bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

No reproductive fitness benefits of dear enemy behaviour 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

Territorial animals often respond less aggressively to neighbours than strangers. This 'dear enemy' effect is hypothesized to be adaptive by reducing unnecessary aggressive interactions with non-threatening individuals. A key prediction of this hypothesis, that individual fitness will be affected by variation in the speed and the 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 to a simulated establishment of a neighbour on an adjacent territory during early stages of breeding, as an assay of individuals' tendencies to form dear enemy relationships. Males reduced their approach to the speaker and sang fewer songs on later playback repetitions. However, only some males exhibited dear enemy behaviour by responding more strongly to a subsequent stranger playback, and when the playback procedure was repeated on a subset of males, there was some indication for consistent differences among individuals in the expression of dear enemy behaviour. We monitored nests and analysed offspring paternity to determine male reproductive success. Individuals that exhibited dear enemy behaviour towards the simulated neighbour did not suffer any costs associated with loss of paternity, but there was also no evidence of reproductive benefits, and no net effect on reproductive fitness. The general ability to discriminate between neighbours and strangers is likely adaptive, but benefits are probably difficult to detect because of the indirect link between individual variation in dear enemy behaviour and reproductive fitness, and because of the complex range of mechanisms affecting relations with territorial neighbours.

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

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

Territoriality is a widespread behaviour that provides important benefits to territory holders such as reducing competition for food or breeding resources, and guarding of mates from extra-pair matings (Stamps 1994; Adams 2001). However, there are many costs to holding territories that arise primarily from the vigilance and aggressive behaviours required to defend and maintain territorial boundaries against intruders (Ydenberg 1984; Mares et al. 2012; Wischhoff et al. 2018). One way to lower these costs is to reduce territorial defence behaviours towards non-threatening individuals and instead focus defence efforts particularly on those individuals that pose a threat of usurping the territory or the resources within (Getty 1987; Temeles 1994). Thus, many species exhibit the 'dear enemy effect', which is defined as individuals showing less aggressive behaviour towards territorial neighbours than towards strangers (Ydenberg et al. 1988; Temeles 1994; Stoddard 1996; Tumulty 2018). If neighbours are not a threat because they hold their own territory, then dear enemy behaviour should benefit individual fitness through reduced costs of unnecessary aggression, facilitating greater investment in foraging and reproduction (Getty 1987; Temeles 1994), as well as defence against truly threatening individuals (Leiser and Itzkowitz 1999). 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, particularly enhancing reproductive success when the territory is used for breeding. However, this key prediction has never been tested.

Selection acts on individual variation, and thus to understand the evolution of the dear enemy effect it is important to determine whether individuals differ consistently in dear enemy behaviour (as a proxy for heritability), and whether these differences have fitness consequences. Repeatable individual variation in dear enemy effect expression is expected because of variation in individuals' cognitive capabilities to

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

recognize familiar neighbours (Reichert and Quinn 2017), because of personality traits such as aggressiveness (Akçay et al. 2014), and due to covariance with other repeatable behavioural traits (Verbeek et al. 1996). Some studies have provided indirect support for fitness benefits of dear enemy behaviour, for instance by demonstrating that individuals holding territories with long-term neighbours have higher reproductive success than those with new neighbours, or that neighbours tend to engage in cooperative behaviours together (Beletsky and Orians 1989; Grabowska-Zhang et al. 2012a, b; Siracusa et al. 2021). However, in these studies it is unknown whether and to what extent individuals discriminated between neighbours and strangers in their aggressive interactions, and therefore whether there was variation in dear enemy behaviour. Thus, while it is clear that the ability and tendency to discriminate neighbours from strangers varies and likely has functional significance, there is still a limited understanding of the selection pressures that may be acting on dear enemy behaviour, especially at the within-species level.

The dear enemy effect is facilitated by cognitive mechanisms that enable individuals to learn some characteristic of their neighbours, allowing them to discriminate between neighbours and strangers (Wiley 2013). These mechanisms – of which more than one may be involved in any given species – range from habituation to a neighbour's signal characteristics and/or their spatial location (Petrinovich and Peeke 1973; Brooks and Falls 1975; Yasukawa 1981; Bee and Gerhardt 2001; Dong and Clayton 2009), associative learning of neighbour characteristics (Richards 1979), to 'true individual recognition' (Johnston and Jernigan 1994; Gheusi et al. 1997; Saeki et al. 2018), and act over time scales ranging from short-term decreases in aggression within a day to persistent recognition of individuals across years (Godard 1991; Tumulty and Sheehan 2020). 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 the cognitive abilities associated with neighbour-

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

stranger discrimination. If this cognitive variation affects reproductive success—for instance if those individuals on breeding territories that quickly learn to recognize their neighbours avoid unnecessary agonistic encounters and can invest more in their offspring—then individuals with superior cognitive ability may have higher fitness and be favoured by selection. Non-cognitive factors also affect the expression of the dear enemy effect. For instance, individuals may differ in aggressiveness or other personality traits (Hyman et al. 2004; Amy et al. 2010; Jacobs et al. 2014), and context-dependent factors such as territory size, breeding status, and the density of neighbouring territory-holders also likely affect responsiveness to neighbours (Werba et al. 2021). 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 (e.g., Bee and Gerhardt 2001), or whether this differs among individuals.

Like many songbirds, great tits (*Parus major*) defend territories around their nest during the breeding season (Gosler 1993). Male great tits do so 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). In great tits, approximately 25-50% of broods contain extrapair young (Lubjuhn et al. 1999; Brommer et al. 2010; Patrick et al. 2012). 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 (Krebs 1971; Falls et al. 1982; McGregor and Avery 1986), and the effect was stronger when the neighbour's song was

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

played from its own territory (McGregor and Avery 1986). Territory defence begins very quickly after settlement, often within hours (Krebs 1971). The speed with which dear enemy relationships are formed is unknown, but given that strong territory defence occurs during a limited breeding season, and that great tits habituate rapidly to song (Krebs 1976), a reduced response to neighbours compared to strangers is likely to occur within a few days.

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 and Avery 1986; Weary and Krebs 1992). There is also among-individual variation in aggressive responses to playback resulting from personality differences (Amy et al. 2010), and personality and reproductive investment were related to the speed with which individuals habituated to a series of song playbacks (Rivera-Gutierrez et al. 2017); this habituation to a familiar stimulus is likely an important mechanism in the establishment of dear enemy relations (Peeke 1984). Great tits trade off the time invested in foraging and territory defence (Ydenberg 1984; Ydenberg and 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).

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. Our overall objective was to examine whether individuals habituated by reducing their response after repeated exposures to the songs of a simulated new neighbour while still maintaining a heightened response to a stranger (i.e.,

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

displayed dear enemy behaviour), whether they differed consistently in this regard, and whether there were any beneficial effects of the tendency to form dear enemy relationships on fitness and reproduction in general. We note that we did not expect these fitness effects to arise because of the individual's actual behaviours towards the simulated neighbour in our playback trials, but rather we assumed that the individual's behaviour in the playback trials was indicative of how it engaged with others in natural territorial interactions, with variation in the tendency to reduce aggression towards a familiar stimulus (as an experimental proxy for the tendency to form dear enemy relationships) leading to effects on reproductive success.

First, we determined if males' responses to playbacks of a simulated neighbour that were broadcast multiple times across three days declined over time, consistent with habituation to the neighbour playback stimulus. Second, we tested whether the response to a subsequent 'stranger' playback was stronger than that to the final neighbour playback, consistent with the reduced aggressive response to a familiar neighbour compared to a stranger that is the hallmark of the dear enemy effect. Third, we estimated the repeatability of several dear enemy behaviour measures to determine if individuals differed intrinsically in these behaviours. We also explored the extent to which the expression of these dear enemy behaviour measures was context dependent, for example due to phenology, which is related to overall levels of nest defence (Hyman 2005; Jin et al. 2021), and the repetition rate of songs during the playback, which affects the rate of habituation towards repeated stimuli (Thompson et al. 1973). Finally, we tested whether there was any evidence for selection or any reproductive benefits of dear enemy behaviour in a number of ways: A) controlling for paternity, we tested the prediction of positive selection on the tendency to perform dear enemy behaviour, that is, that individuals that rapidly reduced their response to neighbours should fare better for a variety of life history traits that are good indicators of reproductive fitness and recruitment to the breeding population in the great tit

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Tinbergen and Boerlijst 1990), in particular clutch size, the number of offspring fledged, and average offspring mass; B) We repeated these analyses including all nestlings (including extra-pair) to test for evidence of any benefits of the tendency to perform dear enemy behaviour at the nest, irrespective of paternity; C) We examined if there was any relationship between observed levels of extra-pair paternity at the social nest and the tendency to perform dear enemy behaviour. We discuss the implications of our experimental results and analyses of reproductive success for our understanding of the dear enemy effect.

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, Fig. S1). Each site contained an array of nestboxes, which are preferentially used by great tits for breeding (East and Perrins 1988), and in which most individuals had been ringed as part of a long-term study (for details see O'Shea et al. 2018). 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.

Playback stimuli

Male great tits sing on territories containing their nesting site. They typically have a repertoire of three to four distinctive song types (Gompertz 1961; McGregor and Krebs 1982a) but usually repeat the same one for several minutes before switching to a different type (Krebs 1976). The playbacks were designed to mimic this repetition of a single song type. Although simulating a larger song repertoire through

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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), 3. Great tits are capable of discriminating between individuals even based on songs of those individuals they have never heard before (Weary and Krebs 1992). Thus, our comparison of birds' responses to a simulated neighbour (recorded song from one individual, played back several times) and a simulated stranger (recorded song from a different individual, played back once) is an appropriate assay of dear enemy behaviour (i.e., the difference in response to a familiar stimulus from a familiar location, the neighbour, and an unfamiliar stimulus from an unfamiliar location, the stranger) and is unlikely to have been perceived by the bird as two different song types from the same individual.

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 and Krebs 1982b)) 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 or ten songs per minute for five minutes, which we refer to as the "low" and "high" stimulus rate treatments, respectively. These two stimulus rates were used to test the prediction that the likelihood of habituation to the neighbour songs depended on the stimulus repetition rate (Thompson and Spencer 1966). A total of 29 song

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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. Great tit males respond readily to playback of song exemplars by exhibiting territorial behaviour (Rivera-Gutierrez et al. 2015; Snijders et al. 2017; Jin et al. 2021).

