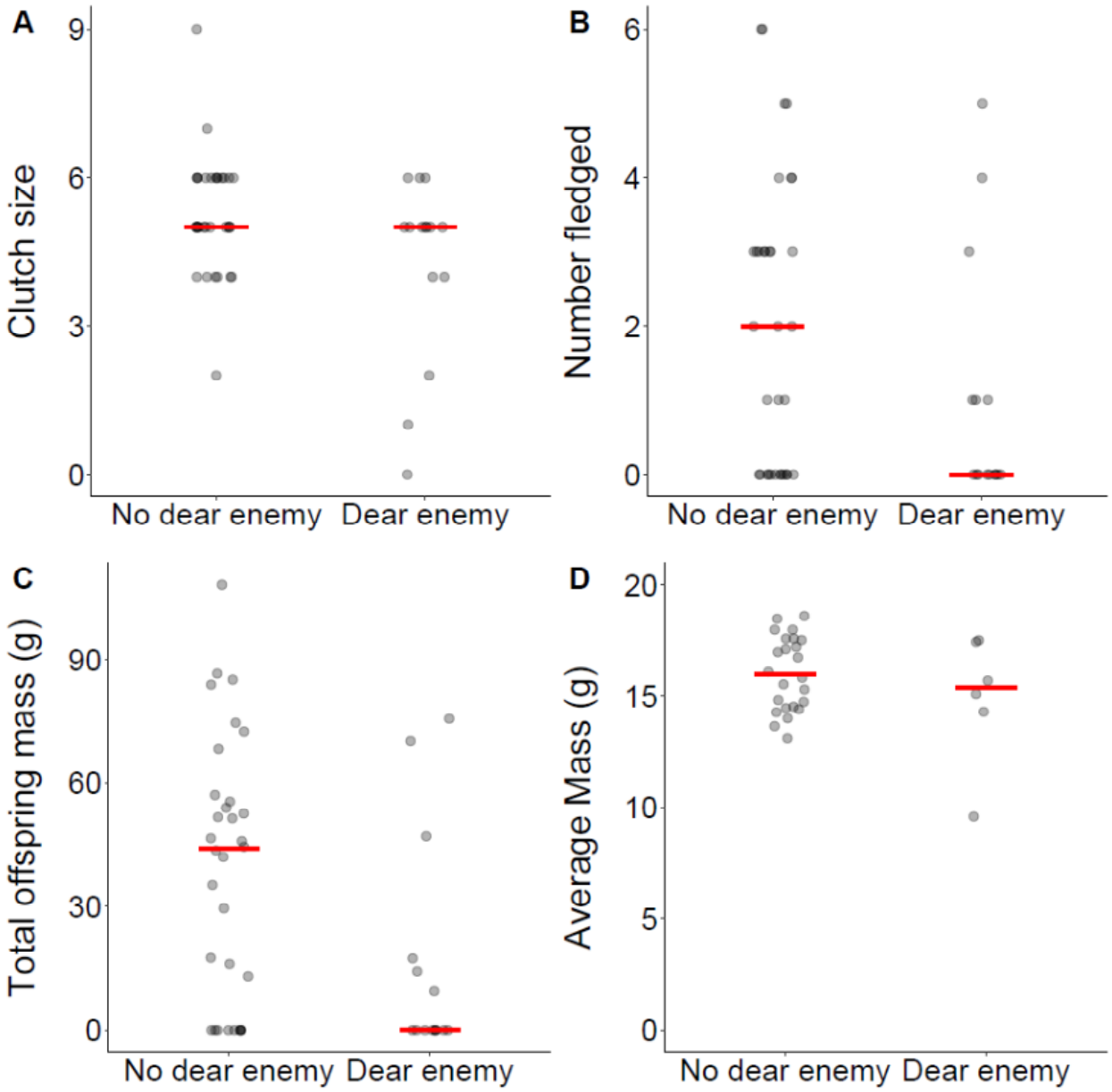


Figure 3: Relationship between whether the individual met the standard criterion (see Methods) for exhibiting the dear enemy effect ("Dear enemy") or not ("No dear enemy") and different measures of reproductive success. Each dot represents the value for an individual nest (points jittered along the x-axis and rendered partially transparent); red horizontal line represents the median. A. Clutch size, B. Number of within-pair offspring successfully fledged, C. Total mass of within-pair offspring at day 15 post hatch, D. Average mass of within-pair offspring at day 15 post hatch (not including nests for which

no offspring survived to day 15). *N* = 48 individuals in each graph, except for C where *N* = 47 and D where *N* = 30.

## DISCUSSION

Male great tits in our study populations exhibited behaviour consistent with the dear enemy effect in response to playbacks of a simulated territorial neighbour. There was an overall decline in song response and approach towards the playback across repeated presentations, consistent with reduced aggression towards an increasingly familiar individual. This reduced aggression was stimulus-specific because in the stranger playback, simulating a different individual, subjects reverted to a strong response to the playback. Our experiments show that individual recognition can arise and be expressed in only three days. However, not all birds exhibited dear enemy behaviour towards the simulated neighbour at the end of the playback trials. This individual variation in response to neighbours had an unexpected association with reproductive output: there was no reproductive benefit for individuals that exhibited dear enemy behaviour in any case. Furthermore, there may in fact have been a reproductive cost paid by these individuals: birds that showed the dear enemy effect had significantly reduced total fledgling biomass and the number of offspring fledged was also lower, if not significantly so. Below we discuss these results in the context of territorial aggression and the evolution of neighbour recognition.

### *Learning mechanisms and individual variation*

The dear enemy effect is widespread in animals (Dimarco et al., 2010; Fox & Baird, 1992; Jaeger, 1981), although most studies demonstrate the effect by comparing responses to one-time playbacks of songs of actual neighbours versus strangers (Lambrechts & Dhondt, 1995). Playback experiments simulating a

new neighbour and monitoring in detail the change in subjects' responses across trials are a potentially powerful, but rarely utilized, tool to study the mechanism of neighbour recognition. There is a cognitive component to the dear enemy effect in great tits, and while we recognize that non-cognitive mechanisms may also contribute to variation in whether an individual responds differently to neighbours and strangers (see below), it is worthwhile to discuss the cognitive mechanism of dear enemy recognition in light of our experimental design and results.

