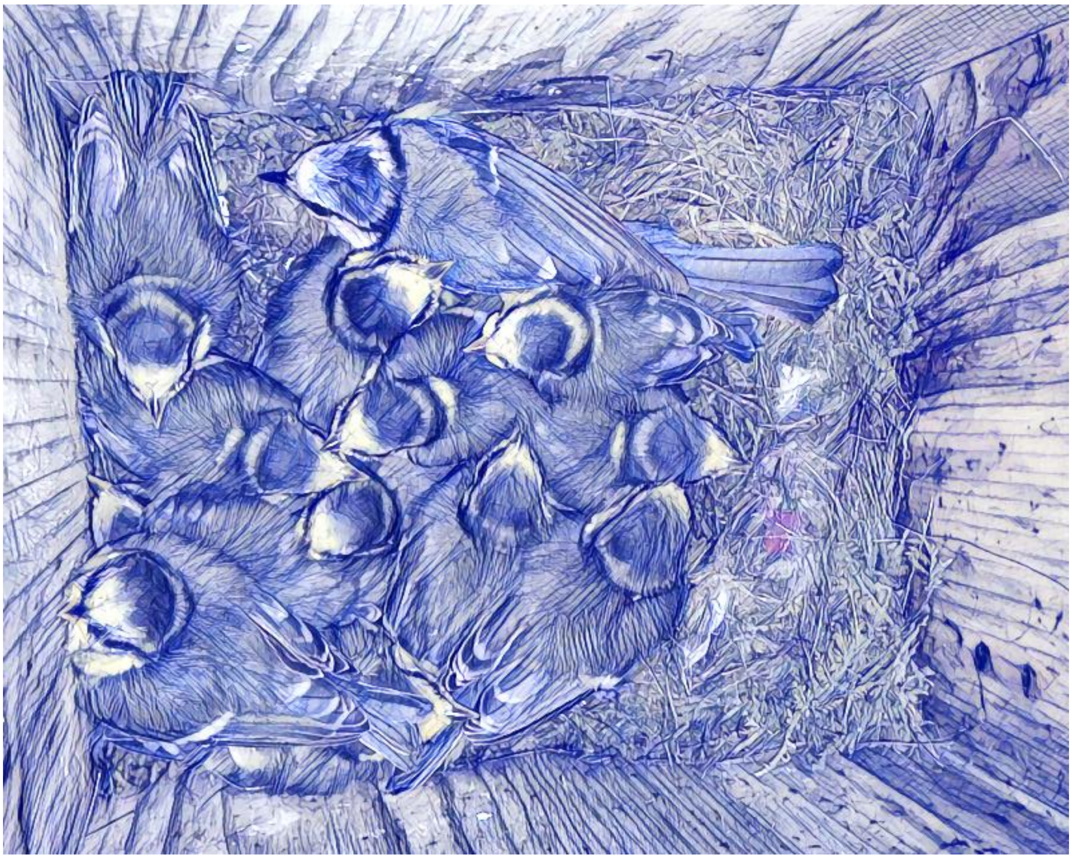
Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades „Doctor rerum naturalium“ (Dr. rer. nat.) an der Fakultät für Biologie der Ludwig-Maximilians-Universität München.

Social interactions: intra- and inter-specific competition between blue and great tits



Dissertation von

Irene Gaona Gordillo

Printed and/or published with the support of the German Academic Exchange Service.

Photo credit: Irene Gaona Gordillo

Photo editor: BeFunky

Social interactions: intra- and inter-specific competition between blue and great tits

Dissertation
durchgeführt an
der Fakultät für Biologie
der Ludwig-Maximilians-Universität München
und dem
Max-Planck-Institut für Biologische Intelligenz
vorgelegt von
Irene Gaona Gordillo
München, February, 2024

Diese Dissertation wurde angefertigt unter der Leitung von Prof. Dr. Niels Dingemanse im Bereich der Verhaltensökologie.

Erstgutachter: Prof. Dr. Niels Dingemanse

Zweitgutachter: Prof. Dr. Richard Merrill

Tag der Abgabe: February 29, 2024

Tag der mündlichen Prüfung: June 19, 2024

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

München, den 29.02.2024 Irene Gaona Gordillo

(Unterschrift)

Erklärung

Hiermit erkläre ich,

- dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.
- dass ich mich anderweitig einer Doktorprüfung ohne Erfolg nicht unterzogen habe.

München, den 29.02.2024 Irene Gaona Gordillo

(Unterschrift)

Table of contents

PEER-REVIEWED PAPER

Sections with the subtitle “Phenotypic integration among species, populations, and sexes” were taken from the following peer-reviewed published paper:

Gaona-Gordillo I, Holtmann B, Mouchet A, Hutfluss A, Sánchez-Tójar A, Dingemanse NJ. 2023. Are animal personality, body condition, physiology and structural size integrated? A comparison of species, populations and sexes, and the value of study replication. Journal of Animal Ecology. 92(9):1707–1718. doi:10.1111/1365-2656.13966.

List of figures

List of tables

List of supplementary material

Abstract

Among-individual differences in behaviour are ubiquitous, yet only over the past two decades has research focused on understanding how and why among-individual differences are generated and maintained. Conceptual models from an adaptive perspective predict that among-individual differences in internal and external state lead to individual differences in behaviour, and correlations between behaviour and other phenotypic traits. Recent conceptual models imply that fluctuations in population density could promote the maintenance of variation in life-history strategies, behaviour ("animal personality"), and other correlated traits. This thesis focused on testing three predictions stemming from the theory explaining animal personality variation as an adaptation to variation in competitive regimes. First, I tested whether patterns of integration of morphological, physiological, and behavioural traits predicted by eco-evolutionary theory were, as expected, generally supported within two bird species (blue tits *Cyanistes caeruleus* vs. great tits *Parus major*), or whether these patterns differed between study populations (Forstenrieder Park vs. Starnberg), and/or between sexes (males vs. females). I further experimentally manipulated the availability of nest boxes suited for blue and great tit breeders to modify the realized breeding density of both species to test other components of the theory. Specifically, I tested for effects of con- and heterospecific competition on life-history decisions of blue and great tits. Finally, I tested whether manipulated breeding densities affected how the two species socially interacted, by studying how both species in each treatment group modified their aggressive behaviour toward con- and heterospecific intruders. I found that there was general and consistent support for the integration between morphology and physiology at different biological levels. However, unexpected discrepancies were observed in the integration of morphology and behaviour, as well as physiology and behaviour. This suggests that species, populations, and sexes respond differently to environmental

conditions, and thus exhibit different patterns of phenotypic integration. Additionally, while the nest box manipulation effectively altered the realized breeding densities of blue and great tits, the reproductive traits that I examined were largely unaffected. Finally, blue and great tit males adjusted their aggressive response based on whether they were confronted by a con- vs. a heterospecific intruder but changes in the realized breeding density did not affect aggressiveness. In conclusion, my thesis implies that changes in the social environmental state might impose species-specific costs and benefits of displaying a behavioural response, resulting in different patterns of phenotypic integration observed at different biological levels. Thus, predictions from life-history theory regarding the integration of life-history and behaviour in response to competition may need to incorporate relevant environmental and ecological effects, particularly in the context of intra- and interspecific interactions, to better our understanding of phenotypic integration and the evolution of life-history in the wild.

CHAPTER 1

1. Introduction

Animals exhibit an overwhelming diversity in life-history, morphological, physiological, and behavioural traits. Therefore, it is not surprising that variation is present among species, among populations, and among individuals (Wilson 1998a). The importance that natural selection has on shaping among-individual differences has long been recognized (Darwin 1859). Theory predicts that stabilizing selection acts against the extreme values of a phenotype and thus reduces (genetic) variation (Darwin 1859; Fisher 1930; Mather 1955). Nevertheless, variation is, in fact, ubiquitous. Then, how are among-individual differences generated and maintained? Counterintuitively, together with mutation (Lande 1975), migration (Felsenstein 1976; Slarkin 1985), and fluctuating selection (Sasaki and Ellner 1997), natural selection itself has been recognized as a potential mechanism that contributes to the generation and maintenance of among-individual differences. When traits are under disruptive selection, the extreme values of a phenotype are favoured and variation is expected to increase (Mather 1955; Rueffler et al. 2006). Because selection can constrain or facilitate among-individual differences, many studies have focused on understanding the patterns of selection, albeit largely with respect to morphological traits (Kingsolver et al. 2001). Much less attention has been given to understanding the maintenance of individual variation in behavioural traits (Kingsolver et al. 2012).

1.1 Evolution of among-individual differences in behaviour

Early ethologists were interested in describing the stereotypic behaviour that could characterize a particular species (Tinbergen 1963; Gordon 2011; Immelmann 2012). Subsequently, many behavioural ecologists aimed to characterize the average value of a given behaviour in the population, and acknowledged among-individual differences, though primarily as noise (Sih and Bell 2008; Gordon 2011; Dingemanse and Dochtermann 2014). Classic behavioural ecology studies considered variation in behaviour to largely result from “reversible” (within-individual) plasticity (Sih et al. 2004). Only over the last few decades, research has started to ask: why do among-individuals differences in behaviour exist, and they perhaps adaptive, and how is this variation maintained?

1.1.1 Understanding among-individual differences: insights from contemporary studies

Among-individual differences in behaviour are, in contemporary studies, largely viewed to come about in one of three ways. First, individuals can differ in their phenotypic mean over many repeated observations, and this is often referred to as “personality” (Réale et al. 2007). Second, they can vary in their level of plasticity (Nussey et al. 2007). Plasticity here refers to the range of phenotypic responses that a single genotype expresses when exposed to different environmental conditions (West-Eberhard 1989). Both components can be quantified using a “reaction norm” framework (Dingemanse et al. 2010). This framework statistically models the range of phenotypic behavioural responses that an individual displays over an environmental gradient by assigning each individual a reaction norm intercept and slope value. When slope values differ between individuals, this is called Individual by Environment interaction or “I×E” (Nussey et al. 2007). I×E can be underpinned by genetic differences in plasticity, called Gene – Environment interaction or G×E (Nussey et al. 2007). This genetic component of plasticity can thus be viewed as a heritable trait and, therefore, can evolve in response to selection

(Scheiner 2002). Third, individuals can differ in their residual within-individual variance (i.e., within-individual deviations from the individual's reaction norm, Westneat et al. 2015). In the past decades, individual differences in within-individual variance were largely ignored because it was, as above, considered random variation, or otherwise assumed to not exist (Biro and Adriaenssens 2013). Nevertheless, heterogeneous residual variance has recently been described as the major source of phenotypic variation (Westneat et al. 2015). These three components of the individual phenotype are, furthermore, either varying independently or in a correlated manner. For example, an individual's behavioural type (reaction norm intercept) and plasticity (reaction norm slope) can be correlated (called "intercept-slope interaction"). In addition, both can covary with level of behavioural predictability (i.e., within-individual residual variance; Biro and Adriaenssens 2013; Westneat et al. 2015) (Figure 1).

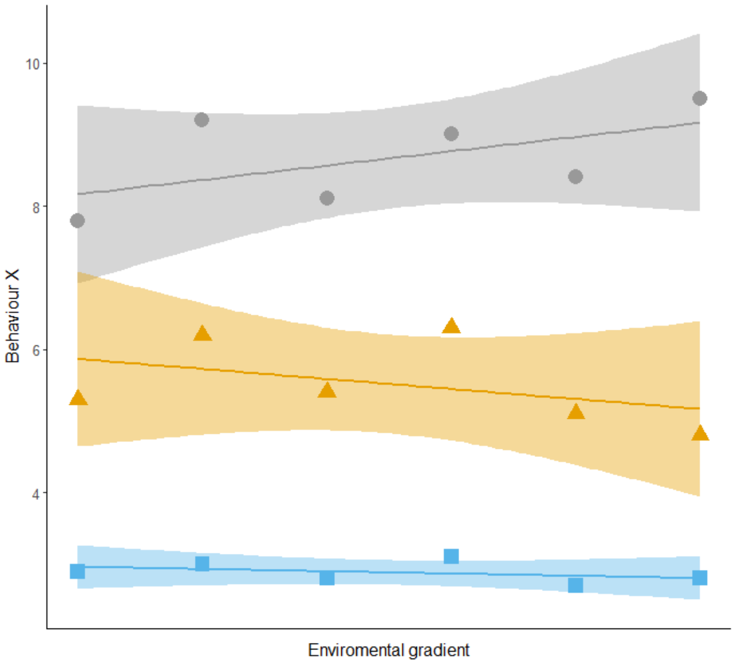


Figure 1. Illustration of the three ways by which behaviour can vary among individuals. The three individuals (different colours) differ in their mean expression (intercept; "personality" variation) of the behaviour X (behavioural type) (y-axis) over an environmental gradient (x-axis) and show individual differences in plasticity (as coloured lines have different slopes). The illustration also depicts individual differences in the behavioural predictability (coloured ribbons): the blue individual is more predictable (i.e., exhibits less residual variance) compared to the orange and grey individuals. Figure inspired by Dingemanse and Wolf (2010).

1.1.1.1 Trait correlations, syndromes and residual variances

Individual differences can also be correlated across functionally distinct behavioural traits, and so can level of plasticity, or residual variance. For example, an individual's mean expression of a behaviour might be correlated with its mean expression for another behaviour (called a "behavioural syndrome"; Sih et al. 2004). Likewise, there can be within-individual correlations between phenotypic traits (i.e., plasticity syndromes; O'Dea et al. 2022; Sheehy and Laskowski 2023) and residual variances (i.e., residual "predictability") among traits. Within-individual correlations occur when a shift in a behaviour from one time point to the next in one behaviour is correlated within that of another (Dingemanse & Dochtermann 2013). Because individuals potentially differ in their behavioural type, plasticity, and behavioural predictability, it should be expected that they could differ in their (behavioural and plasticity) syndromes and residual variances across traits. Over the past decades, there has been a surge of interest in understanding why behavioural syndromes exist (Sih et al. 2004).

1.2 Models for adaptive personality differences

Since the turn of the century, various types of models have been proposed to understand adaptive personality differences. First, state-dependent personality models imply that differences in risk-taking behaviour result from differences in residual reproductive value. Theory predicts that individuals with a relatively fast pace-of-life should invest more heavily in current reproduction (vs. future reproduction) at the cost of reduced survival rates (Houston and McNamara 1999; Wolf et al. 2007); such types of individuals are expected to behave more aggressively and be bolder to achieve this aim (Wolf et al. 2007; Sih et al. 2015). A second type of model incorporates a feedback between the state of the individual and its behaviour. Here the state refers to any feature, both internal (e.g., morphological and physiological traits, Dingemanse and Wolf 2010; Wolf and Weissing 2010) and external (i.e., social and non-social environment, Wolf and Weissing 2010), that affects any behavioural response (Dingemanse and Wolf 2010; Wolf and Weissing 2010). These models postulate that the individual's state can influence what is the optimal behaviour, and that this behaviour can then, in turn, affect its state (Dingemanse and Wolf 2010; Wolf and Weissing 2010).

Over the past decades, it has been proposed that risk-taking behaviours (i.e., behaviours that increase resource acquisition at an increased risk of mortality, predation, or parasitism; Stamps 2007; Wolf et al. 2007; Barber and Dingemanse 2010; Balaban-Feld et al. 2019) could play a major role in mediating life-history trade-offs (Wolf et al. 2007). Theory predicts that bolder phenotypes invest more in current (vs. future) reproduction and, consequently, acquire more resources, but also die young (Wolf et al. 2007; Sih et al. 2015; Laskowski et al. 2021). Even though some studies have described the existence of this link between boldness and pace-of-life, there is accumulating evidence showing conflicting results (Royauté et al. 2018; Moiron et al. 2020). These conflicting results do not suggest that trade-offs between current and future reproduction, hence variation in pace-of-life, do not exist (Laskowski et al. 2021), rather, they demonstrate that variation in pace-of-life might be due to confounds with other factors, particularly ecological conditions (Montiglio et al. 2018).

1.3 Research aims

This thesis centres on the role of the social environment in maintaining variation in life-history and animal personality within and among species. This thesis has three main aims. First, I aimed to test whether patterns of phenotypic integration (Pigliucci 2003) of morphological, physiological, and behavioural traits follow theoretical predictions, thus whether predicted patterns of phenotypic integration between personality and other phenotypic traits are generally supported among and within species, populations, or sexes. My thesis views the social environmental state as a keystone component shaping and maintaining among-individual differences in life-history and behaviour (Wolf and Weissing 2010). Thus, my second aim was to evaluate the role of competitive regimes in maintaining variation in life-history traits among and within species (Lande et al. 2009; Wright et al. 2019). My, third, final aim was to investigate whether social environments cause “reversible” within-individual plasticity, thus, whether among-individual and among-species variation causes variation in behaviour. Each of my research aims will further be detailed in the following sections.

1.3.1 Phenotypic integration among species, populations, and sexes

Based on Gaona-Gordillo et al (2023)

Natural selection often favours particular combinations of phenotypic traits (Lande and Arnold 1983; Brodie et al. 1995; Kingsolver et al. 2001). Such correlational forms of selection can ultimately result in the evolution of genetic correlations (Sinervo and Svensson 2002; Roff and Fairbairn 2012). In garter snakes *Thamnophis ordinoides*, for example, colour pattern and anti-predator behaviour are genetically correlated because individuals with different colour patterns can best avoid predation by also exhibiting a different behaviour when under attack (Brodie 1992; Brodie 1993). Correlational selection thereby may produce optimal trait combinations that organisms require for executing particular functions (Pigliucci and Preston 2004; Klingenberg 2008, 2014). Level of predation risk, intensity of competition for resources, or other ecological conditions specific to a species, population, or sex will thus ultimately shape the patterns of selection observed in nature (Siepielski et al. 2009; Kingsolver et al. 2012). Species-, population-, or sex-specific differences in ecology are therefore expected to result in species-, population-, or sex-specific patterns of phenotypic trait integration (Armbruster and Schwaegerle 1996; Roff and Fairbairn 2012; Armbruster et al. 2014).

Comparing patterns of phenotypic integration among species, populations, and sexes is of general importance if we are to understand whether patterns observed in specific studies are replicable vs. study-specific (Wilson 1998b). In behavioural ecology, for example, there is growing awareness that conclusions regarding the replicability of findings can only be indirectly assessed through the study of publication bias (e.g., Jennions and Møller 2002; Yang et al. 2023) due to lack of study replication, specially ‘exact replication’ (Kelly 2006, 2019; Nakagawa and Parker 2015; Filazzola and Cahill 2021). Here, we take up this challenge. Our objective is to understand whether results are study- (or dataset-) specific vs. replicable across studies (or datasets). We focus on complex patterns of behavioural, physiological, and morphological integration that have previously been predicted by adaptive animal personality theory (Dingemanse and Wolf 2010; Wolf and Weissing 2010; Sih et al. 2015) and subsequently demonstrated empirically for a specific sex (males) of a specific bird species (the great tit *Parus major*; hereafter called “GT”) of a specific population (Starnberg, Bavaria, Germany) (Moiron et al. 2019). We investigate here whether the same structure of phenotypic integration characterizes both sexes (females and males) within the same population of GTs, and whether any patterns of (sex-specific) phenotypic integration also characterize the same and another species (the blue tit *Cyanistes caeruleus*; hereafter called “BT”) studied in another population (the Forstenrieder Park, Munich, Bavaria, Germany). We note that we did not aim to test for effects of *a priori* hypothesized ecological differences between species, population, and/or sexes; this would require many replicates of populations/species differing in ecology (Kelly 2006). We instead aimed to use any observed differences in phenotypic integration as an opportunity to generate ecological hypotheses to be tested by future research.

A major question is why individuals differ in behaviour, and why behavioural traits are often (genetically) correlated with other phenotypic traits (Dall et al. 2004; Wolf and Weissing 2010). This question is of importance as (behavioural) genetic correlations can impose major constraints on micro-evolutionary processes (Dochtermann and Dingemanse 2013). Our replication study, in part, addresses the role of behaviour in phenotypic integration. Particularly, we focused on a risk-taking behaviour (Dall et al. 2004; Sih et al. 2004; Réale et al. 2007; Wolf and Weissing 2010). Risk-taking behaviours are considered as those that facilitate resource acquisition at the expense of increased risk of mortality, predation, or parasitism (Stamps 2007; Wolf et al. 2007; Barber and Dingemanse 2010). Examples are aggressiveness, anti-predator boldness, neophilia, and speed of exploration of novel environments and objects. Meta-analyses have revealed that populations generally exhibit repeatable individual variation in such behaviours (Bell et al. 2009; Holtmann et al. 2017), that different types of risk-taking behaviours are positively correlated among individuals (Garamszegi et al. 2012), and that they are underpinned by tight genetic correlations (Dochtermann 2011).

Behavioural ecologists have thus developed a suite of models seeking to explain when repeatable among-individual variance (or “animal personality”) in risk-taking behaviour may evolve (Dingemanse and Wolf 2010; Wolf and Weissing 2010). Early explanations centred on selection favouring alternative life-histories associated with their specific behavioural adaptations, particularly in the context of pace-of-life: risk-takers would live fast but die young (Stamps 2007; Wolf et al. 2007; Réale et al. 2010). Contemporary explanations reserve a more central role for ecological variation (Dammhahn et al. 2018; Montiglio et al. 2018; Wright et al. 2019; Mouchet et al. 2021). Fluctuating selection induced by ecological variability has come to the foreground, in part, because the mere existence of life-history trade-offs cannot explain the maintenance of variation (Stearns 1992). A specifically appealing explanation is that fast (vs. slow) life-histories are favoured when a focal population is below (vs. at) carrying capacity; variation in the intensity of competition should consequently induce correlational selection favouring the adaptive integration of life-history, morphology, behaviour, and their physiological underpinnings (Wright et al. 2019). The idea is that selection for early reproduction, or large clutch sizes per reproductive attempt (Araya-Ajoy et al. 2018), is favoured when competition is relaxed. Under such conditions, large risk-takers in particular would be able to both monopolize and exploit the available resources required to produce relatively large clutches (Wright et al. 2019). A key component of this “fast” lifestyle is a decreased investment in self-maintenance, which would ultimately result in a shorter lifespan or earlier onset of reproductive senescence (Moiron et al. 2020).

The general prediction that risk-takers are heavier and/or larger has been confirmed by a recent meta-analysis (Niemelä and Dingemanse 2018). Previous work on GTs fully aligned with this meta-analytic result: among males, larger individuals were both heavier, and more explorative and aggressive (Moiron et al. 2019). Simultaneously, the more explorative and aggressive individuals—though larger (thus heavier)—were relatively lean: their body mass was relatively low for their size (Moiron et al. 2019). These findings make sense, as such relatively lean individuals do poorly when breeding densities are high and competition for resources is intense (Both et al. 1999). In the same population that Moiron et al. (2019) studied, the more explorative GTs also produced larger clutches per reproductive attempt, and showed an earlier onset of reproductive senescence (Araya-Ajoy et al. 2016; Dingemanse, Moiron, et al. 2020). Altogether, these findings suggest, first, that GTs with a faster pace-of-life trade-off investments in self-maintenance towards current reproduction and, second, that a suite of morphological, behavioural, and life-history traits are integrated as part of pace-of-life syndromes.

Various empirical studies imply that physiological traits are also integrated into such behaviour-morphology syndromes. For example, in GTs, the less explorative (Carere and Van Oers 2004), and in BTs, the less aggressive (Class and Brommer 2020), individuals have higher breathing rates. This particular behaviour-physiology syndrome is thought to exist because less (vs. more) explorative birds have a reactive (vs. proactive) stress physiology (Groothuis and Carere 2005; Coppens et al. 2010), of which breathing rate is a phenotypic indicator (Carere and Van Oers 2004). The stability of this integration between risk-taking behaviours and physiology is, however, debated as behaviour-physiology correlations may be labile and vary with age, sex, or environmental context (Krams et al. 2014; Class and Brommer 2015; Kluen et al. 2022). This lack of consensus calls for studies investigating whether the integration of behaviour, physiology, and morphology is replicable over—rather than specific to—species, populations, or other characteristics of the study model.

We tackle this question by testing whether individual BT and GT males and females of the same and different populations share the same pattern of phenotypic integration. We repeatedly measured a suite of behavioural (exploration), physiological (breathing rate), and morphological traits (body mass, tarsus length, wing length, bill length) on all individuals breeding in two populations, and did so for multiple years. We then estimated trait correlations among-individuals separately for each unique combination of species, population, and sex (hereafter called "datasets"). BTs were not studied in the Starnberg population, and we thus compared six (rather than eight) datasets. Following Moiron et al. (2019), we use structural equation models (SEMs) to test for the existence of (a) a latent variable (LV) driving observable expressions of size (body mass, tarsus length, wing length, bill length) and other phenotypic traits (exploration and breathing rate), (b) a size-independent (i.e., size-corrected) syndrome between the non-morphological traits (exploration behaviour and breathing rate), and (c) a size-independent effect of body mass on the non-morphological traits (i.e., an effect of body mass not attributable to size). Next, we made use of meta-analytic approaches to specifically ask which of the structural paths included in the SEMs were overall supported (vs. unsupported), and whether the strength and direction of each focal path was the same (vs. different) between our six datasets (Figure 2).

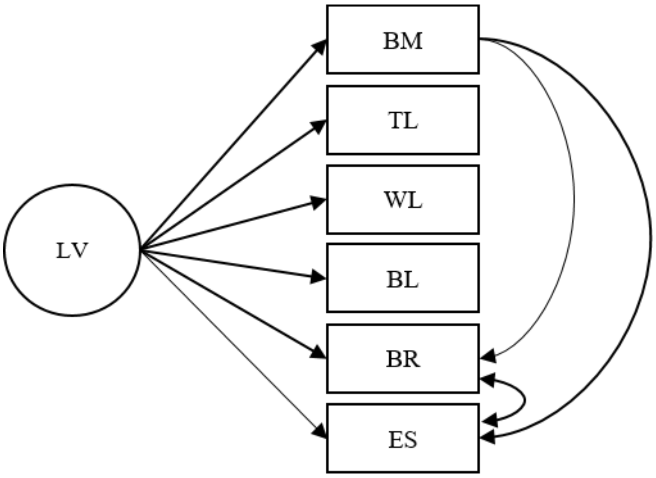


Figure 2. Visualisation of the nine structural paths connecting behaviour, physiology, and morphology in a structural equation model (SEM). The single-headed arrows represent causal relationship between traits; double-headed arrows represent undefined correlations. BM=body mass, TL=tarsus length, WL=wing length, BL=bill length, BR=breathing rate, ES=exploration score, LV= latent variable.

1.3.2 Effects of intra- and interspecific competition on reproductive parameters of blue and great tits

Variation in competition for limited resources constitutes a key driver of variation in life-history traits, whether within or among individuals, populations, or species (Stearns 1992; Roff 1993). For example, in many cavity-breeding bird species, individuals plastically down-regulate their clutch size when breeding densities increase across breeding attempts (Both 1998; Both et al. 2000). Variation in competitive regimes simultaneously has been proposed to be a key driver of fluctuating selection on repeatable differences in life-history decisions (such as clutch size) among-individuals of the same population. For instance, individuals that produce relatively small (vs. large) average clutch sizes over their lifetime exhibit highest life-time reproductive success when breeding densities are high (vs. low) (Sæther et al. 2016). Negative density-dependent clutch size regulation may be adaptive when pairs that breed under high breeding densities produce smaller clutches with fewer—yet more competitive—fledglings (Perrins 1965; Both et al. 2000). Finally, repeatable variation in competitive regimes among populations or species, has been proposed to explain among-population or –species variation in suites of life-history traits (pace-of-life, Ricklefs and Wikelski 2002; Réale et al. 2010; Dammhahn et al. 2018; Wright et al. 2019). Competition thus simultaneously plays a key role in maintaining adaptive variation in life-history decisions at multiple levels of the biological hierarchy.

The vast majority of studies investigating how competition shapes life-history decisions, however, focus solely on competition between members of the same species (called intra-specific competition, Klomp 1964; Dhondt et al. 1990; Both 1998; Nicolaus et al. 2009). This is perhaps not surprising as theory developed in community ecology predicts that competition should be stronger within vs. among species that live in sympatry (Tilman 1994; Chesson 2000; Tokeshi 2009). For example, studies on BTs have demonstrated that changes in conspecific breeding density more strongly affect their reproductive traits (e.g., clutch size, nest success, or fledgling quality or quantity) than changes in the breeding density of competitor species, like GTs (Dhondt 1977; Sasvári et al. 1987; Dhondt and Adriaensen 1999; Møller et al. 2018). Closely related species, however, often overlap in resource preferences, suggesting that competition between members of different species (called inter-specific competition, Crombie 1947) should nevertheless play an important role in shaping life-history evolution, particularly when closely-related species live in sympatry. For example, breeding densities of GTs have been demonstrated to causally affect various reproductive parameters in BTs (Dhondt 1977; Dhondt and Adriaensen 1999; Møller et al. 2018). This is because these two species occur in the same habitat, largely share the same breeding and dietary niche, and thus overall strongly compete for access to the same resources (Minot 1981; Minot and Perrins 1986; Gibb 2008). In such circumstances, understanding the overall role of competitive regimes in shaping life-history decisions requires the simultaneous study of intra- and inter-specific competition (Schlyter and Anderbrant 1993; Harvey et al. 2013; Beard et al. 2018).

A complete understanding of the respective roles of intra- vs. inter-specific competition on life-history decisions, however, requires insight not just in how life-history decisions are on average affected by intra- and inter-specific competitive regimes, but also in whether the strength of each type of competition varies with the social environment (Svensson and Sheldon 1998). For example, the magnitude of negative intra-specific density-dependent clutch size regulation should be expected to be higher when densities of competitor species are higher vs. lower. This is because realized species' niches become narrower when competing species live in sympatry (called character displacement, Brown and Wilson 1956). Consequently, changes in conspecific breeding density should more (vs. less) strongly decrease reproductive output when breeding densities of competitor species are relatively high (vs. low). Surprisingly, studies typically manipulate the amount of intra- or inter-specific competition separately rather than simultaneously (Alatalo 1982). To our knowledge, none of the (seven) previous studies that experimentally tested for intra- and inter-specific competition in BTs and GTs used the full factorial experimental design required to achieve this aim (Table S1, Dhondt 1977; Minot 1981; Sasvári et al. 1987; Sasvári and Orell 1992; Dhondt and Adriaensen 1999; Török and Tóth 1999; Dhondt 2010).

Experiments focusing on intra- and inter-specific competition in the BT-GT species pair have primarily utilized manipulations of numbers of nest boxes of different types to change breeding densities and thereby the strength of competition. A common approach has been to provide boxes where GTs can vs. cannot breed, which can be achieved by manipulating the diameter of the nest box entrance hole (Dhondt and Eyckerman 1980; Dhondt and Adriaensen 1999; Török and Tóth 1999; Dhondt 2010). In the current study we also used this approach. Specifically, we produced "low" vs. "high" densities of nest boxes (per ha) of both BT- and GT-boxes in a full factorial experimental design. As such full factorial designs have not been utilized to date (see above and Table S1), interaction effects have also not previously been studied. Furthermore, we used a study design, where each of the four unique combinations of BT and GT nest box density treatments (BT – GT-box density: low-low, low-high, high-low, high-high) were applied thrice each year (using n=12 subplots per year, detailed §2.1.2). Moreover, we implemented the experiment in each of three years but randomly reallocated subplot treatments across years, and treatments thus varied within subplots across years. This represents a powerful design as treatment was therefore not confounded by subplot-specific characteristics (such as habitat).

We had two main aims. Our first aim was to induce experimental variation in breeding densities of the two species. We expected that an experimental increase in the density of BT-boxes would increase the realized breeding density of BTs, and similarly that an experimental increase in the density of GT-boxes would increase the realized breeding density of GTs. Our second aim was to test whether experimentally-induced changes in the species' realized breeding densities affected the species' reproductive parameters, where we focused on lay date, clutch size, nestling body condition and body size, and the number of fledglings. Previous studies have demonstrated that GTs suffer more strongly—in terms of reproductive performance—from competition with BTs than *vice versa* (Dhondt 1977; Minot 1981; Sasvári et al. 1987). BTs have been proposed to suffer less from inter-specific competition with GTs because of their size: the smaller and lighter BTs can reach invertebrates (e.g., caterpillars and spiders) on outer tree branches where the bigger and heavier GTs cannot forage (Betts 1955; Török 1986; García-Navas et al. 2013). We thus expected GT (vs. BT) reproduction to be more strongly affected by the realized breeding densities of either species.

1.3.3 Aggressive behaviour towards con- vs. heterospecifics in blue and great tits

Individuals can modify their phenotype (e.g., morphology, physiology and behaviour; West-Eberhard, 1989) as a function of the environment (i.e., phenotypic plasticity, Pigliucci, 2001; West-Eberhard, 1989). Particularly, individuals can quickly adapt their behaviour to match the current environmental conditions (i.e., "reversible" within-individual plasticity, Duckworth 2009; Dingemanse, Kazem, et al. 2010). These changes in the expression could further benefit the individual by increasing its fitness (Ghalambor et al. 2010). Assuming that the costs and benefits associated with plasticity do not differ between individuals, we would naïvely expect that individuals do not differ in plasticity. When this is not the case, we expect individual differences in plasticity (Dingemanse and Wolf 2013). Indeed, there is accumulating empirical evidence on insects (Brown and Robinson 2016; Royauté et al. 2019), fish (Urszán et al. 2018), birds (Betini and Norris 2012; Dingemanse et al. 2012), and mammals (Hall and Chalfoun 2019) showing differences in "reversible" within-individual plasticity (Nussey et al. 2007; Dingemanse, Kazem, et al. 2010; Stamps 2016).