Experimental design

On each of three consecutive days we exposed males to three playbacks, performed at 1-hour intervals, of one of the playback stimuli (Fig. 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. Territory settlement generally takes place very rapidly in this species (Krebs 1971), and responses to repeated playback typically decrease over a few days (Krebs et al. 1981; Rivera-Gutierrez et al. 2017), so this interval allowed us to study the initial development of dear enemy behaviour. 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 and Ydenberg 1987), 50 m is a typical nearest-neighbour distance (Krebs 1971), and neighbour-stranger discrimination is typically strongest near the territorial border (Falls and Brooks 1975; Stoddard et al. 1991). There is an inevitable trade-off between the advantages of our approach of broadcasting playbacks from a standard amplitude and distance from the nestbox, and those of

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

estimating males' actual territorial boundaries and placing the speaker at each male's estimated territorial edge. Because we used a fixed playback distance, our playbacks may have been more salient to some individuals than others, depending on whether the standard 25 m playback distance was or was not within their territory. However, although great tits differ in their response to playbacks well outside of their territory [10 m beyond the boundary] and those very close to the centre of their territory [20-30 m within the boundary] (Peake et al. 2001), our setup was unlikely to have corresponded to this situation, and to our knowledge no study has examined whether great tits are capable of discriminating between songs coming from just inside versus just outside territorial boundaries. In contrast, it is well established that song intensity affects the aggressive response of songbirds, including great tits (Brumm and Ritschard 2011; Ritschard et al. 2012; Luther et al. 2016). Furthermore, song amplitude is likely to play a role in the speed of dear enemy learning (Bee 2001). We therefore used a fixed playback distance so that song amplitude could be standardized. 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

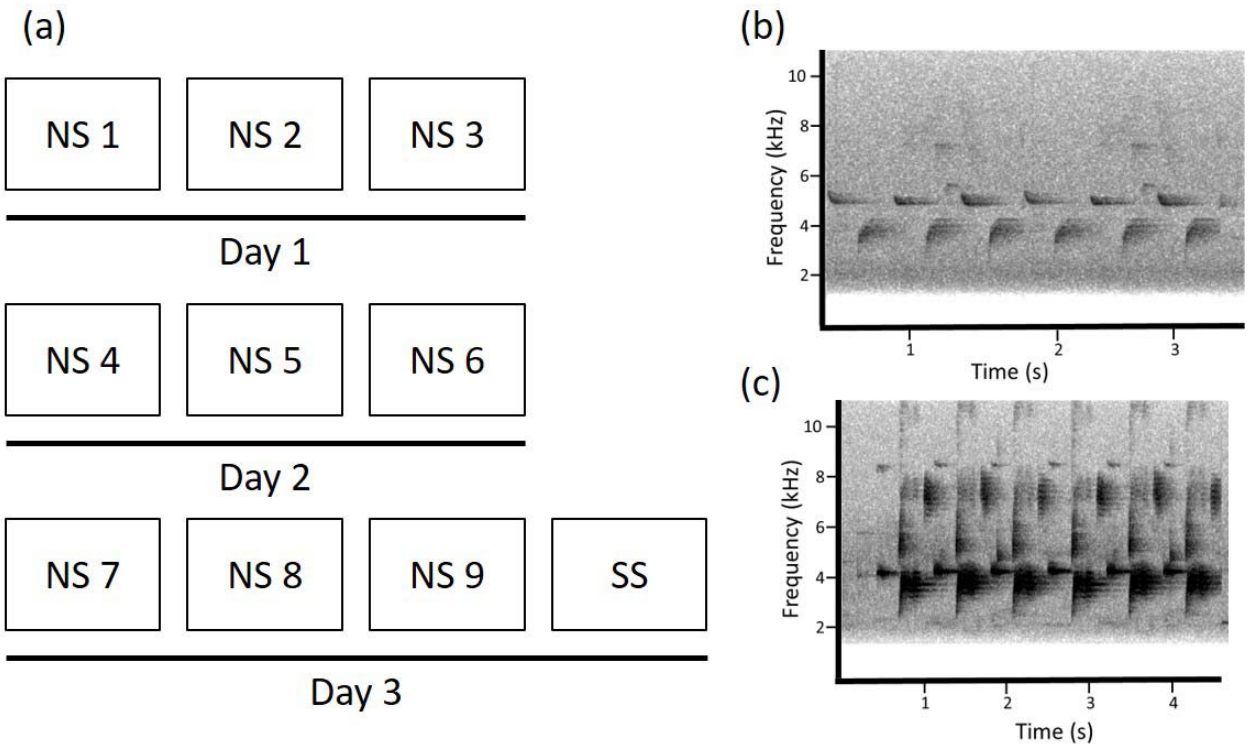


Fig. 1 (a) Timeline of the playback experimental design. We broadcast the neighbour stimulus ('NS') for five minutes per trial (number in each box corresponds to the trial number), for a total of three trials per day with approximately one hour between trials. This procedure was repeated for three consecutive 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 other words that males that had reduced their response to the neighbour retain the capability of an aggressive response towards strangers). In 2017, the entire procedure was repeated with different stimuli after a period of ten days to test the repeatability of dear enemy behaviour. (b) Exemplar playback stimulus. The spectrogram shows a single 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 acoustic variation in great tit songs. An example playback experiment could use the song in (b) for the neighbour stimulus and that in (c) for the stranger stimulus. Note that the spectrograms do not illustrate the equalization of playback amplitude that was used for the trials

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

On the third day, immediately after the response was recorded to 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, and at a 90-degree angle or greater, with respect to the nest box, from the location of the neighbour playbacks). 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 neighbour-stranger discrimination 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 other words, the stranger playback was used to demonstrate that the decline in aggressive response observed across the trials is specific to the neighbour's song. We performed the stranger playback immediately after recording the response to the last neighbour playback for logistical reasons and because doing so is a more conservative test of neighbour-stranger discrimination than if we had delayed the stranger playback by some interval. If an individual was not responsive to its neighbour but then immediately responded to the stranger playback, this is a strong test of the ability to discriminate between the two, whereas any delay between the playbacks increases the chance that any generalized habituation of the aggressive response would decline, potentially leading to misinterpretation of the data. Although it is possible that birds could have mounted an increased aggressive response during the stranger playback session not because they were responding to what they perceived as a stranger but rather because it was a novel situation in that the playback occurred a few minutes, rather than an hour, after the previous playback, we consider this interpretation unlikely because for this to be the case either birds would have had to learn the number of daily playbacks or the time interval between them, or would have had to interpret the playback as a different song type of the neighbour, which is unlikely because great tits are capable of learning individual voice characteristics and then discriminating among

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

them even based on songs they had never heard before (Weary and Krebs 1992) and therefore should not have mistaken our stranger stimulus for another song of the neighbour.

We justify our design of using both a different location and song for the stranger stimulus, because the aim of our experiment was not to identify the specific criteria used by individuals to discriminate neighbours and strangers, i.e., whether discrimination was based on these individuals' song characteristics or location, but rather to determine the consequences of individual differences in tendency to discriminate. Therefore, we used a playback design that increased the opportunities for discrimination between our stimuli by varying both location and song, which is also the most likely scenario for natural neighbour-stranger discrimination because the neighbour would not normally move to a different territory and the stranger would not normally sing from the exact spot that the neighbour was just singing in. We acknowledge that our design does not separate the effects of location and song type by including controls for each. This would not have been feasible given the longitudinal design and limited breeding window in this wild system. Furthermore, although we cannot rule out that subjects interpreted a song type recorded from a different bird and broadcast from a different location as in fact still coming from the original simulated neighbour, as discussed above, great tits discriminate among individual voice characteristics (Weary and Krebs 1992). Therefore, we interpret their responses in the final playback as responses to a simulated stranger, but note that, even under dear enemy relationships, individuals may also respond aggressively to neighbours when they display from a new location (Brindley 1991; Husak and Fox 2003; Lovell and Lein 2005; Dalton et al. 2020).

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, which would indicate intrinsic differences among individuals

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Bell et al. 2009). Therefore, we performed playbacks as above, but with a different song stimulus from a different location, simulating a different new neighbour establishing a territory. A different stranger stimulus was also used in these repeated trials, and the male was given the alternative rate treatment to that which it had been exposed during the first round of playbacks.

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. Playbacks occurred during the early stages of breeding for most birds – between late nest building and egg laying. Due to logistical constraints of running experiments across multiple field sites with respect to the timing of nest checks, seven individuals were tested after the last egg was laid at the nest, including three individuals for which incubation was already started at the nest (the latest start date of the first repetition of the playback trials was 4 days into incubation). To control for the timing of the playback with respect to the breeding cycle (Petrinovich et al. 1976; Mace 1987; Jin et al. 2021), we included the date of the first playback trial relative to the first egg date as a variable in analyses (see below).

Playback procedure

During each playback a single observer was positioned in cover near the playback speaker to monitor the male's behaviour, noting the distance between the bird and the speaker, and making 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 did our best to avoid being observed by the birds during the playback; we consider it unlikely that our presence affected the results because these experiments took place in areas close to busy walking trails and birds were likely habituated to the presence of humans. All playbacks took place in locations where a male was known to occupy a territory

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

because on the days prior to the playbacks it was observed singing near the nestbox and there was evidence of nest building activity. Males were usually visually or acoustically located prior to the start of the playback, and if not, most were sighted during the playback. This was not always possible, however, and we assume that these males initially unseen were within hearing distance of the stimulus during the playback. This was reasonable because males spend the great majority of their time on their territories and when they do leave their territories they only move short distances (Firth et al. 2018); at such distances our playback stimulus amplitude would remain well above the species' masked hearing threshold (Langemann et al. 1998). Nevertheless, it is possible that we did not detect some response vocalizations from birds that were located far from the recorder, although signal to noise ratios of great tit song remain high at least 60 m from the source (Blumenrath and Dabelsteen 2004; Mockford et al. 2011). Joint territory defence by two individuals in response to the playback was not observed in this study, possibly due to the low density of individuals at our study sites. It was not possible to record data blind because our study involved focal animals in the field.

Response variables

We noted the closest approach of the male to the playback speaker, in categories of 5 m (Nelson and Soha 2004). Subjects were not always located at the nest box at the start of the playback, so values greater than 25 m were possible. Individuals that did not respond at all were given a value of "None" for closest approach, which in the ordinal analyses described below was considered the greatest distance (whereas closest approaches from 0-5 m were considered the shortest distance, with successive distance categories ranked according to their 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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.

Criterion for dear enemy behaviour

The typical method for demonstrating the dear enemy effect is to present individuals, in a single playback session, with the signals of an actual established neighbour, followed by a single playback session presenting the signals of an individual that the subject could not have interacted with previously (Brunton et al. 2008; Wei et al. 2011; Battiston et al. 2015). The dear enemy effect is inferred when individuals have a reduced response to the neighbour stimulus compared to the stranger stimulus. In our protocol we simulated the establishment of a new neighbour through a series of playbacks, which essentially served as training sessions to give the subject the opportunity to learn to recognize its neighbour. Therefore, to test whether or not the subject was indeed exhibiting dear enemy behaviour at the end of the necessarily fixed number of trials, we used the typical criterion for testing the dear enemy effect: comparison of the response to a final neighbour playback and a stranger playback. Our criterion (hereafter referred to as the 'standard criterion') that the subject was expressing dear enemy behaviour was that it did not respond in the final (ninth) playback of the simulated neighbour, but then did respond to the subsequent (tenth) playback of a simulated stranger.

Although our binary criterion should theoretically identify individuals that were expressing dear enemy behaviour towards the simulated neighbour, it has its limitations, so we explored additional measures of an individual's change in response towards its neighbour over time. First, in some cases individuals responded to neither the final neighbour playback nor the stranger playback (N = 21). These results are

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

difficult to interpret because they may indicate a general loss of motivation to defend the territory rather than a failure to discriminate neighbours and strangers. We therefore repeated all analyses, defining dear enemy behaviour as above, but only including those birds that responded to the stranger stimulus (referred to as the 'standard criterion without non-responders'; N = 30 of 51 birds meeting this criterion; no individual failed to respond to the stranger playback after responding to the final neighbour playback). Second, for each individual, we extracted slope parameters from a logistic regression of the binary response variable on trial number (not including the stranger playback; referred to as the 'response slope criterion'), and a linear regression of the number of songs on trial number (referred to as the 'song slope criterion'). The aim was to obtain a more quantitative estimate of the change in response across trials that may reveal more variation than our binary criterion. If individuals have developed dear enemy behaviour towards the newly established neighbour, these values would be expected to be negative, indicating a decline in response across trials. However, a decline in response to a neighbour is not sufficient to demonstrate the dear enemy effect, because individuals must continue to respond to strangers. Therefore, we only included individuals that responded to the stranger playback for this second set of analyses (N = 30).