Habituation learning has been demonstrated to mediate dear enemy recognition in many species. For instance, in an extensive series of experiments, playbacks were used to show that the dear enemy effect in bullfrogs (*Rana catesbeiana*) is mediated by habituation learning of specific characteristics of neighbour vocalizations (Bee, 2001, 2003b; Bee & Gerhardt, 2001b, 2001c, 2001a, 2002), and some studies of songbirds have also supported a role of habituation in neighbour recognition (Brooks & Falls, 1975a; Petrinovich & Peeke, 1973). The pattern of response change across trials in our study matches many of the characteristics of habituation (Rankin et al., 2009; Thompson & Spencer, 1966) and suggests that habituation learning is the mechanism behind the initial stages of dear enemy recognition in great tits. First, the response to repeated playback of a simulated neighbour decreased across trials (Figure 2). Second, there was evidence for spontaneous recovery of the response across days, in which the response on the first trial of the day was generally greater than that of the last trial on the previous day (Figure 2). A key characteristic of habituation is that over a series of recoveries and decrements, habituation is potentiated, with the response decrement becoming ever more pronounced (Rankin et al., 2009; Thompson & Spencer, 1966). Although we had only nine testing sessions across three days, comparisons of behaviour on day 2 and day 3 show that the response decrease was much greater on day 3, consistent with a potentiation of habituation effect. Third, the dual-process theory of habituation argues that in addition to the decrement in response caused by habituation, an independent process of

545 sensitization results in an initial, but transient, increase in aggressive response (Groves & Thompson,
546 1970; Petrinovich & Patterson, 1982). Consistent with sensitization, responses tended to be weaker on
547 the first playback trial compared to the subsequent few trials.
548
549 One possible exception for a role of habituation in explaining our results is that habituation is
550 hypothesized to develop faster at higher stimulus repetition rates (Groves & Thompson, 1970;
551 Thompson et al., 1973), but although males in the high rate treatment sang more overall, there was no
552 effect on approach or whether individuals exhibited dear enemy recognition at the end of the playback
553 trials. However, while we varied the rate of repetition of songs within a trial, the interval between
554 successive trials was approximately the same for all birds. Finally, while we argue that our findings are in
555 many ways consistent with a habituation mechanism, as are those in other species (Bee & Gerhardt,
556 2001a; Brooks & Falls, 1975b), this may only apply to the initial stages of neighbour recognition. That
557 neighbour songs can be remembered and discriminated across years (McGregor & Avery, 1986), and
558 that individuals are capable of learning to recognize individual voice characteristics and discriminate
559 among even previously unheard songs (Weary & Krebs, 1992), along with the finding that individuals
560 with strong social associations formed during winter flocking tend to end up on neighbouring breeding
561 territories (Firth & Sheldon, 2016), suggests that other cognitive abilities such as associative learning are
562 also important in long-term neighbour recognition in great tits (see also (Godard, 1991; Richards, 1979)).
563
564 Whatever the mechanism, our results indicate that at an individual level, only some of the males
565 reduced their aggression towards the simulated neighbour and discriminated between it and a
566 simulated stranger. If this variation was due to differences in recognition learning ability, then additional
567 trials per day, or less time in between trials perhaps could have facilitated learning for a greater
568 proportion of the population (Petrinovich & Patterson, 1982), although this approach would have

constrained our ability to test a sufficient number of individuals in the field. Moreover, the simultaneous presentation of a visual stimulus (e.g., taxidermy mount of great tit male), playbacks with interactive elements (Langemann et al., 2000), or playbacks presenting the full song repertoire of a simulated neighbour rather than a single song type (Krebs, 1976; Weary & Krebs, 1992), may provide more intense stimulation, thus facilitating more pronounced learning effects as well as exposing additional individual variation in responses (Araya-Ajoy et al., 2016; Ritschard et al., 2012).

However, there are also many plausible non-cognitive explanations for variation in whether males showed the dear enemy effect. Differences in overall aggressiveness are likely to be particularly important; even if individuals learned to recognize their neighbour, they may maintain an overall more aggressive response. Such variation in aggressiveness is known to affect response to playback in great tits (Araya-Ajoy & Dingemanse, 2014). Males also likely differ in previous experience and this could have affected their response to playback simulating a new neighbour in several ways. For instance, although we assigned playback stimuli randomly and only exposed individuals to songs from different populations, some males may have happened to have been given a playback stimulus more similar acoustically to songs of a previous neighbour, affecting whether they reduced their response over time. Males likely varied in the extent to which they had already interacted with different territorial neighbours in the past, and great tits with more previous exposure to songs of many neighbours are slower to learn to recognize new neighbours (McGregor & Avery, 1986). We do not have data on the songs of individuals' previous neighbours to test these possibilities directly, although we note that there was no difference in dear enemy effect expression between first-year birds and older birds in our study. Other context-based plasticity (e.g. density, progress of the breeding season) can explain variation in the dear enemy effect (Hyman, 2005; Pratt & McLain, 2006; Yoon et al., 2012), and has been reported in one study of great tits (Jin et al., 2020). However, we found no effect of year, date or date relative to the

start of hatching on whether individuals expressed the dear enemy effect. Furthermore, we found some evidence of individual repeatability in the expression of the dear enemy effect despite our small sample size, which suggests that despite context-dependent variation there is some consistent among-individual variation in neighbour recognition behaviour in this species.

*Dear enemy recognition and fitness*

We hypothesized that if the dear enemy effect is adaptive because it reduces time spent in unnecessary aggressive interactions (Getty, 1987; Temeles, 1994), then individuals that reduce the aggression directed towards their neighbours more quickly would have higher reproductive success. This hypothesis was not supported: we found either no effect at all, or an effect in the opposite direction than predicted. Specifically, individuals that expressed the dear enemy effect had lower total chick biomass, but there was no difference in clutch size, average biomass, or number of offspring that fledged. Mass of the chicks is related to parental foraging effort (Schwagmeyer & Mock, 2008) and is an important component of fitness because heavier chicks are more likely to survive after fledging (Tinbergen & Boerlijst, 1990).