Various non-exclusive explanations have been proposed for why individuals differ in reversible plasticity. Theoretical work has shown that the benefits of plasticity can be frequency-dependent, and that positive feedback mechanisms could help reduce its costs (Wolf et al. 2008). Individual differences in plasticity are also expected to occur when individuals differ in perceived level of environmental uncertainty, an aspect of environmental state (Mathot et al. 2012). Indeed, adaptive explanations have focused more generally on how the internal state of an individual, or its social or non-social environment, affects its level of plasticity (Dingemanse and Wolf 2013). Despite the benefits associated with being plastic, costs of plasticity also influence whether plasticity evolves (Scheiner 1993; DeWitt et al. 1998; Wolf et al. 2008; Auld et al. 2010; Dingemanse and Wolf 2013; Stamps 2016). Costs include those associated with time and energy required for sampling and the development and maintenance of structures that are necessary for perceiving, analysing, and expressing (DeWitt et al. 1998). Thus, even if plasticity has benefits it might not evolve if the benefits are outweighed by its costs.

Behavioural traits are extremely labile compared to morphological and physiological traits (Duckworth 2009). This is because behaviour can be adjusted very rapidly in response to changes in environmental conditions (Duckworth 2009), and therefore, exhibit high levels of reversible plasticity (Taborsky and Oliveira 2012; Dingemanse and Wolf 2013). These characteristics make behaviour uniquely suitable for the study of among-individual differences in plasticity. In this study, we focused on risk-taking behaviours; behaviours for which associated costs and benefits are relatively well understood. Risk-taking behaviours facilitate resource acquisition at an increased risk of mortality, predation, or parasitism (Stamps 2007; Wolf et al. 2007; Barber and Dingemanse 2010; Balaban-Feld et al. 2019). Here we focus on risk-taking expressed in the context of territory defence, called aggressiveness. Wild individuals show repeatable and heritable variation plasticity for this trait, which makes it an ideal behaviour to study how individuals can adjust their response based on the social environmental state (Araya-Ajoy and Dingemanse 2017). Particularly, we used BTs and GTs males as model systems because both species display a plastic aggressive response toward con- and heterospecifics that has been described to affect fitness (Krebs et al. 1978; Kempenaers et al. 1995; Araya-Ajoy et al. 2016; Araya-Ajoy and Dingemanse 2017).

BTs and GTs live in sympatry, and partly compete for access to the same resources (Gibb 1954; Minot 1981; Minot and Perrins 1986). Indeed, previous studies have reported negative effects of intra- and interspecific competition on their respective fitness (Dhondt 1977; Sasvári et al. 1987; Dhondt and Adriaensen 1999; Møller et al. 2018). Therefore, the ability of a species to recognize and display a specific aggressive response toward heterospecific intruders, rather than just conspecifics, should be important, especially when the species are competing for the same resources (Reed 1982; Matyjasiak 2005). In other words, we expect phenotypic variation in aggressiveness to exist in wild populations whenever the social environment is heterogeneous because individuals interact both with cons- and heterospecifics. In this study, we therefore tested whether males of BTs and GTs adjust their aggressive response based on whether the intruder is a con- vs. heterospecific (Reed 1982; Matyjasiak 2005; Jin et al. 2021).

Individuals may not just modify their aggressiveness in response to intruder species. Rather, the strength of such responses should depend on other components of the social environment. Specifically, previous studies shown that changes in the density cause changes aggressiveness (Metcalfe and Furness 1987; Johnson et al. 2004; Yoon et al. 2012; Araya-Ajoy and Dingemanse 2017). As part of this study, we thus manipulated the number of nest boxes suitable for BTs and GTs to cause experimental variation in breeding densities. This enabled us test whether the strength of intruder species-related plasticity depended on the strength of competition between BTs and GTs.

1.4 Blue and great tits as model species

BTs and GTs have been studied for many decades, and therefore, much is known about their life-history, breeding ecology, as well as their morphological, physiological, and behavioural adaptations. BTs and GTs are closely-related species that largely share breeding and dietary niches (Gibb 1954; Minot 1981; Minot and Perrins 1986). Where the species live in sympatry, they partly compete for access to the same limited resources. This feature makes them ideal study systems to test for intra- and inter-specific competition on reproductive and behaviour. Importantly, in the wild, these secondary cavity-nesters readily breed in nest boxes. Previous studies have described that the reproductive output of BTs and GTs is sensitive to changes in the realized breeding density (Krebs 1971; Dhondt 1977; Dhondt 2010). Consequently, as both species use nest boxes during the breeding season, one can easily manipulate the availability and type of boxes and thus increase/decrease their realized breeding densities experimentally (§1.3.2). Our ability to attract the birds to nest boxes also means that monitoring, behavioural testing, and collection of reproductive data are relatively easily achieved in the wild.

Both females and males of both species feed their nestlings inside the nest boxes (Cowie and Hinsley 1988). This behaviour makes breeding adults easily captured, banded, and sexed, such that parents can be assigned to each brood. Knowing the identity and sex of each adult allowed us to test whether morphology and behaviour differ between individuals recaptured over multiple years. Furthermore, studies that focused on the dispersal rate of BTs and GTs have labelled them as territorial, especially the males (Könczey et al. 1997; Andreu and Barba 2006). Territorial birds display an aggressive response toward conspecifics and heterospecifics, which functions to defend and secure their territory (Krebs et al. 1978; Kempenaers et al. 1995; Araya-Ajoy and Dingemanse 2017). Quantifying these behaviours allowed me to test whether species differ in their levels of aggression towards con- vs. heterospecific intruders and whether different competition regimens can affect this response (§ 1.3.3). Finally, both species are faithful to their territories, which allowed me to collect repeated measures of the same individual over the years. With these repeated measures, I could estimate patterns of integration both within and among individuals (§ 1.3.1).

CHAPTER 2

2. Materials and methods

2.1 General study area

The data were collected in the Forstenrieder Park, Munich, Germany (48° 2' 49" N, 11° 27' 40" E) during the breeding seasons of 2020 through 2022 (except for data described in § 2.1.3, which were only collected in 2021 and 2022). The vegetation of the park consisted of mixed deciduous forest, primarily composed of beech (*Fagus sylvatica*) and oak (*Quercus robur*). The study site contained 420 nest boxes distributed over six study plots. Plots were established during the autumn of 2018 (A, B, C, D; Figure S1) and 2019 (E, F; Figure S1). All boxes were placed in a ~50-m grid, except for 15 pairs of boxes in plots A, B, and F in 2020; plots A, B, and D in 2021; and A, C, and D in 2022, which were placed ~5 m apart from each other (Table S2; see §2.1.2). All boxes had inner margins of 14.6×12×25 cm^3 but differed in the diameter of the entry hole: 32 mm (“GT-boxes”) vs. 26 mm (“BT-boxes”).

2.1.1 Phenotypic integration among species, populations, and sexes

Based on Gaona-Gordillo et al (2023)

For this section we further used data previously collected from a second study site in Bavaria. This study area consisted of 12 plots situated between the Ammersee and Starnbergersee (south of Starnberg, Germany; 47° 58' N, 11°14' E; hereafter called “Starnberg”) that were monitored from 2010 through 2019. Each plot fitted 50 nest boxes in a regular grid with ~50 m between boxes. For general descriptions of this study area, see Nicolaus et al. (2015). In Starnberg, only GT-boxes were used, that is, only data for GT were collected, which explains why our analyses compared phenotypic integration among six (rather than eight) unique combinations of species, populations, and sexes (Figure 3 in §3.1).

2.1.1.1 Data collection

In both study areas, we inspected nest boxes biweekly from the beginning of April to record lay date, clutch size, and incubation date. Fourteen days following incubation onset, we inspected nest boxes daily until egg hatching (day 0). We captured (and ringed if not previously banded) both parents (day 10-12) with a spring trap set inside the nest box and assayed their exploration behaviour in a cage test (see §2.1.1.1.1). Immediately after the exploration test, the bird was removed from the cage and its breathing rate recorded (Holtmann & Dingemanse 2022). To do so, the focal bird was held in a ringer’s grip laying on its back with its breast and belly visible to the observer (Kluen et al. 2014). In both populations, the number of breaths were counted for a 30s-period. In Starnberg, this procedure was repeated a second time but this data is not used here as we did not collect a repeat measure in the Forstenrieder Park. We further measured body mass, tarsus length, wing length, and bill length. In Forstenrieder Park, we measured body mass using a digital scale (±0.01 g) and wing length as maximum chord (Svensson 1992). In Starnberg, we measured body mass using a Pesola spring balance (±0.25 g) and wing length as the feather length of the 3rd outermost primary (Jenni and Winkler 1989).

Previous studies have shown that these two alternative measures of wing length are highly correlated, thus measuring the same trait (Jenni and Winkler 1989). We determined the breeder’s sex based on the presence/absence of a brood patch (both species), and on the size of the black breast band (GTs), or the hue of crown feathers (BTs). We aged breeders (first-year breeder vs. older) using plumage characteristics (Jenni and Winkler 1994). Following all measurements, we released the focal breeder near its nest box.

2.1.1.1.1 Exploration behaviour

Exploration behaviour was assayed in a small cage representing a smaller-scale field version of the classic novel environment test (Dingemanse et al. 2002; for details of the cage setup, see Stuber et al. 2013). Prior to recording exploration behaviour, the focal individual was placed in a side compartment of the cage for 30 s to allow habituation. We then opened the side door of the compartment to release the bird into the cage without handling, and immediately video-recorded its behaviour for two minutes with a camera (Panasonic HC-V100) placed 1 m from the cage while the observer was out of sight. In Starnberg, videos were later scored using JWatcher v.1.0 (Blumstein et al. 2006) whereas in Forstenrieder Park, we used Boris v.7.9.8 (Friard and Gamba 2016). To do so, we divided the cage into 12 sections consisting of six sections on the front mesh, three floor sections, and three perches (detailed in Fig. 1 in Stuber et al. 2013). Finally, we calculated exploration scores as the number of movements made between the 12 sections within the first two minutes. This behaviour is positively correlated with anti-predator boldness (Stuber et al. 2013) and aggression (Moiron et al. 2019), and was thus viewed as a measure of risk-taking behaviour (cf. Carter et al. 2013).

As part of another experiment in Forstenrieder Park, some individuals (n=60 BTs and n=92 GTs) stayed in the side compartment for up to 120 s (year 2020 only). We analysed a subset of randomly allocated individuals within pairs for which one breeding pair member spent 30 s and the other 120 s in the side compartment (n=6 BT pairs and n=7 GT pairs), which showed no statistically significant effect of the amount of time spent in the side compartment (see Text S1).

All procedures were approved by the Regierung von Oberbayern, Bayern, Germany (permission number: ROB-55.2-2532.Vet_02-17-215). All applicable international, national and/or institutional guidelines for the care and use of animals were followed.

2.1.1.2 Data selection and statistical analysis

GTs normally produce second clutches (Tinbergen 1987; Verhulst 1998) but BTs rarely do so (Dhondt 2012). To facilitate comparisons between the two species, we therefore only analysed data from first broods. We defined first broods as those produced within 30 days after the first egg of each species was laid (within each species within a given population) (Nicolaus et al. 2015). Moreover, we only included first broods for which the identity of the female was known as this helped avoid pseudo-replication caused by replacement broods of unknown females (Starnberg: 1664 out of 2222 GT broods; Forstenrieder Park: 280 out of 397 GT and 210 out of 269 BT broods). Furthermore, Moiron et al. (2019) included aggressiveness in their SEM-analyses of phenotypic integration in male GTs, but we did not because aggressiveness was not assayed in females of either population. The lack of this data for one sex meant that we would not be able to execute our aim (formally comparing the same SEMs across the six unique combinations of species, population, and sex; Figure 3 in §3.1).

2.1.1.2.1 Multivariate mixed-effects modelling

We partitioned variances and covariances among individuals, plots, and field observers for each of the six traits by fitting a multivariate mixed-effects model with the six traits (body mass, tarsus length, wing length, bill length, breathing rate, and exploration score) as the response variables. Response variables were scaled (mean-centred and expressed in SD-units) prior to analyses and modelled with a Gaussian error distribution. We performed six models, one for each unique combination of species (BTs vs. GTs), population (Forstenrieder Park vs. Starnberg), and sex (male vs. female). We included data from 334 adult BTs (400 observations) and 402 adult GTs (511 observations) from Forstenrieder Park, and 2123 adult GTs (3226 observations) from Starnberg. Each model included year as a fixed-effect factor (Forstenrieder Park: n=3; Starnberg: n=10) to account for year-specific temporal effects. Random intercepts were fitted for individual identity, plot (Forstenrieder Park: n=6; Starnberg: n=12), and field observer identity (Forstenrieder Park: n=21; Starnberg: n=56). Following visual inspection of the data, we removed three data points for body mass, two for tarsus length, twelve for wing length, and four for bill length of the Forstenrieder Park dataset. Additionally, we removed one data point for bill length of the Starnberg dataset. As detailed in Text S2, those measurements were all outside the natural range of the focal species, and thus considered faulty data rather than extreme phenotypes. All models were run following the Bayesian framework using the function "MCMCglmm" of the MCMCglmm package v.2.34 (Hadfield 2010) in R v.4.1.2 (R Core Team 2020). The models sampled 650 000 iterations, with a burning interval of 150 000, and a thinning interval of 500. For all multivariate mixed-effects models, we specified an inverse-gamma prior (V=diag(6), nu=1.002) for the residuals and a parameter expanded prior (V=diag(6), nu=6, alpha.mu=diag(0, 6), alpha.V=diag(6)*1000) for the random effects.

2.1.1.2.2 Structural Equation Models

We estimated the strength, sign, and standard error (SE) of each of the nine structural paths associated with each SEM (Figure 3 in §3.1). As model input, we used the estimated among-individual correlation matrix obtained from the multivariate mixed-effects models, where the sample size was assumed to equate the number of individuals. We achieved this by using the "sem" function of the SEM package v.3.1-15 (Fox 2006) in R v.4.1.2 (R Core Team 2020). This procedure was executed separately for each of our six unique combinations of species, population, and sex (Figure 3 in §3.1).

2.1.1.2.3 Meta-analysis

We performed two types of meta-analyses. First, we fitted intercept-only random-effects meta-analyses. We did so separately for each of the nine structural paths that the SEM estimated for each trait (Figure 3 in §3.1). We did so using the "rma" function of the metafor package v.3.81-1 (Viechtbauer 2010) in R v.4.1.2 (R Core Team 2020). The response variable in these meta-analyses was the SEM's point estimate of each focal path, where each of our six unique combinations of species, population, and sex provided one data point. Sampling variance was calculated as the squared of the SE (Nakagawa et al. 2022) associated with each focal SEM estimate. We fitted random intercepts for dataset (i.e., the unique combination of species, population, and sex; n=6). Estimates were considered statistically significant when the 95% confidence intervals (CIs) did not overlap zero. We report I^2 as an estimate of the relative heterogeneity for each of the intercept-only random-effects meta-analyses and corresponds to the unexplained variation across effect sizes—in our case across the six unique combinations of species, population, and sex—that was not attributable to their differences in sample size (i.e., sampling variance). I^2 values of 25%, 50%, and 75% are considered low, moderate, and high levels of heterogeneity, respectively (Higgins et al. 2003).

Second, we performed uni-moderator random-effects meta-regressions for each response variable (see above) that could explain some of the heterogeneity found. Our moderators were species, population, and sex. We calculated χ^2-based significance tests to compare the levels of each moderator, and further report values of R^2_{marginal} as indicators of the amount of heterogeneity (I^2) explained by the moderator in each of the random-effects meta-regressions.

2.1.2 Effects of intra- and interspecific competition on reproductive parameters of blue and great tits

2.1.2.1 Nest box density treatment

We manipulated BT- and GT-box numbers to create “low” vs. “high” densities of each type in a full factorial design. Specifically, we randomly assigned each subplot to a box density treatment where each of the four unique combinations of BT- and GT-box density treatments were applied once within each of three “sets” of subplots (where set refers to four subplots in close proximity; Figure S1). The four box density treatments were randomly re-allocated among the four subplots within each of the three sets across years (i.e., stratified random sampling, Table S2).

For each box type (BT- vs. GT-box), “low” box density areas had 10 boxes, and “high” box density areas had 25 boxes, per subplot (Table S2). Box density treatments were implemented as follows. First, each subplot was permanently fitted with 35 boxes with a 32-mm entrance hole (with ~50 m between adjacent boxes; § 2.1). Second, to create BT-boxes, we inserted a plastic ring into the entrance hole to reduce the entrance hole diameter to 26-mm. To create BT “low” – GT “low” (10 BT-boxes – 10 GT-boxes) box density treatments, 15 of the 35 nest boxes were “closed” with a plastic plug that prevented entry. BT- and GT-boxes were evenly distributed over subplots such that realized box densities of each type did not vary within subplots. To create treatment groups where the box density was “low” for one vs. “high” for the other box type, all 35 boxes were used, and plastic rings utilized as appropriate. For treatment groups with BT “high” - and GT “high” box densities, we also used all 35 boxes but also added 15 additional boxes. Those additional boxes were positioned in close proximity (~5 m from permanently existing boxes) as illustrated in Figure S2. The 15 additional boxes were, as above, also uniformly distributed within the focal subplot. Realized box densities were (mean ± SE) 1.51±0.05 vs. 4.10±0.12 BT-boxes/ha in “low” vs. “high” density treatments, respectively, and 1.58±0.06 vs. 3.91±0.10 GT-boxes/ha. Variation in realized box density resulted from subtle differences in plot shape and size (Figure S1).

2.1.2.2 Data collection

A full description of the data collection protocol is provided by Gaona-Gordillo et al. (2023) (see §2.1.1.1). Briefly, boxes were inspected biweekly from the beginning of April to record lay date, clutch size, and the date of the onset of incubation. Shortly before the expected hatch date, boxes were visited daily until egg hatching (day 0). Ten to twelve days after egg hatching, we captured both parents using a spring trap positioned inside the nest box, and ringed any birds not previously banded. Additionally, we determined the breeder's sex based on plumage characteristic (GTs) or the presence and area of a brood patch (BTs and GTs). When nestlings were 15±1 days (BT) or 14 ±1 days (GT) old, nestlings were ringed and measured (body mass, tarsus length, and wing length). From day 19 onwards, boxes were visited daily to record the number of fledglings. Fledge date was defined as the date where half or all nestlings had fledged. To calculate the final number of fledglings, we visited a nest box until the last fledgling had left the nest box; we recorded the number of chicks that did not fledge (i.e., dead nestlings).

All procedures were approved by the Regierung von Oberbayern, Bayern, Germany (permission numbers: ROB-55.2-2532.Vet_02-17-215 and ROB-55.2-2532.Vet_02-18-179). All applicable international, national and/or institutional guidelines for the care and use of animals were followed.

2.1.2.3 Data selection and statistical analysis

Because BTs rarely produce second clutches (Dhondt 2012), our analyses only included "first clutches", defined as clutches initiated within 30 days after the first egg of the focal species was found in the focal year (Nicolaus et al. 2015). This facilitated species comparisons. Our analyses included all first clutches where a female was found incubating, regardless of the brood outcome; these included nests where female identity was unknown.

2.1.2.3.1 Nest box density treatment effects on breeding densities

To test whether our nest box density manipulations affected the realized breeding densities of BTs and GTs, we constructed a series of univariate generalized linear mixed models (GLMMs) that fitted the number of nest boxes occupied in each focal subplot as the response variable. In total, we fitted six models focusing on different types of nest boxes: models a-c vs. d-f focused on BTs vs. GTs, respectively (Table 2 in § 3.2.1). All models fitted the number of boxes occupied by the focal species per subplot that bred either in (a, d) any type of box, (b, e) BT-boxes, or (c, f) GT-boxes. Analyses were conducted this way to examine whether species competed for access to nest boxes, which would be apparent if BTs chose BT-boxes over GT-boxes when competition for access to boxes was relaxed. All models assumed Poisson errors and were fitted with a log-link function. Models included subplot area size centered on its mean value (6.48 ha) as a covariate; this controlled for differences in area size (Figure S1) among subplot, and simultaneously ensured that model intercepts represented the estimated number of breeders for subplots of average size. To account for temporal effects, we further included year (n=3) as a fixed effect factor, and to test for treatment effects, we fitted both the additive and interactive effect of BT- and GT-box density treatments (low vs. high). Finally, to account for spatial effects, we fitted a random intercept for subplot identity (n=12). The sample size of all models (a-f) was the number of subplots times the number of years (n = 12 subplots × 3 years = 36 data points). Parameter estimates for models a-f are presented both on the latent (Table 2) and data (Table S7) scale. To back-transform estimates, we used the function “exp” (package base, R Core Team 2023). Importantly, our analyses demonstrated no interactive treatment effects (see §3.2), and we therefore reran models a-f after dropping the interaction between BT- and GT-box treatment (see Tables S8 and S9 for estimates on the latent vs data scale, respectively).

2.1.2.3.2 Nest box density treatment versus realized density: effects on reproductive parameters

To fully understand how nest box density treatment affected reproductive parameters, we ran a suite of analyses. Specifically, we asked whether any treatment effects resulted from treatment-induced effects breeding density—an “indirect pathway” (treatment affects density, and density affects reproduction)—vs. another “residual pathway”. A residual pathway could emerge, for example, when breeders used the number of cavities (here: nest boxes) in an area as a proxy for forest habitat quality (Krams et al. 2021). We thus first (1) constructed models that fitted the main and interactive effects of BT- and GT-box density treatment. These models captured the “overall” effect of treatment on reproductive parameters. We then (2) expanded these models by additionally including the linear and quadratic effects of realized BT- and GT-breeding densities, as well as the interaction between the linear components, as covariate predictors. In this way, we controlled for non-linear (main and interaction) effects of realized breeding density. To avoid collinearity among these linear and quadratic components, we fitted orthogonal polynomials (Korner-Nievergelt et al. 2015) using the function “poly” (package stats, R Core Team 2023). In this second set of models, “treatment” tested for any residual (i.e., density-independent) path while realized breeding density tested for any effects of realized breeding density not attributable to treatment (i.e., natural variation in realized breeding density). Finally, we (3) excluded the main and interactive effects of the nest box density treatment from our second set of models, such that realized breeding density tested for the overall (treatment-dependent plus -independent) effect of realized breeding density.

We applied this three-step approach to lay date (n = 666 broods; 269 BT vs. 397 GT broods), clutch size (n = 665 broods; 268 BT vs. 397 GT broods), nestling body condition (defined as tarsus-corrected body mass; n = 3303 nestlings; 1493 BT vs. 1810 GT nestlings), nestling body size (measured as tarsus length; n = 3305 nestlings; 1495 BT vs. 1810 GT nestlings), and the number of fledglings (n = 666 broods; 269 BT vs. 397 GT broods). All analyses we conducted for (a) BT data only and (b) GT data only (Tables S10, S11, S12, S13, S14, S15), and (c) the data of both species combined (Table 3 in § 3.2.2). Models that included both species also fitted the interaction between species and each (main and interaction) effect listed above and below; this tested for species-specific effects.

Next to the fixed effects detailed above, all models also included mean-centered area size as an additional fixed effect covariate (see above). Analyses of both nestling body condition and body size additionally fitted mean-centered nestling age (in days) as a covariate to control for variation in realized age at measurement (target age in BT: 15 days, realized range: 14-16 days; target age in GT: 14 days, range: 14-15 days). The data of 62 nestlings were removed from our analyses of body condition and body size as they were outside the target age range (BT nestlings: 24 measures on day 17, seven on day 18, and one on day 19; GT nestlings: 21 on day 16 and 9 on day 19); we also removed the data of two BT nestlings that were part of a mixed clutch raised by a GT pair. Analyses of nestling body condition further fitted tarsus as a covariate (mean-centered within species), such that effects of body mass represented size-independent effects ("condition").

All models further fitted random intercepts for the unique combination of subplot and year (n = 36 levels) to account for spatiotemporal effects, and female identity (lay date: n = 565, clutch size: n = 564, nestling body condition and body size: n = 385, and number of fledglings = 378 females). For broods where female identity was unknown, a unique identity was assigned. The model of the binary probability of producing fledglings did not include female identity as this prevented model convergence. Analyses of nestling body condition and body size also included random intercepts for brood (n = 464 broods) and observer identity (n = 22 observers).

For all these analyses, we fitted univariate linear mixed models (LMMs) assuming a

Gaussian error distribution. The exception were models with the number of fledglings as the response, where visual inspection of the data revealed an excess of zeros (i.e., failed broods: 77 of 269 BT vs. 132 of 397 GT broods). To avoid analytical complexity, we addressed this issue by, first, analysing the binary probability of producing fledglings (fitting models with a binomial error distribution with the logit link function), and then, for the successful nests, effects on the number of fledglings (fitting fitted models with a Poisson error distribution with a log-link function).

2.1.2.3.3 Model parametrization

All statistical analyses were performed in R 4.3.0 (R Core Team 2023). The GLMMs and LMMS were performed using the "glmer" and "lmer" functions, respectively (package lme4, Bates et al. 2015). Binomial and Poisson models were tested for overdispersion and zero inflation. For overdispersion we used the function "dispersion_glmer" (package blmeco; Korner-Nievergelt et al. 2015), where values >1.4 suggested severe overdispersion (Korner-Nievergelt et al. 2015). To test for zero inflation we used the function "testZeroInflation" (package DHARMa, Hartig 2022). We used the "sim" function (package arm, Gelman and Hill 2006) to obtain 2000 posterior simulations of model parameters, from which we then calculated the mean (β) and associated 95% credible intervals (CIs) for each estimated parameter. Effects of fixed-effects were considered statistically significant in the frequentist sense when the 95% CIs did not overlap zero.

2.1.3 Aggressive behaviour towards con- vs. heterospecifics in blue and great tits

2.1.3.1 Data collection

We inspected the boxes biweekly from the beginning of April to record the lay date, clutch size, the date of the onset of incubation, and determined the species of the breeding pair. During incubation, we performed simulated territorial intrusions to test aggressive behaviour in breeding males. Tests were performed only in first broods. We defined first broods as those initiated within 30 days after the first egg of the focal species was found in the focal year (Nicolaus et al. 2015). Fourteen days after we recorded the onset of incubation, we visited the boxes daily until egg hatching. Ten until twelve days after egg hatching, we captured each adult of the breeding pair using a spring trap placed inside the box. We ringed the focal bird, if it was not previously banded, and determined its sex based on the presence and area of the brood patch (BTs and GTs) and size and hue of the black breast band (GTs).

2.1.3.1.1 Simulated territorial intrusions

Once we recorded the onset of incubation, males were subjected to two simulated territorial intrusions. The first test was performed the following day after we confirmed that the female was incubating and the second test was done on the third day. During both tests, we used visual (i.e., a taxidermic mount) and acoustic (i.e., playback song) stimuli to provoke an aggressive response. Our simulated intrusion treatments included each unique combination of each “type” of “intruder” (con- vs. heterospecific) in a full factorial design (i.e., first – second test: BT – BT, BT – GT, GT – BT, and GT – GT). Within each species, males were allocated to one of our four simulated intrusion treatments in such a way that both species were tested evenly across treatments across years. For each test, one taxidermic mount (i.e., one of seven BT or one of eight vs. GT taxidermic mounts) and one playback song (i.e., one of the 20 BT or one of the 30 GT songs) was randomly assigned to each male.

Prior to the test, we first confirmed that the female was inside the box; if so, we prevented her from exiting the box during the test by placing a closed spring trap inside the box. Next, we placed the randomly assigned mount, fixed to a 1.2 m wooden pole and protected with a green wire mesh, 1 m away from the focal nest box. Next to the wooden pole, we placed on the floor a FOXPRO Shockwave speaker that contained a list of BT playback songs, that were downloaded from Xento-canto (Planqué et al. 2005) (list of playback songs in Text S3), and GT playback songs, that were recorded by Hutfluss et al. (2021) in another population. Next, the observer started the broadcast of the corresponding playback song and moved 15 m away from the box. Once the breeding male entered a 15 m radius from the focal nest box, the observer recorded its behaviours for three minutes. The behaviours included the number of song bouts, alarm calls, dives (i.e., when the focal male approached the mount without landing on it), attacks (i.e., when the focal male landed on the mesh), and nest box landings. Additionally, the observer estimated the minimal approach distance to the mount. Following Araya-Ajoy & Dingemanse (2014), we used the minimal approach distance of the focal male to the mount as our metric to assess the aggressive response. Males that did not arrive within 12 minutes were considered as nonresponsive. Finally, at the end of the test, we removed the spring trap from the box so that the female was able to exit.

All procedures were approved by the Regierung von Oberbayern, Bayern, Germany (permission numbers: ROB-55.2-2532.Vet_02-17-215 and ROB-55.2-2532.Vet_02-18-179). All applicable international, national and/or institutional guidelines for the care and use of animals were followed.

2.1.3.2 Data selection and statistical analysis

Our analyses included broods for which onset of incubation was confirmed previously, regardless of the brood outcome; these included nests where male identity was unknown.

Furthermore, we only included broods where the male was tested twice during incubation (403 broods; 170 BT vs. 233 GT broods).

2.1.3.2.1 Quantifying aggressiveness statistically

We used univariate linear mixed models (LMM) that included the minimal approach distance of the focal male to the taxidermic mount as the response variable. To simplify the interpretation, we multiplied the approach distance with -1, such that higher values represented closer distance to the mount (Araya-Ajoy and Dingemanse 2014). The minimal approach distance was then square-root transformed and modelled following a Gaussian error distribution. Models included data of BT males (Table 4a, b), GT males (Table 4c, d), and males of both species (Table 4e, f). We used a two-step approach to analyse the data. First, we tested the effects of the first simulated intrusion test (Table 4a, c, e). Second, for the second simulated intrusion test, we then tested the effects of the current test and whether these results were (partly) influenced due to carry-over effects of the previous test (Table 4b, d, f). Note that we were unable to perform single analyses with all the data (first and second test) as we could not estimate the carry-over effects of the previous test for the first test. All models included the mount species of the current test (BT vs. GT) as fixed effect. Models assessing the effects of an intrusion test for the second time (Table 4b, d, f) further included the species of the mount of the previous test (BT vs. GT). In the latter models, we included both the additive and interactive effects of the mount species of the current test and the mount species of the previous test (BT vs. GT mount). Models that included data of both species (Table 4e, f) additionally included the species of the focal male (BT vs. GT) as fixed effect and its interaction with each main and interaction effect mentioned above; this tested for species-specific effects. All models then fitted a random intercept for male identity as 13 out of 157 BT individuals, and 17 out of 216 GT individuals were tested in both years. For broods where male identity was unknown, we assigned a unique identity. Models further included random intercepts for a unique combination between subplot and year, mount identity, song identity, and observer identity; for sample sizes see Table 4. Subplot-year accounted for spatial and temporal variation, and mount and song identities accounted for variation attributed to unique features of the both (visual and acoustic) stimuli. Importantly, analyses of the second test showed no interactive effect between the mount species of the current and of the previous test (see § 3.3). We thus reran the models a-f after removing this interaction (Table S16). Furthermore, because we did not detect any carry-over effects of the mount species of the previous test, we also constructed a dataset that combined first and second tests. This combined dataset was further extended to test for nest box density treatment, and realized breeding density effects.

2.1.3.2.2 Nest box density treatment versus realized density effects on aggressiveness

We fitted nine univariate LMMs to test the effects of the nest box density treatment on the male aggressive response toward a simulated intrusion test. All models included the minimal approach distance as response variable (see above). Models included data for BT males (n = 279 observations), GT males (n = 317 observations), and males of both species (n = 596 observations). Following § 2.1.2.3.2, we used a three-step approach to disentangle whether effects observed in the approach distance of the breeding males were due to changes on the realized breeding density that resulted from our nest box density treatment (i.e., indirect pathway) or due to other potential pathways (i.e., residual pathways). First, we constructed models that could capture the overall effect of the nest box density treatment on the minimal approach distance of the males. These models included the additive and interactive effects of the BT- and GT-box density treatment (low vs. high). Second, we fitted models that tested for any effects of our nest box density treatment that were density-independent (i.e., residual pathway) and for any effects of the realized breeding densities that were treatment-independent. To achieve this, we expanded the previous models by adding the mean-centered linear and quadratic effects of the BT- and GT- realized breeding densities. We further included the interaction between the mean-centered linear components of the realized breeding densities of both species. To prevent collinearity between the linear and quadratic terms, we included orthogonal polynomials (Korner-Nievergelt et al. 2015) by using the function "poly" (package stats, R Core Team 2023). Finally, constructed models that tested for the overall effects of the BT and GT realized breeding densities on the minimal approach distance. Here, we reduced our second suite of models by removing the main and interactive effects of the nest box density treatment.

All models further included the species of the mount of the current test (BT vs. GT) as fixed effect. Models that included males of both species additionally included the species of the focal male (BT vs. GT) and its interaction with all the (additive and interactive) effects aforementioned. Finally, we used the same structure of the random intercepts as described in the previous section (see above). For sample sizes, see Table 5.