Breeding data

Breeding data were obtained as part of standard monitoring protocols for the project (O'Shea et al. 2018). We recorded the date when the first egg was laid, the total clutch size, date of hatching and number of fledglings. The identity of the subject male sometimes could be determined by identification of unique colour rings if the male had already been captured prior to the experiment. Some males were also identified using RFID-equipped nestbox entrance doors, which could read the unique passive integrated transponder tag placed on the leg of previously captured males. The age (first year juvenile or

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

adult) of previously captured males was determined from capture records. Males could also be identified when breeding adults were caught at the nest for ringing, ageing (as either first year juveniles or adults), and measurements, 10-12 days after the eggs hatched. Fifteen subjects were not identified because the nest was abandoned prior to trapping (N=1 before eggs laid, N=8 before eggs hatched, N=5 after eggs hatched), or the male could not be caught (N=1). However, our analyses do not rely on knowing the specific identity of the subject, and we can safely assume that no male was recorded in the study more than once based on the timing and distribution of boxes at which playbacks were carried out across the eight sites. Breeding densities are low at our sites and in cases where males could not be identified by colour-rings we nevertheless consider it highly unlikely that more than one individual responded to the playback on different trials. Chicks were weighed at day 15, and we determined the number of fledglings by inspecting the nest for any dead chicks after the breeding attempt was complete.

Paternity analysis

Estimates of male reproductive fitness can be strongly influenced by extra-pair paternity (EPP) (Webster et al. 1995). Although rates of EPP are relatively low in great tits (van Oers et al. 2008; Patrick et al. 2012), they could have altered the relationship between dear enemy behaviour and reproductive fitness, particularly as paternity loss is one of the potential costs of territorial intrusions, and neighbours and strangers may differ in the threat they pose to paternity (Schlicht et al. 2015). EPP levels may also influence selection on males to cooperate via dear enemy effects (Eliassen and Jørgensen 2014). We therefore analyzed males' reproductive success using metrics that excluded any offspring that were identified in a paternity analysis as being extra-pair offspring (note that we did not attempt to quantify males' success at obtaining extra-pair matings at other nests because the small size and fragmented

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

nature of our study sites prevented us from confidently assessing a male's reproductive output away from his own nest). This is the appropriate method to analyse fitness, but we were also interested in the consequences of dear enemy behaviour on parental care behaviour. Therefore, we performed an additional set of analyses on the reproductive success variables in which both within-pair and extra-pair young were included, and present these in the supplement (results were not qualitatively affected by whether or not we excluded extra-pair young).

DNA was obtained from feathers taken from breeding pairs and offspring on their respective dates of capture and ringing. DNA extraction was performed using the protocol of the E.Z.N.A. Tissue DNA Kit (Omega Bio-Tek, Norcross, GA, USA), with the exception of the use of 80 µL elution buffer on a single elution step. Samples were genotyped at eight microsatellite loci, selected based on previously observed variability and utility, as well as potential for multiplexing in a single reaction (Pma69u (k=7 alleles) (Kawano 2003); PmaD22 (k=20), PmaCan1 (k=15), PmaGAn30 (k=5), PmaC25 (k=19), PmaTAGAn86 (k=19), PmaTGAn33 (k=17) and PmaTGAn45 (k=11) (Saladin et al. 2003)). Multiplex PCR was performed in a 3.5 µl total volume, including 1 µl of DNA extract, and 1.75 µl of 2x Top-Bio™ Plain Combi PP Mastermix, with a concentration of 0.03 µM for the Pma69U, PmaCan1, PmaGAn30 and PmaTAGAn86 primers and 0.06 µM for the PmaC25, PmaD22, PmaTGAn33 and PmaTGAn45 primers. The PCR was programmed with an initial denaturation at 95°C (15 min) followed by five cycles of 94°C (30 s), 55°C (90 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 60°C for 30 min. The PCR products were diluted in 14 µL nuclease free water; and run on an Applied Biosystems ABI3500xl DNA analyser using POP-7 polymer with GeneScan™ 600 LIZ™ Dye Size Standard v2.0 (ThermoFisher Scientific). We used GeneMarker version 2.7.0 software (SoftGenetics, Pennsylvania, USA) to determine allele sizes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Paternity was assigned using CERVUS version 3.0.7 software (Kalinowski et al. 2007) using 10,000 cycles, 94 candidate fathers, a 0.02% error rate, two candidate parents and 93% of loci typed as simulation parameters. Only individuals that were successfully genotyped at five or more loci were included in paternity analyses. Individuals were determined to be within-pair offspring if all loci matched those of the social father and social mother combination, or if there was a mismatch at only one of the loci but the social father was identified as the most likely father using critical trio LOD scores returned by the program. Offspring that did not meet these criteria were categorized as extra-pair offspring. We assumed that offspring whose paternity was not determined (9 of 96 fledged offspring and 13 of 111 weighed offspring were of unknown paternity; two offspring fledged but were not weighed or analysed for paternity because they fledged on the day of weighing) were within-pair offspring. The combined exclusion probability for all eight microsatellites was >99.99%. Two of our loci significantly deviated from Hardy-Weinberg equilibrium when the genotypes of all individuals in the analysis were included (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 the family structure of the data.

Data analysis

Except for the analyses of repeatability (see below), all analyses were performed only on the first repetition of the playback trials (i.e., we excluded data from the second repetition of the three-day procedure that was performed in 2017 only). We address the following questions in our analyses:

1. Does the response to the neighbour playback decrease over time? We tested whether males reduced their aggressiveness as they became more exposed to the songs of another male using three different variables, all of which have been demonstrated to be related to the aggressive response in great tits in

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

previous studies (McGregor and Avery 1986; Doutrelant et al. 2000; Amy et al. 2010; Snijders et al. 2017): "response", i.e. whether the male responded at all by singing or approaching (binary), closest approach (ordinal categorical), and number of songs during individual trials. In all cases the trial number (1-9) and stimulus rate treatment (low or high) were entered as fixed effects, and individual identity as a random effect. Response was modelled using a binomial generalized linear mixed model using the glmer function in the lme4 version 1.1-23 package (Bates et al. 2015) in R 4.0.2 software (R Development Core Team 2021). Approach category was modelled using a cumulative link mixed model to account for the ordinal nature of the dependent variable using the clmm function in the ordinal 2019.12-10 package (Christensen 2019). Number of songs was modelled as a poisson variable using the glmmTMB function of the glmmTMB 1.0.2.1 package (Brooks et al. 2017), which accounts for zero inflation in the dataset.

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 dear enemy behaviour is to show that this decrement is specific to the neighbour stimulus. 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 the tendency for individuals to form a dear enemy relationship within the timeframe of our playback design repeatable? We estimated the repeatability of neighbour-stranger discrimination using the rptR package (Nakagawa and 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 meet the standard criterion (i.e., responded to the stranger but not the neighbour stimulus) as a binary dependent variable. One individual was removed from this analysis because it did not respond during

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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.

We examined the context-dependence of dear enemy behaviour using separate models for each of the four different methods of quantifying dear enemy behaviour (see *Criterion for dear enemy recognition*, above). We used generalized linear models for the standard criterion and standard criterion without non-responders, and linear models for the response slope and song slope. In all cases, we tested for effects of year (2017 or 2018), date of the first playback, date of the first playback minus the date the first egg was laid (to account for variation in dear enemy behaviour across the breeding cycle (Jin et al. 2021)), stimulus repetition rate (low or high), and the average number of songs the subject produced across the nine neighbour playbacks, as an estimate of its overall 'aggressiveness'. We performed separate models for male age, with model structure as above but age (first year or adult) as the only factor, because age was only known for a subset of the birds.

4. Is there evidence for a difference in reproductive fitness between individuals that did or did not express dear enemy behaviour? We ran separate models for each of four different metrics of reproductive success, controlling for paternity, and each of the four methods of quantifying 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, *Mustela erminea*, 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In all reproductive success models we included the metric of 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 (similar results on the lack of effects of these variables on reproductive success were found by (O'Shea et al. 2018)). Four measures of reproductive success were analysed as follows: A) 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 therefore excluded from all analyses of reproductive success; B) Number of within-pair fledglings was a Poisson variable, but there were a large number of zero values. We therefore modelled it with a glmmTMB model that accounts for zero inflation. In addition, because of the large number of zero values, we ran a model with the binary dependent variable of whether or not any offspring were fledged from the nest to compare males with successful versus unsuccessful nests; C) Average mass of within-pair offspring on day 15. Offspring biomass is an important determinant of fitness (Tinbergen and Boerlijst 1990), although this variable excludes the large number of subjects (N=18) whose nests failed entirely before day 15 (we had no evidence that these failures were caused by predation (e.g. broken eggs or predated remains), and instead were likely due to either parental abandonment or natural death of all offspring; note that the latter two causes are not readily distinguished), which of course is a severe fitness cost. 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); D) 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 and nest as a random effect.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 then repeated the above analyses including all nestlings (i.e., both within-pair and extra-pair) to determine if there was evidence for any benefit in the social nest of the tendency to perform dear enemy behaviour. We did not reanalyse clutch size because the analyses above also included all eggs, both within- and extra-pair, because we did not determine paternity for all eggs.

5. Finally, we tested for a relationship between dear enemy behaviour and extra-pair paternity using the proportion of extra-pair young at the nest as the dependent variable (binomial; general linear model), and, in separate models, each measure of dear enemy behaviour as a factor.