It is not clear why individuals that did not show dear enemy recognition had higher total offspring biomass. It may be that these individuals were of overall higher quality, and could therefore afford to put more effort into territorial defence while still foraging effectively, which may have resulted in larger and more exclusive territories for subsequent offspring provisioning. Even if there were no quality differences among individuals, if there is variation in territory quality, then individuals on higher quality territories potentially had more to lose (Riechert, 1984), and therefore maintained an aggressive response for longer to ensure they could hold on to the territory. In this case, the expression of the dear

enemy effect is only indirectly related with offspring quality, because the relationship is mediated by joint correlations between both variables and territory quality. One caveat is that our fitness measures only cover a single breeding season. Because of trade-offs between current and future reproduction (Trivers, 1972), fitness effects of dear enemy expression may be more apparent over the lifetime reproductive output and this should be addressed in future studies.

There may be costs associated with rapid reductions in aggressive responses that counteract any benefits of reduced aggression towards neighbours. In particular, individuals that are too quick to reduce their aggression may be ineffective at driving off prospecting individuals, perhaps resulting in compromised territory size and quality. In our experiment, individuals that displayed the dear enemy effect did so after exposure to only one song type and only three days of singing by the simulated newly-arrived neighbour, potentially taking the risk of reducing aggression without fully assessing the individual via its entire song repertoire and over a longer time period (Getty, 1989). It may therefore be beneficial to maintain heightened aggression until territorial boundaries are well established. Individuals that quickly reduce aggression may also be more susceptible to their mates engaging in extrapair copulations with neighbours; we found no such relationship, although a greater sample size may be necessary to detect any effect. Alternatively, the association between poor offspring quality and rapid expression of the dear enemy effect may reflect high costs of the cognitive abilities associated with neighbour recognition learning (Morand-Ferron et al., 2016). Birds that expressed the dear enemy effect very often had failed nests (Figure 3B), although the difference in number of offspring fledged between birds that did and did not show dear enemy behaviour was not significant. Nest failure is generally high in our populations (O'Shea et al., 2018) and although we accounted for the zero values statistically, the results suggest that some of the difference in our measures of reproductive success may be due to a greater susceptibility to nest abandonment in birds showing the dear enemy effect (Figure 3B). Similar

differences in the likelihood of abandonment have been demonstrated in the context of variation in innovative problem solving behaviour in great tits (Cole et al., 2012).

The lack of a reproductive benefit (and potential cost, as indicated by the results for total offspring biomass) for birds showing the dear enemy effect in this study should not necessarily be taken as evidence that selection does not favour the ability to discriminate neighbours from strangers. Dear enemy recognition is well-established in great tits (Falls et al., 1982; Krebs, 1971; McGregor & Avery, 1986), and individuals with a higher proportion of familiar neighbours (i.e., those they had already been a neighbour of in a previous year) had higher reproductive success (Grabowska-Zhang, Wilkin, et al., 2012). Although many birds did not express dear enemy behaviour within the timeframe of the experiment, we expect that with additional exposures, most individuals would have expressed a robust dear enemy effect, as they do towards playbacks of long-established natural neighbours (Krebs, 1971). If individuals that were already showing the dear enemy effect within the short timeframe of the experiment did suffer reduced reproductive success, as some of our results suggest, then we argue that selection may be acting on the rate at which individuals reduce the response to neighbours, favouring males that more gradually reduce their aggressive response. This may or may not involve variation in recognition learning capabilities: as described above, some individuals may have remained aggressive despite having learned to recognize their neighbour.

*Conclusions*

The dear enemy effect is the outcome of dynamic interactions between neighbouring individuals. While progress has been made in understanding between-species differences in responses to territorial neighbours (Bee et al., 2016; Christensen & Radford, 2018; Tumulty & Bee, 2020), the understanding of

the causes and consequences of within-species variation in dear enemy behaviour has lagged behind. Our finding of either no relationship, or in some cases a negative relationship, between dear enemy recognition and reproductive success, along with the moderate repeatability of the expression of the dear enemy effect, indicates that there could be fitness costs to reducing aggression towards neighbours too quickly. Although challenging, additional work along the lines of our study that monitor the development of dear enemy relationships between neighbours and associate this with individual phenotypic characteristics and reproductive success will be essential for understanding a widespread and consequential form of social behaviour in animals.

*Author contributions*. M.S.R. and J.L.Q. conceived of the project, M.S.R., J.M.S.C. and J.O.N. performed the playbacks, M.S.R., G.L.D, I.G.K., J.M.C. and J.O.N. collected breeding data, E.D., C.S. and K.v.O. did the DNA extractions and analyses of paternity, M.S.R. analyzed the data, and M.S.R. wrote the manuscript with input from all authors. Note that middle authors on this manuscript are listed in alphabetical order by surname.

*Acknowledgements*. Karen Cogan assisted with equipment sourcing and purchasing. Maike Foraita assisted with playback trials. Will O'Shea, Paul Whitelaw, Sam Bayley, Iván de la Hera, Amy Cooke, and Jenny Coomes helped with nestbox monitoring during the breeding season. Funding for M.S.R., G.L.D, I.G.K., J.O.N., C.S., J.M.S.C. came from the European Research Council under the European Union's Horizon 2020 Programme (FP7/2007-2013)/ERC Consolidator Grant "EVOLECOCOG" Project No. 617509, awarded to JLQ, and by a Science Foundation Ireland ERC Support Grant 14/ERC/B3118 to JLQ

REFERENCES


Amy, M., Sprau, P., de Goede, P., & Naguib, M. (2010). Effects of personality on territory defence in communication networks: a playback experiment with radio-tagged great tits. *Proceedings of the Royal Society B*, *277*(1700), 3685–3692. https://doi.org/10.1098/rspb.2010.0598