2.1.3.2.3 Model parametrization

We performed all statistical analyses in R 4.3.0 (R Core Team 2023). We used "lmer" function (package lme4, Bates et al. 2015) to perform the LMMs. We then used the "sim" function (package arm, Gelman and Hill 2006) to obtain 2000 posterior simulations of model parameters. We then calculated the mean (β) and associated 95% credible intervals (CIs) for each estimated parameter. We considered that fixed-effects were statistically significant in the frequentist sense when the 95% CIs did not overlap zero.

CHAPTER 3

3. Results

3.1 Phenotypic integration among species, populations, and sexes

Based on Gaona-Gordillo et al (2023)

We found strong evidence for the existence of a latent variable affecting our observable expressions of size. For each of the six unique combinations of species, population, and sex, all four morphological traits (body mass, tarsus length, wing length, and bill length) were positively and statistically significantly connected to the latent variable (Figure 3; Table S3). The only exception was bill length, where the standardized effect of the latent variable was positive in all but significantly supported in only five of the six datasets. The global (meta-analytic) effect sizes for the standardized path coefficient connecting the latent variable to body mass, tarsus length, wing length, and bill length, were, respectively, all positive and strongly supported (due to 95% CIs not overlapping zero) (Table 1; Figure 4). This implies that heavier animals had longer tarsi, longer wings, and longer bills across combinations of species, population, and sex.

In five of the six datasets, the latent variable negatively affected breathing rate (Figure 3; Table S3). The global (meta-analytic) effect size for this standardized effect of the latent variable on breathing rate was negative and strongly supported (due to 95% CIs not overlapping zero) (Table 1; Figure 4). The strength of this path did not differ between datasets as our global meta-analysis showed no statistical support for heterogeneity (Table 1). Larger animals thus generally breathed slower. In five of the six datasets, the residual effect of body mass on breathing rate independent of effect of size was, by contrast, positive. This implies that animals that were relatively heavy for their size, also breathed relatively fast. The global (meta-analytic) effect size for this standardized size-independent effect of body mass on breathing rate was also positive and strongly supported (due to 95% CIs not overlapping zero) (Table 1; Figure 4). The strength of this path did not differ between datasets as there was no statistical support for heterogeneity (Table 1). These results altogether imply that larger animals generally breathe slower, and that animals that are heavy for their size, generally breathe faster.

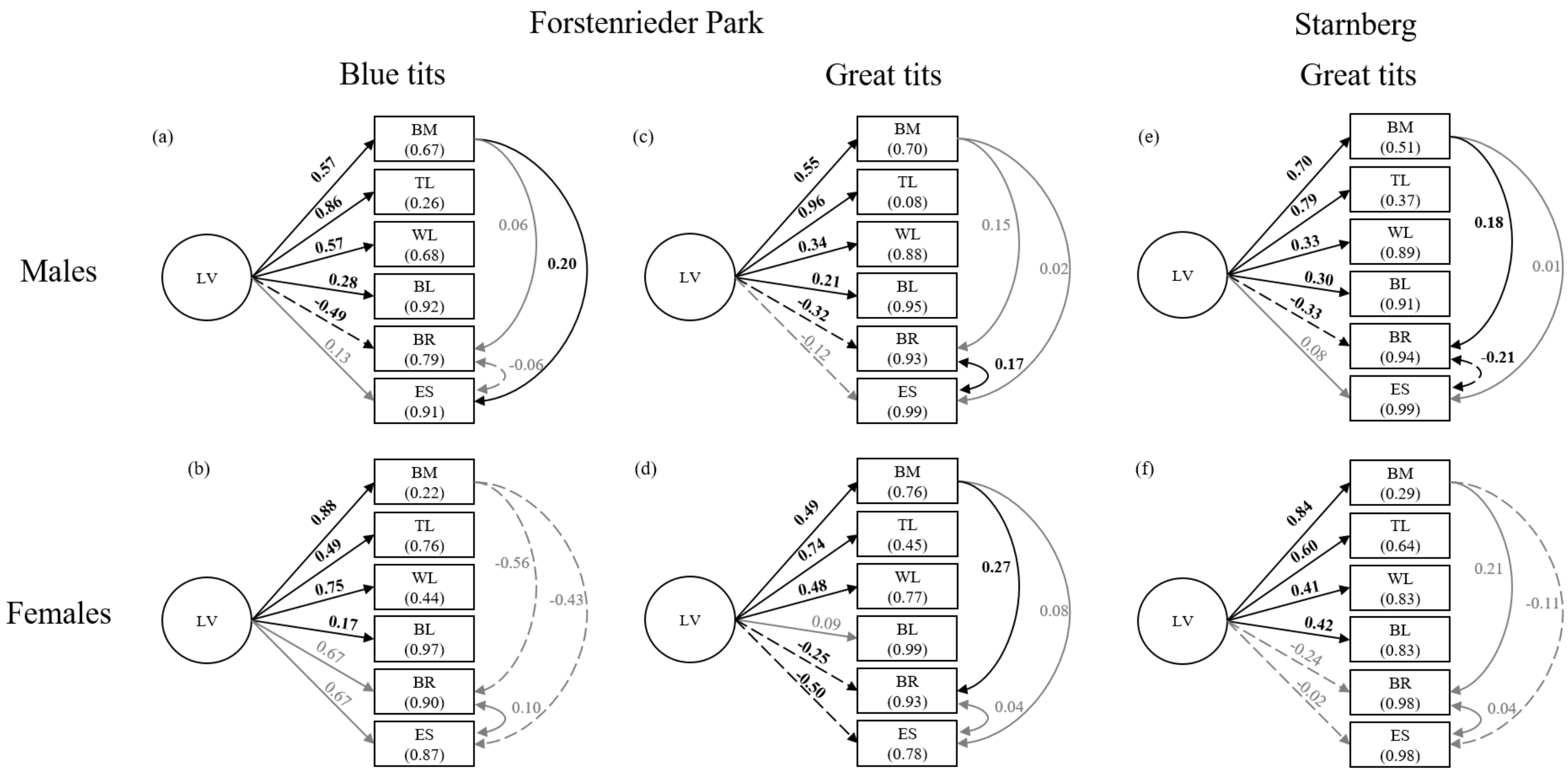


Figure 3. Visualisation of the structural equation model (SEM) fitted for each combination of species, population, and sex. Shown are the point estimates for each of nine structural paths connecting behaviour, physiology, and morphology among-individuals in each dataset. Numbers correspond to standardized path coefficients (single-headed arrows) or residual among-individual correlations (double-headed arrows). Solid vs. dashed arrows are positive vs. negative estimates, respectively. Residual variances are printed inside boxes. Black vs. grey arrows are statistically significant ($p<0.05$) vs. non-significant ($p>0.05$) estimates, respectively. See Table S3 for the uncertainty (SE) and statistical significance of each path. BM=body mass, TL=tarsus length, WL=wing length, BL=bill length, BR=breathing rate, ES=exploration score, LV= latent variable.

Table 1. Global (meta-analytic) effect sizes of each of nine hypothesized structural paths with 95% confidence interval (CI), 95% predictability interval (PI), and among-dataset heterogeneity (I^2). Single-headed arrows (→) represent standardized path coefficients; double-headed arrows (↔) represent correlations.

Path	Global effect size (95% CI)	95%PI	Heterogeneity		
			I^2	χ^2_5	p
Latent variable → Body mass	0.68 (0.56, 0.81)	(0.38, 0.99)	83.65%	25.91	<0.001
Latent variable → Tarsus length	0.72 (0.60, 0.85)	(0.42, 1.03)	84.90%	29.39	<0.001
Latent variable → Wing length	0.47 (0.35, 0.60)	(0.17, 0.77)	86.50%	27.26	<0.001
Latent variable → Bill length	0.26 (0.16, 0.36)	(0.03, 0.49)	77.90%	25.10	<0.001
Latent variable → Breathing rate	-0.32 (-0.40, -0.23)	(-0.40, -0.23)	0.00%	7.64	0.18
Latent variable → Exploration score	-0.04 (-0.26, 0.17)	(-0.54, 0.46)	82.10%	23.99	<0.001
Body mass → Breathing rate	0.17 (0.10, 0.25)	(0.10, 0.25)	0.10%	5.39	0.37
Body mass → Exploration score	0.03 (-0.04, 0.09)	(-0.04, 0.09)	0.00%	6.47	0.26
Breathing rate ↔ Exploration score	0.00 (-0.11, 0.12)	(-0.27, 0.28)	87.00%	51.24	<0.001

The latent variable, however, did not consistently affect exploration score. This link was significantly heterogeneous, while the global (meta-analytic) effect did not differ from zero (due to 95% CIs overlapping zero) (Table 1; Figure 4). In other words, in some datasets, the larger animals were more explorative, while in other datasets, they were less explorative (Figure 3). There was also no overall support for a residual effect of body mass on exploration score that was independent of the effect of size, and neither was this link heterogeneous (Table 1). That is, animals that were relatively heavy for their size were not more/less explorative. Indeed, only in one out of six datasets was this link statistically supported (Figure 3; Table S3).

Finally, we found strong evidence for heterogeneity in the correlation between behaviour and physiology that was attributable to their respective covariances with size, while the overall (meta-analytic) effect size was not different from zero (due to 95% CIs overlapping zero) (Table 1). This was because this size-corrected correlation between breathing rate and exploration score was significantly positive in some but significantly negative in other datasets (Figure 3; Table S3).

In summary, we found strong and consistent support for both size- and condition-related physiology, as larger animals breathed slower, but animals heavy for their size breathed faster. By contrast, relationships between size and exploration behaviour, and between physiology and exploration behaviour were strongly heterogeneous and not overall supported.

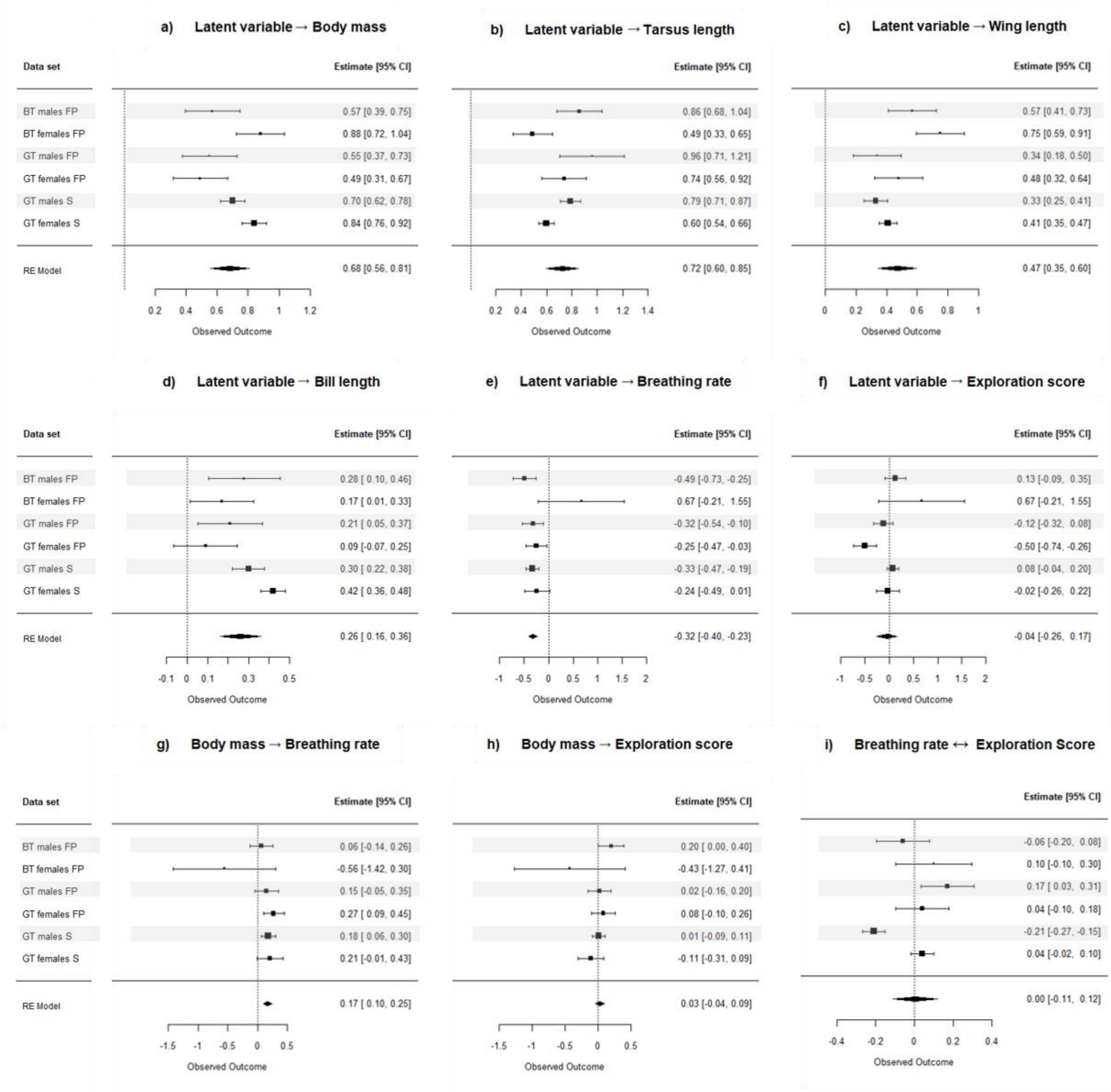


Figure 4. Forest plots of the global (meta-analytic) effects for standardized paths coefficients (a-h) and correlations (i) with 95% confidence intervals (for model estimates see Table 1). BT=blue tit, GT=great tit, M=male, F=female, FP=Forstenrieder Park, S=Starnberg, RE Model=global effect size.

The uni-moderator random-effects meta-regressions (Tables S4-S6) provided no support for main effects of species, population, or sex on integration between size and the non-morphological traits (exploration score and breathing rate). These factors neither explained variation in the effect of mass on the non-morphological traits independent of size, nor in the covariance between non-morphological traits independent of size. The heterogeneity between datasets in integration between size and exploration behaviour, or between exploration behaviour and breathing rate (Table 1) was therefore attributable to other dataset-specific factors of unknown origin. These analyses did, by contrast, support effects of species, population, and sex on how the latent variable affected certain observable expressions of size. Specifically, wing length was more closely reflective of the latent variable (“size”) in BTs compared to GTs (Table S4), bill length was more closely reflective of the latent variable in Forstenrieder Park compared to Starnberg (Table S5), and tarsus length was more closely reflective of the latent variable in males compared to females (Table S6).

3.2 Effects of intra- and interspecific competition on reproductive parameters of blue and great tits

3.2.1 Nest box density treatment effects on breeding densities

Experimental increases in BT- and GT-box numbers increased—as predicted—the tits' breeding densities (Figure 5). When GT-box numbers were low, the number of BT-boxes provided in a subplot positively affected the total number of boxes occupied by BT pairs (main effect of BT-box number; Table 2a); this effect did not change when the number of GT-boxes was high (as 95% CIs for the interaction BT- × GT-box number overlapped zero; Table 2a). Similarly, when the number of BT-boxes was low, the number of GT-boxes provided in a subplot positively affected the total number of boxes occupied by GT pairs (main effect of GT-box; Table 2d); this effect did not change when the number of BT-boxes was high (interaction BT- × GT-box; Table 2d). To understand the overall effects of BT- and GT-box numbers, we re-ran our models after removing the non-significant BT- × GT-box number interaction term. This showed that BT-box numbers—but not GT-box numbers—affected BT numbers (Table S8a), and that GT-box numbers—but not BT-box numbers—affected GT numbers (Table S8d). Thus, for both species, the number of breeding pairs was affected by the number of "conspecific"—not "heterospecific"— boxes provided.

The back-transformed estimates of our (Poisson) models (Table S9) showed that our 2.5-fold increase in the number of BT-boxes (i.e., from "low" = 10 boxes to "high" = 25 boxes) increased the number of BT pairs by a factor of 1.57 (Table S9a); the same increment in GT-boxes increased the number of GT pairs by a similar factor (1.80; Table S9d). As upper 95% CIs of either effect did not include 2.5 (Table S8), we conclude that relatively more boxes remained unoccupied in areas with higher nest box numbers. When BT-box numbers were low, the estimated breeding density of BTs was 4.81 (intercept value in Table S9a) divided by 6.48 (the average area size) = 0.74 pairs/ha (95% CI = 0.47-1.17). When BT-box numbers were high, the estimated breeding density of BTs was 1.16 pair/ha (0.74 × 1.57). When GT-box numbers were low, the estimated breeding density of GTs was 6.89 (intercept value in Table S9d) divided by 6.48 (the average area size) = 1.06 pairs/ha (95% CI = 0.84-1.40). When GT-box numbers were high, the estimated breeding density of GTs was 1.91 pair/ha (1.06 × 1.80).

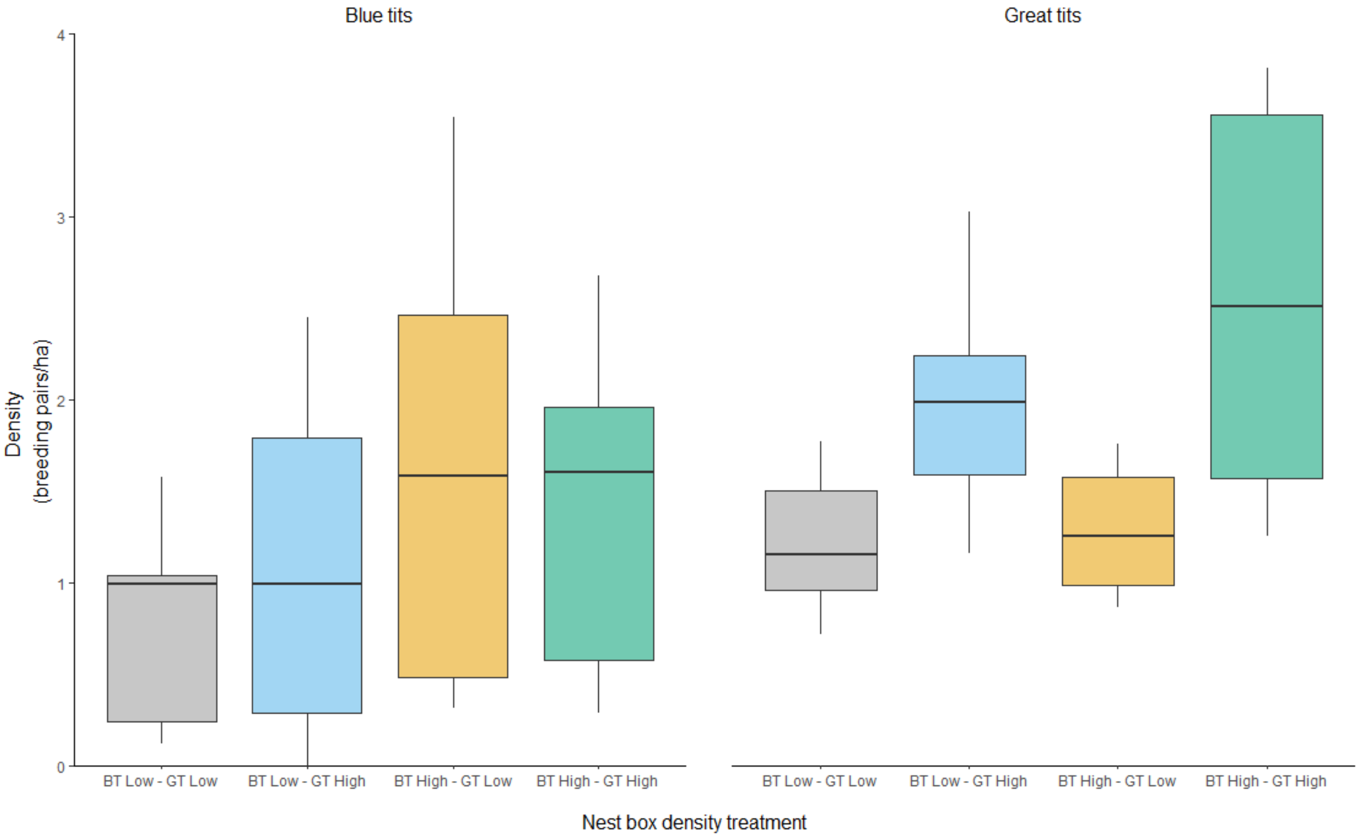


Figure 5. Boxplots showing realized breeding densities for blue tits (left) and great tits (right) for each unique combination of blue tit (BT) and great tit (GT) box density treatment effects (Low vs. High). The boxplots illustrate five summary statistics: the central line (black bar) of the box represents median value, box limits represent lower and upper quartiles upper (25th and 75th percentile, respectively), and whiskers represent 1.5 × the interquartile range (IQR) below and above the lower and upper quartiles, respectively.

Table 2. Estimated effect sizes and 95% credible intervals (CIs) for predictors of the number of blue tits (a-c) and great tits (d-f) breeding in 12 subplots located in the Forstenrieder Park, Munich, Germany. We analyse numbers of breeders in (a, d) any type of box, (b, e) blue tit (BT) boxes and (c, f) great tit (GT) boxes. Predictor variables are subplot area size (in ha), year (2020, 2021, 2022), the number (no) of BT-boxes (Low vs. High), the no of GT-boxes (Low vs. High), and their interaction. Parameters were estimated using GLMMs following a Poisson error distribution and a log-link function. Estimated effect sizes and 95% CIs are shown on the latent scale. The sample size was the number of subplots times the number of years (n=12 subplots × 3 years = 36 data points).

	Blue tits breeding in			Great tits breeding in		
	(a) Any type of box	(b) BT-boxes	(c) GT-boxes	(d) Any type of box	(e) BT-boxes	(f) GT-boxes
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	1.45 (0.97, 1.94)	1.32 (0.83, 1.83)	-1.56 (-3.14, 0.05)	1.98 (1.69, 2.26)	-0.37 (-1.49, 0.72)	1.89 (1.60, 2.19)
Area [b]	-0.45 (-0.83, -0.08)	-0.41 (-0.80, -0.01)	-0.78 (-1.40, -0.17)	-0.06 (-0.21, 0.09)	-0.27 (-0.81, 0.28)	-0.05 (-0.19, 0.09)
Year 2021	0.15 (-0.12, 0.46)	0.20 (-0.11, 0.52)	-0.03 (-0.80, 0.73)	0.11 (-0.13, 0.35)	-0.39 (-1.55, 0.73)	0.13 (-0.12, 0.38)
Year 2022	0.00 (-0.29, 0.29)	0.08 (-0.26, 0.41)	-0.39 (-1.23, 0.48)	0.15 (-0.09, 0.38)	-0.19 (-1.28, 0.90)	0.16 (-0.08, 0.41)
BT-box no [c]	0.61 (0.23, 0.99)	0.63 (0.25, 1.02)	0.79 (-0.96, 2.43)	-0.01 (-0.34, 0.31)	0.19 (-1.01, 1.41)	-0.03 (-0.38, 0.31)
GT-box no [d]	0.20 (-0.34, 0.72)	-0.13 (-0.73, 0.44)	2.49 (0.90, 4.03)	0.52 (0.20, 0.84)	-1.63 (-3.78, 0.64)	0.57 (0.24, 0.89)
BT × GT-box no [e]	-0.50 (-1.12, 0.14)	-0.22 (-0.88, 0.46)	-1.99 (-3.80, -0.09)	0.12 (-0.30, 0.56)	1.34 (-1.23, 3.76)	0.12 (-0.30, 0.58)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Subplot ID	0.34 (0.17, 0.56)	0.40 (0.20, 0.64)	0.11 (0.04, 0.21)	0.02 (0.01, 0.04)	0.00 (0.00, 0.00)	0.02 (0.01, 0.03)
Residual [f]	0.21 (0.13, 0.32)	0.24 (0.15, 0.36)	1.75 (0.67, 3.18)	0.13 (0.10, 0.17)	0.89 (0.40, 1.69)	0.14 (0.11, 0.18)

[a] Reference category: estimate is for the number of occupied nest boxes in the “Low” BT- and “Low” GT-box density treatment for the year 2020 for areas of average size.

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Estimate is the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[d] Estimate is the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[e] Estimate is the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[f] Defined as ln(1/exp(βo)+1) (following Nakagawa & Schielzeth, 2010).

3.2.1.1 Box types used by BTs and GTs

BTs can readily breed in GT-boxes, and small GTs can also breed in BT-boxes. To understand the mechanism by which our manipulations (detailed § 2.1.2.3.1) affected realized breeding densities, we therefore also tabulated the numbers of breeders per species and box type. Over the entire study, we found 269 BT breeding pairs: 231 (85%) in BT-boxes and 38 (14%) in GT-boxes. Furthermore, we found 397 GT breeding pairs: 18 (4%) in BT-boxes and 379 (95%) in GT-boxes. Even though BTs were three times more likely to breed in the "heterospecific" (GT) box type compared to GTs, both species mainly bred in their "own" box type. This suggests that effects of BT-box numbers on the number of BT breeders were largely due to BTs preferably choosing BT-boxes. This was indeed the case: when we re-ran our analyses focusing on BTs breeding in BT-boxes vs. GT-boxes, we found that BT-box numbers affected the number of BTs breeding in BT-boxes (Tables 2b, S8b) but not the number of BTs breeding in GT-boxes (Tables 2c, S8c). However, when BT-box numbers were low, more BTs bred in GT-boxes when the number of GT-boxes was high (Tables 2c, S8c). By contrast, when BT-box numbers were high, BTs did not use more GT-boxes when their numbers were high (negative interaction BT- × GT-box number; Table 2c). This suggests that GT-boxes were only used by BTs when BT-boxes were provided in low numbers. For GTs, such conditional effects were not detected: GTs simply bred more in GT-boxes when their numbers were high (Tables 2f, S8f) and neither treatment affected how many GTs bred in BT-boxes (Table 2e, S8e).

3.2.2 Nest box density treatment versus realized density: effects on reproductive parameters

3.2.2.1 Overall nest box density treatment effects

In BTs, reproductive parameters (lay date, clutch size, nestling body condition, nestling body size, and the number of fledglings in successful nests) were neither affected by main nor by interactive effects of BT- and GT-box numbers (all 95% CIs overlapped zero, Table 3a, d, g, j, and p where BT was set as the reference species). The only exception was the binary probability to produce fledglings (Table 3m). Specifically, when the number of GT-boxes was low, an increase in BT-box numbers led to an increased probability that BTs produced fledglings (main effect of BT box number in Table 3m). This effect was, notably, only moderately supported as 95% CIs overlapped zero. By contrast, when the number of GT-boxes was high, an increase in BT-box numbers led to a significantly less positive response (BT- × GT-box number interaction Table 3m). As the BT- × GT-box number interaction term represents the difference in the effect of adding BT-boxes between the low and high GT-box number treatment groups, these results altogether imply that adding BT-boxes had a positive effect when there were few GT-boxes but that this effect disappeared when there were many GT-boxes (as illustrated in Figure 7). We also found little evidence for species differences in treatment effects. That is, two-way interactions between species and BT-, GT-, or BT- × GT-box numbers were generally not supported (Table 3), though there were two exceptions. First, when BT-box numbers were low, an increase in GT-box number from low to high, resulted in GTs producing significantly smaller clutches relative to BTs (GT-box × Species; Table 3d, Figure 6). As detailed above for the probability to produce fledglings, these species differences were caused by BTs—not GTs—responding to increases in GT-box number when BT-box numbers were low. Indeed, the estimate for BTs (main effect of GT-box treatment; Table 3d) was moderately positively (though not significantly) supported. Altogether, these analyses imply that when BT-box numbers were low, BTs—but not GTs—increased their clutch sizes when GT-box numbers were increased from low to high, and that this effect disappeared when BT-box numbers were high (Figure 6). This interpretation was supported when we performed the analyses separately for the two species; this showed a similarly positive point estimate for the effect of GT-box treatment in BTs (Table S11a) that did not exist in GTs (Table S11d). Second, the very same effect was present for the binary probability to produce fledglings, where, again, the same negative effect in GTs compared to BTs seemed to be caused by a moderately supported positive (again, not significant) effect in BTs that did not exist in GTs (GT-box number × Species; Table 3m, Figure 7). This interpretation was again fully supported when we performed the analyses separately for BTs (Table S14a) and GTs (Table S14d). In summary, significant treatment effects were not present in species-specific analyses, while significant species differences simultaneously existed in treatment effects. This altogether suggests that great caution is required in interpreting these effects and that firm conclusions cannot be drawn with the data at hand.

Table 3. Estimated effect sizes and 95% credible intervals (CIs) for predictors of lay date (a-c), clutch size (d-f), nestling body condition (g-i), nestling body size (j-l), the binary probability to produce fledglings (m-o), and the number of fledglings (in nests with at least one fledgling; p-r) for models fitting (main and interaction effects of) the number (no) of blue tit (BT) and great tit (GT) boxes (nest box density treatment; a, d, g, j, m, p), realized breeding density (c, f, i, l, o, r) or both (b, e, h, k, n, q). We further fit subplot area size, tarsus (analyses of nestling body condition only), age (analyses of nestling body condition and size only), and the interaction between mentioned fixed effects and species. Parameters were estimated using LMMs with a Gaussian error distribution (lay date, clutch size, nestling body condition and body size), GLMMs with a binomial error distribution (binary probability to produce fledglings), or GLMMs with a Poisson error distribution (number of fledglings models).