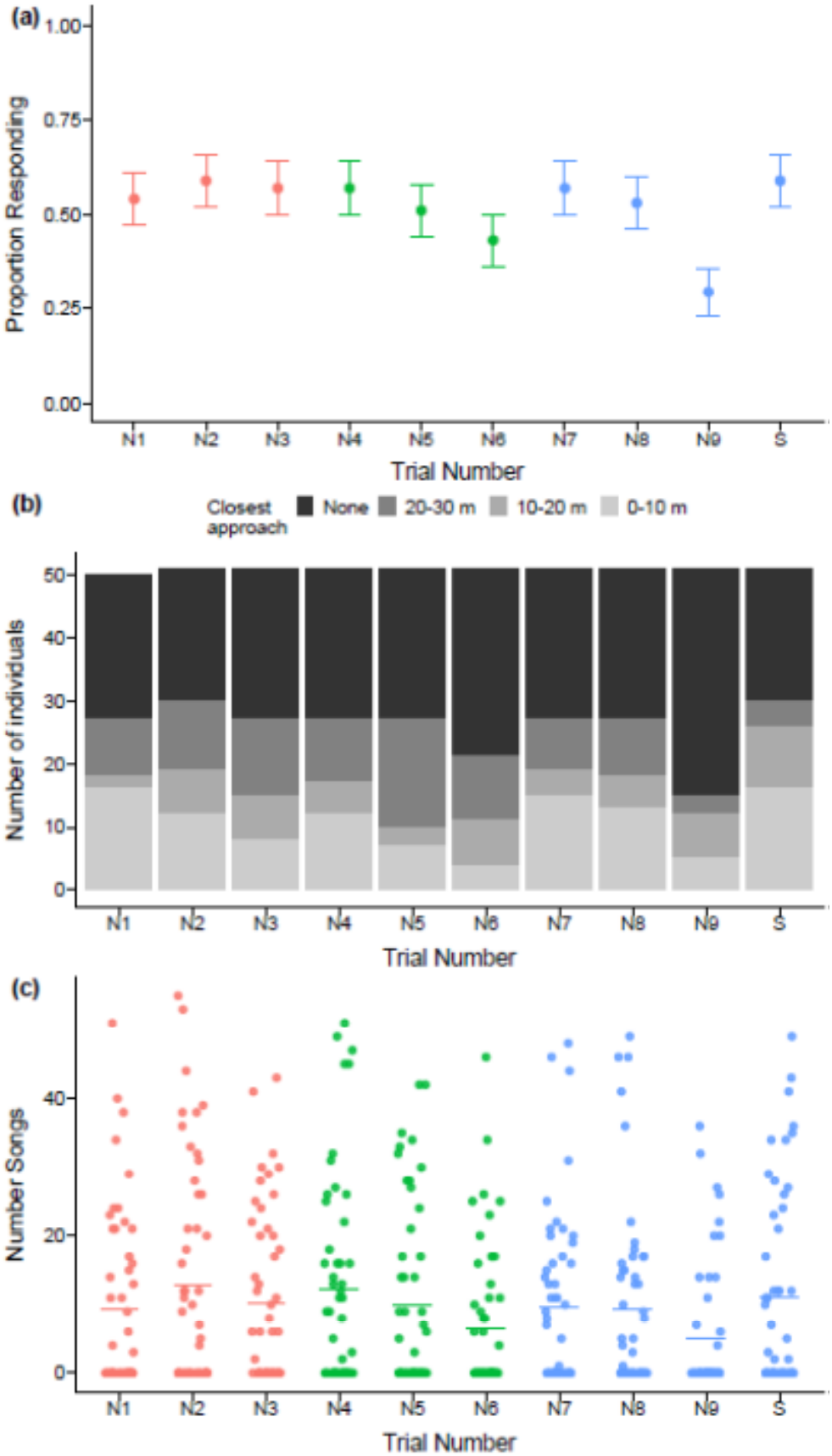
bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RESULTS

Change in response to the neighbour playback

The likelihood of a response (singing and/or moving towards the speaker) significantly decreased across the neighbour playback trials (Table 1; Fig. 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. Closest approach 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; Fig. 2b). Likewise, the number of songs produced by males during the playback decreased across trials (Table 1; Fig. 2c). There was no effect of stimulus rate on response or closest approach, 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).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



579

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 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 (red = day 1, green = day 2, blue = day 3). (b) The closest approach, in categories of 10 m (note that in the statistical analyses we used categories of 5 m, but for ease of visualization we use broader categories here). None corresponds to trials in which the bird did not move at all towards the playback speaker. (c) 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1 Tests of change in response to neighbour playback across trials. Output is from a generalized linear mixed model for response (binary, yes or no), a cumulative link mixed model for closest approach (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 (estimate = 0). Trial number was entered as a numerical variable, with values from one (first trial with neighbour stimulus) to nine (last trial with neighbour stimulus). *N* = 51 individuals

Variable	Factor	Estimate (SE)	z	P
Response	Intercept	0.53 (0.27)		
	Trial number	−0.10 (0.04)	−2.52	0.01
	Rate (high)	0.04 (0.28)	0.13	0.89
Closest approach	Trial number	0.11 (0.04)	3.10	0.002
	Rate (high)	−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 (high)	0.23 (0.12)	1.89	0.058

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 (Fig. 2A; Table 2). However, there was a significant difference in both the closest approach to the speaker and in the 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 (Fig. 2b, c; Table 2). Thus, the decline in the response was specific to the neighbour's song (stimulus), which suggests dear enemy discrimination rather than a general decrease in aggression. Stimulus rate did not affect the responses when only these two trials were considered (Table 2).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2 The effects of stimulus specificity and stimulus rate on response to neighbour playback. Output is from models as in Table 1, but now only including the response to the final neighbour playback and the stranger playback; estimate shown for the stranger trial (10) and the reference category (estimate = 0) is the final neighbour playback (trial 9). The parameter estimate for the high stimulus rate is shown and the reference category is the low rate. *N* = 51 individuals

Dependent variable	Factor	Estimate (SE)	z	P
Response	Intercept	−2.54 (2.51)		
	Stimulus (stranger)	3.59 (3.33)	1.08	0.28
	Stimulus rate (high)	−0.12 (1.42)	−0.08	0.94
Closest approach	Stimulus (stranger)	−2.09 (0.56)	−3.73	<0.001
	Stimulus rate (high)	0.13 (0.79)	0.17	0.87
Number of songs	Intercept	−0.04 (0.74)		
	Stimulus (stranger)	0.30 (0.08)	3.56	<0.001
	Stimulus rate (high)	0.04 (0.92)	0.04	0.97

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Repeatability and context dependence of dear enemy behaviour

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 (Table S2; 15 of 30 when non-responders were removed i.e. standard criterion without non-responders, 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). There was a trend for significant repeatability in the standard criterion for 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.

Whether an individual met the standard criterion for dear enemy behaviour (whether including or excluding non-responders) did not depend on the date of testing, the date relative to the day on which the first egg was laid, the year, the average number of songs produced by the subject across the neighbour playbacks, or the stimulus rate (Table S3). These variables also did not affect the slope of the response or number of songs across trials (Table S3). 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; standard criterion without non-responders: −0.47±0.97, z = −0.48, P = 0.63, N = 18). However, there was an effect of age on the slope of number of songs across trials, with older birds having a less negative slope than first year

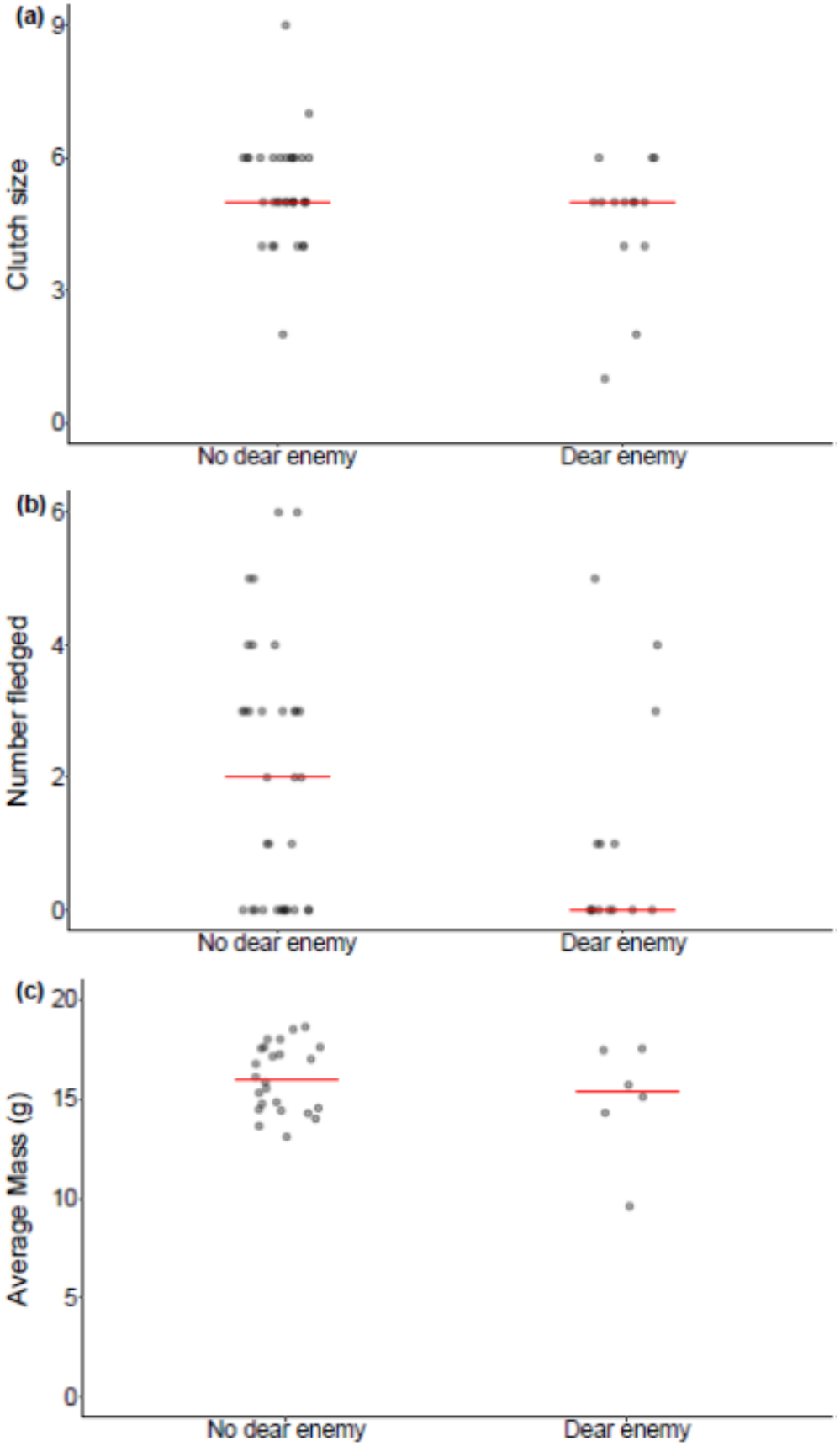
bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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).

Reproductive success and dear enemy behaviour

There was no effect of whether an individual exhibited dear enemy behaviour (standard criterion) on the number of eggs laid in its nest (Fig. 3a; Table S4). Sixteen offspring were identified as extra-pair young and were excluded from the analyses of number of offspring fledged and offspring mass. There was also no effect of whether an individual exhibited dear enemy behaviour (standard criterion) on the number of within-pair offspring that fledged (Fig. 3b; Table S4), or on the binary variable of whether or not any offspring were fledged (Table S4). 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 (Fig. 3c; Table S4) or on the individual mass of within-pair offspring (Table S4). When the standard criterion without non-responders, the slope of responses, or the slope of the number of songs across trials was used as the response variable, there were again no significant effects on any of the reproductive success measures (Table S4). When we repeated the above analyses but this time including data from both within-pair and extra-pair young, we again found no relationship between dear enemy behaviour and any measure of reproductive success (Table S5). Finally, there was no difference between individuals that did or did not show dear enemy behaviour (standard criterion) in the proportion of fledglings that were extrapair (Estimate = 0.20 ± 0.71, P = 0.78, N = 29 nests).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



664

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 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) Average mass of within-pair offspring at day 15 post hatch (not including nests for which no offspring survived to day 15). N = 47 individuals in each graph, except for (c) where N = 30

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DISCUSSION

Male great tits in our study populations exhibited an overall decline in song response and approach towards the neighbour playback stimulus 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. This is consistent with dear enemy behaviour because at the end of the trials, individuals were less aggressive towards the neighbour than the stranger. Our experiments show that neighbour-stranger discrimination can arise and be expressed in only three days, although this was not the case for all individuals. Against predictions, there was no fitness benefit for individuals that exhibited dear enemy behaviour, and no reproductive benefit of any kind that we could detect. Below we discuss these results in the context of territorial aggression and the evolution of neighbour-stranger discrimination.