Araya-Ajoy, Y. G., & Dingemanse, N. J. (2014). Characterizing behavioural "characters": an evolutionary framework. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1776), 20132645. https://doi.org/10.1098/rspb.2013.2645

Araya-Ajoy, Y. G., Kuhn, S., Mathot, K. J., Mouchet, A., Mutzel, A., Nicolaus, M., Wijmenga, J. J., Kempenaers, B., & Dingemanse, N. J. (2016). Sources of (co)variation in alternative siring routes available to male great tits (*Parus major*). *Evolution*, *70*, 2308–2321. https://doi.org/10.1111/EVO.13024

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal Of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Battiston, M. M., Wilson, D. R., Graham, B. A., Kovach, K. A., & Mennill, D. J. (2015). Rufous-and-white wrens *Thryophilus rufalbus* do not exhibit a dear enemy effects towards conspecific or heterospecific competitors. *Current Zoology*, *61*(1), 23–33. https://doi.org/10.1093/czoolo/61.1.23

Bee, M. A. (2001). Habituation and sensitization in bullfrogs (*Rana catesbiana*): testing the dual-process theory of habituation. *Journal of Comparative Psychology*, *115*(3), 307–316.

Bee, M. A. (2003a). A test of the "dear enemy effect" in the strawberry dart-poison frog (*Dendrobates pumilio*). *Behavioral Ecology and Sociobiology*, *54*(6), 601–610. https://doi.org/10.1007/s00265-003-0657-5

Bee, M. A. (2003b). Experience-based plasticity of acoustically evoked aggression in a territorial frog. *Journal of Comparative Physiology A*, *189*(6), 485–496. https://doi.org/10.1007/s00359-003-0420-4

Bee, M. A., & Gerhardt, H. C. (2001a). Habituation as a mechanism of reduced aggression between neighboring territorial male bullfrogs (*Rana catesbeiana*). *Journal of Comparative Psychology*, *115*(1), 68–82. https://doi.org/10.1037/0735-7036.115.1.68

Bee, M. A., & Gerhardt, H. C. (2001b). Neighbor-stranger discrimination by territorial male bullfrogs (*Rana catesbiana*): II. Perceptual basis. *Animal Behaviour*, *62*, 1141–1150. https://doi.org/10.1006/anbe.2001.1852

Bee, M. A., & Gerhardt, H. C. (2001c). Neighbour-stranger discrimination by territorial male bullfrogs (*Rana catesbeiana*): I. Acoustic basis. *Animal Behaviour*, *62*, 1129–1140. https://doi.org/10.1006/anbe.2001.1851

Bee, M. A., & Gerhardt, H. C. (2002). Individual voice recognition in a territorial frog (*Rana catesbiana*). *Proceedings of the Royal Society B: Biological Sciences*, *269*, 1443–1448. https://doi.org/10.1098/rspb.2002.2041

Bee, M. A., Reichert, M. S., & Tumulty, J. P. (2016). Assessment and recognition of rivals in anuran contests. *Advances in the Study of Behavior*, *48*, 161–249. https://doi.org/10.1016/bs.asb.2016.01.001

Beletsky, L. D., & Orians, G. H. (1989). Familiar neighbors enhance breeding success in birds. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(20), 7933–7936. https://doi.org/10.1073/pnas.86.20.7933

Blurton Jones, N. G. (1968). Observations and experiments on causation of threat displays of the great tit (*Parus major*). *Animal Behaviour Monographs*, *1*, 73–158. https://doi.org/10.1016/S0066-1856(68)80002-0

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R Journal*, *9*(2), 378–400.

https://doi.org/10.32614/rj-2017-066

Brooks, R. J., & Falls, J. B. (1975). Individual recognition by song in white-throated sparrows. I. Discrimination of songs of neighbors and strangers. *Canadian Journal of Zoology*, *53*(7), 879–888.

Bruintjes, R., Lynton-Jenkins, J., Jones, J. W., & Radford, A. N. (2016). Out-group threat promotes within-group affiliation in a cooperative fish. *American Naturalist*, *187*(2), 274–282. https://doi.org/10.1086/684411

Brunton, D. H., Evans, B., Cope, T., & Ji, W. (2008). A test of the dear enemy hypothesis in female New Zealand bellbirds (*Anthornis melanura*): female neighbors as threats. *Behavioral Ecology*, *19*(4), 791–798. https://doi.org/10.1093/beheco/arn027

Christensen, C., & Radford, A. N. (2018). Dear enemies or nasty neighbors? Causes and consequences of variation in the responses of group-living species to territorial intrusions. *Behavioral Ecology*, *29*(5), 1004–1013. https://doi.org/10.1093/beheco/ary010

Christensen, R. H. B. (2019). *ordinal - regression models for ordinal data. R package version 2019.12-10*. https://cran.r-project.org/package=ordinal

Cole, E. F., Morand-Ferron, J., Hinks, A. E., & Quinn, J. L. (2012). Cognitive ability influences reproductive life history variation in the wild. *Current Biology*, *22*(19), 1808–1812. https://doi.org/10.1016/j.cub.2012.07.051

Courvoisier, H., Camacho-Schlenker, S., & Aubin, T. (2014). When neighbours are not “dear enemies”: A study in the winter wren, *Troglodytes troglodytes*. *Animal Behaviour*, *90*, 229–235. https://doi.org/10.1016/j.anbehav.2014.02.001

Dimarco, R. D., Farji-Brener, A. G., & Premoli, A. C. (2010). Dear enemy phenomenon in the leaf-cutting ant *Acromyrmex lobicornis*: Behavioral and genetic evidence. *Behavioral Ecology*, *21*(2), 304–310. https://doi.org/10.1093/beheco/arp190

Doutrelant, C., Leitao, A., Otter, K., & Lambrechts, M. M. (2000). Effect of blue tit song syntax on great tit

territorial resposiveness - An experimental test of the character shift hypothesis. *Behavioral Ecology and Sociobiology*, *48*(2), 119–124. https://doi.org/10.1007/s002650000220