	Lay date (n= 666 broods)			Clutch size (n= 665 broods)			Nestling body condition (n= 3303 nestlings)		
	(a) Treatment	(b) Treatment and realized breeding density	(c) Realized breeding density	(d) Treatment	(e) Treatment and realized breeding density	(f) Realized breeding density	(g) Nest box density treatment	(h) Treatment and realized breeding density	(i) Realized breeding density
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	17.47 (14.99, 19.83)	17.86 (14.24, 21.39)	17.99 (16.56, 19.43)	9.59 (9.12, 10.04)	9.50 (8.78, 10.23)	9.70 (9.43, 9.99)	10.61 (10.15, 11.1)	10.79 (10.14, 11.45)	10.84 (10.6, 11.09)
Area [b]	1.11 (-0.24, 2.35)	1.22 (-0.69, 3.26)	1.10 (-0.74, 2.93)	-0.32 (-0.59, -0.05)	-0.26 (-0.68, 0.15)	-0.12 (-0.5, 0.27)	0.07 (-0.16, 0.31)	0.11 (-0.26, 0.44)	0.11 (-0.21, 0.44)
Age [c]	-	-	-	-	-	-	-0.39 (-0.93, 0.15)	-0.44 (-0.96, 0.08)	-0.44 (-0.98, 0.07)
Tarsus [d]	-	-	-	-	-	-	0.75 (0.66, 0.83)	0.75 (0.66, 0.83)	0.75 (0.66, 0.83)
Nest box density treatment									
BT-box no [e]	0.82 (-2.25, 4.11)	0.82 (-2.95, 4.45)	-	0.12 (-0.50, 0.71)	-0.04 (-0.79, 0.65)	-	0.18 (-0.40, 0.78)	0.08 (-0.55, 0.71)	-
GT-box no [f]	-0.88 (-4.11, 2.35)	-0.84 (-5.91, 4.07)	-	0.59 (-0.01, 1.24)	0.58 (-0.40, 1.50)	-	0.06 (-0.52, 0.67)	-0.02 (-0.93, 0.89)	-
BT- × GT-box no [g]	1.14 (-3.55, 5.38)	0.42 (-4.89, 5.94)	-	-0.69 (-1.49, 0.11)	-0.42 (-1.45, 0.60)	-	-0.08 (-0.89, 0.74)	0.06 (-0.88, 0.95)	-
Realized breeding density [h]									
BT-density	-	-3.87 (-51.84, 46.01)	0.50 (-41.69, 43.64)	-	4.83 (-5.04, 14.75)	5.01 (-3.98, 13.84)	-	-4.09 (-23.26, 14.91)	-3.37 (-21.04, 14.26)
BT- density squared	-	-6.61 (-46.48, 33.23)	-8.59 (-44.55, 26.06)	-	1.17 (-6.94, 9.16)	1.49 (-5.69, 8.54)	-	5.83 (-10.17, 23.34)	5.63 (-8.61, 19.65)
GT-density	-	26.87 (-47.27, 98.20)	22.7 (-24.32, 72.01)	-	-5.63 (-19.62, 8.44)	-1.60 (-10.55, 7.37)	-	8.67 (-19.22, 37.21)	9.42 (-8.83, 27.04)
GT-density squared	-	19.31 (-23.06, 61.04)	27.12 (-6.41, 59.08)	-	-1.20 (-8.50, 5.90)	-4.54 (-10.28, 1.26)	-	-0.06 (-15.35, 15.80)	1.10 (-11.2, 12.87)
BT- × GT-density	-	-0.86 (-3.53, 1.84)	-0.93 (-3.43, 1.59)	-	0.21 (-0.26, 0.71)	0.15 (-0.33, 0.61)	-	-0.27 (-0.73, 0.18)	-0.28 (-0.72, 0.13)
Species [i]	2.88 (1.11, 4.71)	1.82 (-0.71, 4.38)	1.82 (0.81, 2.81)	-0.90 (-1.50, -0.28)	-0.81 (-1.67, 0.07)	-1.20 (-1.54, -0.86)	5.57 (5.09, 6.08)	5.83 (5.16, 6.53)	5.37 (5.13, 5.62)
Area × Species [i]	1.00 (0.03, 1.99)	0.40 (-1.09, 1.90)	0.35 (-1.01, 1.79)	0.37 (0.01, 0.69)	0.17 (-0.31, 0.68)	0.04 (-0.43, 0.50)	-0.07 (-0.31, 0.18)	0.10 (-0.25, 0.47)	0.02 (-0.31, 0.34)
Age × Species [i]	-	-	-	-	-	-	0.91 (0.19, 1.67)	1.01 (0.27, 1.74)	0.94 (0.21, 1.65)
Tarsus × Species [i]	-	-	-	-	-	-	0.31 (0.21, 0.41)	0.31 (0.21, 0.41)	0.31 (0.21, 0.41)
Nest box density treatment									
BT-box no × Species [j]	-0.50 (-2.92, 1.83)	0.15 (-2.43, 2.82)	-	-0.43 (-1.21, 0.35)	-0.26 (-1.18, 0.64)	-	-0.27 (-0.89, 0.36)	-0.35 (-1.05, 0.37)	-
GT-box no × Species [k]	-1.30 (-3.70, 1.03)	-0.27 (-3.83, 3.21)	-	-0.78 (-1.59, -0.03)	-0.70 (-1.84, 0.50)	-	-0.33 (-0.92, 0.29)	-0.73 (-1.74, 0.22)	-
BT- × GT-box no × Species [l]	1.14 (-1.83, 4.15)	0.17 (-3.66, 3.87)	-	0.90 (-0.13, 1.93)	0.58 (-0.69, 1.89)	-	0.40 (-0.44, 1.20)	0.54 (-0.41, 1.57)	-
Realized breeding density [m]									
BT-density × Species	-	-8.83 (-43.64, 24.2)	-6.95 (-37.94, 24.65)	-	-8.90 (-21.17, 3.49)	-9.76 (-20.21, 0.82)	-	6.62 (-13.01, 25.43)	6.38 (-11.9, 24.68)
BT-density squared × Species	-	0.94 (-28.56, 30.48)	0.75 (-25.41, 26.92)	-	1.52 (-8.26, 10.96)	0.30 (-8.31, 8.79)	-	-1.75 (-19.61, 15.98)	-7.43 (-22.55, 8.25)
GT-density × Species	-	-22.55 (-70.36, 26.12)	-23.80 (-54.35, 7.03)	-	7.47 (-9.89, 23.92)	2.27 (-8.19, 13.12)	-	11.45 (-17.82, 40.70)	-0.84 (-18.86, 16.92)
GT-density squared × Species	-	1.24 (-25.79, 29.87)	3.39 (-18.5, 25.01)	-	3.04 (-6.12, 12.17)	6.27 (-0.61, 13.20)	-	-4.75 (-20.78, 10.88)	1.77 (-10.52, 13.82)
BT- × GT-density × Species	-	1.31 (-0.37, 3.13)	1.30 (-0.41, 2.96)	-	-0.37 (-0.98, 0.21)	-0.28 (-0.84, 0.31)	-	-0.09 (-0.54, 0.37)	-0.06 (-0.50, 0.38)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Female ID	8.88 (7.82, 10.06)	8.87 (7.76, 10.12)	8.80 (7.69, 9.98)	1.68 (1.5, 1.88)	1.7 (1.52, 1.92)	1.68 (1.49, 1.88)	0.16 (0.14, 0.19)	0.18 (0.15, 0.21)	0.17 (0.14, 0.20)
Subplot-year ID	10.25 (6.96, 14.11)	10.31 (6.76, 14.41)	10.05 (6.64, 14.12)	0.00 (0.00, 0.00)	0.01 (0.00, 0.01)	0.02 (0.01, 0.03)	0.12 (0.08, 0.18)	0.12 (0.07, 0.17)	0.11 (0.07, 0.17)
Brood ID	-	-	-	-	-	-	0.90 (0.80, 1.02)	0.89 (0.78, 1.00)	0.90 (0.79, 1.01)
Observer ID	-	-	-	-	-	-	0.02 (0.01, 0.03)	0.02 (0.01, 0.03)	0.02 (0.01, 0.03)
Residual	13.62 (12.23, 15.17)	13.78 (12.37, 15.4)	13.70 (12.28, 15.27)	1.26 (1.14, 1.41)	1.25 (1.12, 1.39)	1.25 (1.12, 1.39)	0.63 (0.60, 0.66)	0.63 (0.60, 0.66)	0.63 (0.60, 0.66)

[a] Reference category: estimate is for BTs in areas of average size breeding in the “Low-Low” (BT-GT) box density treatment (all models except c, f, i, l, o, r) under average realized breeding densities (all models except a, d, g, j, m, p) for nestlings with species-average tarsus length (models g-i) at nestling age 15 (BTs) or 14 (GTs) (models g-l).

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Effect of Age (days); age was centered around nestling age 15 (BTs) or 14 (GTs).

[d] Effect of within-species mean-centered tarsus.

[e] Estimate for BTs: the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[f] Estimate for BTs: the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[g] Estimate for BTs: the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[h] Estimate for BTs: the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

[i] Estimate is the difference between BTs (reference) and GTs (main effect), or this difference in the effect of area size, age, or tarsus (interactive effects).

[j] Estimate is the difference between BTs (reference) and GTs in the effect described in footnote e.

[k] Estimate is the difference between BTs (reference) and GTs in the effect described in footnote f.

[l] Estimate is the difference between BTs (reference) and GTs in the effect described in footnote g.

[m] Estimates are differences between BTs (reference) and GTs in effects described in footnote h.

Table 3. *Continued*

	Nestling body size (n= 3305 nestlings)			Binary probability to produce fledglings (n= 666 broods)			Number of fledglings (n= 457 broods)		
	(j) Treatment	(k) Treatment and realized breeding density	(l) Realized breeding density	(m) Treatment	(n) Treatment and realized breeding density	(o) Realized breeding density	(p) Treatment	(q) Treatment and realized breeding density	(r) Realized breeding density
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	16.62 (16.35, 16.9)	16.65 (16.31, 17.02)	16.54 (16.39, 16.7)	0.37 (-0.28, 1.02)	0.65 (-0.31, 1.67)	0.95 (0.53, 1.36)	2.00 (1.85, 2.14)	2.10 (1.90, 2.29)	2.05 (1.98, 2.12)
Area [b]	0.06 (-0.07, 0.19)	0.15 (-0.03, 0.34)	0.17 (0.00, 0.33)	0.17 (-0.21, 0.56)	0.40 (-0.19, 0.98)	0.44 (-0.10, 1.01)	-0.05 (-0.12, 0.02)	-0.02 (-0.12, 0.09)	-0.01 (-0.11, 0.08)
Age [c]	0.10 (-0.19, 0.41)	0.15 (-0.15, 0.45)	0.15 (-0.14, 0.44)	-	-	-	-	-	-
Nest box density treatment									
BT-box no [e]	-0.18 (-0.51, 0.15)	-0.27 (-0.63, 0.08)	-	0.85 (-0.08, 1.73)	0.79 (-0.24, 1.75)	-	0.00 (-0.18, 0.19)	-0.04 (-0.25, 0.18)	-
GT-box no [f]	-0.01 (-0.34, 0.32)	-0.06 (-0.55, 0.43)	-	0.96 (-0.04, 1.91)	0.22 (-1.10, 1.57)	-	0.12 (-0.08, 0.31)	-0.02 (-0.30, 0.25)	-
BT- × GT-box no [g]	0.12 (-0.33, 0.56)	0.27 (-0.23, 0.77)	-	-1.29 (-2.58, -0.01)	-0.89 (-2.37, 0.66)	-	-0.11 (-0.34, 0.13)	-0.05 (-0.32, 0.24)	-
Realized breeding density [h]									
BT-density	-	9.36 (-0.89, 19.75)	7.33 (-1.65, 16.80)	-	2.79 (-11.60, 16.90)	7.55 (-5.19, 20.11)	-	-0.17 (-2.31, 1.98)	-0.07 (-2.15, 1.95)
BT- density squared	-	-1.47 (-10.10, 6.77)	-1.80 (-9.15, 5.85)	-	2.14 (-9.50, 13.94)	1.82 (-8.70, 12.11)	-	0.55 (-1.15, 2.27)	0.30 (-1.20, 1.81)
GT-density	-	-2.84 (-17.62, 11.83)	-0.88 (-10.02, 8.11)	-	4.38 (-16.42, 26.38)	1.68 (-12.06, 14.79)	-	1.91 (-1.2, 5.12)	1.42 (-0.59, 3.48)
GT-density squared	-	0.49 (-7.57, 8.50)	-0.32 (-6.56, 6.04)	-	-10.98 (-22.16, 0.09)	-9.81 (-18.96, -0.40)	-	-0.85 (-2.5, 0.76)	-0.88 (-2.24, 0.42)
BT- × GT-density	-	0.02 (-0.23, 0.26)	0.05 (-0.19, 0.27)	-	0.09 (-0.61, 0.77)	0.02 (-0.69, 0.71)	-	-0.05 (-0.18, 0.09)	-0.06 (-0.19, 0.08)
Species [i]	2.69 (2.43, 2.94)	2.98 (2.61, 3.33)	2.83 (2.71, 2.96)	0.68 (-0.10, 1.45)	0.30 (-0.91, 1.47)	-0.32 (-0.78, 0.16)	-0.13 (-0.31, 0.05)	-0.19 (-0.43, 0.05)	-0.14 (-0.22, -0.05)
Area × Species [i]	-0.14 (-0.26, -0.01)	-0.10 (-0.29, 0.09)	-0.16 (-0.33, 0.01)	-0.15 (-0.59, 0.30)	-0.41 (-1.05, 0.24)	-0.46 (-1.08, 0.16)	0.05 (-0.04, 0.14)	0.01 (-0.11, 0.13)	0.01 (-0.11, 0.13)
Age × Species [i]	0.33 (-0.07, 0.72)	0.33 (-0.08, 0.74)	0.30 (-0.10, 0.7)	-	-	-	-	-	-
Nest box density treatment									
BT-box no × Species [j]	0.19 (-0.15, 0.51)	0.12 (-0.23, 0.50)	-	-0.88 (-1.99, 0.23)	-0.81 (-2.10, 0.51)	-	0.05 (-0.18, 0.28)	0.08 (-0.18, 0.34)	-
GT-box no × Species [k]	-0.05 (-0.38, 0.27)	-0.36 (-0.87, 0.15)	-	-1.26 (-2.34, -0.16)	-0.65 (-2.27, 0.92)	-	-0.07 (-0.29, 0.14)	0.05 (-0.3, 0.37)	-
BT- × GT-box no × Species [l]	-0.16 (-0.58, 0.29)	-0.09 (-0.61, 0.41)	-	1.03 (-0.42, 2.47)	0.62 (-1.24, 2.42)	-	-0.05 (-0.34, 0.24)	-0.07 (-0.42, 0.28)	-
Realized breeding density [m]									
BT-density × Species	-	-6.51 (-17.08, 4.01)	-3.29 (-12.88, 6.50)	-	-1.36 (-17.55, 14.54)	-3.81 (-18.46, 10.83)	-	-0.46 (-3.02, 2.23)	-0.28 (-2.81, 2.22)
BT-density squared × Species	-	8.19 (-0.37, 17.18)	4.22 (-3.75, 12.08)	-	3.58 (-9.41, 17.37)	1.15 (-10.86, 13.21)	-	-0.58 (-2.96, 1.67)	-0.36 (-2.30, 1.63)
GT-density × Species	-	17.92 (3.12, 33.52)	6.85 (-2.44, 15.61)	-	-6.60 (-30.21, 16.90)	-11.32 (-25.77, 3.74)	-	-1.30 (-4.95, 2.55)	-1.24 (-3.65, 1.15)
GT-density squared × Species	-	-4.25 (-12.25, 3.78)	0.90 (-5.64, 7.10)	-	6.61 (-5.96, 19.89)	7.60 (-2.72, 17.72)	-	0.65 (-1.45, 2.82)	0.65 (-1.06, 2.30)
BT- × GT-density × Species	-	-0.24 (-0.48, 0.01)	-0.27 (-0.50, -0.04)	-	0.37 (-0.46, 1.19)	0.40 (-0.39, 1.20)	-	-0.04 (-0.20, 0.13)	-0.04 (-0.20, 0.12)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Female ID	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	NE [n]	NE [n]	NE [n]	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Subplot-year ID	0.04 (0.03, 0.07)	0.03 (0.02, 0.05)	0.03 (0.02, 0.05)	0.19 (0.12, 0.28)	0.11 (0.07, 0.17)	0.15 (0.09, 0.22)	0.01 (0.00, 0.01)	0.00 (0.00, 0.01)	0.01 (0.00, 0.01)
Brood ID	0.32 (0.28, 0.36)	0.31 (0.28, 0.35)	0.31 (0.28, 0.35)	-	-	-	-	-	-
Observer ID	0.04 (0.03, 0.06)	0.04 (0.03, 0.06)	0.04 (0.03, 0.07)	-	-	-	-	-	-
Residual	0.33 (0.31, 0.35)	0.33 (0.32, 0.35)	0.33 (0.31, 0.35)	$\pi^{2/3}$ [o]	$\pi^{2/3}$ [o]	$\pi^{2/3}$ [o]	0.13 (0.15, 0.11) [o]	0.12 (0.14, 0.1) [o]	0.12 (0.13, 0.11) [o]

[a] Reference category: estimate is for BTs in areas of average size breeding in the “Low-Low” (BT-GT) box density treatment (all models except c, f, i, l, o, r) under average realized breeding densities (all models except a, d, g, j, m, p) for nestlings with species-average tarsus length (models g-i) at nestling age 15 (BTs) or 14 (GTs) (models g-l).

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Effect of Age (days); age was centered around nestling age 15 (BTs) or 14 (GTs).

[d] Effect of within-species mean-centered tarsus.

[e] Estimate for BTs: the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[f] Estimate for BTs: the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[g] Estimate for BTs: the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[h] Estimate for BTs: the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

[i] Estimate is the difference between BTs (reference) and GTs (main effect), or this difference in the effect of area size, age, or tarsus (interactive effects).

[j] Estimate is the difference between BTs (reference) and GTs in the effect described in footnote e.

[k] Estimate is the difference between BTs (reference) and GTs in the effect described in footnote f.

[l] Estimate is the difference between BTs (reference) and GTs in the effect described in footnote g.

[m] Estimates are differences between BTs (reference) and GTs in effects described in footnote h.

[n] Female ID was excluded as this prevented model convergence (NE).

[o] Residual variance fixed to $\pi^{2/3}$ (following Nakagawa & Schielzeth, 2010) for the models of the binary probability of producing fledglings (models m-o) or defined as ln(1/exp(βo)+1) (following Nakagawa & Schielzeth, 2010) for the models of number of fledglings (models p-r).

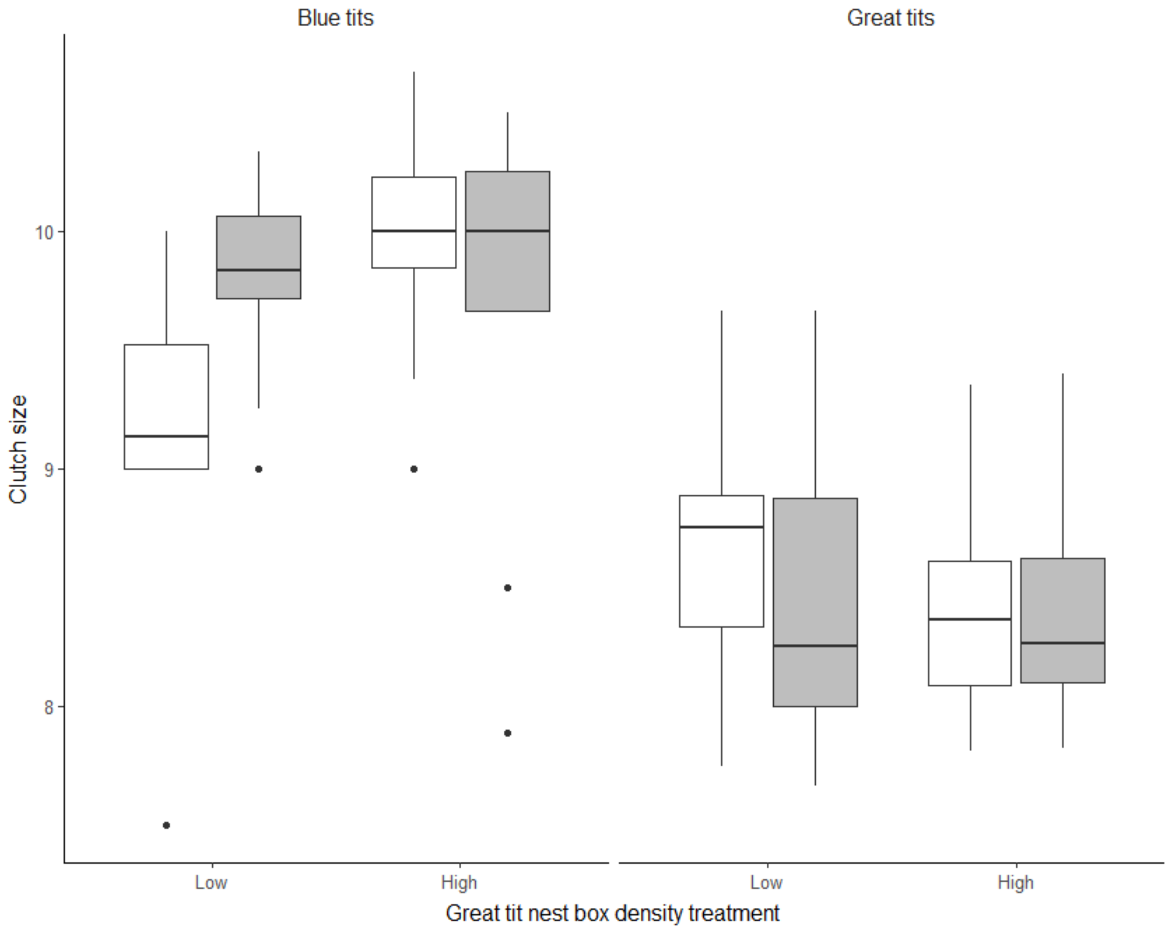


Figure 6. Boxplots showing the clutch sizes for blue (left) and great tits (right) breeding in “Low” vs. “High” great tit box density treatments. Boxplots in white vs. grey are for “Low” vs. “High” blue tit-box density treatment groups. The boxplots illustrate five summary statistics: the central line (black bar) of the box represents median value, box limits represent lower and upper quartiles upper (25th and 75th percentile, respectively), and whiskers represent 1.5 × the interquartile range (IQR) below and above the lower and upper quartiles, respectively. Outliers are considered as values more than 1.5 times the IQR and marked with black dots.

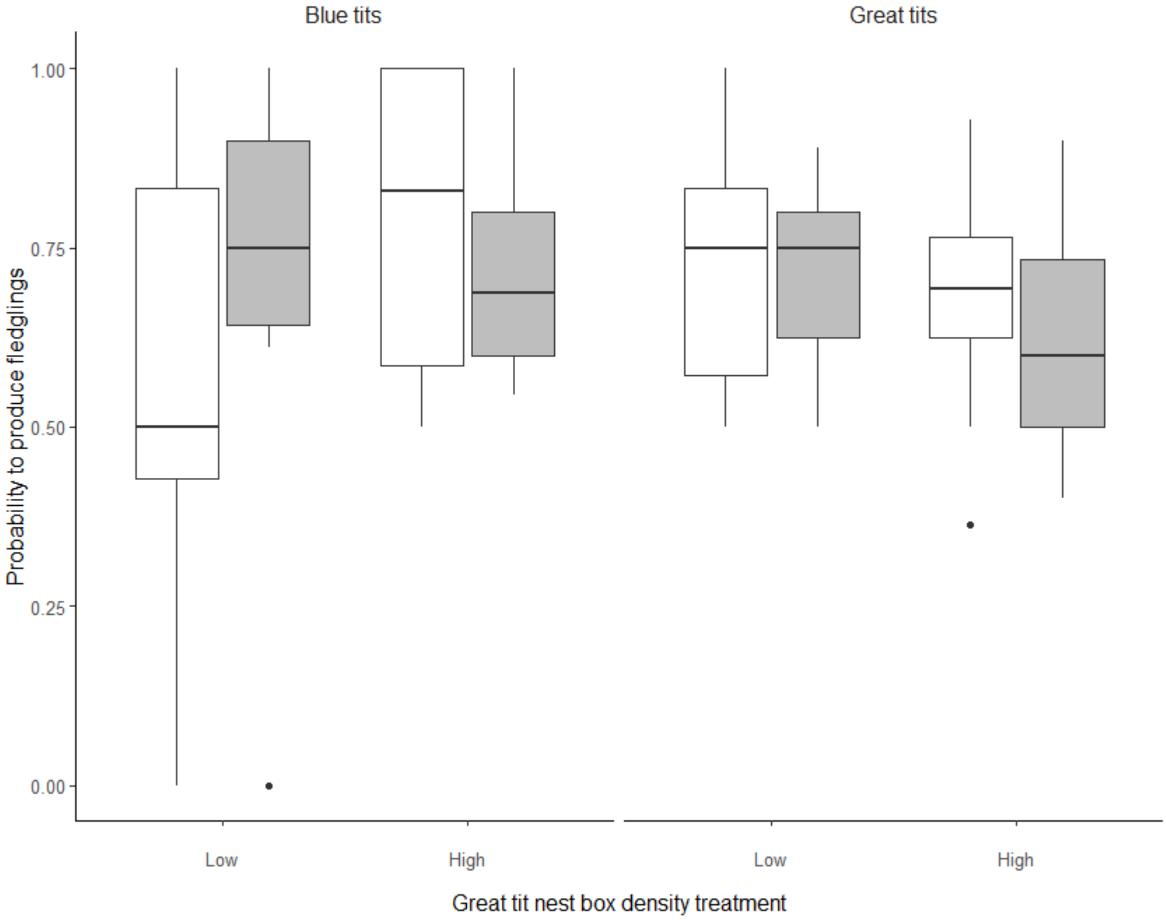


Figure 7. Boxplots showing the binary probability to produce fledglings for blue (left) and great tits (right) breeding in “Low” vs. “High” great tit-box density treatments. Boxplots in white vs. grey are for “Low” vs. “High” blue tit-box density treatment groups. The boxplots illustrate five summary statistics: the central line (black bar) of the box represents median value, box limits represent lower and upper quartiles upper (25th and 75th percentile, respectively), and whiskers represent 1.5 × the interquartile range (IQR) below and above the lower and upper quartiles, respectively. Outliers are considered as values more than 1.5 times the IQR and marked with black dots.

3.2.2.2 Residual (density-independent) treatment effects

Expanding our models by additionally including the realized (linear, quadratic, and interactive) effects of breeding densities of BTs and GTs did not lead to any major changes in estimated effect sizes for the effects of BT-, GT-, or BT- × GT-box number (Table 3b, e, h, k, n, and q). For example, the (non-significant) positive effect of GT-box numbers on BT clutch size when BT-numbers were low had a similar effect size when comparing models where effects of realized breeding densities were controlled for (Table 3e) vs. not controlled for (Table 3d). This suggests that those species-specific increases in clutch size did not exist because of an indirect pathway (where treatment affected density, and density in turn affected clutch size). For treatment effects on the probability of producing fledglings (detailed above), all relevant effects sizes (point estimates for GT-, BT- × GT-, and GT-box number × Species in Table 3m vs. 3n) did drop substantially (i.e., were closer to zero), suggesting those may have been mediated via treatment effect on realized densities.

3.2.2.3 Overall effects of realized breeding density

As a final step, we modelled the overall effects of realized breeding density. This was achieved by modelling reproductive parameters as a function of linear, quadratic, and interactive effects of BT- and GT-densities without controlling for box number treatment effects. In BTs, we found very little evidence for effects of intra- or inter-specific density (Table 3c, f, i, l, o, and r). We only found a significant negative quadratic effect of GT-realized breeding densities (effect of GT squared density; Table 3o) on the probability to produce fledglings. We found the same effect size of this variable (GT squared density) when treatments were controlled for (Table 3n), implying this effect was unrelated to treatment. The effects of intra- or inter-specific breeding densities also did not differ between the two species as two-way interactions between species and linear, quadratic, or interactive effects of BT and GT breeding densities were not supported. The only exception was nestling body size, which was more negatively affected by the interaction between BT- × GT breeding density in GTs compared to BTs (BT- × GT-density × Species interaction; Table 3l). Species-specific analyses suggested that this negative effect existed in GTs (BT- × GT-density effect: Table S13f) but not in BTs (Table S13c). A plot of this interactive effect showed that the realized breeding density of GTs negatively affected nestling body size of GTs when BT densities were high but not when BT densities were low (Figure 8).

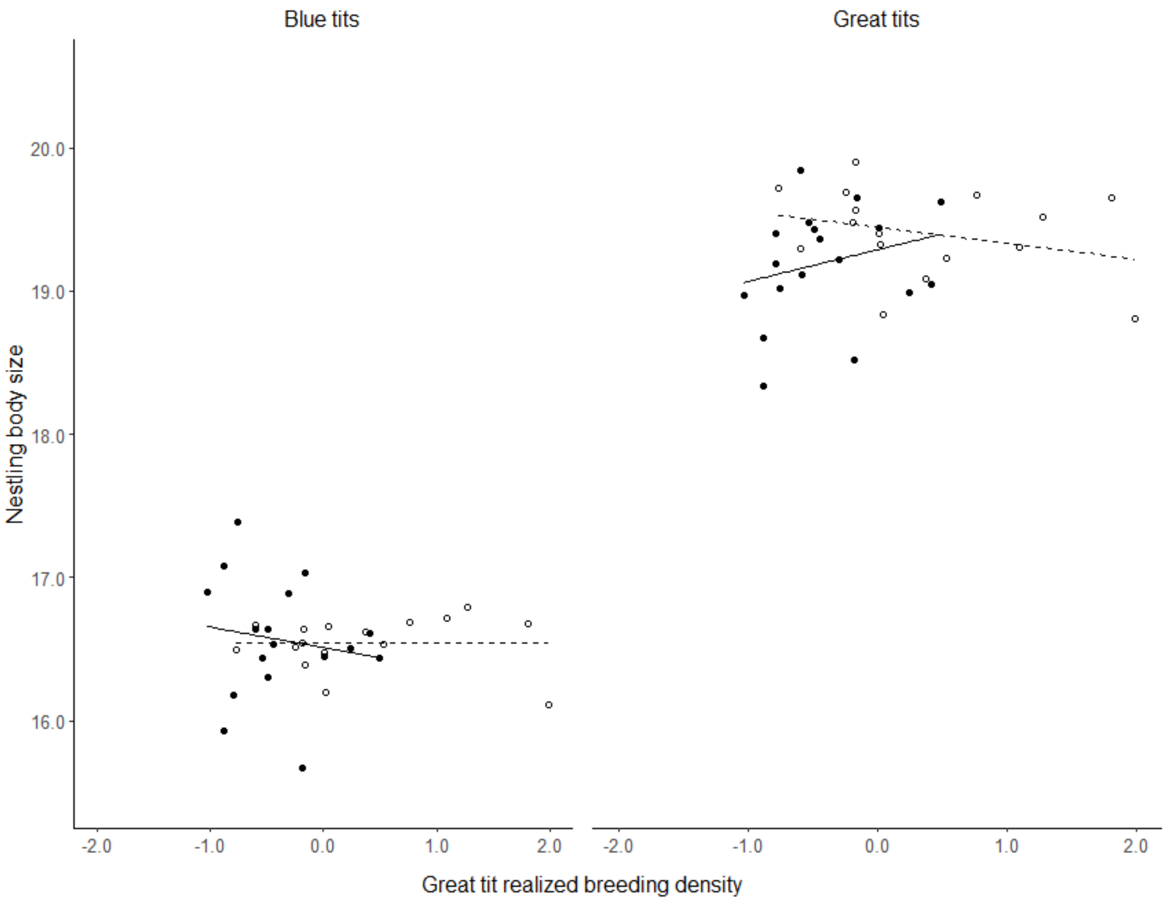


Figure 8. Illustration of the interaction effect between (mean-centered) realized breeding densities of blue and great tits on the nestling body size (tarsus length) for blue tits (left) and great tits (right). For illustrational purposes only we plotted realized great tit breeding density on the x-axis and then categorized birds breeding under the 50% lowest (solid points and regression lines) vs. 50% highest (open points and dashed regression lines) realized blue tit breeding densities. Dots are mean values for nestling body size for each unique combination of subplot and year.

3.2.3 Area size effects

The size of the areas varied unintentionally across subplots despite our effort to provide nest boxes in a regular grid; this was due to the positioning of forest roads, gaps, and edges (Figure S1). Consequently, in our models, we included the mean-centered value of area size (average value: 6.48 ha). Doing so resulted in various biological patterns worth reporting. First, BTs produced smaller clutches in subplots where the nest boxes were fitted over a larger area (Table 3d). Second, GTs were more strongly affected than BTs by area size (interactions Area × Species). Particularly, GTs started laying later (Table 3a), produced bigger clutches (Table 3d) and smaller nestlings (Table 3j) with an increase in unit area size relative to BTs. Third, the number of BT (not GT) breeding pairs was significantly lower in relatively larger areas (compare Table 2a vs Table 2d). Altogether, this suggests that area size was linked with perceived and realized habitat quality and that GTs were more strongly affected by a subplot's habitat quality compared to BTs.

3.3 Aggressive behaviour towards con- vs. heterospecifics in blue and great tits

3.3.1 Simulated territorial intrusions effects on aggressive behaviour

The minimal approach distance to a cons- vs. a heterospecific mount differed between BT and GT males. Overall, BT males approached conspecific mounts closer than heterospecific ones. BT males tested for the first time approached a conspecific (BT) mount closer (main effect of current test, Table 4a) compared to a heterospecific (GT) mount (Figure 9). Similarly, during the second test, when BT males were first tested using a conspecific (BT) mount, individuals more closely approached a conspecific (BT) compared to a heterospecific (GT) mount (though the effect was not significant; Table 4b). Similar to the BT males tested for the first time, GT approached a conspecific mount (GT) close than a heterospecific (BT) one (though the effect was not significant; Table 4c). Nevertheless, during the second test, GT males that had previously between tested using a heterospecific (BT) mount, approached less closely to a conspecific (GT) mount (though this effect was not significant; Table 4d). To understand the overall effects of the species of the current test and the previous test (during the second test), we reran our models after removing non-significant interactions. The same patterns were observed in BT (main effect current test; Table S16a, b) and GT males (main effect current test; Table S16c, d).

We observed similar patterns when we ran the models that included both species combined (Table S16e, f). First, BTs (vs. GTs) approached a BT mount closer (intercept vs. main effect of species; Table S16e, f). Second, during the second test we detected a significant effect of mount species on the minimal approach distance of BT males. Differences in significance probably occurred because this analyses used twice the amount of data compared to first-test and second-test analyses. Interestingly, these models detected species-specific differences on the minimal approach distance between a conspecific and a heterospecific mounts. This means that the species differed in intruder species-related plasticity. GT males had similar approach distances to both conspecific (GT) and a heterospecific (BT) mounts, but BT approached conspecific (BT) mounts closer compared to the heterospecific (GT) mounts (interaction effect current test × GT; Table S16e, f, Figure 9). We did not detect any effects of the nest box density treatment nor of the realized breeding densities of either species on minimal approach distance (Table 5).

Table 4. Estimated effect sizes and 95% credible intervals (CIs) for predictors of the minimal approach distance of blue tit (BT) males, great tit (GT) males, and males of both species during the first test (models a, c and e, respectively) and during the second test (models b, d, and f, respectively) to a con- vs. a heterospecific mount. Estimates are shown in a latent scale. Parameters were estimated using LMMs with a Gaussian error distribution.