Learning mechanisms and individual variation

Although complex cognitive mechanisms such as 'true' individual recognition (Tibbetts and Dale 2007) are sometimes involved in dear enemy behaviour, and more complex associations between neighbours may arise, perhaps requiring long-term memory, such as coalitions or associations outside the breeding season (McGregor and Avery 1986; Grabowska-Zhang et al. 2012a; Firth and Sheldon 2016), in many species habituation learning of some characteristic of neighbours is a major contributor to the reduced aggression towards those individuals that is characteristic of the dear enemy effect (Petrinovich and Peeke 1973; Brooks and Falls 1975; Peeke 1984; Bee and Gerhardt 2001; Dong and Clayton 2009). We showed that at least some individuals distinguished simulated neighbours from strangers, and that the

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pattern of response change across trials matches many of the characteristics of habituation (Thompson and Spencer 1966; Rankin et al. 2009).

First, the response to repeated playback of a simulated neighbour decreased across trials (Table 1). 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 (Fig. 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 (Thompson and Spencer 1966; Rankin et al. 2009). Although our nine testing sessions were conducted across just 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. Note that the dear enemy effect does not necessarily require that aggression towards neighbours is eliminated entirely, and indeed increased responsiveness to neighbours at the beginning of the day seems to be common in birds (Briefer et al. 2008; Foote et al. 2008) and may serve to maintain the stability of the territorial boundary. Third, the dual-process theory of habituation argues that in addition to the decrement in response caused by habituation, an independent process of sensitization results in an initial, but transient, increase in aggressive response (Groves and Thompson 1970; Petrinovich and Patterson 1982). Consistent with sensitization, responses tended to be weaker on the first playback trial compared to the subsequent few trials. Importantly, the expression of habituation can be context-dependent through the formation of associations between the habituating stimulus (e.g. the neighbour playback song) and a context (e.g. the location at which the neighbour playback was performed) (Jordan et al. 2000; Uribe-Bahamonde et al. 2019); thus, although our experimental design did not involve independent manipulations of stimulus location and song type, the association between these two variables does not rule out a habituation mechanism for the stimulus-specific reduction in aggression we observed.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Against habituation being the mechanism underlying the observed decline in response, habituation was not stronger at higher song repetition rate, as might have been expected. However, most effects of stimulation interval on habituation have been demonstrated based on variation in the intervals between successive playbacks (rather than between successive songs within a playback), which were approximately the same for the birds in our study, so this prediction may not apply (Groves and Thompson 1970; Thompson et al. 1973; Bee 2001). Taken together, our results suggest that habituation learning is one of the key mechanisms behind the initial stages of dear enemy behaviour in great tits, which was expected, given the general importance of habituation in territorial aggression in songbirds and its specific well-established role in dear enemy recognition (Peeke 1984; Ydenberg et al. 1988; Dong and Clayton 2009). Nevertheless, natural dear enemy relationships in these species almost certainly involve other cognitive mechanisms as well and elucidating the relative contributions of these mechanisms to individual variation in dear enemy behaviour is a worthy target of future research. Furthermore, it is possible that different individuals of the same species use different cognitive mechanisms for individual recognition, and this could have influenced their pattern of response and whether or not they exhibited dear enemy behaviour in this study (Gokcekus et al. 2021).

Non-cognitive mechanisms could also explain the variation in dear enemy behaviour we observed. Differences in average aggressiveness can affect response to playback in great tits (Araya-Ajoy and Dingemanse 2014), and could have played a role, although we found no effect of the average number of songs in response to the neighbour playbacks (a plausible measure of aggressiveness) on whether subjects exhibited dear enemy behaviour. Variation in other personality traits such as explorativeness is also known to relate to variation in habituation speed in great tits (Rivera-Gutierrez et al. 2017) and could therefore have driven some of the variation observed in our study. Given that we did our

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

experiment in a wild population, males could have differed in experience that influenced their response to playbacks. Specifically, males varied in the extent to which they had already interacted with different territorial neighbours in the past. Great tits with more prior exposure to songs of many neighbours are slower to learn to recognize new neighbours (McGregor and Avery 1986), and indeed we found that older birds showed a slower decrement in response and number of songs across the trials than did first year birds, although this did not result in differences between the age groups in whether they differed in response to the final neighbour playback versus the stranger playback. Factors such as density and progress of the breeding season often explain variation in the dear enemy effect (Hyman 2005; Pratt and McLain 2006; Yoon et al. 2012), including in great tits (Jin et al. 2021). However, we found no evidence for this in our study: there was no effect of year, date, or date relative to the start of hatching on whether individuals expressed the dear enemy effect. Whatever the exact mechanisms involved, the individual repeatability in dear enemy behaviour we report indicates that we captured intrinsic differences among individuals in how they discriminated between familiar neighbours and strangers.

We found behaviour consistent with the dear enemy effect because over time individuals responded less to the stimulus simulating the establishment of a new neighbour than to a stimulus that simulated an unknown stranger. It could be objected that subjects did not actually recognize the playback as a true neighbour, and the other as a different stranger bird, and that therefore the stimulus-specific response decrement we observed is not the same as the dear enemy effect. In the strict sense, even playback designs using songs of true neighbours do not necessarily overcome this objection, because even demonstrating a stimulus- and location-specific response to a true neighbour's songs does not give definitive evidence that this stimulus was recognized as the territorial neighbour, rather than being a stimulus that the individual had simply habituated to. We justify our approach and interpretation because 1) we were interested in individual variation in the cognitive processes involved in establishing

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

dear enemy relationships, which cannot be studied with established neighbours. Given the frequency of habituation as a mechanism for dear enemy recognition and the very rapid settlement of territorial relations and habituation to song in great tits (Krebs et al. 1981; Peeke 1984; Rivera-Gutierrez et al. 2017), our design was appropriate. 2) Although no playback can ever replicate the interactions that take place between two birds, we replicated the essential features of territorial neighbours – an increasingly familiar song from the same location on an adjacent territory – and those of strangers – an unfamiliar song from a different location. Song familiarity and location are likely the two most important elements in dear enemy behaviour in great tits, and the behaviour we observed was consistent with the expression of the dear enemy effect. 3) Our simulated neighbour approach reduces some of the confounding effects associated with using playbacks of an established neighbour such as individual variation in experience with that neighbour (Falls et al. 1982; Grabowska-Zhang et al. 2012a). Nevertheless, additional studies of dear enemy behaviour in this species that examine its expression over longer time periods and incorporating additional stimulus dimensions and long-term observations of interacting individuals would be valuable.

Dear enemy recognition and fitness

We hypothesized that if the dear enemy effect is adaptive because it reduces time spent in unnecessary aggressive interactions or leads to other beneficial interactions with neighbours (Getty 1987; Temeles 1994), then individuals that reduce the aggression directed towards their neighbours more quickly would have higher reproductive benefits leading to increased reproductive fitness. There was good reason to expect these predictions to be met in our system because male great tits are known to trade off foraging and territory defence (Ydenberg 1984). We reiterate that our experiment was based on the premise that individual variation in the reduction in aggression to the neighbour stimulus and

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

subsequent response to the stranger stimulus served as a proxy for individual variation in the tendency to form dear enemy relationships in natural interactions, and it was the latter that we proposed to drive any fitness effects. Our hypothesis was not supported because there was no difference in clutch size, average biomass, or number of offspring that fledged between individuals that did or did not express the dear enemy effect. This was true in terms of the measure most closely related to male fitness (when extra-pair paternity was excluded) but it was also true for the overall reproductive success at the nest (when extra-pair young were included), for which detectable effects should be most apparent. In a field experiment of this nature on a species with a complex multi-stage reproductive strategy, inevitably many processes could explain this result, and we highlight those that we suggest particularly deserve attention in future research.

First, 'enemies may not always be dear' (Muller and Manser 2007; Courvoisier et al. 2014; Moser-Purdy et al. 2017), and there may be costs associated with reducing aggression. For example it may lead to prospective individuals settling, leading to reduced territory size and quality (Getty 1989), or to reduced rates of provisioning offspring by the males (Sillett et al. 2004). Reduced aggression by males could have also led to the cost of their mates engaging in extrapair copulations with neighbours, though we found no evidence of this, nor did a previous study of song sparrows, *Melospiza melodia* (Krippel et al. 2017). We note, however, that reduced aggression could also have led to reduced extra-pair paternity if it facilitated mate guarding, though we also found no evidence for this. Second, the lack of any observed benefit to dear enemy behaviour could be because it traded off against other fitness-related behaviours. For instance, in great tits, a positive association between problem solving performance and both clutch size and number of fledglings was masked by a positive association with nest desertion leading to no net reproductive benefits (Cole et al. 2012). We observed high levels of nest failure due to either abandonment or natural death of offspring among birds that expressed dear enemy behaviour, which

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

may be an indication of such a cost, although the likelihood of nest failure was not significantly different from birds that did not express dear enemy behaviour. However, perhaps a more likely reason for no observed effect is that our fitness measures were for a single breeding season, but typically selection varies over time (Siepielski et al. 2009). Similarly, we examined the initial stages of formation of dear enemy relationships and it may be that variation in the strength and specificity of relationships with neighbours over the course of the whole breeding season is more consequential for fitness, although these temporal dynamics and their relationship with fitness have not been explored. Third, although our measures of dear enemy behaviour had a high (albeit non-significant) repeatability coefficient, suggesting the possibility of intrinsic differences among individuals, in reality our measure was likely to have been influenced by other factors—for example temporary environmental effects—making it more difficult to identify links between dear enemy behaviour and variation in fitness (Zsebők et al. 2017). And finally, alterations to the experimental design, such as using a taxidermic mount to simulate visual cues (Ritschard et al. 2012; Araya-Ajoy et al. 2016) or carrying out the playbacks for more than three days, may have revealed additional individual variation that is more directly linked to dear enemy behaviour.

Dear enemy behaviours are undoubtedly adaptive and may well be under current selection in many systems, especially in species such as the great tit where territorial behaviour is important (Krebs 1971; Falls et al. 1982; McGregor and Avery 1986; Grabowska-Zhang et al. 2012b). Despite the intrinsic differences among individuals in their tendency to reduce the response to a familiar stimulus and then respond to a novel stimulus, none of these behaviours showed any positive or negative association with a variety of life history traits. However, revealing hypothetical links on behavioural traits such as the dear enemy effect that are likely underpinned by a variety of mechanisms, including cognitive mechanisms that themselves influence other functional behaviours, is likely to be more challenging than

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for behaviours more closely related to reproduction, and to require a multivariate and longer-term approach across different ecological conditions (Morand-Ferron et al. 2016).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgments 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. Three anonymous reviewers provided helpful comments on a previous draft of this article.

Author contributions MSR and JLQ conceived of the project, MSR, JMSC and JON performed the playbacks, MSR, GLD, IGK, JMC and JON collected breeding data, ED, CS and KvO did the DNA extractions and analyses of paternity, MSR analyzed the data, and MSR wrote the manuscript with input from all authors. Note that middle authors on this manuscript are listed in alphabetical order by surname.

Data availability statement Raw data associated with this project are available at https://doi.org/10.6084/m9.figshare.17148338.v1.

Funding Support for MSR, GLD, IGK, JON, CS, JMSC 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.

Conflict of interest The authors declare no competing interests.