East, M. L., & Perrins, C. M. (1988). The effect of nestboxes on breeding populations of birds in broadleaved temperate woodlands. *Ibis*, *130*(3), 393–401. https://doi.org/10.1111/j.1474-919X.1988.tb08814.x

Eliassen, S., & Jørgensen, C. (2014). Extra-pair mating and evolution of cooperative neighbourhoods. *PLoS ONE*, *9*(7), e99878. https://doi.org/10.1371/journal.pone.0099878

Falls, J. B., & Brooks, R. J. (1975). Individual recognition by song in white-throated sparrows. II. Effects of location. *Canadian Journal of Zoology*, *53*(10), 1412–1420. https://doi.org/10.1139/z75-210

Falls, J. B., Krebs, J. R., & McGregor, P. K. (1982). Song matching in the great tit (*Parus major*): The effect of similarity and familiarity. *Animal Behaviour*, *30*(4), 997–1009. https://doi.org/10.1016/S0003-3472(82)80188-7

Firth, J. A., & Sheldon, B. C. (2016). Social carry-over effects underpin trans-seasonally linked structure in a wild bird population. *Ecology Letters*, *19*(11), 1324–1332. https://doi.org/10.1111/ele.12669

Fox, S. F., & Baird, T. A. (1992). The dear enemy phenomenon in the collared lizard, *Crotaphytus collaris*, with a cautionary note on experimental methodology. *Animal Behaviour*, *44*(4), 780–782. https://doi.org/10.1016/S0003-3472(05)80306-9

Getty, T. (1987). Dear enemies and the prisoner's dilemma: Why should territorial neighbors form defensive coalitions? *American Zoologist*, *27*, 327–336.

Getty, T. (1989). Are dear enemies in a war of attrition? *Animal Behaviour*, *37*(2), 337–339. https://doi.org/10.1016/0003-3472(89)90125-5

Giraldeau, L. A., & Ydenberg, R. C. (1987). The center-edge effect: The result of a war of attrition between territorial contestants? *The Auk*, *104*(3), 535–538. https://doi.org/10.1111/j.1365-263X.1994.tb00128.x

Godard, R. (1991). Long-term memory of individual neighbours in a migratory songbird. *Nature*, *354*, 293–295.

Gompertz, T. (1961). The vocabulary of the Great Tit. *British Birds*, *54*, 369–399.

Gosler, A. G. (1993). *The great tit*. Hamlyn.

Grabowska-Zhang, A. M., Sheldon, B. C., & Hinde, C. A. (2012). Long-term familiarity promotes joining in neighbour nest defence. *Biology Letters*, *8*(4), 544–546. https://doi.org/10.1098/rsbl.2012.0183

Grabowska-Zhang, A. M., Wilkin, T. A., & Sheldon, B. C. (2012). Effects of neighbor familiarity on reproductive success in the great tit (*Parus major*). *Behavioral Ecology*, *23*(2), 322–333. https://doi.org/10.1093/beheco/arr189

Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychol Rev*, *77*(5), 419–450.

Henningsen, A. (2019). *censReg: Censored regression (Tobit) models. R package version 0.5-30.* https://cran.r-project.org/package=censReg

Hill, C. E., Akçay, Ç., Campbell, S. E., & Beecher, M. D. (2011). Extrapair paternity, song, and genetic quality in song sparrows. *Behavioral Ecology*, *22*(1), 73–81. https://doi.org/10.1093/beheco/arq171

Hinde, R. A. (1956). The biological significance of the territories of birds. *Ibis*, *98*(3), 340–369. https://doi.org/10.1111/j.1474-919X.1956.tb01419.x

Hyman, J. (2005). Seasonal variation in response to neighbors and strangers by a territorial songbird. *Ethology*, *111*(10), 951–961. https://doi.org/10.1111/j.1439-0310.2005.01104.x

Jacobs, C. G. C., van Overveld, T., Careau, V., Matthysen, E., Adriaensen, F., & Slabbekoorn, H. (2014). Personality-dependent response to field playback in great tits: Slow explorers can be strong responders. *Animal Behaviour*, *90*, 65–71. https://doi.org/10.1016/j.anbehav.2014.01.016

Jaeger, R. G. (1981). Dear enemy recognition and the costs of aggression between salamanders. *The*

*American Naturalist*, *117*(6), 962–974. http://www.jstor.org/stable/2460574

Jin, L., Liang, J., Fan, Q., Yu, J., Sun, K., & Wang, H. (2020). Male great tits (*Parus major*) adjust dear enemy effect expression in different breeding stages. *Journal of Ornithology*. https://doi.org/10.1007/s10336-020-01815-3

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, *16*(5), 1099–1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x

Kawano, K. M. (2003). Isolation of polymorphic microsatellite markers in the great tit (*Parus major minor*). *Molecular Ecology Notes*, *3*(2), 314–315. https://doi.org/10.1046/j.1471-8286.2003.00431.x

Krebs, J. R. (1971). Territory and breeding density in the great tit , *Parus major* L . *Ecology*, *52*(1), 2–22.

Krebs, J. R. (1976). Habituation and song repertoires in the great tit. *Behavioral Ecology and Sociobiology*, *1*(2), 215–227. https://doi.org/10.1007/BF00299198

Krebs, J. R. (1977). Song and territory in the great tit Parus major. In B. Stonehouse & C. Perrins (Eds.), *Evolutionary Ecology* (pp. 47–62). Palgrave.

Krebs, J. R. (1982). Territorial defence in the great tit (Parus major): Do residents always win? *Behavioral Ecology and Sociobiology*, *11*(3), 185–194. https://doi.org/doi.org/10.1007/BF00300061

Lambrechts, M. M., & Dhondt, A. A. (1995). Individual voice discrimination in birds. *Current Ornithology*, *12*, 115–139. https://doi.org/10.1007/978-1-4615-1835-8_4

Langemann, U., Tavares, J. P., Peake, T. M., & McGregor, P. K. (2000). Response of great tits to escalating patterns of playback. *Behaviour*, *137*, 451–471. https://doi.org/10.1163/156853900502178

Martin, T. E. (1987). Food as a limit on breeding birds: a life-history perspective. *Annual Review of Ecology and Systematics*, *18*, 453–487.