	Blue tits		Great tits		Both species	
	(a) First test (n = 137 observations)	(b) Second test (n = 142 observations)	(c) First test (n = 151 observations)	(d) Second test (n = 166 observations)	(e) First test (n = 288 observations)	(f) Second test (n = 308 observations)
Fixed effects	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	-1.72 (-2.17, -1.29)	-1.56 (-2.17, -0.97)	-2.83 (-3.41, -2.26)	-2.72 (-3.26, -2.23)	-1.77 (-2.24, -1.30)	-1.45 (-1.93, -0.98)
Mount species						
Current test [b]	-0.54 (-1.03, -0.04)	-0.55 (-1.30, 0.19)	0.17 (-0.54, 0.87)	-0.17 (-0.76, 0.45)	-0.64 (-1.22, -0.04)	-0.69 (-1.29, -0.05)
Previous test [c]	-	-0.30 (-0.85, 0.25)	-	-0.14 (-0.68, 0.42)	-	-0.36 (-0.95, 0.23)
Current- × previous-test [d]	-	0.17 (-0.69, 1.05)	-	0.42 (-0.36, 1.18)	-	0.22 (-0.64, 1.11)
Focal species						
GT [e]	-	-	-	-	-1.05 (-1.51, -0.62)	-1.39 (-2.00, -0.83)
Mount species × focal species						
Current test × GT [f]	-	-	-	-	0.78 (0.15, 1.38)	0.64 (-0.19, 1.43)
Previous test × GT [g]	-	-	-	-	-	0.32 (-0.51, 1.18)
Current- × previous-test × GT [h]	-	-	-	-	-	0.00 (-1.14, 1.19)
Random effects	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Male ID	0.00 (0.00, 0.00)	0.41 (0.30, 0.54)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.12 (0.09, 0.14)
Subplot-year ID	0.00 (0.00, 0.00)	0.01 (0.00, 0.02)	0.22 (0.11, 0.37)	0.07 (0.04, 0.12)	0.11 (0.06, 0.17)	0.00 (0.00, 0.00)
Mount ID	0.05 (0.02, 0.10)	0.12 (0.05, 0.21)	0.22 (0.1, 0.38)	0.02 (0.01, 0.04)	0.14 (0.06, 0.24)	0.05 (0.02, 0.10)
Song ID	0.00 (0.00, 0.00)	0.29 (0.17, 0.42)	0.20 (0.13, 0.30)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.02 (0.01, 0.03)
Observer ID	0.25 (0.11, 0.46)	0.27 (0.13, 0.47)	0.00 (0.00, 0.00)	0.24 (0.12, 0.39)	0.08 (0.04, 0.14)	0.14 (0.07, 0.24)
Residual	1.72 (1.34, 2.17)	0.81 (0.64, 1.02)	1.30 (1.04, 1.65)	1.44 (1.15, 1.79)	1.62 (1.37, 1.91)	1.42 (1.22, 1.67)
Sample size	n	n	n	n	n	n
Male ID	128	131	140	157	268	288
Subplot-year ID	24	23	24	24	24	24
Mount ID	14	14	13	14	14	15
Song ID	46	47	49	49	50	50
Observer ID	16	16	16	15	17	16

[a] Reference category: estimate is for the minimal approach distance of a (BT; both species models) male presented with a BT mount during the current test (all modes), when the previous mount was a BT (second test models). Approach distance was multiplied by "-1" prior to the analyses.

[b] Estimate (for BT males; both species models) is the effect during the current test of a GT mount relative to a BT mount (all models) when the previous mount species was a BT (second test models).

[c] Estimate (for BT males; both species models) is the effect during the previous test of a GT mount relative to a BT mount when the current test mount species is a BT (second test models).

[d] Estimate (for BT males; both species models) is the effect during the current test of a GT mount relative to a BT mount when the previous mount species was a GT expressed as a difference from the same effect observed when the previous mount species was a BT.

[e] Estimate is the difference between BTs (reference) and GTs (main effect).

[f] Estimate is the difference between BTs (reference) and GTs described in the footnote b.

[g] Estimate is the difference between BTs (reference) and GTs described in the footnote c.

[h] Estimate is the difference between BTs (reference) and GTs described in the footnote d.

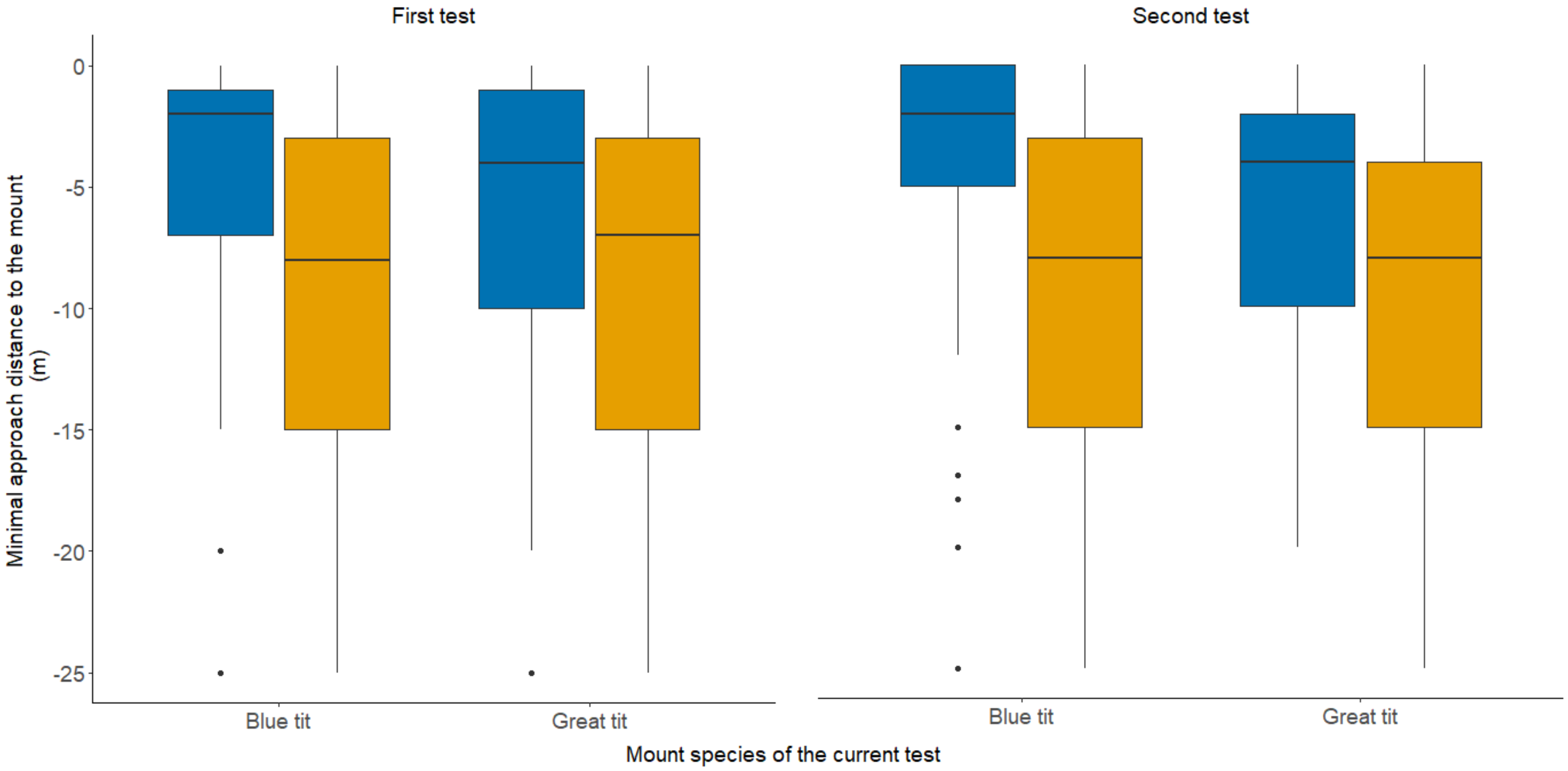


Figure 9. Boxplots showing the effects of the mount species used in the current test (x-axis) on the minimal approach distance to the mount (multiplied with -1) (y-axis) of blue tit (blue boxplots) and great tit (yellow boxplots) males. Boxplots show the distance to the mount during the first (left panel) and second test (right panel). The boxplots provide a visual representation of five summary statistics. The central line within the box (black bar) illustrates the median value. The limits of each box correspond to the lower and upper quartiles, which represent the 25th and 75th percentiles, respectively. The whiskers extend up to 1.5 times the interquartile range (IQR) below and above the lower and upper quartiles. Any data points that fall outside this range and exceed 1.5 times the IQR are considered outliers (black dots).

Table 5. Estimated effect sizes and 95% credible intervals (CIs) for predictors of male's minimal approach distance to the mount in the current test of blue tits (a-c), great tits (d-f), and both species (g-i) for models fitting (main and interaction effects of) the number (no) of blue tit (BT) and great tit (GT) boxes (nest box density treatment; a, d, g), realized breeding density (c, f, i) or both (b, e, h). Estimates are shown in a latent scale. Parameters were estimated using LMMs with a Gaussian error distribution.

	Blue tits (n = 279 observations)			Great tits (n = 317 observations)			Both species (n = 596 observations)		
	(a) Treatment	(b) Treatment and realized breeding density	(c) Realized breeding density	(d) Treatment	(e) Treatment and realized breeding density	(f) Realized breeding density	(g) Nest box density treatment	(h) Treatment and realized breeding density	(i) Realized breeding density
Fixed effects	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	-1.66 (-2.21, -1.14)	-1.45 (-2.1, -0.77)	-1.67 (-2.01, -1.32)	-2.74 (-3.32, -2.15)	-2.56 (-3.31, -1.80)	-2.84 (-3.29, -2.40)	-1.60 (-2.13, -1.06)	-1.39 (-2.06, -0.76)	-1.63 (-2.00, -1.28)
Mount species									
Current test [b]	-0.53 (-0.96, -0.11)	-0.54 (-0.94, -0.15)	-0.54 (-0.94, -0.13)	0.08 (-0.33, 0.51)	0.10 (-0.30, 0.51)	0.09 (-0.32, 0.51)	-0.59 (-0.96, -0.21)	-0.60 (-0.97, -0.24)	-0.60 (-0.96, -0.23)
Nest box density treatment									
BT-box no [c]	0.03 (-0.49, 0.56)	-0.05 (-0.73, 0.61)	-	-0.20 (-0.86, 0.46)	-0.42 (-1.24, 0.35)	-	-0.03 (-0.57, 0.54)	-0.05 (-0.71, 0.62)	-
GT-box no [d]	-0.10 (-0.68, 0.52)	-0.23 (-1.01, 0.54)	-	-0.14 (-0.77, 0.50)	-0.39 (-1.27, 0.51)	-	-0.09 (-0.68, 0.52)	-0.23 (-0.96, 0.53)	-
BT- × GT-box no [e]	-0.04 (-0.79, 0.73)	-0.24 (-1.34, 0.92)	-	0.22 (-0.63, 1.11)	0.54 (-0.59, 1.64)	-	-0.04 (-0.82, 0.74)	-0.24 (-1.28, 0.86)	-
Realized breeding density [f]									
BT-density	-	0.79 (-3.97, 5.70)	0.30 (-3.70, 4.27)	-	3.34 (-4.16, 10.64)	3.30 (-2.24, 8.81)	-	0.73 (-7.50, 8.29)	1.23 (-5.98, 7.74)
BT- density squared	-	-0.08 (-5.35, 5.36)	-1.14 (-5.39, 3.11)	-	1.88 (-4.39, 8.16)	0.70 (-4.09, 5.66)	-	-0.63 (-8.33, 6.80)	-2.32 (-9.30, 4.79)
GT-density	-	4.21 (-5.78, 14.42)	0.99 (-4.95, 6.59)	-	0.23 (-11.42, 11.63)	-0.83 (-8.11, 6.66)	-	6.81 (-6.70, 20.47)	1.86 (-7.07, 10.74)
GT-density squared	-	0.34 (-5.19, 6.00)	0.60 (-3.24, 4.59)	-	-3.46 (-10.05, 3.34)	-2.56 (-7.53, 2.31)	-	-0.32 (-7.88, 7.13)	0.12 (-5.58, 5.70)
BT- × GT-density	-	-0.07 (-0.52, 0.35)	-0.13 (-0.52, 0.26)	-	-0.04 (-0.49, 0.41)	0.01 (-0.42, 0.45)	-	-0.12 (-0.55, 0.28)	-0.17 (-0.56, 0.22)
Focal species									
GT [g]	-	-	-	-	-	-	-1.21 (-1.77, -0.61)	-1.10 (-1.87, -0.39)	-1.19 (-1.51, -0.85)
Current test × GT [h]	-	-	-	-	-	-	0.72 (0.29, 1.14)	0.75 (0.36, 1.16)	0.73 (0.30, 1.13)
Nest box density treatment									
BT-box no × Species [i]	-	-	-	-	-	-	-0.15 (-0.89, 0.56)	-0.37 (-1.21, 0.47)	-
GT-box no × Species [j]	-	-	-	-	-	-	0.01 (-0.71, 0.71)	-0.21 (-1.13, 0.71)	-
BT- × GT-box no × Species [k]	-	-	-	-	-	-	0.26 (-0.66, 1.18)	0.78 (-0.51, 2.04)	-
Realized breeding density [l]									
BT-density × Species	-	-	-	-	-	-	-	3.13 (-6.73, 12.95)	2.34 (-5.58, 10.26)
BT-density squared × Species	-	-	-	-	-	-	-	5.43 (-4.27, 15.18)	5.17 (-3.19, 13.54)
GT-density × Species	-	-	-	-	-	-	-	-2.24 (-18.45, 13.28)	0.57 (-9.36, 10.87)
GT-density squared × Species	-	-	-	-	-	-	-	-4.30 (-13.28, 4.51)	-3.34 (-10.2, 3.33)
BT- × GT-density × Species	-	-	-	-	-	-	-	-0.01 (-0.53, 0.48)	0.09 (-0.37, 0.54)
Random effects	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Male ID	0.21 (0.16, 0.28)	0.19 (0.14, 0.25)	0.20 (0.15, 0.26)	0.25 (0.20, 0.31)	0.26 (0.20, 0.32)	0.26 (0.20, 0.32)	0.28 (0.23, 0.33)	0.28 (0.24, 0.33)	0.27 (0.23, 0.32)
Subplot-year ID	0.02 (0.01, 0.03)	0.07 (0.04, 0.12)	0.04 (0.02, 0.06)	0.16 (0.08, 0.26)	0.17 (0.09, 0.29)	0.15 (0.08, 0.24)	0.04 (0.02, 0.06)	0.04 (0.02, 0.07)	0.04 (0.02, 0.06)
Mount species ID	0.06 (0.03, 0.11)	0.05 (0.02, 0.10)	0.06 (0.02, 0.10)	0.07 (0.03, 0.13)	0.07 (0.03, 0.13)	0.07 (0.03, 0.12)	0.04 (0.01, 0.06)	0.03 (0.01, 0.06)	0.03 (0.01, 0.06)
Song ID	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.01 (0.01, 0.02)	0.00 (0.00, 0.00)	0.01 (0.01, 0.02)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Observer ID	0.12 (0.06, 0.21)	0.12 (0.06, 0.21)	0.12 (0.06, 0.21)	0.22 (0.12, 0.33)	0.21 (0.11, 0.32)	0.21 (0.11, 0.32)	0.13 (0.07, 0.21)	0.13 (0.07, 0.21)	0.13 (0.06, 0.21)
Residual	1.44 (1.23, 1.71)	1.46 (1.23, 1.74)	1.46 (1.23, 1.72)	1.18 (1.01, 1.38)	1.18 (1.01, 1.39)	1.18 (1.01, 1.38)	1.35 (1.20, 1.51)	1.35 (1.21, 1.52)	1.35 (1.2, 1.51)
Sample size	n			n			n		
Male ID	153			194			347		
Subplot-year ID	24			24			24		
Mount species ID	15			14			15		
Song ID	50			50			50		
Observer ID	16			16			17		

[a] Reference category: estimate is for the minimal approach distance of a (BT; both species models) male presented with a BT mount during the current test (all modes) in the "Low-Low" (BT-GT) box density treatment (all models except c, f, i) under average realized breeding densities (all models except a, d, g). Approach distance was multiplied by "-1" prior to the analyses.

[b] Estimate (for BT males; both species models) is the effect during the current test of a GT mount relative to a BT mount.

[c] Estimate (for BT males; both species models): the effect of the "High" BT-box treatment relative to the "Low" BT-box treatment when GT-box numbers were "Low".

[d] Estimate (for BT males; both species models): the effect of the "High" GT-box treatment relative to the "Low" GT-box treatment when BT-box numbers were "Low".

[e] Estimate (for BT males; both species models): the effect of the "High" BT-box treatment relative to the "Low" BT-treatment for the "High" GT-treatment expressed as a difference from the same effect observed in the "Low" GT-treatment.

[f] Estimate (for BT males; both species models): the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

[g] Estimate is the difference between BTs (reference) and GTs (main effect).

[h] Estimate is the difference between BTs (reference) and GTs described in the footnote b.

[i] Estimate is the difference between BTs (reference) and GTs described in the footnote c.

[j] Estimate is the difference between BTs (reference) and GTs described in the footnote d.

[k] Estimate is the difference between BTs (reference) and GTs described in the footnote e.

[l] Estimate is the difference between BTs (reference) and GTs described in the footnote f.

CHAPTER 4

4. Discussion

Why do individuals differ in morphology, physiology, and behaviour? Why are these traits often integrated into syndromes? How and what can contribute to maintain among-individual differences in behaviour? Over the past decades, behavioural ecologists have developed theories that predict why individuals vary in their behaviour. Particularly, from an adaptive perspective, theory postulates that individuals differ in their behaviour because they differ in their internal state, and then are expected to adjust their behaviour in an adaptive way (Dingemanse and Wolf 2010; Wolf and Weissing 2010; Sih et al. 2015). From this perspective, individuals with a higher metabolism and better condition should be risk-takers (i.e., bolder and aggressive). These bolder and more aggressive individuals can then gain access and secure (higher-quality) resources that would allow them to maintain their condition. Contrary, individuals in poor condition are then expected to be risk-adverse and do not have access to ample resources (Sih et al. 2015). Within the same adaptive framework, contemporary models have further considered the role that the (social and non-social) environmental state should have in shaping and maintaining among-individual differences in behaviour (Dingemanse and Wolf 2010; Wolf and Weissing 2010; Sih et al. 2015).

This thesis aimed to test various elements of adaptive theory. First, I aimed to test whether patterns of phenotypic integration predicted by the theory characterized two different species (BTs vs. GTs), populations (Forstenrieder Park vs. Starnberg), and sexes (females vs.

males). Then, I tested whether various features of the social environmental state also played a role in shaping and maintain among-individual differences. Particularly, my second aim was to testing whether changes in the con- and heterospecific realized breeding density affected reproductive parameters in BTs and GTs. Finally, my third aim was to investigate whether competitor species causes plasticity in behaviour, and whether those affects vary with con- and heterospecific breeding density. The results obtained for each of research aims will be discussed in detail in the following sections.

4.1 Phenotypic integration among species, populations, and sexes

Based on Gaona-Gordillo et al (2023)

This study was designed to test whether patterns of phenotypic integration between morphology, behaviour, and physiology were generally representative of tit populations vs. specific to species, population, or the sex under study. We therefore assessed and compared patterns of phenotypic integration among species (BTs vs. GTS), populations (Forstenrieder Park vs. Starnberg), and sexes (males vs. females). We found that certain components, specifically, the integration of physiological and morphological traits, was highly conserved across datasets. This conservation occurs when the strength of the trait correlations is qualitatively similar between datasets (Dingemanse, Barber, et al. 2020). In all cases, i) morphological traits were positively correlated, which supported the notion that each represented an "observable" reflection of the animals' size. Each dataset further exhibited strong support for the existence of ii) size-dependent physiology, and iii) size-independent effect of body mass on breathing rate. Larger animals breathed slower, but animals that were heavy for their size breathed faster. Our analyses implied that these structures of phenotypic integration were generally representative rather than specific to characteristics of the study species, population, or sex.

The structure of phenotypic trait integration was thus shared among species, populations, and sexes, but this was only true for the integration of physiology and morphology. Any links with behaviour were, by contrast, dataset-specific. In some datasets, we found evidence for size-dependent behaviour. Similarly, in some datasets, behaviour and physiology were structured into syndromes independent of size or size-corrected body mass. Importantly, in both cases, the characteristic pattern of phenotypic integration found in one specific combination of species, population, and sex, was not predictive of patterns of phenotypic integration in other datasets. This was because neither pattern of phenotypic integration was overall supported, while both patterns of phenotypic integration varied qualitatively and significantly across datasets. Our uni-moderator random-effects meta-regressions implied that this was not because patterns of behavioural integration differed between species, population, or sexes. This suggests that population differences in study design (see §2.1.1) did not underpin this form of heterogeneity and, thus, that ecological conditions specific to unique combinations of species, populations, and/or sexes are likely drivers of heterogeneous patterns of phenotypic integration between behaviour and morphology, and behaviour and physiology observed in natural tit populations. The specific ecological drivers of the correlational selection pressures that might shape the evolution of these diverse patterns of integration between behaviour and other phenotypic traits now requires study. For example, differences in the strength of competition or risk of predation specific to a unique combination of species, population, or sex might explain this variation. Identifying such drivers goes beyond the scope of this study, as this would require many replicates species and populations for each sex as well as quantitative information on ecological variation between these major biological categories. For other patterns of phenotypic integration, studies of the role of ecological variation may perhaps not be warranted. Specifically, the size-independent relationship between body mass and exploration was neither heterogeneous nor supported overall. This conclusion, notably, assumes that our specific study species and populations were more broadly representative of the average tit species and population.

Patterns of (variation in) phenotypic integration are ultimately underpinned by (variation in) the relative contribution of genetic and environmental effects on trait correlations (Searle 1961). Our finding that correlations between observable expressions of size (tarsus length, bill length, wing length, and body mass) were overall highly preserved suggests that patterns of genetic correlations or early-life environmental effects are largely shared among species, populations, and sexes. However, we did find some differences in how specific components of morphology was integrated differently across species (wing length), populations (bill length) or sexes (tarsus). A next step would thus be to unravel the quantitative genetics background of size- and condition-dependent physiology and to study whether the pattern of genetic integration of these traits facilitates or constraints adaptive evolution (Dochtermann and Dingemanse 2013). This would ultimately require not just an understanding of both the quantitative genetics underpinning (e.g., Dingemanse, Barber et al. 2020) but also insights into the patterns of (correlational) selection in the wild (Lande and Arnold 1983; Schluter 1996).

This combination of quantitative genetics and selection studies will be particularly fruitful when applied to the integration of (exploration) behaviour. From a proximate perspective, there are multiple non-exclusive explanations for why the phenotypic integration between behaviour and morphology, and behaviour and physiology varied among datasets. This is evident when inspecting the mathematical equation demonstrating how different variance components contribute to observed correlations (detailed in Text S4). Specifically, the among-individual correlation r_I between two focal traits that we used as input in our SEMs can differ between datasets because they differ in the amount of additive genetic variance, the influence of permanent environmental effects, genetic trait correlations, or permanent environmental trait correlations. Genetic correlations can result from selection on trait integration ("correlational selection") causing linkage-disequilibrium. Differences in trait correlations among species, populations, or sexes may thus reflect variation in selection pressures (Roff and Fairbairn 2012). Permanent environmental correlations, by contrast, can occur when the two focal traits are underpinned by different—yet correlated—environmental factors, and thus change when correlations between environmental factors are different between unique combinations of species, populations or sexes. These examples imply that a productive next step in furthering our understanding of differences in phenotypic integration among datasets is to estimate the variance components that shape observed correlation structures (Kruuk 2004). In stickleback fish, for example, behavioural correlations were stronger in populations living in the presence of predators (Dingemanse, Dochtermann, et al. 2010) but genetic correlations did not differ between the types of population, neither were they affected by experience with predators during ontogeny (Dingemanse, Barber, et al. 2020). This implied that differences in behavioural correlations were caused by differences in permanent-environmental correlations between predator-naïve and predator-sympatric populations (Dingemanse, Barber, et al. 2020). Such approaches might also be applied to understand why allometric relationships (the integration of observable expression of size) differed somewhat between species, populations, and sexes.

Our study demonstrates the value of study replication in behavioural ecology. We asked whether the same biological patterns of phenotypic integration characterized data collected from the same species and sex across populations, and from different species within the same population. Our analyses demonstrated that certain aspects of phenotypic integration were common to all datasets, which implies that such patterns likely also characterize BT and GT males and females from other populations. The ability to draw such conclusions is an important benefit of replication studies. Similarly, we have demonstrated that the sign and strength of the link with exploration behaviour is highly dataset-specific and not due to main effects of species, population, or sex. In particular, based on a focal study of BTs or GTs, we cannot predict how animal personality is related to structural size or body condition in other populations of those species if populations differ in ecology. An important obstacle in interpreting meta-analytical differences and similarities is that datasets often differ in methodology (Nakagawa and Santos 2012). In our case, body mass was measured with different devices (digital scale vs. Pesola spring balance) and wing length was measured slightly different (maximum chord vs. feather length, see §2.1.1), which could explain why their respective links with the latent variable reflective of “size” were somewhat heterogeneous across datasets. However, we did not find population differences in path coefficients involving body mass or wing length. This finding corroborates our assumption that our alternative measures of wing length measured the same trait, as did our alternative measures of body mass. We otherwise used exactly the same methodology to measure our physiological and behavioural traits. Another obstacle in comparing published datasets is that estimates obtained from the literature are typically derived from differently structured models. We avoided this problem by fitting the same structure to all six datasets (Figure 3). This approach, however, does constrain our ability to compare our findings with previously published results coming from different analyses. Specifically, in male GTs from the Starnberg population, (Moiron et al. 2019) found evidence for size- and condition-dependent behaviour. Our analyses of the same dataset implied that neither link was present. The difference in results of analyses that were based on the very same dataset can likely be explained by our decision to use breathing rate in the current study rather than aggression as in (Moiron et al. 2019). We took this decision because the former was measured for all datasets (i.e. sexes) whereas the latter was not (see §2.1.1). This could mean that among-individual correlations between exploration behaviour and size and between exploration behaviour and size-independent body mass vary depending on whether correlations with aggression vs. breathing rate are accounted for. This underlines the need for methodical robustness in analyses as an insightful tool for testing the generality of published results.

4.2 Effects of intra- and interspecific competition on reproductive parameters of blue and great tits

This study manipulated the number of BT and GT nest boxes to experimentally affect inter- and intra-specific competitive regimes, and thereby test for the strength of intra- and inter-specific competition induced by density-dependent processes during reproduction. We expected—and found—that experimental increases in a species' nest box type increased its realized breeding density. We further expected—but did not find—that treatment-induced increases in realized breeding density negatively affected key reproductive parameters of the two species. Specifically, we had expected that treatment-induced increases in realized breeding densities would lead to increases in competition for resources, whether within or among species, and hence decreases in reproductive performance (Dhondt 1977; Dhondt and Adriaensen 1999). We finally expected—but did not find—that the strength of negative density-dependence was stronger within species (intra-specific competition) vs. among species (inter-specific competition), and that the latter form of competition was stronger in GTs compared to BTs (Dhondt 1977; Minot 1981; Sasvári et al. 1987; Dhondt 2010). Our experiment also did not provide evidence in favour of this prediction. These findings are overall surprising as the occurrence of negative density-dependence of reproduction is generally well established (Dhondt 1977; Minot 1981; Sasvári et al. 1987; Both 1998; Dhondt 2010). Our discussion therefore develops explanations for this apparent mismatch between theory, previous experiments, and our data. Based on our data and a literature review (discussed below), we postulate that our realized breeding densities were within the natural range, but our populations did not show detectable competition for resources during the breeding season, leading to lack of experimental evidence for negative density-dependent reproduction. Our study thus suggests that the effect of realized breeding density on reproduction is non-linear in nature and changes in terms of strength or sign with population density.

Our approach to manipulating competition within and among BTs and GTs was to produce nest boxes that were identical in all respects except for the diameter of the entrance hole (Dhondt and Eyckerman 1980; Dhondt and Adriaensen 1999; Török and Tóth 1999; Dhondt 2010). BT-boxes had entrance holes that were too narrow for most GTs to enter, which explains why manipulations of the number of BT-boxes in a subplot did not affect the number of GT breeders. Indeed, we detected very few GTs breeding in BT-boxes. GT-boxes instead had entrance holes that allowed both species to enter, which explains why manipulations of the number of GT-boxes did affect the number of GT breeders but also why we found interactive treatment effects on the number of BT breeders occupying BT- vs. GT-boxes. That is, when we offered more BT-boxes, the density of BTs increased because BTs chose those surplus BT-boxes over GT-boxes for breeding. The number of BTs breeding in GT-boxes increased when we offered more GT-boxes but only when there were few BT-boxes in the focal subplot. This pattern of occupancy implies that nest boxes were limiting when offered in low supply, and that BTs either lost competition for GT-boxes, and/or that the benefits of occupying a GT-box normally do not outweigh the costs. This latter explanation seems fitting as we had some cases where GTs took over a BT clutch laid in a GT-box, resulting in two mixed clutches. In some of those cases, we found dead BTs in the nest boxes next to breeding GTs, which strongly suggests that BTs were killed by GTs during box take-overs. Overall, this pattern of treatment- and species-specific occupancy implies that breeding cavities and their location in given habitat are a key resource worth fighting for in our study area. BTs can somewhat escape this competition with GTs over nesting sites by actively choosing nest-boxes with small entrance holes.

While our treatments were highly effective in simultaneously manipulating breeding densities of both species, our treatments showed no clear experimental evidence for negative-density dependent reproduction. This lack of treatment effects was most obvious in GTs, where we found no (interactive) treatment effects on reproduction. In BTs, we found moderate—rather than strong—support for interactive treatment effects when there was relatively little competition, that were thus possibly compatible with Allee effects (positive effects of breeding density; reviewed in Stephens et al. 1999). That is, in BTs, the number of GT-boxes appeared to positively affect certain reproductive parameters (i.e., clutch size, probability to produce fledglings) when the number of BT-boxes was low, while this effect disappeared when the number of BT-boxes was high. We should emphasize that we did not find strong supported for most of these patterns as 95% credible intervals largely overlapped zero in analyses of BT data (Tables S10, S11, S12, S13, S14, S15). For the probability to produce fledglings, effect sizes associated with treatment effects shrank when we statistically controlled for realized breeding densities, suggesting that treatment-induced increases in realized breeding density positively—rather than negatively—affected BT reproduction when breeding densities were low (i.e., an Allee effect; reviewed in Stephens et al. 1999). Altogether, these findings suggest that our study populations exhibited competitive regimes at the lower end of the natural range, where competition for resources among breeders is either relaxed (GTs), or where breeders benefit from con- or heterospecifics (BTs).

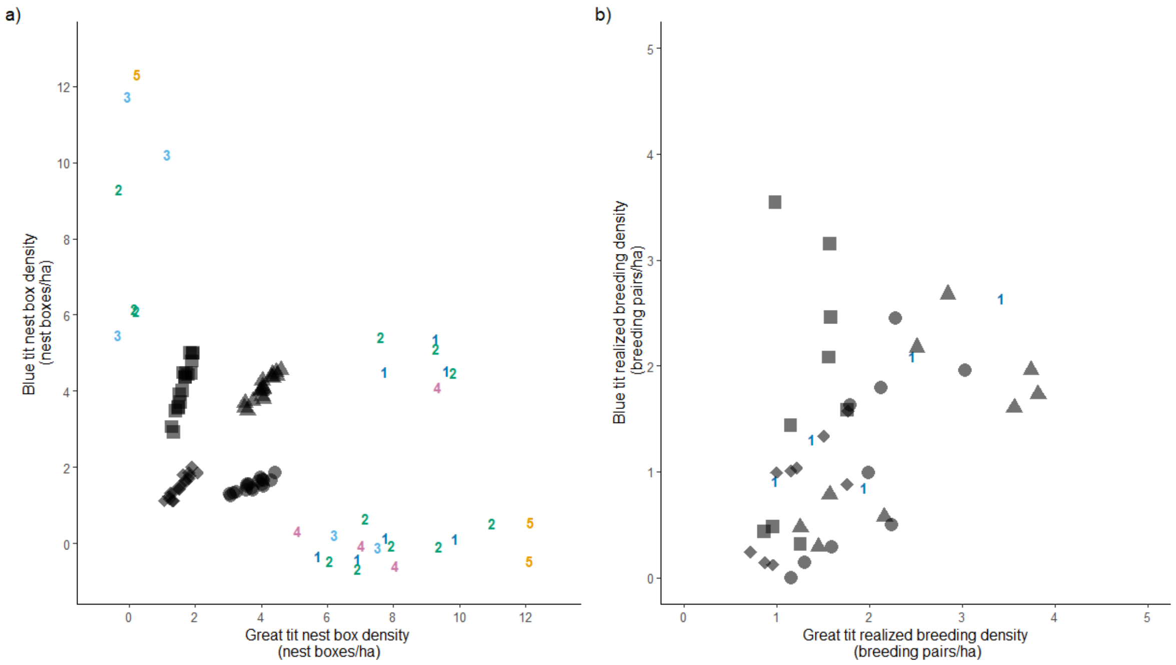


Figure 10. Jitter plot in panel a (left panel) shows the nest box densities (boxes/ha) of "blue tit" (BT) boxes (y-axis) and "great tit" (GT) boxes (x-axis) offered in our study (black symbols) and in published (numbers) studies where both types of boxes were fitted. Jitter plot in panel b (right panel) shows the realized breeding density of BTs (y-axis) and GTs (x-axis) in our study (black symbols) and in published studies (numbers). Each number represents a separate (sub)plot within the same study. For our study, we generated treatment-specific symbols: black diamond = BT "Low"– GT "Low"; black circle = BT "Low" – GT "High"; black squares = BT "High" – GT "Low"; black triangle = BT "High" BT – GT "High", where each symbol represents the density of a given plot in a given year. Numbers refer to the following papers: ^{1}Dhondt (2010), ^{2}Dhondt and Andriaensen (1999), ^{3}Dhondt and Eyckerman (1980), ^{4}Stenseth et al. (2015), ^{5}Török and Tóth (1999). Note that Török and Tóth (1999) reported that they fitted boxes "over 12 boxes ha^{-1}", and for which we printed the lower range (12 boxes/ha) in panel a. Note further that only Dhondt (2010) provided information on the mean realized breeding density which we printed in panel b.