Ethical Approval 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

Adams ES (2001) Approaches to the study of territory size and shape. Annu Rev Ecol Syst 32:277–303. https://doi.org/10.1146/annurev.ecolsys.32.081501.114034

Akçay Ç, Campbell SE, Beecher MD (2014) Individual differences affect honest signalling in a songbird. Proc R Soc Lond B 281:20132496

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. Proc R Soc Lond B 277:3685–3692. https://doi.org/10.1098/rspb.2010.0598

Araya-Ajoy YG, Dingemanse NJ (2014) Characterizing behavioural "characters": an evolutionary framework. Proc R Soc Lond B 281:20132645. https://doi.org/10.1098/rspb.2013.2645

Araya-Ajoy YG, Kuhn S, Mathot KJ, Mouchet A, Mutzel A, Nicolaus M, Wijmenga JJ, Kempenaers B, Dingemanse NJ (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 BM, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48. https://doi.org/10.18637/jss.v067.i01

Battiston MM, Wilson DR, Graham BA, Kovach KA, Mennill DJ (2015) Rufous-and-white wrens *Thryophilus rufalbus* do not exhibit a dear enemy effects towards conspecific or heterospecific competitors. Curr Zool 61:23–33. https://doi.org/10.1093/czoolo/61.1.23

Bee MA (2001) Habituation and sensitization in bullfrogs (*Rana catesbiana*): testing the dual-process theory of habituation. J Comp Psychol 115:307–316

Bee MA, Gerhardt HC (2001) Habituation as a mechanism of reduced aggression between neighboring territorial male bullfrogs (*Rana catesbeiana*). J Comp Psychol 115:68–82. https://doi.org/10.1037/0735-7036.115.1.68

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bee MA, Reichert MS, Tumulty JP (2016) Assessment and recognition of rivals in anuran contests. Adv Stud Behav 48:161–249. https://doi.org/10.1016/bs.asb.2016.01.001

Beletsky LD, Orians GH (1989) Familiar neighbors enhance breeding success in birds. P Natl Acad Sci USA 86:7933–7936. https://doi.org/10.1073/pnas.86.20.7933

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783. https://doi.org/10.1016/j.anbehav.2008.12.022

Blumenrath SH, Dabelsteen T (2004) Degradation of great tit (*Parus major*) song before and after foliation: Implications for vocal communication in a deciduous forest. Behaviour 141:935–958

Blurton Jones NG (1968) Observations and experiments on causation of threat displays of the great tit (*Parus major*). Anim Behav Monogr 1:73–158. https://doi.org/10.1016/S0066-1856(68)80002-0

Briefer E, Rybak F, Aubin T (2008) When to be a dear enemy: flexible acoustic relationships of neighbouring skylarks, *Alauda arvensis*. Anim Behav 76:1319–1325. https://doi.org/10.1016/j.anbehav.2008.06.017

Brindley EL (1991) Response of European robins to playback of song: neighbour recognition and overlapping. Anim Behav 41:503–512. https://doi.org/10.1016/S0003-3472(05)80853-X

Brommer JE, Alho JS, Biard C et al (2010) Passerine extrapair mating dynamics: a bayesian modeling approach comparing four species. Am Nat 176:178–187. https://doi.org/10.1086/653660

Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Mächler M, Bolker BM (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J 9:378–400. https://doi.org/10.32614/rj-2017-066

Brooks RJ, Falls JB (1975) Individual recognition by song in white-throated sparrows. I. Discrimination of songs of neighbors and strangers. Can J Zool 53:879–888

Brumm H, Ritschard M (2011) Song amplitude affects territorial aggression of male receivers in chaffinches. Behav Ecol 22:310–316. https://doi.org/10.1093/beheco/arq205

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Brunton DH, 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. Behav Ecol 19:791–798. https://doi.org/10.1093/beheco/arn027

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

Cole EF, Morand-Ferron J, Hinks AE, Quinn JL (2012) Cognitive ability influences reproductive life history variation in the wild. Curr Biol 22:1808–1812. 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*. Anim Behav 90:229–235. https://doi.org/10.1016/j.anbehav.2014.02.001

Dalton B, Settle R, Medley K, Mathis A (2020) When neighbors cheat: a test of the dear enemy phenomenon in southern red-backed salamanders. Behav Ecol Sociobiol 74:56. https://doi.org/10.1007/s00265-020-02838-9

Dong S, Clayton DF (2009) Habituation in songbirds. Neurobiol Learn Mem 92:183–188. https://doi.org/10.1016/j.nlm.2008.09.009

Doutrelant C, Leitao A, Otter K, Lambrechts MM (2000) Effect of blue tit song syntax on great tit territorial resposiveness - An experimental test of the character shift hypothesis. Behav Ecol Sociobiol 48:119–124. https://doi.org/10.1007/s002650000220

East ML, Perrins CM (1988) The effect of nestboxes on breeding populations of birds in broadleaved temperate woodlands. Ibis 130:393–401. 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:e99878. https://doi.org/10.1371/journal.pone.0099878

Falls JB, Brooks RJ (1975) Individual recognition by song in white-throated sparrows. II. Effects of location. Can J Zool 53:1412–1420. https://doi.org/10.1139/z75-210

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Falls JB, Krebs JR, McGregor PK (1982) Song matching in the great tit (*Parus major*): The effect of similarity and familiarity. Anim Behav 30:997–1009. https://doi.org/10.1016/S0003-3472(82)80188-7

Firth JA, Sheldon BC (2016) Social carry-over effects underpin trans-seasonally linked structure in a wild bird population. Ecol Lett 19:1324–1332. https://doi.org/10.1111/ele.12669

Firth JA, Verhelst BL, Crates RA, Garroway CJ, Sheldon BC (2018) Spatial, temporal and individual-based differences in nest-site visits and subsequent reproductive success in wild great tits. J Avian Biol 49:e01740. https://doi.org/10.1111/jav.01740

Foote JR, Fitzsimmons LP, Mennill DJ, Ratcliffe LM (2008) Male chickadees match neighbors interactively at dawn: support for the social dynamics hypothesis. Behav Ecol 19:1192–1199. https://doi.org/10.1093/beheco/arn087

Getty T (1987) Dear enemies and the prisoner's dilemma: Why should territorial neighbors form defensive coalitions? Am Zool 27:327–336

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

Gheusi G, Goodall G, Dantzer R (1997) Individually distinctive odours represent individual conspecifics in rats. Anim Behav 53:935–944. https://doi.org/10.1006/anbe.1996.0314

Giraldeau LA, Ydenberg RC (1987) The center-edge effect: The result of a war of attrition between territorial contestants? Auk 104:535–538. 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

Gokcekus S, Firth JA, Regan C, Sheldon BC (2021) Recognising the key role of individual recognition in social networks. Trends Ecol Evol 36:1024–1035. https://doi.org/10.1016/j.tree.2021.06.009

Gompertz T (1961) The vocabulary of the Great Tit. Brit Birds 54:369–399

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gosler AG (1993) The great tit. Hamlyn, London

Grabowska-Zhang AM, Sheldon BC, Hinde CA (2012a) Long-term familiarity promotes joining in neighbour nest defence. Biol Lett 8:544–546. https://doi.org/10.1098/rsbl.2012.0183

Grabowska-Zhang AM, Wilkin TA, Sheldon BC (2012b) Effects of neighbor familiarity on reproductive success in the great tit (*Parus major*). Behav Ecol 23:322–333. https://doi.org/10.1093/beheco/arr189

Groves PM, Thompson RF (1970) Habituation: A dual-process theory. Psychol Rev 77:419–450

Hill CE, Akçay Ç, Campbell SE, Beecher MD (2011) Extrapair paternity, song, and genetic quality in song sparrows. Behav Ecol 22:73–81. https://doi.org/10.1093/beheco/arq171

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

Husak JF, Fox SF (2003) Adult male collared lizards, *Crotaphytus collaris*, increase aggression towards displaced neighbours. Anim Behav 65:391–396. https://doi.org/10.1006/anbe.2003.2058

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

Hyman J, Hughes M, Searcy WA, Nowicki S (2004) Individual variation in the strength of territory defense in male song sparrows: Correlates of age, territory tenure, and neighbor aggressiveness. Behaviour 141:15–27. https://doi.org/10.1163/156853904772746574

Jacobs CGC, 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. Anim Behav 90:65–71. https://doi.org/10.1016/j.anbehav.2014.01.016

Jin L, Liang J, Fan Q, Yu J, Sun K, Wang H (2021) Male great tits (*Parus major*) adjust dear enemy effect expression in different breeding stages. J Ornithol 162:221–229. https://doi.org/10.1007/s10336-020-01815-3

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Johnston RE, Jernigan P (1994) Golden hamsters recognize individuals, not just individual scents. Anim Behav 48:129–136. https://doi.org/10.1006/anbe.1994.1219

Jordan WP, Strasser HC, McHale L (2000) Contextual control of long-term habituation in rats. J Exp Psychol Anim B 26:323–339. https://doi.org/10.1037/0097-7403.26.3.323

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x

Kawano KM (2003) Isolation of polymorphic microsatellite markers in the great tit (*Parus major minor*). Mol Ecol Notes 3:314–315. https://doi.org/10.1046/j.1471-8286.2003.00431.x

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

Krebs JR (1976) Habituation and song repertoires in the great tit. Behav Ecol Sociobiol 1:215–227. https://doi.org/10.1007/BF00299198

Krebs JR (1977) Song and territory in the great tit *Parus major*. In: Stonehouse B, Perrins C (eds) Evolutionary Ecology. Palgrave, London, pp 47–62

Krebs JR (1982) Territorial defence in the great tit (*Parus major*): Do residents always win? Behav Ecol Sociobiol 11:185–194. https://doi.org/doi.org/10.1007/BF00300061

Krebs JR, Ashcroft R, van Orsdol K (1981) Song matching in the great tit *Parus major* L. Anim Behav 29:918–923

Krippel J, Ballentine B, Hyman J (2017) Reproductive consequences of aggression in a territorial songbird. Ethology 123:261–269. https://doi.org/10.1111/eth.12588

Langemann U, Gauger B, Klump GM (1998) Auditory sensitivity in the great tit: perception of signals in the presence and absence of noise. Anim Behav 56:763–769. https://doi.org/10.1006/anbe.1998.0879

Leiser JK, Itzkowitz M (1999) The benefits of dear enemy recognition in three-contender convict cichlid

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(*Cichlasoma nigrofasciatum*) contests. Behaviour 136:983–1003. https://doi.org/10.1163/156853999501685

Lovell SF, Lein MR (2005) Individual recognition of neighbors by song in a suboscine bird, the alder flycatcher *Empidonax alnorum*. Behav Ecol Sociobiol 57:623–630. https://doi.org/10.1007/s00265-004-0884-4

Lubjuhn T, Strohbach S, Bruen J, Gerken T, Epplen JT (1999) Extra-pair paternity in great tits (*Parus major*): A long term study. Behaviour 136:1157–1172

Luther DA, Danner R, Danner J, Gentry K, Derryberry EP (2016) The relative response of songbirds to shifts in song amplitude and song minimum frequency. Behav Ecol 28:391-397. https://doi.org/10.1093/beheco/arw172

Mace R (1987) The dawn chorus in the great tit *Parus major* is directly related to female fertility. Nature 330:745–746. https://doi.org/10.1038/330745a0

Mares R, Young AJ, Clutton-Brock TH (2012) Individual contributions to territory defence in a cooperative breeder: weighing up the benefits and costs. Proc R Soc Lond B 279:3989–3995. https://doi.org/10.1098/rspb.2012.1071