McGregor, P. K. (1993). Signalling in territorial systems: a context for individual identification, ranging

and eavesdropping. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *340*, 237–244.

McGregor, P. K., & Avery, M. I. (1986). The unsung songs of great tits (*Parus major*): learning neighbours' songs for discrimination. *Behavioral Ecology and Sociobiology*, *18*(4), 311–316. https://doi.org/10.1007/BF00300009

McGregor, P. K., & Krebs, J. R. (1982a). Mating and song types in the great tit. *Nature*, *297*, 60–61. https://doi.org/10.1038/297060a0

McGregor, P. K., & Krebs, J. R. (1982b). Song types in a population of great tits (*Parus major*): Their distribution, abundance and acquisition by individuals. *Behaviour*, *79*(2/4), 126–152.

Morand-Ferron, J., Cole, E. F., & Quinn, J. L. (2016). Studying the evolutionary ecology of cognition in the wild: a review of practical and conceptual challenges. *Biological Reviews*, *91*, 367–389. https://doi.org/10.1111/brv.12174

Muller, C. A., & Manser, M. B. (2007). "Nasty neighbours" rather than "dear enemies" in a social carnivore. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1612), 959–965.

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews*, *85*(4), 935–956. https://doi.org/10.1111/j.1469-185X.2010.00141.x

Nelson, D. A., & Soha, J. A. (2004). Male and female white-crowned sparrows respond differently to geographic variation in song. *Behaviour*, *141*, 53–69. https://doi.org/10.1163/156853904772746600

O'Shea, W., O'Halloran, J., & Quinn, J. L. (2018). Breeding phenology, provisioning behaviour, and unusual patterns of life history variation across an anthropogenic heterogeneous landscape. *Oecologia*, *188*(4), 953–964. https://doi.org/10.1007/s00442-018-4155-x

Patrick, S. C., Chapman, J. R., Dugdale, H. L., Quinn, J. L., & Sheldon, B. C. (2012). Promiscuity, paternity

and personality in the great tit. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1734), 1724–1730. https://doi.org/10.1098/rspb.2011.1820

Petrinovich, L., & Patterson, T. L. (1982). Field studies of habituation: V. Evidence for a two-factor, dual-process system. *Journal of Comparative and Physiological Psychology*, *96*(2), 284–296. https://doi.org/10.1037/h0077868

Petrinovich, L., & Peeke, H. V. S. (1973). Habituation to territorial song in the white-crowned sparrow (*Zonotrichia leucophrys*). *Behavioral Biology*, *8*(6), 743–748. https://doi.org/10.1016/S0091-6773(73)80116-6

Pratt, A. E., & McLain, D. K. (2006). How dear is my enemy: Intruder-resident and resident-resident encounters in male sand fiddler crabs (Uca pugilator). *Behaviour*, *143*, 597–617. https://doi.org/10.1163/156853906776759501

R Development Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., Coppola, G., Geyer, M. A., Glanzman, D. L., Marsland, S., McSweeney, F. K., Wilson, D. A., Wu, C. F., & Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, *92*(2), 135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Richards, D. G. (1979). Recognition of neighbors by associative learning in rufous-sided towhees. *The Auk*, *96*(4), 688–693. https://doi.org/10.1093/auk/96.4.688

Riechert, S. E. (1984). Games spiders play. III: Cues underlying context-associated changes in agonistic behaviour. *Animal Behaviour*, *32*(1), 1–15. https://doi.org/10.1016/S0003-3472(84)80318-8

Ritschard, M., van Oers, K., Naguib, M., & Brumm, H. (2012). Song amplitude of rival males modulates the territorial behaviour of great tits during the fertile period of their mates. *Ethology*, *118*(2), 197–

202. https://doi.org/10.1111/j.1439-0310.2011.01999.x

Saladin, V., Bonfils, D., Binz, T., & Richner, H. (2003). Isolation and characterization of 16 microsatellite loci in the Great Tit *Parus major*. *Molecular Ecology Notes*, *3*(4), 520–522. https://doi.org/10.1046/j.1471-8286.2003.00498.x

Schlicht, L., Valcu, M., & Kempenaers, B. (2015). Spatial patterns of extra-pair paternity: Beyond paternity gains and losses. *Journal of Animal Ecology*, *84*(2), 518–531. https://doi.org/10.1111/1365-2656.12293

Schwagmeyer, P. L., & Mock, D. W. (2008). Parental provisioning and offspring fitness: size matters. *Animal Behaviour*, *75*(1), 291–298. https://doi.org/10.1016/j.anbehav.2007.05.023

Siracusa, E. R., Boutin, S., Dantzer, B., Lane, J. E., Coltman, D. W., & McAdam, A. G. (2021). Familiar neighbors, but not relatives, enhance fitness in a territorial mammal. *Current Biology*, *31*, 438–445. https://doi.org/10.1016/j.cub.2020.10.072

Snijders, L., van Oers, K., & Naguib, M. (2016). Sex-specific responses to territorial intrusions in a communication network: evidence from radio-tagged great tits. *Ecology and Evolution*, *November 2016*, 918–927. https://doi.org/10.1002/ece3.2686

Stoddard, P. K. (1996). Vocal recognition of neighbors by territorial passerines. In D. E. Kroodsma & E. H. Miller (Eds.), *Ecology and Evolution of Acoustic Communication in Birds* (pp. 356–34). Cornell University Press.

Stoddard, P. K., Beecher, M. D., Horning, C. L., & Campbell, S. E. (1991). Recognition of individual neighbors by song in the song sparrow, a species with song repertoires. *Behavioral Ecology and Sociobiology*, *29*(3), 211–215. https://doi.org/10.1007/BF00166403

Temeles, E. (1990). Northern harriers on feeding territories respond more aggressively to neighbors than to floaters. *Behavioral Ecology and Sociobiology*, *26*, 57–63.