As a post-hoc test of this idea, we compared our realized nest box densities with other studies that applied the same study design (Figure 10a; see also Table S1). This showed that our realized nest box densities were at the lower to intermediate range for both species (Figure 10a). However, our realized breeding densities spanned over the entire range (Figure 10b); this latter conclusion was, notably, based on the comparison of our data with a single study (Dhondt 2010) as none of the other four studies associated with our literature review (Table S1; Figure 10a) provided estimates of realized breeding densities. To acquire more information on whether our realized breeding densities were relatively low or not, we also compared—for GTs—our realized breeding densities with those published as part of another study (Mouchet et al. 2021; A. Mouchet, pers comm). Specifically, realized breeding densities for a Dutch (Lauwersmeer) and German (Seewiesen) population, where boxes were placed in an identical grid (with 50 m between adjacent boxes), spanned over the same range (Figure S3). These comparisons overall imply that our realized breeding densities were similar to studies reporting density-dependent effects (see citations above). Overall, we therefore suggest that the lack of evidence for intra- and inter-specific competition in our study implies that our population was far below carrying capacity, perhaps because it was regulated by factors outside the breeding season, such that competition for resources during breeding was much relaxed. This interpretation also implies that other metrics than realized breeding density should probably be applied to quantify the strength of competition (e.g., McAdam et al. 2022), and that the pattern of density-dependent reproduction is nonlinear in wild tit populations.

Our study spanned over multiple years and study plots, resulting in substantial natural variation in breeding densities, which our analyses also examined. For example, our analyses of the sources of variation in breeding density (Table 2) also demonstrated variation among subplots after controlling for temporal (year) and treatment effects, suggesting spatial variation in breeding density. Our three-step analyses of the data, notably, enables us to draw conclusions about how natural variation in breeding densities affected reproductive parameters. That is, our first set of models tested for treatment effects, our second set of models additionally fitted realized breeding densities, while our third set of models included realized breeding densities but not treatment. Effects of realized breeding density in our second set of models thus tested for effects of realized breeding density not induced by treatment, that is, natural rather than experimental effects of realized breeding density, while our third set of models tested for the combined effects of natural and experimental realized breeding density effects. Altogether, these analyses neither demonstrated effects of natural realized breeding densities on BT nor on GT reproduction, thereby corroborating our experimental results discussed above. Nestling body size in GTs was the only exception, where we found interactive effects of BT- and GT-realized breeding densities. Specifically, GTs fledged with shorter tarsi when the realized breeding density of GTs was relatively high but only in a situation where the realized breeding density of BTs was also high (Figure 8). As we did not detect such effects for clutch size or lay date, this finding may suggest that GTs—but not BTs—experience inter-specific competition for resources required for nestling growth (rather than egg production) (Minot 1981). As we did not detect causal effects of treatment-induced variation in realized breeding density, these non-experimental relationships may, however, not be causally related to realized breeding density. The findings may instead result from variation in habitat unrelated to competitive regimes. The species-specific effects of area size on realized densities and reproductive parameters suggest that such explanations should be considered seriously.

In conclusion, this study effectively altered realized breeding densities of BTs and GTs in a full factorial design (Figure 5), and within the natural range reported by other studies (Figure 10). We manipulated the number of BT- and GT-boxes to achieve this aim. Patterns of settlement in BT- vs. GT-boxes implied that the two species competed for access to breeding cavities, and that GTs were the more competitive species (Minot and Perrins 1986). Our manipulations were subtle, in the sense that both our “low” and “high” nest box densities were low when compared to previous studies (Figure 10), while our literature comparisons implied realized breeding densities were within the normal range reported for the species. Our analyses demonstrated no strong causal effects on reproduction though BTs may have experienced Allee effects when competition for resources was relaxed. Our findings overall imply that realized breeding density did not causally affect reproductive parameters of either species and that patterns of negative density-dependent reproduction thus only occur in environments with stronger competition for resources during breeding.

4.3 Aggressive behaviour towards con- vs. heterospecifics in blue and great tits

Labile traits, such as behavioural traits, can be adjusted over short time periods to match current environmental conditions (Duckworth, 2009). Accordingly, these traits are expected to be highly plastic and exhibit reversible plasticity (Dingemanse & Wolf, 2013; Taborsky & Oliveira, 2012). In our study, we focused on one particular plastic and heritable risk-taking behaviour: male territorial aggression (Araya-Ajoy & Dingemanse, 2017). During incubation, we tested the effects of the social environment (i.e., the identity of the intruder, and the realized breeding density; see § 2.1.3) on aggressiveness displayed by BT and GT males during simulated intrusions. We found that males adjusted their response based on whether they were exposed to a con- vs. heterospecific intruder, implying that the social environment matters. We did not, however, find that BT and GT males plastically adjusted their aggressiveness to changes in the realized breeding density of BTs and GTs, implying that all but certain components of social environments matter.

Male territorial aggression can increase the fitness of a focal male if it increases the access and quality of resources (Brown, 1969). Nevertheless, the benefits obtained from displaying an aggressive response can be outweighed by the costs related to energy or time spent (e.g., reduced time left for incubation, nestling provisioning, or other aspects of parental care; Duckworth 2006a) and the risk of injury (Brown, 1964). For competing species, the costs and benefits of displaying an aggressive response are typically asymmetrical (Martin et al., 2017; Peiman & Robinson, 2010). Particularly in closely related species that live in sympatry and partly compete for the same limited resources, the frequency of agonistic encounters is expected to be higher (Crombie, 1947; Leighton et al., 2024). Differences in body mass between the competing species can further affect the costs of displaying an aggressive response. The costs of inter-specific aggression for the subordinate smaller species are larger, as it is more likely to lose an agonistic interaction with a bigger dominant species (Martin &

Ghalambor, 2014). In that sense, when species differ in the value they give to the same resource, the benefits of an aggressive display are also expected to be species-specific (Enquist & Leimar, 1987; Peiman & Robinson, 2010).

In this context, the effective recognition of con- and heterospecific intruders might help reduce the costs if males can adjust their aggressive response based on whether the intruder is a con- vs. a heterospecific (Lehtonen et al., 2010; Leighton et al., 2024; Tibbetts & Dale, 2007). Our species-specific results first showed that even though BT males approached an intruder closer than GT males, males of both species approached closer to con- than heterospecific intruders. These results suggest that in our population, intraspecific competition over the same limited resource was stronger, hence males are more aggressive towards conspecifics (Klomp, 1964; Peiman & Robinson, 2010). In our population, the number of nest boxes was not a limited resource, and males in areas with more nest boxes did not defend more aggressively (Table 5). Our results rather suggest that males could have behaved more aggressively towards a conspecific intruder when they were defending a high-quality territory and/or a mating partner. Future studies assessing paternity loss could help support this idea. Finally, the results also suggest that both species were able to distinguish between con- and heterospecific intruders. Particularly, this recognition can further benefit the smaller subordinate species as it will help avoid costly agonistic interaction with the bigger dominant species (Leighton et al., 2024; Martin & Ghalambor, 2023). Contrary to expectations, territorial BT males approached a heterospecific intruder closer compared to territorial GT males. These results could suggest that because GTs can easily overtake the nest boxes occupied by BTs (Dhondt & Eyckerman, 1980; Kempenaers & Dhondt, 1991), BT males benefit more from an aggressive display to reduce the likelihood of GTs approaching and taking over the BTs nest. In that sense, for BTs living in sympatry with GTs, the plastic response displayed by BT males may reflect strategic decision-making that helps reduce the costs of potential threats from dominant GT males.

Our nest box density manipulation was effective in generating variation in the realized breeding densities of BTs and GTs (§ 3.2.1). We had expected that males would behave more aggressively in areas with higher realized breeding densities (Araya-Ajoy & Dingemanse, 2017). Previous studies have reported effects of the number of nest boxes and/or of the realized breeding density on the aggressiveness in passerines (Araya-Ajoy & Dingemanse, 2017; Duckworth, 2006b); we were, by contrast, unable to detect any such effects. If these responses were adaptive, our results suggest that the benefits of increased aggressiveness in higher density areas are outweighed by the costs; thus, we did not detect density-dependent adjustment of the aggressiveness (Both & Visser, 2003). Accordingly, we suggest that temporal and spatial variation in resources in our population might not have been the limiting factor triggering male territorial aggression (Both & Visser, 2003). Our results rather suggest that other environmental factors (e.g., habitat quality, predation risk) could be contributing to the costs and benefits associated with displaying an aggressive response.

In conclusion, our study contributes to our understanding of “reversible” within-individual plasticity in aggressiveness at the species level. Our results demonstrated that individual differences in plasticity partly arise because species differ in the level of plasticity towards con- and heterospecifics intruders. In the context of among-individual differences, our results highlight the importance of the link between the social environmental state and behavioral response. By examining the plastic responses of BTs and GTs, we provide insights into the relevance of intra- and interspecific social interactions for understanding population dynamics in the wild.

4.4 Summarizing discussion

The pace-of-life syndrome hypothesis postulates that life-history trade-offs can be resolved differently between species, populations, or sexes (Ricklefs and Wikelski 2002; Réale et al. 2010; Dammhahn et al. 2018; Hämäläinen et al. 2018; Montiglio et al. 2018). Adaptive state-dependent theory postulates that individuals with a relatively faster pace-of-life should invest more heavily in current reproduction at the cost of reduced survival rates (Sæther 1988; Roff 1993; McNamara and Houston 1996; Houston and McNamara 1999; Wolf et al. 2007), and are expected to be bolder and more aggressive (Wolf et al. 2007; Sih et al. 2015). Thus, the allocation of resources between reproduction and survival results in patterns of covariation among life-history, physiology and behaviour, which can be placed along a "slow-fast" continuum (Ricklefs and Wikelski 2002; Dammhahn et al. 2018; Healy et al. 2019). With my thesis I first aimed to understand whether the same ecological mechanism is responsible for shaping the integration between morphology, physiology, and behaviour at different biological levels (Montiglio et al. 2018; Laskowski et al. 2021). My results showed that species, populations, and sexes shared the same patterns of phenotypic integration between morphology and physiology. These results suggest that the ecological factors underpinning the association between these types of traits are the same across species, populations, and sexes. Nonetheless, we did not find the same patterns of integration between morphology and behaviour and physiology and behaviour. The lack of a general pattern of integration between these traits at different biological levels suggests that multiple ecological mechanisms must be at play. My study exemplify that pace-of-life syndromes are not fixed, but may evolve in response to ecological conditions characteristic to specific biological levels.

Contemporary models are thus considering the major role that environmental state has in shaping and maintaining among-individual differences. In the context of personality research, a particularly suitable conceptual model proposed by MacArthur and Wilson (1967), theoretical worked by Engen et al. (2013), and conceptualized by Wright et al. (2019) describes how fluctuations in the population density could promote different patterns of density-dependent selection that could contribute to maintain variation in life-history strategies and differences in behaviour at different biological levels. My second and third aim were to test whether changes in the realized breeding density of con- and heterospecifics affected reproductive parameters and behavioural traits in BTs and GTs. For this, I used a full factorial design to disentangle whether the effects on life-history decisions and behavioural responses were due to changes in the intra-specific and/or inter-specific competition and whether the strength of the effects changed with the social environment. The results showed that overall there were no density-dependent effects on the reproductive parameters nor changes in the aggressiveness in BTs and GTs. My results suggest that in our populations more birds did not mean (stronger) competition. Nevertheless, when I tested BTs and GTs for their aggressive response toward intruders, I found species-species responses. These results further suggest that species react differently to con- vs. heterospecific intruders possibly to reduce the costs of displaying an aggressive response (Tibbetts and Dale 2007; Lehtonen et al. 2010; Leighton et al. 2024).

Overall, my thesis implies that different patterns of phenotypic integration observed at different biological levels are probably due to species-specific costs and benefits of displaying a behavioural response imposed by changes in social environmental state.

CHAPTER 5

5. Conclusions

Conceptual models for adaptive animal personalities predict a link between the (internal or external) state of an individual and its behaviour, leading to the evolution of suites of integrated phenotypic traits, thus different types of individuals in the same population. A prominent idea is that alternative types are favoured when competition is strong vs. weak. I expected that species, populations, and the sexes would share the same pattern of phenotypic integration. This was, however, not the case. We thus need to develop explanations for why different patterns of phenotypic integration are observed at different biological levels. Furthermore, I expected that reproductive performance was affected by both intra- and interspecific competitive regimes. While my research partly confirmed this expectation, the weak effects of my nest box density treatments induced by my experimental design suggest the need for stronger, or other types of experimental designs, to fully test this idea. Third, contemporary models predict that the social environment could be an important source of variation in behaviour not only because of how selection maintains alternative behavioural strategies in different competitive regimes, but also because individuals might plastically modify their behaviour in response to the species with which they interact. This was indeed the case. Species differed in plasticity towards con- vs. heterospecific intruders. Overall, my findings combined imply that life-history theory developed to predict the integration of life-history and behaviour as a function of competition may need to be expanded or modified to include relevant environmental and ecological effects to thereby better understand phenotypic integration and life-history evolution in the wild.

6. Supplementary material

Table S1. Previous studies that tested the effects of intra- vs. inter-specific competition of blue tits (BTs) and great tits (GTs) on reproductive parameters. BT-boxes only= boxes with a reduced entrance hole diameter of 26-mm, GT-boxes only= boxes with an entrance hole diameter of 32-mm.

Reference	Manipulation of the number and/or type of nest boxes	Experimental design within plot [a]	Full factorial design [b]
Dhondt (1977)	No	GT-boxes only	No
Dhondt (2010)	Yes	GT-boxes only BT- plus GT-boxes	No
Dhondt and Adriaensen (1999)	Yes	BT-boxes only GT-boxes only BT- plus GT-boxes	No
Minot (1981)	No	GT-boxes only	No
Sasvári et al. (1987)	No	GT-boxes only	No
Sasvári and Orell (1992)	No	GT-boxes only	No
Török and Tóth (1999)	Yes	BT-boxes only GT-boxes only	No

1 The experimental design of some plots, within the same study, changed across year, either in number or type of nest boxes provided (Dhondt and Adriaensen 1999; Dhondt 2010).

2 The manipulation of the number of BT- and GT-boxes (low vs. high) was done independently to study the additive and interactive effects of both types of nest boxes.

Figure S1. Map showing the dimensions and distribution of 12 study subplots located in the Forstenrieder Park, Munich, Germany. The different colours show the sets of four subplots that were in close proximity used to randomly assign each subplot to one of the four possible nest box density treatments (see § 2). To estimate each subplot's area (in ha), we used the geographic information system software QGIS v. 3.28.3. To do so, we first established the limits of the polygon using the outermost nest boxes of each subplot. We then used the function "$area" from the Field Calculator to obtain each surface measurement. Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org.

Table S2. Subplot specific nest box density treatments applied in the Forstenrieder Park for three consecutive years. The four unique combination of density treatments (L: low vs. H: high, for blue tits: BT boxes and great tits: GT boxes) are shown thrice per year among the plots (A-F). Additionally, the table shows the total number of nest boxes per subplot and plot.

Year	Plot	Subplot	Treatment		Number of boxes per subplot	Number boxes per plot
			BT boxes	GT boxes		
2020	A	A1	H	L	35	85
		A2	H	H	50	
	B	B1	H	H	50	85
		B2	L	H	35	
	C	C1	H	L	35	55
		C2	L	L	20	
	D	D1	L	H	35	70
		D2	H	L	35	
	E	E1	L	H	35	55
		E2	L	L	20	
	F	F1	L	L	20	70
		F2	H	H	50	
2021	A	A1	L	H	35	85
		A2	H	H	50	
	B	B1	H	H	50	85
		B2	L	H	35	
	C	C1	L	L	20	55
		C2	H	L	35	
	D	D1	H	H	50	85
		D2	H	L	35	
	E	E1	L	L	20	55
		E2	H	L	35	
	F	F1	L	H	35	55
		F2	L	L	20	
2022	A	A1	H	H	50	85
		A2	H	L	35	
	B	B1	L	L	20	55
		B2	L	H	35	
	C	C1	H	L	35	85
		C2	H	H	50	
	D	D1	H	H	50	70
		D2	L	L	20	
	E	E1	L	H	35	55
		E2	L	L	20	
	F	F1	L	H	35	70
		F2	H	L	35	

Text S1. We tested if the time spent in the side compartment of the exploration cage might influence the exploration scores. We, thus, ran a sensitivity analysis for a subset of randomly allocated pairs for which one partner spent 30 s and the other partner 120 s in the side compartment (n = 6 blue tit pairs and n = 7 great tit pairs). We fitted a linear mixed model (LMM) assuming a Gaussian error distribution. Exploration score was included as response variable. Species (great tit vs blue tit) and time spent in the side compartment (30 s vs. 120 s) were fitted as fixed factors. We also fitted random intercepts for nest box identity (n = 13). We did not find any statistically significant effect of species (β = 3.011, 95 % credible intervals = -14.962, 20.459) or the time spent on the side compartment (β = -11.321, 95 % credible intervals = -23.745, 1.617). These results suggest that differences in acclimation time did not affect exploration scores.

Text S2. After visual inspection of the data from the blue tit adults of the Forstenrieder Park, we removed two observations of body mass (15.20 and 15.38 gr; natural range 9.2-12.4 gr; Blondel et al. 2002; Kullberg et al. 2002; Perrier et al. 2018), two observations of tarsus length (18.6 and 18.9 mm; maximum value 18.4 mm in 15 days old nestlings; Nord and Nilsson 2011), eight values of wing length (>71 mm; maximum value found 69 mm but measured to the nearest of 1mm; (Furness and Furness 2016), and two observations of bill (10.5 and 10.8 mm; natural range 6.4-10.3 mm; Blondel et al. 2002; Perrier et al. 2018). Furthermore, we used the 10 years of data on body mass collected on great tits in Starnberg to assess the data from the great tit adults of the Forstenrieder Park. We removed one observation of body mass (11.63 gr; range 13.5-20.6 gr). Additionally, we removed four observations of wing length (>81 mm; maximum value 81 mm; Nowakowski and Rowiński 1996), and two observations of bill length (9.0 and 14.1 mm; natural range 9.15-13.33mm; Matthysen et al. 1991). Moreover, we removed one observation of bill length of great tits of Starnberg (8.8 mm; minimum value 9.15 mm; Matthysen et al. 1991).

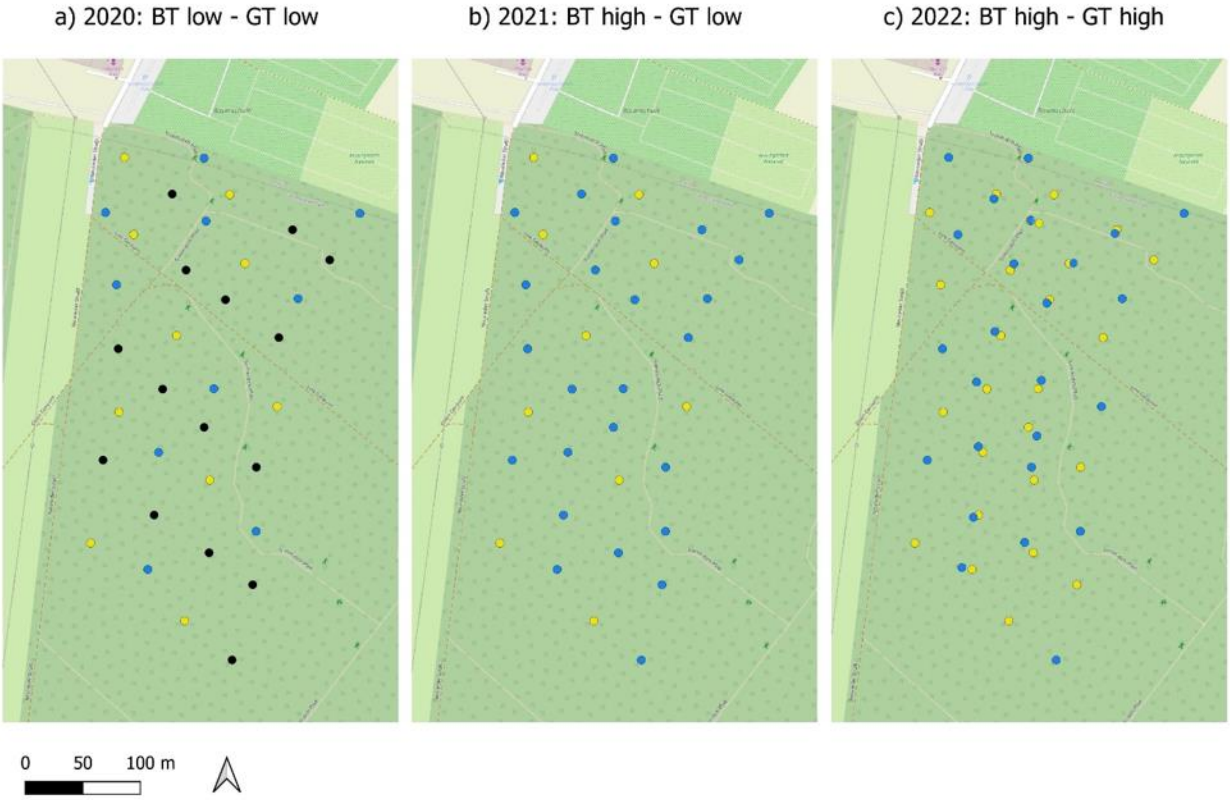


Figure S2. Maps showing an example (subplot C2) of realized positions of different box types for the years 2020, 2021, and 2022. Panel (a) shows how the blue tit (BT) "low" - great tit (GT) "low" nest box density treatment (i.e., 10 BT-, 10-GT and 15 "closed"-boxes) was implemented. Panel (b) shows how the BT "high" - GT "low" nest box density treatment (i.e., 25 BT-, 10 GT-boxes) was implemented. Panel (c) shows how the BT "high" vs. GT "high" nest box density treatment (i.e., 25 BT-, 25 GT-boxes) was implemented. Blue circles = BT-boxes, yellow circles = GT-boxes, black circles = closed-boxes. Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org.

Table S3. For each unique combination of species, population, and sex, we ran a structural equation model that estimated all nine hypothesized paths. The full model was fitted using the among-individual correlation matrix among all six mean- and variance-standardized traits. We printed each path's standardized estimate, SE, and p-value. LV = latent variable, BM = body mass, TL = tarsus length, WL = wing length, BL = bill length, BR = breathing rate, and ES = exploration score. Bold values represent paths that were statistically significant ($p < 0.05$). We present two types of parameters: standardized path coefficients (→) and correlations (↔).

Paths	Forstenrieder Park								Starnberg			
	Blue tits				Great tits				Great tits			
	Male		Female		Male		Female		Male		Female	
	Estimate (SE)	p	Estimate (SE)	p	Estimate (SE)	p	Estimate (SE)	p	Estimate (SE)	p	Estimate (SE)	p
LV → BM	**0.57 (0.09)**	**<0.001**	**0.88 (0.08)**	**<0.001**	**0.55 (0.09)**	**<0.001**	**0.49 (0.09)**	**<0.001**	**0.70 (0.04)**	**<0.001**	**0.84 (0.04)**	**<0.001**
LV → TL	**0.86 (0.09)**	**<0.001**	**0.49 (0.08)**	**<0.001**	**0.96 (0.13)**	**<0.001**	**0.74 (0.09)**	**<0.001**	**0.79 (0.04)**	**<0.001**	**0.60 (0.03)**	**<0.001**
LV → WL	**0.57 (0.08)**	**<0.001**	**0.75 (0.08)**	**<0.001**	**0.34 (0.08)**	**<0.001**	**0.48 (0.08)**	**<0.001**	**0.33 (0.04)**	**<0.001**	**0.41 (0.03)**	**<0.001**
LV → BL	**0.28 (0.09)**	**<0.001**	**0.17 (0.08)**	**0.04**	**0.21 (0.08)**	**0.01**	0.09 (0.08)	0.25	**0.30 (0.04)**	**<0.001**	**0.42 (0.03)**	**<0.001**
LV → BR	**-0.49 (0.12)**	**<0.001**	0.67 (0.45)	0.14	**-0.32 (0.11)**	**<0.001**	**-0.25 (0.11)**	**0.02**	**-0.33 (0.07)**	**<0.001**	-0.24 (0.13)	0.06
LV → ES	0.13 (0.11)	0.23	0.67 (0.45)	0.14	-0.12 (0.10)	0.22	**-0.50 (0.12)**	**<0.001**	0.08 (0.06)	0.22	-0.02 (0.12)	0.86
BM →BR	0.06 (0.10)	0.55	-0.56 (0.44)	0.20	0.15 (0.10)	0.16	**0.27 (0.09)**	**<0.001**	**0.18 (0.06)**	**<0.001**	0.21 (0.11)	0.06
BM →ES	**0.20 (0.10)**	**0.04**	-0.43 (0.43)	0.32	0.02 (0.09)	0.79	0.08 (0.09)	0.42	0.01 (0.05)	0.89	-0.11 (0.10)	0.27
BR↔ES	-0.06 (0.07)	0.42	0.10 (0.10)	0.33	**0.17 (0.07)**	**0.02**	0.04 (0.07)	0.56	**-0.21 (0.03)**	**<0.001**	0.04 (0.03)	0.19

Table S4. Estimates, 95% confidence intervals (CI) and the amount of heterogeneity ($R^2_{marginal}$) explained by the species (moderator) for each of the nine hypothesized structural paths.

Path	Species		Variance explained by species		
	Blue tits	Great tits			
	Estimate (95%CI)	Estimate (95%CI)	$R^2_{marginal}$	χ^2_1	p
Latent variable → Body mass	0.73 (0.48, 0.98)	0.66 (0.49, 0.83)	0.00%	0.21	0.65
Latent variable → Tarsus length	0.67 (0.42, 0.91)	0.75 (0.58, 0.93)	0.00%	0.29	0.59
Latent variable → Wing length	0.66 (0.54, 0.79)	0.38 (0.32, 0.45)	90.84%	14.63	<0.001
Latent variable → Bill length	0.22 (0.03, 0.42)	0.27 (0.15, 0.40)	0.00%	0.17	0.68
Latent variable → Breathing rate	-0.41 (-0.64, -0.19)	-0.30 (-0.40, -0.21)	22.17%	0.81	0.40
Latent variable → Exploration score	0.24 (-0.20, 0.68)	-0.13 (-0.37, 0.11)	2.86%	2.04	0.15
Body mass → Breathing rate	0.03 (-0.16, 0.22)	0.19 (0.12, 0.28)	0.00%	2.53	0.11
Body mass → Exploration score	0.17 (-0.02, 0.36)	0.01 (-0.06, 0.08)	50.71%	2.38	0.12
Breathing rate ↔ Exploration score	0.01 (-0.22, 0.24)	0.00 (-0.15, 0.15)	0.00%	0.01	0.94

Table S5. Estimates, 95% confidence intervals (CI) and the amount of heterogeneity ($R^2_{marginal}$) explained by the population (moderator) for each of the nine hypothesized structural paths.

Path	Population		Variance explained by population		
	Forstenrieder Park	Starnberg			
	Estimate (95%CI)	Estimate (95%CI)	$R^2_{marginal}$	χ^2_1	p
Latent variable → Body mass	0.63 (0.47, 0.79)	0.77 (0.57, 0.97)	4.19%	1.22	0.27
Latent variable → Tarsus length	0.75 (0.56, 0.93)	0.69 (0.46, 0.92)	0.00%	0.13	0.72
Latent variable → Wing length	0.54 (0.39, 0.68)	0.37 (0.19, 0.55)	21.72%	1.96	0.16
Latent variable → Bill length	0.18 (0.09, 0.28)	0.37 (0.27, 0.46)	69.75%	6.77	0.001
Latent variable → Breathing rate	-0.32 (-0.45, -0.20)	-0.31 (-0.43, -0.19)	74.46%	0.03	0.87
Latent variable → Exploration score	-0.09 (-0.39, 0.21)	0.03 (-0.34, 0.41)	0.00%	0.24	0.62
Body mass → Breathing rate	0.16 (0.05, 0.26)	0.19 (0.08, 0.29)	31.60%	0.16	0.69
Body mass → Exploration score	0.09 (-0.02, 0.19)	-0.10 (-0.10, 0.07)	41.13%	2.03	0.15
Breathing rate ↔ Exploration score	0.06 (-0.08, 0.20)	-0.09 (-0.25, 0.09)	13.37%	1.69	0.19

Table S6. Estimates, 95% confidence intervals (CI) and the amount of heterogeneity (R^2_{marginal}) explained by the sex (moderator) for each of the nine hypothesized structural paths.

Path	Sex		Variance explained by sex		
	Females	Males			
	Estimate (95%CI)	Estimate (95%CI)	$R^2_{marginal}$	χ^2_1	p
Latent variable → Body mass	0.75 (0.57, 0.92)	0.62 (0.44, 0.79)	4.05%	1.11	0.29
Latent variable → Tarsus length	0.60 (0.52, 0.68)	0.83 (0.73, 0.92)	89.67%	12.36	<0.001
Latent variable → Wing length	0.54 (0.36, 0.71)	0.41 (0.23, 0.59)	0.00%	1.00	0.32
Latent variable → Bill length	0.24 (0.09, 0.40)	0.27 (0.11, 0.42)	0.00%	0.03	0.87
Latent variable → Breathing rate	-0.22 (-0.38, -0.05)	-0.14 (-0.34, 0.05)	38.08%	2.15	0.14
Latent variable → Exploration score	-0.16 (-0.48, 0.17)	0.03 (-0.24, 0.30)	10.62%	0.78	0.38
Body mass → Breathing rate	0.23 (0.09, 0.36)	0.15 (0.06, 0.24)	64.17%	0.88	0.35
Body mass → Exploration score	-0.02 (-0.18, 0.13)	0.06 (-0.06, 0.17)	0.00%	0.65	0.42
Breathing rate ↔ Exploration score	0.06 (-0.11, 0.22)	-0.05 (-0.21, 0.12)	0.00%	0.75	0.39

Table S7. Estimated effect sizes and 95% credible intervals (CIs) for predictors of the number of blue tits (a-c) and great tits (d-f) breeding in 12 subplots located in the Forstenrieder Park, Munich, Germany. We analyse numbers of breeders in (a, d) any type of box, (b, e) blue tit (BT) boxes and (c, f) great tit (GT) boxes. Predictor variables are subplot area size (in ha), year (2020, 2021, 2022), the number (no) of BT-boxes (Low vs. High), the no of GT-boxes (Low vs. High), and their interaction. Parameters were estimated using GLMMs following a Poisson error distribution and a log-link function. Estimated effect size and 95% CIs were back-transformed from the latent scale using the function "exp" (package base, R Core Team 2023). Parameters are thus shown in data scale. The sample size was the number of subplots times the number of years (n=12 subplots × 3 years = 36 data points).

	Blue tits breeding in			Great tits breeding in		
	(a) Any type of box	(b) BT-boxes	(c) GT-boxes	(d) Any type of box	(e) BT-boxes	(f) GT-boxes
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	4.26 (2.61, 6.89)	3.74 (2.25, 6.11)	0.21 (0.04, 1.05)	7.17 (5.42, 9.58)	0.67 (0.23, 1.93)	6.62 (4.9, 9.03)
Area [b]	0.63 (0.43, 0.93)	0.66 (0.45, 0.99)	0.46 (0.26, 0.84)	0.94 (0.81, 1.09)	0.77 (0.45, 1.35)	0.95 (0.83, 1.09)
Year 2021	1.15 (0.86, 1.55)	1.22 (0.89, 1.67)	0.97 (0.46, 2.03)	1.11 (0.87, 1.42)	0.72 (0.23, 2.36)	1.14 (0.89, 1.48)
Year 2022	1.00 (0.74, 1.36)	1.08 (0.79, 1.49)	0.68 (0.31, 1.54)	1.16 (0.90, 1.48)	0.85 (0.29, 2.56)	1.17 (0.91, 1.51)
BT-box no [c]	1.84 (1.23, 2.69)	1.88 (1.23, 2.75)	2.18 (0.41, 12.3)	0.99 (0.70, 1.42)	1.21 (0.38, 3.86)	0.97 (0.66, 1.40)
GT-box no [d]	1.22 (0.71, 2.14)	0.89 (0.49, 1.58)	12.06 (2.56, 59.74)	1.68 (1.22, 2.32)	0.21 (0.02, 1.93)	1.79 (1.30, 2.48)
BT- × GT-box no [e]	0.61 (0.33, 1.14)	0.79 (0.41, 1.6)	0.14 (0.02, 0.93)	1.13 (0.72, 1.70)	3.49 (0.31, 42.52)	1.14 (0.73, 1.77)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Subplot ID	1.42 (1.20, 1.77)	1.49 (1.22, 1.92)	1.12 (1.04, 1.23)	1.02 (1.01, 1.04)	1.00 (1.00, 1.00)	1.02 (1.01, 1.03)
Residual [f]	1.23 (1.15, 1.38)	1.27 (1.16, 1.45)	5.75 (1.95, 24.05)	1.14 (1.11, 1.19)	2.48 (1.52, 5.26)	1.15 (1.12, 1.21)

[a] Reference category: estimate is for the number of occupied nest boxes in the "Low" BT- and "Low" GT-box density treatment for the year 2020 for areas of average size.