Martin TE (1987) Food as a limit on breeding birds: a life-history perspective. Annu Rev Ecol Syst 18:453–487

McGregor PK, Avery MI (1986) The unsung songs of great tits (*Parus major*): learning neighbours' songs for discrimination. Behav Ecol Sociobiol 18:311–316. https://doi.org/10.1007/BF00300009

McGregor PK, Krebs JR (1982a) Song types in a population of great tits (*Parus major*): Their distribution, abundance and acquisition by individuals. Behaviour 79:126–152

McGregor PK, Krebs JR (1982b) Mating and song types in the great tit. Nature 297:60–61. https://doi.org/10.1038/297060a0

Mockford EJ, Marshall RC, Dabelsteen T (2011) Degradation of rural and urban great tit song: Testing

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

transmission efficiency. PLoS ONE 6: e28242. https://doi.org/10.1371/journal.pone.0028242

Morand-Ferron J, Cole EF, Quinn JL (2016) Studying the evolutionary ecology of cognition in the wild: a review of practical and conceptual challenges. Biol Rev 91:367–389. https://doi.org/10.1111/brv.12174

Moser-Purdy C, Macdougall-Shackleton EA, Mennill DJ (2017) Enemies are not always dear: male song sparrows adjust dear enemy effect expression in response to female fertility. Anim Behav 126:17–22. https://doi.org/10.1016/j.anbehav.2017.01.009

Muller CA, Manser MB (2007) “Nasty neighbours” rather than “dear enemies” in a social carnivore. Proc R Soc Lond B 274:959–965

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

Nelson DA, Soha JA (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 JL (2018) Breeding phenology, provisioning behaviour, and unusual patterns of life history variation across an anthropogenic heterogeneous landscape. Oecologia 188:953–964. https://doi.org/10.1007/s00442-018-4155-x

Patrick SC, Chapman JR, Dugdale HL, Quinn JL, Sheldon BC (2012) Promiscuity, paternity and personality in the great tit. Proc R Soc Lond B 279:1724–1730. https://doi.org/10.1098/rspb.2011.1820

Peake TM, Terry AM, McGregor PK, Dabelsteen T (2001) Male great tits eavesdrop on simulated male-to-male vocal interactions. Proc R Soc Lond B 268:1183–1187. https://doi.org/10.1098/rspb.2001.1648

Peeke HVS (1984) Habituation and the maintenance of territorial boundaries. In: Peeke HVS, Petrinovich L (eds) Habituation, sensitization, and behavior. Academic Press, New York, pp 393–421

Petrinovich L, Patterson TL (1982) Field studies of habituation: V. Evidence for a two-factor, dual-process

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

system. J Comp Physiol Psychol 96:284–296. https://doi.org/10.1037/h0077868

Petrinovich L, Patterson T, Peeke HVS (1976) Reproductive condition and the response of white-crowned sparrows (*Zonotrichia leucophrys nuttalli*) to song. Science 191:206–207. https://doi.org/10.1126/science.191.4223.206

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

Pratt AE, McLain DK (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, Vienna, Austria, http://www.R-project.org

Rankin CH, Abrams T, Barry RJ et al (2009) Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. Neurobiol Learn Mem 92:135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Reichert MS, Quinn JL (2017) Cognition in contests: Mechanisms, ecology, and evolution. Trends Ecol Evol 32:773–785. https://doi.org/10.1016/j.tree.2017.07.003

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

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 118:197–202. https://doi.org/10.1111/j.1439-0310.2011.01999.x

Rivera-Gutierrez HF, Martens T, Pinxten R, Eens M (2017) Learning speed is affected by personality and reproductive investment in a songbird. PLoS ONE 12:e0185410. https://doi.org/10.1371/journal.pone.0185410

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rivera-Gutierrez HF, Pinxten R, Eens M (2015) Songbirds never forget: long-lasting behavioural change triggered by a single playback event. Behaviour 152:1277–1290. https://doi.org/10.1163/1568539X-00003278

Saeki T, Sogawa S, Hotta T, Kohda M (2018) Territorial fish distinguish familiar neighbours individually. Behaviour 155:279–293. https://doi.org/10.1163/1568539X-00003489

Saladin V, Bonfils D, Binz T, Richner H (2003) Isolation and characterization of 16 microsatellite loci in the Great Tit *Parus major*. Mol Ecol Notes 3:520–522. 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. J Anim Ecol 84:518–531. https://doi.org/10.1111/1365-2656.12293

Siepielski AM, DiBattista JD, Carlson SM (2009) It's about time: the temporal dynamics of phenotypic selection in the wild. Ecol Lett 12:1261–1276. https://doi.org/10.1111/j.1461-0248.2009.01381.x

Sillett TS, Rodenhouse NL, Holmes RT (2004) Experimentally reducing neighbor density affects reproduction and behavior of a migratory songbird. Ecology 85:2467–2477. https://doi.org/10.1890/03-0272

Siracusa ER, Boutin S, Dantzer B, Lane JE, Coltman DW, McAdam AG (2021) Familiar neighbors, but not relatives, enhance fitness in a territorial mammal. Curr Biol 31:438–445. https://doi.org/10.1016/j.cub.2020.10.072

Snijders L, van Oers K, Naguib M (2017) Sex-specific responses to territorial intrusions in a communication network: Evidence from radio-tagged great tits. Ecol Evol 7:918–927. https://doi.org/10.1002/ece3.2686

Stamps J (1994) Territorial behavior: Testing the assumptions. Adv Stud Behav 23:173–232. https://doi.org/10.1016/S0065-3454(08)60354-X

Stoddard PK (1996) Vocal recognition of neighbors by territorial passerines. In: Kroodsma DE, Miller EH

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(eds) Ecology and Evolution of Acoustic Communication in Birds. Cornell University Press, Ithaca, NY, pp 356–374

Stoddard PK, Beecher MD, Horning CL, Campbell SE (1991) Recognition of individual neighbors by song in the song sparrow, a species with song repertoires. Behav Ecol Sociobiol 29:211–215. https://doi.org/10.1007/BF00166403

Temeles E (1994) The role of neighbors in territorial systems: when are they “dear enemies”? Anim Behav 47:339–350. https://doi.org/10.1006/anbe.1994.1047

Thompson RF, Groves PM, Teyler TJ, Roemer RA (1973) A dual-process theory of habituation: theory and behavior. In: Peeke HVS, Herz MJ (eds) Habituation I: Behavioral Studies. Academic Press, New York, pp 239–271

Thompson RF, Spencer WA (1966) Habituation: A model phenomenon for the study of neuronal substrates of behavior. Psychol Rev 73:16–43

Tibbetts EA, Dale J (2007) Individual recognition: it is good to be different. Trends Ecol Evol 22:529–537. https://doi.org/10.1016/j.tree.2007.09.001

Tinbergen J, Boerlijst M (1990) Nestling weight and survival in individual great tits (*Parus major*). J Anim Ecol 59:1113–1127. https://doi.org/10.2307/5035

Tumulty JP (2018) Dear enemy effect. In: Vonk J, Shackelford TK (eds) Encyclopedia of Animal Cognition and Behavior. Springer, Cham, https://doi.org/10.1007/978-3-319-47829-6_693-1

Tumulty JP, Sheehan MJ (2020) What drives diversity in social recognition mechanisms? Front Ecol Evol 7:517. https://doi.org/10.3389/fevo.2019.00517

Uribe-Bahamonde YE, Becerra SA, Ponce FP, Vogel EH (2019) A quantitative account of the behavioral characteristics of habituation: The Sometimes Opponent Processes model of stimulus processing. Front Psychol 10:504. https://doi.org/10.3389/fpsyg.2019.00504

van Oers K, Drent PJ, Dingemanse NJ, Kempenaers B (2008) Personality is associated with extrapair

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

paternity in great tits, *Parus major*. Anim Behav 76:555–563. https://doi.org/10.1016/j.anbehav.2008.03.011

Verbeek MEM, Boon A, Drent PJ (1996) Exploration, aggressive behaviour and dominance in pair-wise confrontations of juvenile male great tits. Behaviour 133:945–963. https://doi.org/10.1017/CBO9781107415324.004

Weary DM, Krebs JR (1992) Great tits classify songs by individual voice characteristics. Anim Behav 43:283–287. https://doi.org/10.1016/S0003-3472(05)80223-4

Webster MS, Pruett-Jones S, Westneat DF, Arnold SJ (1995) Measuring the effects of pairing success, extra-pair copulations and mate quality on the opportunity for sexual selection. Evolution 49:1147–1157. 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. J Ornithol 152:431–438. https://doi.org/10.1007/s10336-010-0609-6

Werba JA, Stuckert AM, Edwards M, McCoy MW (2021) Stranger Danger: A meta-analysis of the dear enemy hypothesis. Behav Process 194:104542. https://doi.org/10.1016/j.beproc.2021.104542

Wiley RH (2013) Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. Biol Rev 88:179–195. https://doi.org/10.1111/j.1469-185X.2012.00246.x

Wischhoff U, Marques-Santos F, Manica LT, Roper JJ, Rodrigues M (2018) Parenting styles in white-rumped swallows (*Tachycineta leucorrhoa*) show a trade-off between nest defense and chick feeding. Ethology 124:623–632. https://doi.org/10.1111/eth.12770

Yasukawa K (1981) Song repertoires in the red-winged blackbird (*Agelaius phoeniceus*) - A test of the Beau-Geste hypothesis. Anim Behav 29:114–125

Ydenberg RC (1984) The conflict between feeding and territorial defence in the great tit. Behav Ecol

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sociobiol 15:103–108. https://doi.org/10.1007/BF00299376

Ydenberg RC, Giraldeau LA, Falls JB (1988) Neighbours, strangers, and the asymmetric war of attrition. Anim Behav 36:343–347

Ydenberg RC, Krebs JR (1987) The tradeoff between territorial defence and foraging in the great tit (*Parus major*). Am Zool 27:337–346. https://doi.org/10.1093/icb/27.2.337

Yoon J, Sillett TS, Morrison SA, Ghalambor CK (2012) Breeding density, not life history, predicts interpopulation differences in territorial aggression in a passerine bird. Anim Behav 84:515–521. https://doi.org/10.1016/j.anbehav.2012.05.024

Zsebők S, Herczeg G, Blázi G, Laczi M, Nagy G, Szász E, Markó G, Török J, Garamszegi LZ (2017) Short- and long-term repeatability and pseudo-repeatability of bird song: sensitivity of signals to varying environments. Behav Ecol Sociobiol 71:154. https://doi.org/10.1007/s00265-017-2379-0

Table S1. Study site locations and summary statistics.