Temeles, E. (1994). The role of neighbors in territorial systems: when are they “dear enemies”? *Animal*

*Behaviour*, *47*(2), 339–350. https://doi.org/10.1006/anbe.1994.1047

Thompson, R. F., Groves, P. M., Teyler, T. J., & Roemer, R. A. (1973). A dual-process theory of habituation: theory and behavior. In H. V. S. Peeke & M. J. Herz (Eds.), *Habituation I: Behavioral Studies* (pp. 239–271). Academic Press.

Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, *73*(1), 16–43.

Tinbergen, J., & Boerlijst, M. (1990). Nestling weight and survival in individual great tits (*Parus major*). *The Journal of Animal Ecology*, *59*(3), 1113–1127. https://doi.org/10.2307/5035

Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual Selection and the Descent of Man, 1871-1971*. Heinemann.

Tumulty, J. P. (2018). Dear enemy effect. In J. Vonk & T. K. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior*. Springer. https://doi.org/10.1007/978-3-319-47829-6

Tumulty, J. P., & Bee, M. A. (2021). Ecological and social drivers of neighbor recognition and the dear enemy effect in a poison frog. *Behavioral Ecology*, *32*(1), 138–150. https://doi.org/10.1093/beheco/araa113

Tumulty, J. P., Pašukonis, A., Forester, J. D., Ringler, M., Hödl, W., & Bee, M. A. (2018). Brilliant-thighed poison frogs do not use acoustic identity information to treat territorial neighbours as dear enemies. *Animal Behaviour*, *141*, 203–220. https://doi.org/10.1016/j.anbehav.2018.05.008

van Oers, K., Drent, P. J., Dingemanse, N. J., & Kempenaers, B. (2008). Personality is associated with extrapair paternity in great tits, *Parus major*. *Animal Behaviour*, *76*(3), 555–563. https://doi.org/10.1016/j.anbehav.2008.03.011

Weary, D. M., & Krebs, J. R. (1992). Great tits classify songs by individual voice characteristics. *Animal Behaviour*, *43*(2), 283–287. https://doi.org/10.1016/S0003-3472(05)80223-4

Webster, M. S., Pruett-Jones, S., Westneat, D. F., & Arnold, S. J. (1995). Measuring the effects of pairing

success, extra-pair copulations and mate quality on the opportunity for sexual selection. *Evolution*, *49*(6), 1147–1157. https://doi.org/10.1111/j.1558-5646.1995.tb04441.x.

Wei, M., Lloyd, H., & Zhang, Y. (2011). Neighbour–stranger discrimination by Yellow-bellied Tit *Parus venustulus*: evidence for the "dear-enemy" effect. *Journal of Ornithology*, *152*(2), 431–438. https://doi.org/10.1007/s10336-010-0609-6

Wiley, R. H. (2013). Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. *Biological Reviews*, *88*, 179–195. https://doi.org/10.1111/j.1469-185X.2012.00246.x

Ydenberg, R. C. (1984). The conflict between feeding and territorial defence in the great tit. *Behavioral Ecology and Sociobiology*, *15*(2), 103–108. https://doi.org/10.1007/BF00299376

Ydenberg, R. C., Giraldeau, L. A., & Falls, J. B. (1988). Neighbours, strangers, and the asymmetric war of attrition. *Animal Behaviour*, *36*(2), 343–347.

Ydenberg, R. C., & Krebs, J. R. (1987). The tradeoff between territorial defence and foraging in the great tit (*Parus major*). *American Zoologist*, *27*, 337–346. https://doi.org/10.1093/icb/27.2.337

Yoon, J., Sillett, T. S., Morrison, S. A., & Ghalambor, C. K. (2012). Breeding density, not life history, predicts interpopulation differences in territorial aggression in a passerine bird. *Animal Behaviour*, *84*(3), 515–521. https://doi.org/10.1016/j.anbehav.2012.05.024

946 Table S1. Study site locations and summary statistics.

| Site Name | Coordinates | N DE behaviour/N tested | Clutch Size | N Fledged | Total Mass (g) | Average Mass (g) |
|---|---|---|---|---|---|---|
| Ballinphellic (BP) | 51.840546, −8.630297 | 0/1 | 6 | 5 | 83.8 | 16.8 |
| Castle Bernard (CB) | 51.741568, −8.767775 | 2/9 | 5.0 (1.3) | 1.4 (1.8) | 33.5 (38.7) | 14.4 (3.1) |
| Dukes Wood (DW) | 51.785965, −8.751366 | 2/9 | 4.4 (1.1) | 1.7 (1.7) | 29.3 (28.0) | 16.6 (1.9) |
| Dunderrow (DD) | 51.719555, −8.600512 | 0/4* | 5 | 0 | 45.9 | 15.3 |
| Garrettstown (GT) | 51.655159, −8.616456 | 0/1 | 6 | 6 | 108.0 | 18.0 |
| Innishannon (IN) | 51.763415, −8.664943 | 2/9 | 5.9 (1.5) | 1.0 (1.6) | 19.2 (20.8) | 15.4 (1.6) |
| Kilbrittain (KB) | 51.671290, −8.682011 | 8/16 | 4.6 (2.1) | 2.1 (2.0) | 33.6 (31.4) | 15.9 (1.6) |
| Shippool (SP) | 51.737780, −8.630792 | 1/2 | 4.5 (0.7) | 0.5 (0.7) | 8.8 (12.4) | 17.6 |

947 Site names and coordinates as in O'Shea (2017). Values for the reproductive success variables are means (±standard deviation, not reported
948 when only one individual was sampled at a site). Clutch size includes all eggs whether or not they were identified as extrapair; number fledged,
949 total mass and average mass were calculated excluding any young identified as extrapair. Failed nests in which no offspring survived to day 15
950 (aside from the three excluded from Dunderrow, see below) were given a value of 0 for total mass, but were not included in calculations of
951 average mass. Initials for the site name are given that correspond to the site variable in the raw data file. N DE behaviour refers to the number of
952 individuals exhibiting dear enemy behaviour under the standard criterion.