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Estimate is the effect of the "High" BT-box treatment relative to the "Low" BT-box treatment when GT-box numbers were "Low".

[d] Estimate is the effect of the "High" GT-box treatment relative to the "Low" GT-box treatment when BT-box numbers were "Low".

[e] Estimate is the effect of the "High" BT-box treatment relative to the "Low" BT-treatment for the "High" GT-treatment expressed as a difference from the same effect observed in the "Low" GT-treatment.

[f] Defined as ln(1/exp(βo)+1) (following Nakagawa and Schielzeth 2010).

Table S8. Estimated effect sizes and 95% credible intervals (CIs) for predictors of the number of blue tits (a-c) and great tits (d-f) breeding in 12 subplots located in the Forstenrieder Park, Munich, Germany. We analyse numbers of breeders in (a, d) any type of box, (b, e) blue tit (BT) boxes and (c, f) great tit (GT) boxes. Predictor variables are subplot area size (in ha), year (2020, 2021, 2022), the number (no) of BT-boxes (Low vs. High), and the no of GT-boxes (Low vs. High). Parameters were estimated using GLMMs following a Poisson error distribution and a log-link function. Estimated effect sizes and 95% CIs are shown on the latent scale. The sample size was the number of subplots times the number of years (n=12 subplots × 3 years = 36 data points).

	Blue tits breeding in			Great tits breeding in		
	(a) Any type of box	(b) BT-boxes	(c) GT-boxes	(d) Any type of box	(e) BT-boxes	(f) GT-boxes
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	1.58 (1.14, 2.03)	1.37 (0.89, 1.86)	-0.65 (-1.66, 0.39)	1.93 (1.67, 2.20)	-0.60 (-1.62, 0.42)	1.85 (1.60, 2.12)
Area [b]	-0.44 (-0.80, -0.07)	-0.41 (-0.82, -0.00)	-0.74 (-1.37, -0.11)	-0.06 (-0.21, 0.08)	-0.26 (-0.78, 0.28)	-0.05 (-0.19, 0.09)
Year 2021	0.13 (-0.17, 0.42)	0.20 (-0.12, 0.53)	-0.21 (-0.97, 0.54)	0.11 (-0.14, 0.34)	-0.33 (-1.45, 0.81)	0.13 (-0.12, 0.38)
Year 2022	-0.01 (-0.30, 0.29)	0.08 (-0.24, 0.40)	-0.52 (-1.35, 0.34)	0.15 (-0.10, 0.40)	-0.18 (-1.26, 0.90)	0.17 (-0.09, 0.42)
BT-box no [c]	0.45 (0.13, 0.76)	0.56 (0.23, 0.90)	-0.58 (-1.35, 0.22)	0.06 (-0.16, 0.28)	0.54 (-0.49, 1.52)	0.04 (-0.20, 0.27)
GT-box no [d]	-0.12 (-0.49, 0.22)	-0.28 (-0.66, 0.12)	1.36 (0.39, 2.33)	0.58 (0.35, 0.81)	-0.68 (-1.68, 0.28)	0.63 (0.40, 0.86)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Subplot ID	0.37 (0.20, 0.60)	0.41 (0.20, 0.67)	0.22 (0.08, 0.42)	0.02 (0.01, 0.04)	0.00 (0.00, 0.00)	0.02 (0.01, 0.03)
Residual [e]	0.19 (0.12, 0.28)	0.23 (0.14, 0.34)	1.07 (0.52, 1.84)	0.13 (0.10, 0.17)	1.04 (0.51, 1.80)	0.15 (0.11, 0.18)

[a] Reference category: estimate is for the number of occupied nest boxes in the “Low” BT- and “Low” GT-box density treatment for the year 2020 for areas of average size.

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Estimate is the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[d] Estimate is the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[e] Defined as $\ln (1/\exp(\beta o)+1)$ (following Nakagawa and Schielzeth 2010).

Table S9. Estimated effect sizes and 95% credible intervals (CIs) for predictors of the number of blue tits (a-c) and great tits (d-f) breeding in 12 subplots located in the Forstenrieder Park, Munich, Germany. We analyse numbers of breeders in (a, d) any type of box, (b, e) blue tit (BT) boxes and (c, f) great tit (GT) boxes. Predictor variables are subplot area size (in ha), year (2020, 2021, 2022), the number (no) of BT-boxes (Low vs. High), and the no of GT-boxes (Low vs. High). Parameters were estimated using GLMMs following a Poisson error distribution and a log-link function. Estimated effect size and 95% CIs were back-transformed from the latent scale using the function "exp" (package base, R Core Team 2023). Parameters are shown in data scale. The sample size was the number of subplots times the number of years (n=12 subplots × 3 years = 36 data points).

	Blue tits breeding in			Great tits breeding in		
	(a) Any type of box	(b) BT-boxes	(c) GT-boxes	(d) Any type of box	(e) BT-boxes	(f) GT-boxes
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	4.81 (3.10, 7.54)	3.94 (2.46, 6.49)	0.51 (0.19, 1.35)	6.89 (5.26, 8.85)	0.54 (0.19, 1.55)	6.30 (4.90, 8.33)
Area [b]	0.64 (0.44, 0.95)	0.66 (0.45, 0.97)	0.47 (0.25, 0.86)	0.94 (0.80, 1.08)	0.77 (0.44, 1.31)	0.95 (0.82, 1.09)
Year 2021	1.15 (0.84, 1.54)	1.22 (0.88, 1.67)	0.81 (0.37, 1.73)	1.11 (0.86, 1.42)	0.73 (0.23, 2.32)	1.14 (0.88, 1.46)
Year 2022	0.99 (0.74, 1.35)	1.09 (0.79, 1.49)	0.60 (0.26, 1.36)	1.16 (0.90, 1.49)	0.84 (0.28, 2.44)	1.19 (0.92, 1.52)
BT-box no [c]	1.57 (1.14, 2.18)	1.77 (1.26, 2.44)	0.56 (0.26, 1.20)	1.06 (0.84, 1.32)	1.72 (0.63, 4.62)	1.05 (0.83, 1.31)
GT-box no [d]	0.88 (0.61, 1.30)	0.75 (0.51, 1.11)	4.01 (1.52, 10.28)	1.80 (1.45, 2.25)	0.5 (0.19, 1.31)	1.90 (1.51, 2.39)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Subplot ID	1.46 (1.22, 1.84)	1.51 (1.23, 1.95)	1.25 (1.09, 1.52)	1.02 (1.01, 1.04)	1.00 (1.00, 1.00)	1.02 (1.01, 1.03)
Residual [e]	1.21 (1.13, 1.32)	1.26 (1.15, 1.40)	2.97 (1.73, 6.30)	1.14 (1.12, 1.19)	2.86 (1.65, 6.23)	1.16 (1.12, 1.21)

[a] Reference category: estimate is for the number of occupied nest boxes in the "Low" BT- and "Low" GT-box density treatment for the year 2020 for areas of average size.

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Estimate is the effect of the "High" BT-box treatment relative to the "Low" BT-box treatment when GT-box numbers were "Low".

[d] Estimate is the effect of the "High" GT-box treatment relative to the "Low" GT-box treatment when BT-box numbers were "Low".

[e] Defined as ln(1/exp(βo)+1) (following Nakagawa and Schielzeth 2010).

Table S10. Estimated effect sizes and 95% credible intervals (CIs) for predictors of lay date for blue tits (BT) (models a-c) and great tits (GT) (models d-f) for models fitting (main and interaction effects of) the number (no) of BT- and GT-boxes (nest box density treatment; a, d), realized breeding density (c, f) or both (b, e). We further fit subplot area size as fixed effect. Parameters were estimated using LMMs assuming a Gaussian error distribution.

	Lay date					
	Blue tits (n= 269 broods)			Great tits (n= 397 broods)		
	(a) Treatment	(b) Treatment and realized breeding density	(c) Realized breeding density	(d) Treatment	(e) Treatment and realized breeding density	(f) Realized breeding density
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	17.05 (14.53, 19.59)	17.37 (13.46, 21.53)	18.17 (16.47, 19.88)	20.82 (18.8, 22.85)	20.05 (17.04, 23.08)	20.19 (18.88, 21.51)
Area [b]	0.95 (-0.40, 2.35)	1.23 (-0.98, 3.43)	1.09 (-0.83, 3.15)	2.22 (1.10, 3.32)	1.73 (0.11, 3.31)	1.52 (0.10, 2.97)
Nest box density treatment						
BT-box no [c]	1.38 (-2.20, 4.89)	1.70 (-2.40, 5.85)	-	0.21 (-2.64, 3.24)	0.86 (-2.23, 4.16)	-
GT-box no [d]	-0.52 (-4.09, 3.22)	-0.28 (-5.91, 5.43)	-	-2.38 (-5.10, 0.43)	-0.99 (-4.90, 3.03)	-
BT- × GT-box no [e]	0.87 (-4.08, 5.66)	-0.08 (-5.83, 5.94)	-	2.66 (-1.29, 6.55)	0.82 (-3.78, 5.55)	-
Realized breeding density [f]						
BT-density	-	-7.22 (-39.84, 26.97)	-3.14 (-34.21, 27.29)	-	-4.53 (-36.90, 28.95)	-1.96 (-28.75, 25.54)
BT- density squared	-	-11.95 (-39.91, 15.86)	-10.98 (-35.40, 12.31)	-	-5.43 (-27.15, 16.47)	-6.01 (-26.03, 14.18)
GT-density	-	12.46 (-35.36, 59.26)	11.99 (-19.04, 41.92)	-	1.47 (-44.31, 46.96)	0.69 (-30.31, 34.27)
GT-density squared	-	12.44 (-14.96, 39.74)	17.23 (-3.48, 39.09)	-	19.37 (-8.65, 47.16)	26.60 (3.53, 48.77)
BT- × GT-density	-	-1.10 (-3.80, 1.62)	-1.13 (-3.86, 1.60)	-	0.38 (-1.88, 2.65)	0.25 (-1.85, 2.47)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Female ID	11.33 (9.71, 13.21)	11.37 (9.66, 13.36)	11.35 (9.63, 13.34)	5.73 (4.76, 6.86)	5.79 (4.79, 6.89)	5.64 (4.63, 6.73)
Subplot-Year ID	11.44 (7.35, 16.46)	11.7 (7.24, 17.57)	11.39 (7.10, 17.18)	8.28 (5.12, 12.33)	8.20 (5.03, 12.11)	7.97 (4.86, 11.61)
Residual	5.86 (4.96, 6.91)	5.93 (4.96, 7.04)	5.95 (5.01, 7.04)	20.28 (17.69, 23.17)	20.29 (17.67, 23.47)	20.41 (17.79, 23.25)

[a] Reference category: estimate is for tits in areas of average size breeding in the “Low-Low” (BT-GT) box density treatment (all models except c, f) under average realized breeding densities (all models except a, d).

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[d] Estimate for the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[e] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[f] Estimate for the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

Table S11. Estimated effect sizes and 95% credible intervals (CIs) for predictors of clutch size for blue tits (BT) (models a-c) and great tits (GT) (models d-f) for models fitting (main and interaction effects of) the number (no) of BT- and GT-boxes (nest box density treatment; a, d), realized breeding density (c, f) or both (b, e). We further fit subplot area size as fixed effect. Parameters were estimated using LMMs assuming a Gaussian error distribution.

	Clutch size					
	Blue tits (n= 268 broods)			Great tits (n= 397 broods)		
	(a) Treatment	(b) Treatment and realized breeding density	(c) Realized breeding density	(d) Treatment	(e) Treatment and realized breeding density	(f) Realized breeding density
Intercept [a]	9.53 (9.06, 10.02)	9.51 (8.71, 10.29)	9.75 (9.44, 10.03)	8.68 (8.25, 9.09)	8.69 (8.05, 9.34)	8.52 (8.23, 8.80)
Area [b]	-0.32 (-0.59, -0.05)	-0.28 (-0.72, 0.13)	-0.14 (-0.52, 0.23)	0.06 (-0.17, 0.30)	-0.05 (-0.39, 0.29)	-0.05 (-0.38, 0.27)
Nest box density treatment						
BT-box no [c]	0.12 (-0.50, 0.74)	-0.04 (-0.74, 0.64)	-	-0.27 (-0.85, 0.30)	-0.25 (-0.91, 0.40)	-
GT-box no [d]	0.59 (-0.06, 1.27)	0.63 (-0.45, 1.67)	-	-0.19 (-0.75, 0.34)	-0.11 (-0.96, 0.69)	-
BT- × GT-box no [e]	-0.70 (-1.52, 0.13)	-0.45 (-1.47, 0.63)	-	0.23 (-0.50, 0.97)	0.15 (-0.74, 1.07)	-
Realized breeding density [f]						
BT-density	-	3.32 (-2.04, 8.51)	3.37 (-1.69, 8.11)	-	-3.22 (-9.68, 3.52)	-3.59 (-9.72, 2.23)
BT- density squared	-	0.62 (-4.28, 5.51)	0.94 (-3.29, 5.10)	-	1.76 (-3.20, 6.59)	1.21 (-3.18, 5.25)
GT-density	-	-3.76 (-13.42, 5.65)	-0.9 (-6.04, 4.36)	-	1.72 (-7.21, 10.60)	0.86 (-5.33, 6.68)
GT-density squared	-	-0.47 (-5.01, 3.97)	-2.76 (-6.27, 0.82)	-	2.12 (-3.63, 7.95)	1.96 (-2.49, 6.37)
BT- × GT-density	-	0.23 (-0.25, 0.73)	0.14 (-0.29, 0.60)	-	-0.15 (-0.60, 0.29)	-0.11 (-0.56, 0.33)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Female ID	1.86 (1.55, 2.22)	1.77 (1.47, 2.15)	1.72 (1.44, 2.07)	1.69 (1.46, 1.93)	1.75 (1.52, 2.00)	1.74 (1.52, 1.99)
Subplot-Year ID	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.08 (0.05, 0.12)	0.09 (0.06, 0.14)	0.08 (0.05, 0.12)
Residual	1.28 (1.08, 1.54)	1.37 (1.14, 1.64)	1.41 (1.19, 1.68)	1.10 (0.96, 1.27)	1.06 (0.92, 1.22)	1.06 (0.92, 1.22)

[a] Reference category: estimate is for tits in areas of average size breeding in the “Low-Low” (BT-GT) box density treatment (all models except c, f) under average realized breeding densities (all models except a, d).

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[d] Estimate for the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[e] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[f] Estimate for the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

Table S12. Estimated effect sizes and 95% credible intervals (CIs) for predictors of nestling body condition for blue tits (BT) (models a-c) and great tits (GT) (models d-f) for models fitting (main and interaction effects of) the number (no) of BT- and GT-boxes (nest box density treatment; a, d), realized breeding density (c, f) or both (b, e). We further fit subplot area size, tarsus, and age as fixed effects. Parameters were estimated using LMMs assuming a Gaussian error distribution.

	Nestling body condition					
	Blue tits (n= 1493 nestlings)			Great tits (n= 1810 nestlings)		
	(a) Treatment	(b) Treatment and realized breeding density	(c) Realized breeding density	(d) Treatment	(e) Treatment and realized breeding density	(f) Realized breeding density
Intercept [a]	10.7 (10.31, 11.08)	10.86 (10.21, 11.54)	10.81 (10.56, 11.05)	16.19 (15.74, 16.59)	16.63 (16.07, 17.25)	16.22 (15.97, 16.48)
Area [b]	0.07 (-0.14, 0.26)	0.09 (-0.25, 0.43)	0.07 (-0.21, 0.34)	0.02 (-0.20, 0.23)	0.20 (-0.12, 0.53)	0.12 (-0.18, 0.42)
Age [c]	-0.38 (-0.78, 0.02)	-0.42 (-0.83, -0.01)	-0.42 (-0.84, -0.01)	0.52 (-0.06, 1.15)	0.58 (-0.02, 1.17)	0.49 (-0.10, 1.08)
Tarsus [d]	0.74 (0.68, 0.80)	0.74 (0.68, 0.80)	0.74 (0.68, 0.80)	1.06 (0.99, 1.14)	1.06 (0.99, 1.13)	1.06 (0.99, 1.13)
Nest box density treatment						
BT-box no [e]	0.10 (-0.38, 0.60)	-0.03 (-0.66, 0.63)	-	-0.10 (-0.67, 0.46)	-0.26 (-0.92, 0.38)	-
GT-box no [f]	0.03 (-0.47, 0.55)	-0.11 (-0.97, 0.74)	-	-0.28 (-0.83, 0.28)	-0.74 (-1.55, 0.06)	-
BT- × GT-box no [g]	-0.05 (-0.70, 0.63)	0.11 (-0.82, 1.04)	-	0.35 (-0.41, 1.13)	0.59 (-0.33, 1.51)	-
Realized breeding density [h]						
BT-density	-	-0.08 (-11.74, 11.32)	-0.28 (-9.98, 9.05)	-	0.01 (-13.46, 14.11)	2.01 (-10.57, 13.91)
BT- density squared	-	5.00 (-5.19, 14.69)	4.60 (-3.18, 12.38)	-	3.02 (-7.43, 13.31)	-1.01 (-9.80, 8.61)
GT-density	-	6.26 (-10.41, 23.45)	5.10 (-5.21, 15.10)	-	15.66 (-4.34, 35.53)	7.00 (-5.04, 19.14)
GT-density squared	-	-0.54 (-9.52, 9.13)	0.22 (-6.70, 7.08)	-	-3.60 (-15.3, 8.04)	2.11 (-7.20, 11.36)
BT- × GT-density	-	-0.22 (-0.66, 0.20)	-0.20 (-0.60, 0.17)	-	-0.37 (-0.81, 0.07)	-0.35 (-0.78, 0.10)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Female ID	0.12 (0.09, 0.16)	0.13 (0.10, 0.17)	0.13 (0.10, 0.16)	0.20 (0.16, 0.24)	0.24 (0.19, 0.29)	0.23 (0.19, 0.29)
Subplot-Year ID	0.10 (0.06, 0.16)	0.17 (0.10, 0.27)	0.12 (0.07, 0.18)	0.14 (0.08, 0.23)	0.14 (0.08, 0.22)	0.14 (0.08, 0.22)
Brood ID	0.48 (0.40, 0.58)	0.46 (0.38, 0.55)	0.47 (0.39, 0.57)	1.17 (1.00, 1.36)	1.10 (0.94, 1.28)	1.11 (0.95, 1.30)
Observer ID	0.03 (0.01, 0.04)	0.02 (0.01, 0.03)	0.02 (0.01, 0.03)	0.03 (0.01, 0.04)	0.02 (0.01, 0.04)	0.02 (0.01, 0.04)
Residual	0.32 (0.30, 0.35)	0.32 (0.30, 0.35)	0.32 (0.30, 0.35)	0.89 (0.83, 0.95)	0.89 (0.83, 0.94)	0.89 (0.83, 0.95)

[a] Reference category: estimate is for tits in areas of average size breeding in the “Low-Low” (BT-GT) box density treatment (all models except c, f) under average realized breeding densities (all models except a, d) for nestlings with species-average tarsus length at nestling age 15 (BTs) or 14 (GTs).

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Age (days) was centered around nestling age 15 (BTs) or 14 (GTs).

[d] Tarsus was mean-centered within each species.

[e] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[f] Estimate for the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[g] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[h] Estimate for the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

Table S13. Estimated effect sizes and 95% credible intervals (CIs) for predictors of nestling body size for blue tits (BT) (models a-c) and great tits (GT) (models d-f) for models fitting (main and interaction effects of) the number (no) of BT- and GT-boxes (nest box density treatment; a, d), realized breeding density (c, f) or both (b, e). We further fit subplot area size and age as fixed effects. Parameters were estimated using LMMs assuming a Gaussian error distribution.

	Nestling body size					
	Blue tits (n= 1495 nestlings)			Great tits (n= 1810 nestlings)		
	(a) Treatment	(b) Treatment and realized density	(c) Realized density	(d) Treatment	(e) Treatment and realized density	(f) Realized density
Intercept [a]	16.63 (16.39, 16.86)	16.71 (16.37, 17.05)	16.58 (16.44, 16.73)	19.3 (19.06, 19.55)	19.62 (19.3, 19.93)	19.35 (19.19, 19.5)
Area [b]	0.05 (-0.05, 0.15)	0.14 (-0.01, 0.29)	0.14 (0.01, 0.27)	-0.11 (-0.23, 0.02)	0.02 (-0.14, 0.18)	-0.01 (-0.17, 0.14)
Age [c]	0.12 (-0.12, 0.36)	0.14 (-0.09, 0.38)	0.14 (-0.12, 0.39)	0.43 (0.11, 0.74)	0.47 (0.15, 0.78)	0.45 (0.14, 0.75)
Nest box density treatment						
BT-box no [d]	-0.18 (-0.44, 0.08)	-0.28 (-0.57, 0.02)	-	0.02 (-0.30, 0.35)	-0.12 (-0.46, 0.22)	-
GT-box no [e]	-0.04 (-0.31, 0.23)	-0.08 (-0.48, 0.31)	-	-0.06 (-0.36, 0.24)	-0.43 (-0.85, -0.00)	-
BT- × GT-box no [f]	0.17 (-0.17, 0.51)	0.33 (-0.07, 0.74)	-	-0.06 (-0.48, 0.37)	0.13 (-0.34, 0.61)	-
Realized breeding density [g]						
BT-density	-	5.35 (0.31, 10.49)	3.55 (-0.73, 7.97)	-	-0.28 (-7.95, 6.84)	2.23 (-4.34, 8.68)
BT- density squared	-	-1.21 (-5.67, 3.10)	-1.11 (-4.74, 2.39)	-	4.36 (-1.04, 9.94)	1.50 (-3.11, 6.01)
GT-density	-	-2.46 (-10.28, 4.84)	-0.81 (-5.40, 3.97)	-	11.53 (1.76, 21.70)	4.04 (-2.15, 10.53)
GT-density squared	-	-0.05 (-4.22, 4.10)	-0.28 (-3.35, 3.01)	-	-3.26 (-9.49, 2.67)	0.23 (-4.78, 5.11)
BT- × GT-density	-	0.04 (-0.15, 0.23)	0.08 (-0.10, 0.26)	-	-0.20 (-0.42, 0.03)	-0.21 (-0.43, 0.00)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Female ID	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Subplot-Year ID	0.02 (0.01, 0.03)	0.02 (0.01, 0.03)	0.01 (0.01, 0.02)	0.05 (0.03, 0.08)	0.04 (0.02, 0.05)	0.04 (0.02, 0.06)
Brood ID	0.18 (0.15, 0.21)	0.18 (0.15, 0.22)	0.19 (0.16, 0.22)	0.41 (0.36, 0.48)	0.41 (0.35, 0.48)	0.41 (0.36, 0.48)
Observer ID	0.06 (0.04, 0.09)	0.06 (0.04, 0.09)	0.06 (0.04, 0.09)	0.03 (0.02, 0.05)	0.03 (0.02, 0.05)	0.03 (0.02, 0.05)
Residual	0.25 (0.23, 0.27)	0.25 (0.23, 0.26)	0.25 (0.23, 0.26)	0.40 (0.37, 0.43)	0.40 (0.38, 0.43)	0.40 (0.38, 0.43)

[a] Reference category: estimate is for tits in areas of average size breeding in the “Low-Low” (BT-GT) box density treatment (all models except c, f) under average realized breeding densities (all models except a, d) for nestlings the age of 15 (BTs) or 14 (GTs).

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Age (days) was centered around nestling age 15 (BTs) or 14 (GTs).

[d] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[e] Estimate for the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[f] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[g] Estimate for the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

Table S14. Estimated effect sizes and 95% credible intervals (CIs) for predictors of the binary probability to produce fledglings for blue tits (BT) (models a-c) and great tits (GT) (models d-f) for models fitting (main and interaction effects of) the number (no) of BT- and GT-boxes (nest box density treatment; a, d), realized breeding density (c, f) or both (b, e). We further fit subplot area as fixed effect. Parameters were estimated using GLMMs assuming a binomial error distribution with the logit link function. Parameters are shown in the latent scale.

	Probability to produce fledglings					
	Blue tits (n= 269 broods)			Great tits (n= 397 broods)		
	(a) Treatment	(b) Treatment and realized density	(c) Realized density	(d) Treatment	(e) Treatment and realized density	(f) Realized density
Intercept [a]	0.37 (-0.29, 1.05)	0.77 (-0.24, 1.79)	1.14 (0.70, 1.60)	1.01 (0.45, 1.56)	0.90 (0.20, 1.59)	0.56 (0.29, 0.83)
Area [b]	0.15 (-0.25, 0.54)	0.41 (-0.19, 0.98)	0.42 (-0.13, 0.99)	0.03 (-0.25, 0.30)	-0.02 (-0.37, 0.34)	-0.03 (-0.37, 0.30)
Nest box density treatment						
BT-box no [c]	0.86 (-0.10, 1.78)	0.79 (-0.16, 1.80)	-	0.01 (-0.76, 0.80)	0.03 (-0.79, 0.89)	-
GT-box no [d]	1.01 (0.05, 1.95)	0.24 (-1.11, 1.59)	-	-0.28 (-0.95, 0.43)	-0.41 (-1.32, 0.52)	-
BT- × GT-box no [e]	-1.37 (-2.63, -0.08)	-0.92 (-2.44, 0.51)	-	-0.33 (-1.28, 0.62)	-0.34 (-1.42, 0.71)	-
Realized breeding density [f]						
BT-density	-	2.55 (-4.78, 9.92)	4.92 (-2.16, 12.39)	-	-0.75 (-7.83, 6.50)	1.66 (-4.78, 8.08)
BT- density squared	-	1.20 (-5.92, 8.23)	0.98 (-5.61, 7.36)	-	3.64 (-2.08, 9.27)	2.10 (-2.69, 7.21)
GT-density	-	3.36 (-9.41, 16.16)	1.57 (-6.51, 9.53)	-	-1.68 (-10.89, 7.55)	-7.33 (-13.68, -1.19)
GT-density squared	-	-6.30 (-13.05, 0.06)	-5.95 (-11.37, -0.53)	-	-3.27 (-9.43, 3.07)	-1.47 (-6.26, 3.27)
BT- × GT-density	-	0.06 (-0.63, 0.74)	-0.01 (-0.69, 0.69)	-	0.48 (-0.03, 1.00)	0.40 (-0.06, 0.88)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Female ID [g]	NE	NE	NE	NE	NE	NE
Subplot-Year ID	0.17 (0.10, 0.26)	0.05 (0.03, 0.08)	0.13 (0.08, 0.20)	0.05 (0.03, 0.08)	0.00 (0.00, 0.00)	0.01 (0.01, 0.01)
Residual [h]	$\Pi^2/^3$	$\Pi^2/^3$	$\Pi^2/^3$	$\Pi^2/^3$	$\Pi^2/^3$	$\Pi^2/^3$

[a] Reference category: estimate is for tits in areas of average size breeding in the “Low-Low” (BT-GT) box density treatment (all models except c, f) under average realized breeding densities (all models except a, d).

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[d] Estimate for the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[e] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[f] Estimate for the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

[g] Female ID was excluded as this prevented model convergence (NE).

[h] Residual variance fixed to $\pi^{2/3}$ (following Nakagawa and Schielzeth 2010).

Table S15. Estimated effect sizes and 95% credible intervals (CIs) for predictors of the number of fledglings for blue tits (BT) (models a-c) and great tits (GT) (models d-f) for models fitting (main and interaction effects of) the number (no) of BT- and GT-boxes (nest box density treatment; a, d), realized breeding density (c, f) or both (b, e). We further fit subplot area size as fixed effect. Parameters were estimated using GLMMs following a Poisson error distribution and a log-link function. Parameters are shown in the latent scale.

	Number of fledglings					
	Blue tits (n= 192 broods)			Great tits (n= 265 broods)		
	(a) Treatment	(b) Treatment and realized density	(c) Realized density	(d) Treatment	(e) Treatment and realized density	(f) Realized density
Intercept [a]	2.00 (1.86, 2.14)	2.10 (1.90, 2.31)	2.06 (1.99, 2.12)	1.86 (1.75, 1.98)	1.91 (1.77, 2.06)	1.92 (1.85, 1.98)
Area [b]	-0.05 (-0.12, 0.01)	-0.02 (-0.12, 0.09)	-0.01 (-0.11, 0.08)	0.00 (-0.06, 0.06)	-0.01 (-0.09, 0.07)	0.00 (-0.08, 0.07)
Nest box density treatment						
BT-box no [c]	0.00 (-0.17, 0.17)	-0.04 (-0.24, 0.15)	-	0.06 (-0.11, 0.23)	0.04 (-0.14, 0.21)	-
GT-box no [d]	0.11 (-0.06, 0.28)	-0.02 (-0.29, 0.24)	-	0.05 (-0.10, 0.20)	0.02 (-0.19, 0.23)	-
BT- × GT-box no [e]	-0.10 (-0.33, 0.12)	-0.04 (-0.31, 0.24)	-	-0.16 (-0.37, 0.06)	-0.11 (-0.34, 0.12)	-
Realized breeding density [f]						
BT-density	-	0.10 (-1.00, 1.18)	0.06 (-0.97, 1.13)	-	-0.43 (-1.82, 0.86)	-0.25 (-1.49, 1.03)
BT- density squared	-	0.28 (-0.73, 1.26)	0.14 (-0.80, 1.04)	-	-0.02 (-1.07, 1.06)	-0.06 (-0.97, 0.86)
GT-density	-	1.22 (-0.57, 3.01)	0.94 (-0.21, 2.11)	-	0.46 (-1.42, 2.35)	0.15 (-1.12, 1.39)
GT-density squared	-	-0.51 (-1.42, 0.40)	-0.56 (-1.29, 0.20)	-	-0.18 (-1.37, 1.02)	-0.21 (-1.19, 0.73)
BT- × GT-density	-	-0.06 (-0.18, 0.07)	-0.07 (-0.18, 0.05)	-	-0.09 (-0.20, 0.02)	-0.10 (-0.21, 0.01)
	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Female ID	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Subplot-Year ID	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.01 (0.00, 0.01)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Residual [g]	0.13 (0.15, 0.11)	0.12 (0.14, 0.09)	0.12 (0.13, 0.11)	0.14 (0.16, 0.13)	0.14 (0.16, 0.12)	0.14 (0.15, 0.13)

[a] Reference category: estimate is for tits in areas of average size breeding in the “Low-Low” (BT-GT) box density treatment (all models except c, f) under average realized breeding densities (all models except a, d).

[b] Effect of mean-centered area size (in ha). Mean area size was 6.48 ha.

[c] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-box treatment when GT-box numbers were “Low”.

[d] Estimate for the effect of the “High” GT-box treatment relative to the “Low” GT-box treatment when BT-box numbers were “Low”.

[e] Estimate for the effect of the “High” BT-box treatment relative to the “Low” BT-treatment for the “High” GT-treatment expressed as a difference from the same effect observed in the “Low” GT-treatment.

[f] Estimate for the effects of realized breeding densities (pairs/ha; mean-centered) of BT and GT (linear and quadratic effects, and the interaction between the linear effects).

[g] Defined as $\ln(1/\exp(\beta_o)+1)$ (following Nakagawa and Schielzeth 2010).

Text S3. List of blue tit playback songs downloaded from Xeno-canto (Planqué et al. 2005):

Buhl, J. (2015) *Cyanistes caeruleus* in Germany, XC268505, www.xeno-canto.org

Buhl, J. (2017) *Cyanistes caeruleus* in Germany, XC354752, XC354754, XC354778, XC368870, and XC354809, www.xeno-canto.org

Buhl, J. (2018) *Cyanistes caeruleus* in Germany, XC399405, XC399406, XC399392, and XC399407, www.xeno-canto.org

Deroussen, F. (2020) *Cyanistes caeruleus* in France, XC533328, www.xeno-canto.org

Fischer, J. (2015) *Cyanistes caeruleus* in Switzerland, XC246955, www.xeno-canto.org

Fischer, J. (2015) *Cyanistes caeruleus* in Switzerland, XC412012, www.xeno-canto.org

Matusiak, J. (2020) *Cyanistes caeruleus* in Austria XC530818, www.xeno-canto.org

Risch, S. (2019) *Cyanistes caeruleus* in Germany, XC513785 and XC513786, www.xeno-canto.org

Risch, S. (2021) *Cyanistes caeruleus* in Germany, XC615122, www.xeno-canto.org

Schnabel, T. (2016) *Cyanistes caeruleus* in Germany, XC367169, www.xeno-canto.org

Schnabel, T. (2017) *Cyanistes caeruleus* in France, XC384601 and XC384605, www.xeno-canto.org

Table S16. Estimated effect sizes and 95% credible intervals (CIs) for predictors of the minimal approach distance of blue tit (BT) males, great tit (GT) males, and males of both species during the first test (models a, c and e, respectively) and during the second test (models b, d, and f, respectively) to a con- vs. a heterospecific mount. Estimates are shown in a latent scale. Parameters were estimated using LMMs with a Gaussian error distribution.