Site Name	Coordinates	N DE behaviour/N tested	Clutch Size	N Fledged	Average Mass (g)
Ballinphellic (BP)	51.840546, −8.630297	0/1	6	5	16.8
Castle Bernard (CB)	51.741568, −8.767775	2/9	5.0 (1.3)	1.4 (1.8)	14.4 (3.1)
Dukes Wood (DW)	51.785965, −8.751366	2/9	4.4 (1.1)	1.7 (1.7)	16.6 (1.9)
Dunderrow (DD)	51.719555, −8.600512	0/4*	5	0	15.3
Garrettstown (GT)	51.655159, −8.616456	0/1	6	6	18.0
Innishannon (IN)	51.763415, −8.664943	2/9	5.9 (1.5)	1.0 (1.6)	15.4 (1.6)
Kilbrittain (KB)	51.671290, −8.682011	8/16**	4.9 (1.3)	2.2 (2.0)	15.9 (1.6)
Shippool (SP)	51.737780, −8.630792	1/2	4.5 (0.7)	0.5 (0.7)	17.6

Site names and coordinates as in O'Shea 2017. See also Figure S1. Values for the reproductive success variables are means (±standard deviation, not reported when only one individual was sampled at a site). Clutch size includes all eggs whether or not they were identified as extrapair; number fledged and average mass were calculated excluding any young identified as extrapair. 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 individuals exhibiting dear enemy behaviour under the standard criterion.

*Three of these individuals at Dunderrow were excluded from the analyses of reproductive success variables because of widespread stoat predation of nests at the site in 2018. The numbers reported here are from the one nest that was included.
**One individual from Kilbrittain that had no eggs in the nest was excluded from the calculations of reproductive success variables.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

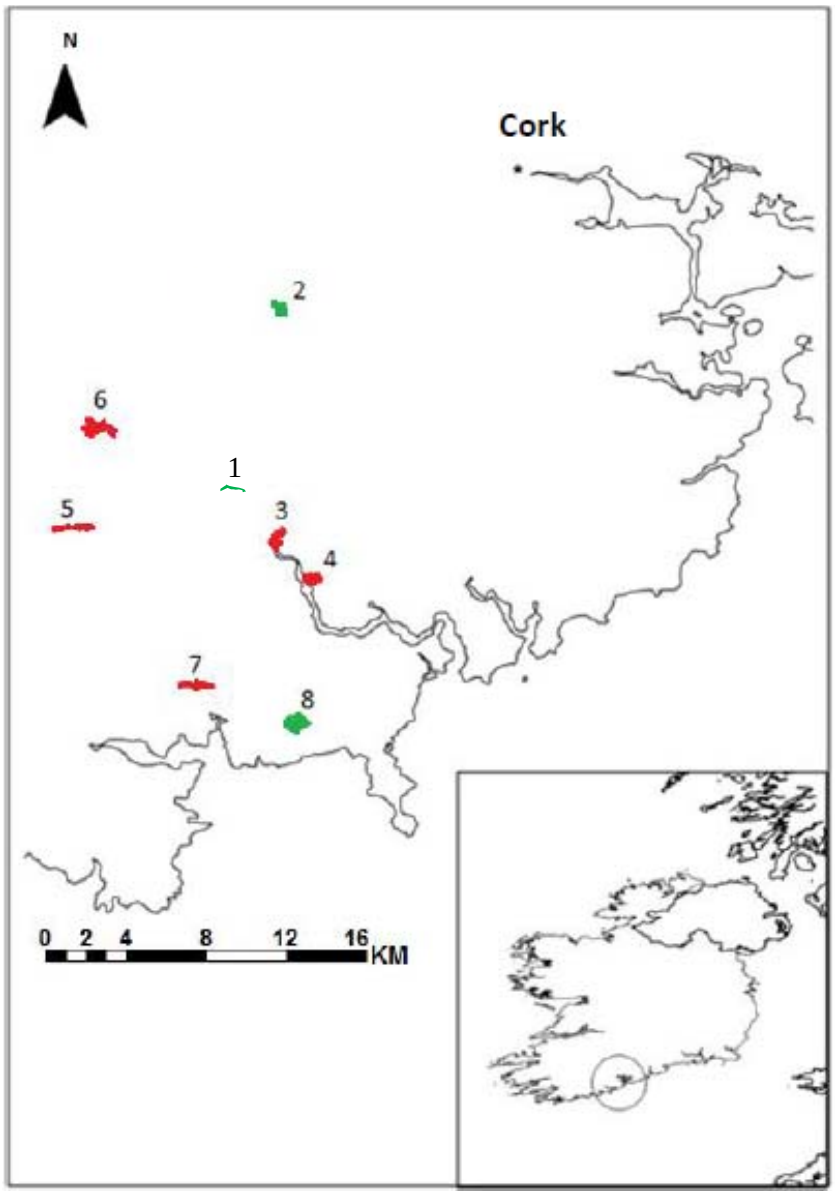


Figure S1. Map of the study sites, modified from O'Shea (2017). Numbers correspond to the individual sites (see also Table S1): 1. Innishannon, 2. Ballinphellic, 3. Shippool, 4. Dunderrow, 5. Castle Bernard, 6. Dukes Wood, 7. Kilbrittain, 8. Garrettstown.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S2. Response on last neighbour trial versus response on stranger trial

	Responded stranger trial	Did not respond stranger trial
Responded last neighbour trial	15	0
Did not respond last neighbour trial	15	21

The table shows the number of individuals for each combination of responding or not responding to the last neighbour trial and to the subsequent stranger trial.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S3. Effects of context on dear enemy behaviour.

Dear enemy criterion	Factor	Estimate	SE	Test statistic	P
Standard criterion	Intercept	1049	1668	0.63	0.53
	Year	−0.52	0.83	−0.63	0.53
	Date	0.07	0.05	1.33	0.18
	Stimulus rate (high)	0.58	0.70	0.82	0.41
	Date relative to first egg	−0.01	0.06	−0.18	0.86
	Average song number	−0.07	0.06	−1.13	0.26
Standard without nonresponders	Intercept	5701	3478	1.64	0.10
	Year	−2.83	1.73	−1.64	0.10
	Date	0.15	0.10	1.52	0.13
	Stimulus rate (high)	0.87	1.08	0.81	0.42
	Date relative to first egg	−0.02	0.08	−0.26	0.80
	Average song number	−0.25	0.12	−2.09	0.04
Response slope	Intercept	12.37	73.49	0.17	0.87
	Year	−0.01	0.04	−0.17	0.87
	Date	0.001	0.002	0.46	0.65
	Stimulus rate (high)	0.01	0.03	0.29	0.77
	Date relative to first egg	0.002	0.002	1.05	0.3
	Average song number	0.002	0.002	1.00	0.33
Song slope	Intercept	485	2262	0.21	0.83
	Year	−0.24	1.12	−0.22	0.83
	Date	0.08	0.07	1.13	0.27
	Stimulus rate (high)	−0.70	0.81	−0.86	0.40
	Date relative to first egg	0.01	0.06	0.12	0.91
	Average song number	0.05	0.05	0.96	0.35

Results are from a generalized linear model of whether or not the individual met the standard criterion for expressing dear enemy behaviour (where the test statistic is *z*), or from linear models for the slope variables (where the test statistic is *t*). The reference category for stimulus rate is the low rate stimulus. The variable "average song number" refers to the average number of songs produced by the bird over the nine total neighbour trials. Date is coded as the day number of the year on which the first playback was performed for that bird, with 1 March being day 1. Date relative to first egg refers to which day the first playback was performed for that bird relative to when the first egg was laid in the nest (so day zero would indicate that the first playback was performed on the same day that the first egg was laid in the nest). $N = 50$ individuals for the standard criterion (one individual not included because there was never an egg laid at the nest), 29 individuals for standard without nonresponders, and 25 individuals for the slope variables.

1208 **Table S4**. Full models of dear enemy behaviour and reproductive success, including brood size effects when included in models.
1209

Reproductive measure	Dear enemy behaviour measure	Estimate(SE)	test statistic	P	N
Clutch size	Standard criterion	−0.14 (0.15)	−0.98	0.33	47
	Standard without nonresponders	−0.15 (0.17)	−0.85	0.40	28
	Response slope	0.59 (1.29)	0.46	0.65	28
	Song slope	0.008 (0.043)	0.19	0.85	28
Number fledged	Standard criterion	−0.42 (0.35)	−1.22	0.22	47
	Standard without nonresponders	−0.33 (0.37)	−0.90	0.37	28
	Response slope	3.12 (2.13)	1.47	0.14	28
	Song slope	0.04 (0.07)	0.58	0.56	28
Fledged offspring yes/no	Standard criterion	−0.85 (0.65)	−1.30	0.19	47
	Standard without nonresponders	−0.58 (0.76)	−0.75	0.45	28
	Response slope	−1.48 (5.73)	−0.26	0.80	28
	Song slope	0.04 (0.19)	0.19	0.85	28
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	95/30
	Brood size	0.20 (0.26)	0.75	0.46	
	Standard without nonresponders	0.77 (0.99)	0.78	0.45	49/16
	Brood size	0.38 (0.31)	1.24	0.24	
	Response slope	−7.00 (8.81)	−0.79	0.44	49/16
	Brood size	0.57 (0.44)	1.28	0.22	
	Song slope	0.01 (0.30)	0.05	0.96	49/16
	Brood size	0.29 (0.47)	0.62	0.55	

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Output is from models with the reproductive measure as the dependent variable (all measures except for clutch size 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 (for the analyses of individual offspring mass they refer to the number of individual offspring/number of subject males) 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S5. Full models of dear enemy behaviour and reproductive success. Unlike Table S4, these analyses include both within-pair and extra-pair young.

Reproductive measure	Dear enemy behaviour measure	Estimate(SE)	test statistic	P	N
Number fledged	Standard criterion	−0.45 (0.30)	−1.50	0.13	47
	Standard without nonresponders	−0.36 (0.33)	−1.08	0.28	28
	Response slope	2.75 (2.02)	1.36	0.17	28
	Song slope	0.06 (0.07)	0.80	0.42	28
Fledged offspring yes/no	Standard criterion	−0.69 (0.65)	−1.07	0.29	47
	Standard without nonresponders	−0.59 (0.77)	−0.76	0.45	28
	Response slope	−2.75 (5.84)	−0.47	0.64	28
	Song slope	0.04 (0.19)	0.23	0.82	28
Average offspring biomass (g)	Standard criterion	−0.77 (0.86)	−0.89	0.38	31
	Brood size	0.31 (0.26)	1.19	0.25	
	Standard without nonresponders	0.69 (1.00)	0.69	0.50	17
	Brood size	0.47 (0.31)	1.52	0.15	
	Response slope	−5.89 (8.62)	−0.68	0.51	17
	Brood size	0.57 (0.40)	1.45	0.17	
	Song slope	0.11 (0.29)	0.38	0.71	17
	Brood size	0.28 (0.41)	0.67	0.51	
Individual offspring mass (g)	Standard criterion	−0.70 (0.83)	−0.85	0.41	111/31
	Brood size	0.22 (0.25)	0.89	0.38	
	Standard without nonresponders	0.75 (0.91)	0.83	0.42	59/17
	Brood size	0.38 (0.29)	1.33	0.21	
	Response slope	−5.68 (7.99)	−0.71	0.49	59/17
	Brood size	0.48 (0.38)	1.26	0.23	
	Song slope	0.06 (0.26)	0.23	0.82	59/17
	Brood size	0.23 (0.38)	0.59	0.56	

Output is from models with the reproductive measure as the dependent variable (all measures include both extra-pair and within−pair young), and the dear enemy behaviour measure as a factor. Results for clutch size are not reported here because the identical analysis is already

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

presented in Table S4 (we did not remove extra-pair young from the measure of clutch size). 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 (for the analyses of individual offspring mass they refer to the number of individual offspring/number of subject males) 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.28.441816; this version posted June 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary References

O'Shea, W. 2017. The evolutionary ecology of personality and life history variation in an anthropogenic heterogeneous landscape. PhD Thesis, University College Cork.