953 *Three of these individuals at Dunderrow were excluded from the analyses of reproductive success variables because of widespread stoat

954 predation of nests at the site in 2018. The numbers reported here are from the one nest that was included.

955

956 Table S2. Effects of context on dear enemy behaviour.

957

| Dear enemy criterion | Factor | Estimate (SE) | Test statistic | P |
|---|---|---|---|---|
| Standard criterion | Year | −0.73 (0.81) | −0.91 | 0.37 |
| | Date | 0.08 (0.05) | 1.50 | 0.13 |
| | Stimulus rate (high) | −0.45 (0.67) | −0.66 | 0.51 |
| | Date relative to first egg | −0.002 (0.06) | −0.03 | 0.98 |
| Standard without nonresponders | Year | −2.47 (1.32) | −1.87 | 0.062 |
| | Date | 0.15 (0.07) | 1.95 | 0.051 |
| | Stimulus rate (high) | 0.32 (0.85) | 0.38 | 0.70 |
| | Date relative to first egg | 0.001 (0.068) | 0.02 | 0.98 |
| Response slope | Year | −0.003 (0.036) | −0.08 | 0.94 |
| | Date | 0.001 (0.002) | 0.25 | 0.81 |
| | Stimulus rate (high) | 0.008 (0.026) | 0.32 | 0.75 |
| | Date relative to first egg | 0.002 (0.002) | 0.97 | 0.34 |
| Song slope | Year | −0.14 (1.12) | −0.13 | 0.90 |
| | Date | 0.06 (0.07) | 0.95 | 0.35 |
| | Stimulus rate (high) | −0.68 (0.81) | −0.84 | 0.41 |
| | Date relative to first egg | 0.003 (0.06) | 0.04 | 0.97 |

958

959 Results are from a generalized linear model of whether or not the individual met the standard criterion
960 for expressing dear enemy behaviour (where the test statistic is *z*), or from linear models for the slope
961 variables (where the test statistic is *t*). The reference category for stimulus rate is the low rate stimulus.
962 *N* = 50 individuals for the standard criterion (one individual not included because there was never an egg

963 laid at the nest), 29 individuals for standard without nonresponders, and 25 individuals for the slope

964 variables.

965 Table S3. Full models of dear enemy behaviour and reproductive success, including brood size effects when included in models.

966

| Reproductive measure | Dear enemy behaviour measure | Estimate(SE) | test statistic | P | N |
|---|---|---|---|---|---|
| Clutch size | Standard criterion | −0.21 (0.15) | −1.45 | 0.15 | 48 |
| | Standard without nonresponders | −0.21 (0.17) | −1.26 | 0.21 | 29 |
| | Response slope | 1.06 (1.29) | 0.83 | 0.41 | 29 |
| | Song slope | 0.009 (0.043) | 0.20 | 0.84 | 29 |
| Number fledged | Standard criterion | −0.45 (0.35) | −1.27 | 0.21 | 48 |
| | Standard without nonresponders | −0.35 (0.38) | −0.93 | 0.35 | 29 |
| | Response slope | 3.19 (2.14) | 1.49 | 0.14 | 29 |
| | Song slope | 0.04 (0.07) | 0.58 | 0.56 | 29 |
| Total biomass (g) | Standard criterion | −38.2 (15.4) | −2.49 | **0.01** | 47 |
| | Standard without nonresponders | −35.5 (18.2) | −1.95 | 0.051 | 29 |
| | Response slope | 139.0 (135.3) | 1.03 | 0.30 | 29 |
| | Song slope | 3.84 (4.57) | 0.69 | 0.49 | 29 |
| Average offspring biomass (g) | Standard criterion | −0.78 (0.94) | −0.83 | 0.41 | 30 |
| | Brood size | 0.28 (0.27) | 1.06 | 0.30 | |

| | | | | | |
|---|---|---|---|---|---|
| | Standard without nonresponders | 0.69 (1.11) | 0.62 | 0.55 | 16 |
| | Brood size | 0.46 (0.34) | 1.35 | 0.2 | |
| | Response slope | −7.00 (9.64) | −0.73 | 0.48 | 16 |
| | Brood size | 0.62 (0.46) | 1.36 | 0.2 | |
| | Song slope | 0.09 (0.33) | 0.27 | 0.79 | 16 |
| | Brood size | 0.28 (0.49) | 0.58 | 0.57 | |
| Individual offspring mass (g) | Standard criterion | −0.70 (0.92) | −0.76 | 0.45 | 30 |
| | Brood size | 0.20 (0.26) | 0.75 | 0.46 | |
| | Standard without nonresponders | 0.77 (0.99) | 0.78 | 0.45 | 16 |
| | Brood size | 0.38 (0.31) | 1.24 | 0.24 | |
| | Response slope | −7.00 (8.81) | −0.79 | 0.44 | 16 |
| | Brood size | 0.57 (0.44) | 1.28 | 0.22 | |
| | Song slope | 0.01 (0.30) | 0.05 | 0.96 | 16 |
| | Brood size | 0.29 (0.47) | 0.62 | 0.55 | |

Output is from models with the reproductive measure as the dependent variable (all measures include only within-pair young), and the dear enemy behaviour measure as a factor. Models for average and individual offspring mass included brood size as an additional factor (brood size

never had a significant effect; results reported in Table S3). Test statistic is *z* for analyses of clutch size and number fledged, and *t* for the remaining reproductive measures. Sample sizes refer to the number of subject males included and vary because some variables excluded individuals that failed to respond to both the final neighbour playback and the stranger playback ("nonresponders"), and because for many nests all offspring died before weighing on day 15. For the standard criteria, the reference category was individuals that did not show dear enemy behaviour. Statistically significant effects are shown in bold.

975 Supplementary References

976

977 O'Shea, W. 2017. The evolutionary ecology of personality and life history variation in an anthropogenic

978 heterogeneous landscape. PhD Thesis, University College Cork.

979