	Blue tits		Great tits		Both species	
	(a) First test (n = 137 observations)	(b) Second test (n = 142 observations)	(c) First test (n = 151 observations)	(d) Second test (n = 166 observations)	(e) First test (n = 288 observations)	(f) Second test (n = 308 observations)
Fixed effects	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Intercept [a]	-1.72 (-2.14, -1.28)	-1.62 (-2.17, -1.05)	-2.81 (-3.37, -2.27)	-2.84 (-3.30, -2.39)	-1.77 (-2.23, -1.31)	-1.49 (-1.91, -1.07)
Mount species						
Current test [b]	-0.54 (-1.04, -0.03)	-0.46 (-1.1, 0.17)	0.17 (-0.53, 0.84)	0.07 (-0.34, 0.51)	-0.63 (-1.26, -0.03)	-0.60 (-1.11, -0.12)
Previous test [c]	-	-0.23 (-0.66, 0.20)	-	0.07 (-0.31, 0.45)	-	-0.26 (-0.69, 0.16)
Focal species						
GT [d]	-	-	-	-	-1.05 (-1.49, -0.65)	-1.42 (-1.93, -0.92)
Mount species × focal species						
Current test × GT [e]	-	-	-	-	0.77 (0.20, 1.36)	0.68 (0.11, 1.25)
Previous test × GT [f]	-	-	-	-	-	0.35 (-0.22, 0.89)
Random effects	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)	σ^2 (95% CI)
Male ID	0.00 (0.00, 0.00)	0.41 (0.30, 0.54)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.11 (0.09, 0.13)
Subplot-year ID	0.00 (0.00, 0.00)	0.01 (0.00, 0.02)	0.22 (0.11, 0.37)	0.04 (0.02, 0.07)	0.11 (0.05, 0.18)	0.00 (0.00, 0.00)
Mount ID	0.05 (0.02, 0.10)	0.12 (0.05, 0.22)	0.21 (0.10, 0.38)	0.02 (0.01, 0.05)	0.14 (0.07, 0.24)	0.05 (0.02, 0.10)
Song ID	0.00 (0.00, 0.00)	0.31 (0.19, 0.46)	0.20 (0.12, 0.3)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.02 (0.02, 0.04)
Observer ID	0.26 (0.11, 0.46)	0.27 (0.12, 0.49)	0.00 (0.00, 0.00)	0.22 (0.11, 0.37)	0.08 (0.04, 0.15)	0.14 (0.07, 0.24)
Residual	1.72 (1.34, 2.19)	0.79 (0.63, 0.99)	1.30 (1.03, 1.63)	1.47 (1.19, 1.83)	1.62 (1.38, 1.91)	1.42 (1.22, 1.66)
Sample size	n	n	n	n	n	n
Male ID	128	131	140	157	268	288
Subplot-year ID	24	23	24	24	24	24
Mount ID	14	14	13	14	14	15
Song ID	46	47	49	49	50	50
Observer ID	16	16	16	15	17	16

1 Reference category: estimate is for the minimal approach distance of a (BT; both species models) male presented with a BT mount during the current test (all modes). Approach distance was multiplied by "-1" prior to the analyses.
2 Estimate (for BT males; both species models) is the effect during the current test of a GT mount relative to a BT mount.
3 Estimate (for BT males; both species models) is the effect during the previous test of a GT mount relative to a BT mount (second test models).
4 Estimate is the difference between BTs (reference) and GTs (main effect).
5 Estimate is the difference between BTs (reference) and GTs described in the footnote b.
6 Estimate is the difference between BTs (reference) and GTs described in the footnote c.

Text S4. From a proximate perspective, there are multiple non-exclusive explanations for why the phenotypic integration between behaviour and morphology, and behaviour and physiology, varied significantly among datasets. This is evident when inspecting the mathematical equation demonstrating how different variance components contribute to observed correlations. Specifically, the among-individual correlation (r_I) between two focal traits (A and B) that we used as input in our SEMs, is underpinned by six distinct variance components:

$$r_I = r_G \sqrt{\left(\frac{V_{G_A}}{V_{G_A}+V_{PE_A}}\right)\left(\frac{V_{G_B}}{V_{G_B}+V_{PE_B}}\right)} + r_{PE} \sqrt{\left(\frac{V_{PE_A}}{V_{G_A}+V_{PE_A}}\right)\left(\frac{V_{PE_B}}{V_{G_B}+V_{PE_B}}\right)} \qquad \text{(Eqn. S1)}$$

where r_G and r_{PE} represent the genetic and permanent environmental correlations, respectively, and V_G and V_{PE} the phenotypic variance in the focal trait (A, B) attributable to additive genetic and permanent environmental effects, respectively (Searle 1961; Dingemanse and Dochtermann 2013). r_I can differ between datasets (species, populations, sexes) for the following reasons. First, assume that neither r_G nor r_{PE} differs between datasets but one is positive and the other is negative. Changes in the sign of r_I then result from differences between species, populations, or sexes in the amount of additive genetic variance or the influence of permanent environmental effects on the traits' expression. Second, assume instead that the magnitude of V_G and V_{PE} does not differ between species, populations, or sexes. Differences in the sign of r_I can then only result from differences in r_G and r_{PE}.

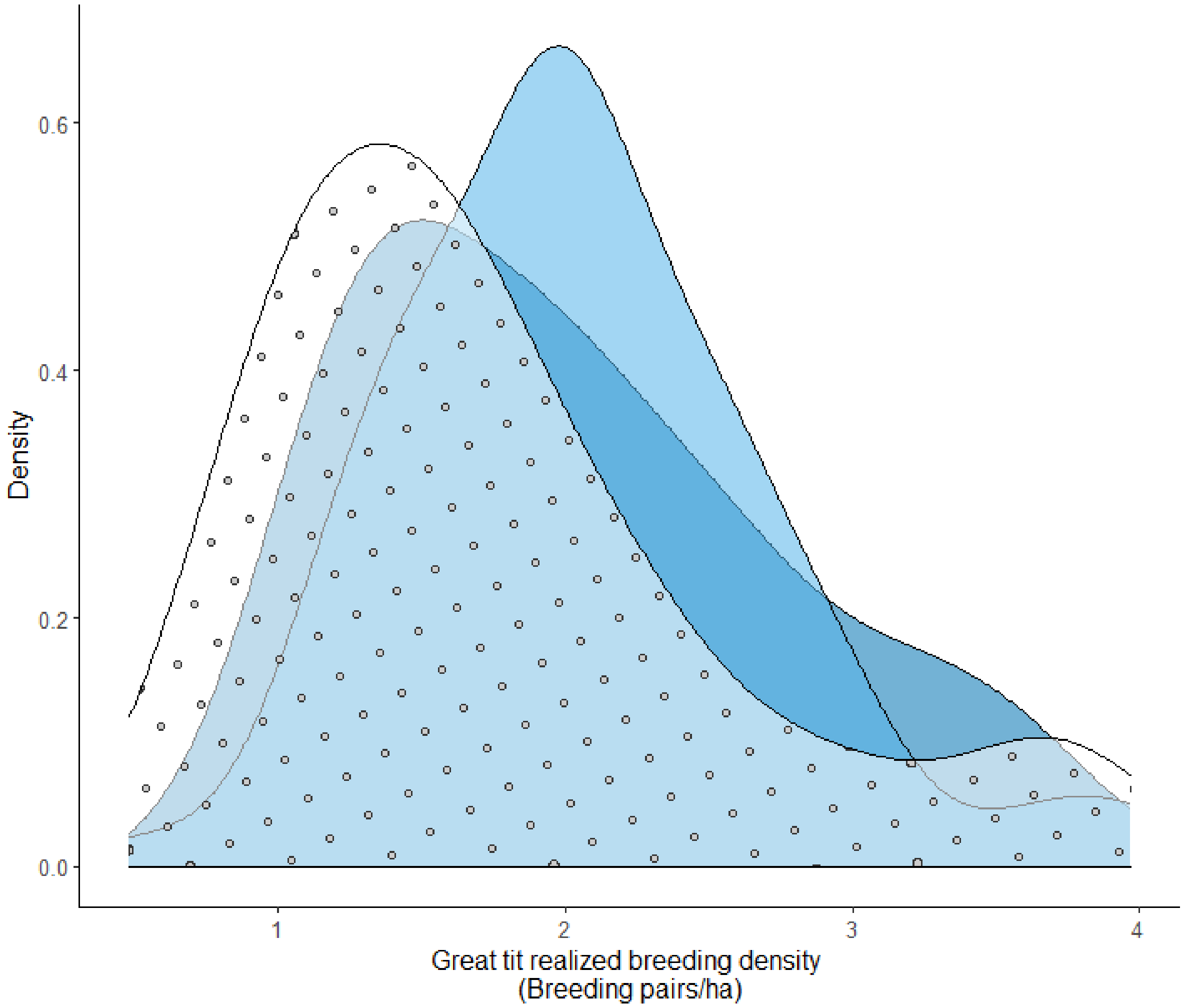


Figure S3. Density plot showing the great tit realized breeding density (x-axis) in our study (white density plot with open circles) and in a study published by Mouchet et al. (2021) (density plots in shades of blue). Each shade of blue represents different populations: Lauwersmeer (dark blue) and Starnberg (light blue) populations. The setup of the nest boxes in the three populations was identical; nest boxes were placed in a grid with 50 m between adjacent boxes.

7. References

Alatalo R V. 1982. Evidence for interspecific competition among European tits Parus spp.: a review. Ann Zool Fennici. 19(4):309–317.

Andreu J, Barba E. 2006. Breeding dispersal of great tits Parus major in a homogeneous habitat: effects of sex, age, and mating status. Ardea. 94(1):45.

Araya-Ajoy YG, Bolstad GH, Brommer J, Careau V, Dingemanse NJ, Wright J. 2018. Demographic measures of an individual's "pace of life": fecundity rate, lifespan, generation time, or a composite variable? Behav Ecol Sociobiol. 72(5):1–14. doi:10.1007/s00265-018-2477-7.

Araya-Ajoy YG, Dingemanse NJ. 2014. Characterizing behavioural "characters": An evolutionary framework. Proceedings of the Royal Society B: Biological Sciences. 281(1776):20132645. doi:10.1098/rspb.2013.2645.

Araya-Ajoy YG, Dingemanse NJ. 2017. Repeatability, heritability, and age-dependence of seasonal plasticity in aggressiveness in a wild passerine bird. van de Pol M, editor. Journal of Animal Ecology. 86(2):227–238. doi:10.1111/1365-2656.12621. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2656.12621.

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 (N Y). 70(10):2308–2321. doi:10.1111/evo.13024. https://academic.oup.com/evolut/article/70/10/2308/6851951.

Armbruster WS, Pélabon C, Bolstad GH, Hansen TF. 2014. Integrated phenotypes: understanding trait covariation in plants and animals. Philosophical Transactions of the Royal Society B: Biological Sciences. 369(1649):20130245. doi:10.1098/rstb.2013.0245.

Armbruster WS, Schwaegerle KE. 1996. Causes of covariation of phenotypic traits among populations. J Evol Biol. 9(3):261–276. doi:10.1046/j.1420-9101.1996.9030261.x.

Auld JR, Agrawal AA, Relyea RA. 2010. Re-evaluating the costs and limits of adaptive phenotypic plasticity. Proceedings of the Royal Society B: Biological Sciences. 277(1681):503–511. doi:10.1098/rspb.2009.1355. https://royalsocietypublishing.org/doi/10.1098/rspb.2009.1355.

Balaban-Feld J, Mitchell WA, Kotler BP, Vijayan S, Tov Elem LT, Rosenzweig ML, Abramsky Z. 2019. Individual willingness to leave a safe refuge and the trade-off between food and safety: a test with social fish. Proceedings of the Royal Society B: Biological Sciences. 286(1907):20190826. doi:10.1098/rspb.2019.0826. https://royalsocietypublishing.org/doi/10.1098/rspb.2019.0826.

Barber I, Dingemanse NJ. 2010. Parasitism and the evolutionary ecology of animal personality. Philosophical Transactions of the Royal Society B: Biological Sciences.

365(1560):4077–4088. doi:10.1098/rstb.2010.0182. https://royalsocietypublishing.org/doi/10.1098/rstb.2010.0182.

Bates D, Mächler M, Bolker BM, Walker SC. 2015. Fitting linear mixed-effects models using lme4. J Stat Softw. 67(1):1–48. doi:10.18637/jss.v067.i01.

Beard KH, Choi RT, Meiners JM. 2018. Competition and coexistence in plant communities : intraspecific competition is stronger than interspecific competition. Ecol Lett. 21(9):1319–1329. doi:10.1111/ele.13098.

Bell AM, Hankison SJ, Laskowski KL. 2009. The repeatability of behaviour: a meta-analysis. Anim Behav. 77(4):771–783. doi:10.1016/j.anbehav.2008.12.022.Bell.

Betini GS, Norris DR. 2012. The relationship between personality and plasticity in tree swallow aggression and the consequences for reproductive success. Anim Behav. 83(1):137–143. doi:10.1016/j.anbehav.2011.10.018.

Betts MM. 1955. The food of titmice in oak woodland. J Anim Ecol. 24(2):282. doi:10.2307/1715.

Biro PA, Adriaenssens B. 2013. Predictability as a personality trait: consistent differences in intraindividual behavioral variation. Am Nat. 182(5):621–629. doi:10.1086/673213. https://www.journals.uchicago.edu/doi/10.1086/673213.

Blondel J, Perret P, Anstett MC, Thébaud C. 2002. Evolution of sexual size dimorphism in birds: Test of hypotheses using blue tits in contrasted Mediterranean habitats. J Evol Biol. 15(3):440–450. doi:10.1046/j.1420-9101.2002.00397.x.

Blumstein DT, Daniel JC, Evans CS. 2006. JWatcher 1.0: An introductory user’s guide. (January):211.

Both C. 1998. Density dependence of clutch size: habitat heterogeneity or individual adjustment? Journal of Animal Ecology. 67(4):659–666. doi:10.1046/j.1365-2656.1998.00227.x. https://besjournals.onlinelibrary.wiley.com/doi/10.1046/j.1365-2656.1998.00227.x.

Both C, Tinbergen JM, Visser ME. 2000. Adaptive density dependence of avian clutch size. Ecology. 81(12):3391. doi:10.2307/177502. http://www.jstor.org/stable/177502?origin=crossref.

Both C, Visser ME. 2003. Density dependence, territoriality, and divisibility of resources: from optimality models to population processes. Am Nat. 161(2):326–336. doi:10.1086/346098. https://www.journals.uchicago.edu/doi/10.1086/346098.

Both C, Visser ME, Verboven N. 1999. Density-dependent recruitment rates in great tits: The importance of being heavier. Proceedings of the Royal Society B: Biological Sciences. 266(1418):465–469. doi:10.1098/rspb.1999.0660.

Brodie ED. 1992. Correlational selection for color pattern and antipredator behavior in the garter snake Thamnophis ordinoides. Evolution (N Y). 46(5):1284–1298. doi:10.1111/j.1558-5646.1992.tb01124.x.

Brodie ED. 1993. Homogeneity of the genetic variance-covariance matrix for antipredator traits in two natural populations of the garter snake Thamnophis ordinoides. Evolution (N Y). 47(3):844–854. doi:10.1111/j.1558-5646.1993.tb01238.x.

Brodie ED, Moore AJ, Janzen FJ. 1995. Visualizing and quantifying natural selection. Trends Ecol Evol. 10(8):313–318. doi:10.1016/S0169-5347(00)89117-X.

Brown AL, Robinson BW. 2016. Variation in behavioural plasticity regulates consistent individual differences in Enallagma damselfly larvae. Anim Behav. 112:63–73. doi:10.1016/j.anbehav.2015.11.018.

Brown JL. 1964. The evolution of diversity in avian territorial systems. Wilson Bull. 76(2):160–169. https://www.jstor.org/stable/4159278.

Brown JL. 1969. Territorial behavior and population regulation in birds: a review and re-evaluation. Wilson Bull. 81(3):293–329. https://www.jstor.org/stable/4159863.

Brown WL, Wilson EO. 1956. Character displacement. Syst Zool. 5(2):49–64. doi:10.2307/2411924.

Carere C, Van Oers K. 2004. Shy and bold great tits (Parus major): body temperature and breath rate in response to handling stress. Physiol Behav. 82(5):905–912. doi:10.1016/j.physbeh.2004.07.009.

Carter AJ, Feeney WE, Marshall HH, Cowlishaw G, Heinsohn R. 2013. Animal personality: what are behavioural ecologists measuring? Biological Reviews. 88(2):465–475. doi:10.1111/brv.12007.

Chesson P. 2000. Mechanisms of maintenance of species diversity. Annu Rev Ecol Syst. 31(1):343–366. doi:10.1146/annurev.ecolsys.31.1.343.

Class B, Brommer J. 2020. Contrasting multilevel relationships between behavior and body mass in blue tit nestlings. Behavioral Ecology. 31(3):702–710. doi:10.1093/beheco/araa014.

Class B, Brommer JE. 2015. A strong genetic correlation underlying a behavioural syndrome disappears during development because of genotype–age interactions. Proceedings of the Royal Society B: Biological Sciences. 282(1809):20142777. doi:10.1098/rspb.2014.2777.

Coppens CM, De Boer SF, Koolhaas JM. 2010. Coping styles and behavioural flexibility: towards underlying mechanisms. Philosophical Transactions of the Royal Society B: Biological Sciences. 365(1560):4021–4028. doi:10.1098/rstb.2010.0217.

Cowie RJ, Hinsley SA. 1988. Feeding ecology of great tits (Parus major) and blue tits (Parus caeruleus), breeding in suburban gardens. J Anim Ecol. 57(2):611. doi:10.2307/4928. https://www.jstor.org/stable/4928?origin=crossref.

Crombie AAC. 1947. Interspecific competition. Journal of Animal Ecology. 16(1):44–73. doi:https://doi.org/10.2307/1506.

Dall SRX, Houston AI, McNamara JM. 2004. The behavioural ecology of personality: consistent individual differences from an adaptive perspective. Ecol Lett. 7(8):734–739. doi:10.1111/j.1461-0248.2004.00618.x.

Dammhahn M, Dingemanse NJ, Niemelä PT, Réale D. 2018. Pace-of-life syndromes: a framework for the adaptive integration of behaviour, physiology and life history. Behav Ecol Sociobiol. 72(3):1–8. doi:10.1007/s00265-018-2473-y.

Darwin C. 1859. On the origin of species. London.

DeWitt TJ, Sih A, Wilson DS. 1998. Costs and limits of phenotypic plasticity. Trends Ecol Evol. 13(2):77–81. doi:10.1016/S0169-5347(97)01274-3. https://linkinghub.elsevier.com/retrieve/pii/S0169534797012743.

Dhondt AA. 1977. Interspecific competition between great and blue tit. Nature. 268(5620):521–523. doi:10.1038/268521a0.

Dhondt AA. 2010. Effects of competition on great and blue tit reproduction: intensity and importance in relation to habitat quality. Journal of Animal Ecology. 79(1):257–265. doi:10.1111/j.1365-2656.2009.01624.x. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/j.1365-2656.2009.01624.x.

Dhondt AA. 2012. Interspecific competition in birds. Oxford, UK: Oxford University Press.

Dhondt AA, Adriaensen F. 1999. Experiments on competition between great and blue tit: effects on blue tit reproductive success and population processes. Ostrich. 70(1):39–48. doi:10.1080/00306525.1999.9639748.

Dhondt AA, Eyckerman R. 1980. Competition between the great tit and the blue tit outside the breeding season in field experiments. Ecology. 61(6):1291–1296. doi:10.2307/1939036. https://esajournals.onlinelibrary.wiley.com/doi/10.2307/1939036.

Dingemanse N, Dochtermann N. 2014. Individual behaviour: behavioural ecology meets quantitative genetics. In: Quantitative genetics in the wild. p. 54–67.

Dingemanse NJ, Barber I, Dochtermann NA. 2020. Non-consumptive effects of predation: does perceived risk strengthen the genetic integration of behaviour and morphology in stickleback? Ecol Lett. 23(1):107–118. doi:10.1111/ele.13413.

Dingemanse NJ, Both C, Drent PJ, Van Oers K, Van Noordwijk AJ. 2002. Repeatability and heritability of exploratory behaviour in great tits from the wild. Anim Behav. 64(6):929–938. doi:10.1006/anbe.2002.2006.

Dingemanse NJ, Bouwman KM, van de Pol M, van Overveld T, Patrick SC, Matthysen E, Quinn JL. 2012. Variation in personality and behavioural plasticity across four populations of the great tit Parus major. Journal of Animal Ecology. 81(1):116–126. doi:10.1111/j.1365-2656.2011.01877.x.

Dingemanse NJ, Dochtermann NA. 2013. Quantifying individual variation in behaviour: mixed-effect modelling approaches. Journal of Animal Ecology. 82(1):39–54. doi:10.1111/1365-2656.12013.

Dingemanse NJ, Dochtermann N, Wright J. 2010. A method for exploring the structure of behavioural syndromes to allow formal comparison within and between data sets. Anim Behav. 79(2):439–450. doi:10.1016/j.anbehav.2009.11.024.

Dingemanse NJ, Kazem AJN, Réale D, Wright J. 2010. Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol Evol. 25(2):81–89. doi:10.1016/j.tree.2009.07.013. https://linkinghub.elsevier.com/retrieve/pii/S0169534709002432.

Dingemanse NJ, Moiron M, Araya-Ajoy YG, Mouchet A, Abbey-Lee RN. 2020. Individual variation in age-dependent reproduction: fast explorers live fast but senesce young? Journal of Animal Ecology. 89(2):601–613. doi:10.1111/1365-2656.13122.

Dingemanse NJ, Wolf M. 2010. Recent models for adaptive personality differences: a review. Philosophical Transactions of the Royal Society B: Biological Sciences. 365(1560):3947–3958. doi:10.1098/rstb.2010.0221.

Dingemanse NJ, Wolf M. 2013. Between-individual differences in behavioural plasticity within populations: causes and consequences. Anim Behav. 85(5):1031–1039. doi:10.1016/j.anbehav.2012.12.032. https://linkinghub.elsevier.com/retrieve/pii/S0003347213000092.

Dochtermann NA. 2011. Testing Cheverud’s conjecture for behavioral correlations and behavioral syndromes. Evolution (N Y). 65(6):1814–1820. doi:10.1111/j.1558-5646.2011.01264.x.

Dochtermann NA, Dingemanse NJ. 2013. Behavioral syndromes as evolutionary constraints. Behavioral Ecology. 24(4):806–811. doi:10.1093/beheco/art002.

Duckworth RA. 2006a. Behavioral correlations across breeding contexts provide a mechanism for a cost of aggression. Behavioral Ecology. 17(6):1011–1019. doi:10.1093/beheco/arl035.

Duckworth RA. 2006b. Aggressive behaviour affects selection on morphology by influencing settlement patterns in a passerine bird. Proceedings of the Royal Society B: Biological Sciences. 273(1595):1789–1795. doi:10.1098/rspb.2006.3517.

Duckworth RA. 2009. The role of behavior in evolution: a search for mechanism. Evol Ecol. 23(4):513–531. doi:10.1007/s10682-008-9252-6. http://link.springer.com/10.1007/s10682-008-9252-6.

Engen S, Lande R, Sæther BE. 2013. A quantitative genetic model of r- and K-selection in a fluctuating population. American Naturalist. 181(6):725–736. doi:10.1086/670257.

Enquist M, Leimar O. 1987. Evolution of fighting behaviour: the effect of variation in resource value. J Theor Biol. 127(2):187–205. doi:10.1016/S0022-5193(87)80130-3. https://linkinghub.elsevier.com/retrieve/pii/S0022519387801303.

Felsenstein J. 1976. The theoretical population genetics of variable selection and migration. Annu Rev Genet. 10(1):253–280. doi:10.1146/annurev.ge.10.120176.001345. https://www.annualreviews.org/doi/10.1146/annurev.ge.10.120176.001345.

Filazzola A, Cahill JF. 2021. Replication in field ecology: identifying challenges and proposing solutions. Methods Ecol Evol. 12(10):1780–1792. doi:10.1111/2041-210X.13657.

Fisher RA. 1930. The genetical theory of natural selection. Oxford: Clarendon Press. https://www.biodiversitylibrary.org/bibliography/27468.

Fox J. 2006. Structural equation modeling with the sem package in R. Structural Equation Modeling. 13(3):465–486. doi:10.1207/s15328007sem1303_7.

Friard O, Gamba M. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods Ecol Evol. 7(11):1325–1330. doi:10.1111/2041-210X.12584.

Furness EN, Furness RW. 2016. Influences of feather wear and age on wing lengths of Blue Tits Cyanistes caeruleus. Ringing and Migration. 31(2):102–106. doi:10.1080/03078698.2016.1260838.

Gaona-Gordillo I, Holtmann B, Mouchet A, Hutfluss A, Sánchez-Tójar A, Dingemanse NJ. 2023. Are animal personality, body condition, physiology and structural size integrated? A comparison of species, populations and sexes, and the value of study replication. Journal of Animal Ecology. 92(9):1707–1718. doi:10.1111/1365-2656.13966. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2656.13966.

Garamszegi LZ, Markó G, Herczeg G. 2012. A meta-analysis of correlated behaviours with implications for behavioural syndromes: mean effect size, publication bias, phylogenetic effects and the role of mediator variables. Evol Ecol. 26(5):1213–1235. doi:10.1007/s10682-012-9589-8.

García-Navas V, Ferrer ES, Sanz JJ. 2013. Prey choice, provisioning behaviour, and effects of early nutrition on nestling phenotype of titmice. Ecoscience. 20(1):9–18. doi:10.2980/20-1-3545.

Gelman A, Hill J. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge, UK: Cambridge University Press.

Ghalambor CK, Angeloni LM, Carroll SP. 2010. Behavior as phenotypic plasticity. In: Westneat E, Fox C, editors. Evolutionary behavioral ecology. Oxford University Press. p. 90–107.

Gibb J. 1954. Feeding ecology of tits, with notes on treecreeper and goldcrest. Ibis. 96(4):513–543. doi:10.1111/j.1474-919X.1954.tb05476.x. https://onlinelibrary.wiley.com/doi/10.1111/j.1474-919X.1954.tb05476.x.

Gibb J. 2008. Feeding ecology of tits, with notes on treecreeper and goldcrest. Ibis. 96(4):513–543. doi:10.1111/j.1474-919X.1954.tb05476.x.

Gordon DM. 2011. The fusion of behavioral ecology and ecology. Behavioral Ecology. 22(2):225–230. doi:10.1093/beheco/arq172. https://academic.oup.com/beheco/article-lookup/doi/10.1093/beheco/arq172.

Groothuis TGG, Carere C. 2005. Avian personalities: characterization and epigenesis. Neurosci Biobehav Rev. 29(1 SPEC. ISS.):137–150. doi:10.1016/j.neubiorev.2004.06.010.

Hall LE, Chalfoun AD. 2019. Behavioural plasticity modulates temperature-related constraints on foraging time for a montane mammal. Dantzer B, editor. Journal of Animal Ecology. 88(3):363–375. doi:10.1111/1365-2656.12925. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2656.12925.

Hämäläinen A, Immonen E, Tarka M, Schuett W. 2018. Evolution of sex-specific pace-of-life syndromes: causes and consequences. Behav Ecol Sociobiol. 72(3). doi:10.1007/s00265-018-2466-x.

Hartig F. 2022. DHARMa: residual diagnostics for hierarchical (multi-Level/ mixed) regression models.

Harvey JA, Poelman EH, Tanaka T. 2013. Intrinsic inter- and intraspecific competition in parasitoid wasps. Annu Rev Entomol. 58:333–351. doi:10.1146/annurev-ento-120811-153622.

Healy K, Ezard T, Jones O, Salguero-Gómez R, Buckley Y. 2019. Animal life history is shaped by the pace of life and the distribution of age-specific mortality and reproduction. Nat Ecol Evol. 3(8):1217–1224. doi:doi:10.1038/s41559-019-0938-7.

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. 2003. Measuring inconsistency in meta-analyses. Br Med J. 327(7414):557–560. doi:10.1136/bmj.327.7414.557.

Holtmann B, Lagisz M, Nakagawa S. 2017. Metabolic rates, and not hormone levels, are a likely mediator of between-individual differences in behaviour: a meta-analysis. Funct Ecol. 31(3):685–696. doi:10.1111/1365-2435.12779.

Houston A, McNamara J. 1999. Models of adaptive behaviour: an approach based on state. Cambridge, UK: Cambridge University Press.

Hutfluss A, Rohr VA, Scheidt S, Steinbichl L, Bermúdez-Cuamatzin E, Slabbekoorn H, Dingemanse NJ. 2021. Does song overlap signal aggressiveness? An experimental study with repeated measures in free-ranging great tits. Anim Behav. 179:199–211. doi:10.1016/j.anbehav.2021.07.006. https://linkinghub.elsevier.com/retrieve/pii/S0003347221002207.

Immelmann K. 2012. Introduction to ethology. Springer US.

Jenni L, Winkler R. 1989. The feather-length of small passerines: a measurement for wing-length in live birds and museum skins. Bird Study. 36(1):1–15. doi:10.1080/00063658909476996.

Jenni L, Winkler R. 1994. Moult and aging of european passerines. London, UK: Academic Press.

Jennions MD, Møller AP. 2002. Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. Proceedings of the Royal Society B: Biological Sciences. 269(1486):43–48. doi:10.1098/rspb.2001.1832.

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(1):221–229. doi:10.1007/s10336-020-01815-3. https://link.springer.com/10.1007/s10336-020-01815-3.

Johnson CA, Grant JWA, Giraldeau LA. 2004. The effect of patch size and competitor number on aggression among foraging house sparrows. Behavioral Ecology. 15(3):412–418. doi:10.1093/beheco/arh026.

Kelly CD. 2006. Replicating empirical research in behavioral ecology: how and why it should be done but rarely ever is. Q Rev Biol. 81(3):221–236. doi:10.1086/506236.

Kelly CD. 2019. Rate and success of study replication in ecology and evolution. PeerJ. 2019(9). doi:10.7717/peerj.7654.

Kempenaers B, Dhondt AA. 1991. Competition between blue and great tit for roosting sites in winter: an aviary experiment. Ornis Scandinavica. 22(1):73. doi:10.2307/3676624. https://www.jstor.org/stable/3676624?origin=crossref.

Kempenaers B, Verheyen GR, Dhondt A. 1995. Mate guarding and copulation behaviour in monogamous and polygynous blue tits: do males follow a best-of-a-bad-job strategy? Behav Ecol Sociobiol. 36:33–42.

Kingsolver JG, Diamond SE, Siepielski AM, Carlson SM. 2012. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. Evol Ecol. 26(5):1101–1118. doi:10.1007/s10682-012-9563-5.

Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, Hoang A, Gibert P, Beerli P. 2001. The strength of phenotypic selection in natural populations. American Naturalist. 157(3):245–261. doi:10.1086/319193.

Klingenberg CP. 2008. Morphological integration and developmental modularity. Annu Rev Ecol Evol Syst. 39:115–132. doi:10.1146/annurev.ecolsys.37.091305.110054.

Klingenberg CP. 2014. Studying morphological integration and modularity at multiple levels: concepts and analysis. Philosophical Transactions of the Royal Society B: Biological Sciences. 369(1649):33–35. doi:10.1098/rstb.2013.0249.

Klomp H. 1964. Intraspecific competition and the regulation of insect numbers. Annu Rev Entomol. 9(1):17–40. doi:https://doi.org/10.1146/annurev.en.09.010164.000313.

Kluen E, Rönkä K, Thorogood R. 2022. Prior experience of captivity affects behavioural responses to ‘novel’ environments. PeerJ. 10:e13905. doi:10.7717/peerj.13905.

Kluen E, Siitari H, Brommer JE. 2014. Testing for between individual correlations of personality and physiological traits in a wild bird. Behav Ecol Sociobiol. 68(2):205–213. doi:10.1007/s00265-013-1635-1.

Könczey R, Tóth L, Török J. 1997. Site fidelity of great and blue tits in the Pilis-Visegrád mountains. Opusc Zool Bp. 29–30:103–111.

Korner-Nievergelt F, Roth T, von Felten S, Guelat J, Almasi B, Korner-Nievergelt P. 2015. Bayesian data analysis in ecology using linear models with R, BUGS and Stan. Academic Press.

Krams IA, Vrublevska J, Sepp T, Abolins-Abols M, Rantala MJ, Mierauskas P, Krama T. 2014. Sex-specific associations between nest defence, exploration and breathing rate in breeding pied flycatchers. Ethology. 120(5):492–501. doi:10.1111/eth.12222.

Krams R, Krama T, Brūmelis G, Elferts D, Strode L, Dauškane I, Luoto S, Šmits A, Krams IA. 2021. Ecological traps: evidence of a fitness cost in a cavity-nesting bird. Oecologia. 196(3):735–745. doi:10.1007/s00442-021-04969-w.

Krebs J, Ashcroft R, Webber M. 1978. Song repertoires and territory defence in the great tit. Nature. 271(5645):539–542. doi:10.1038/271539a0. https://www.nature.com/articles/271539a0.

Krebs JR. 1971. Territory and breeding density in the great tit, Parus major L. Ecology. 52(1):2–22. doi:10.2307/1934734. http://doi.wiley.com/10.2307/1934734.

Kruuk LEB. 2004. Estimating genetic parameters in natural populations using the “animal model.” Philosophical Transactions of the Royal Society B: Biological Sciences. 359(1446):873–890. doi:10.1098/rstb.2003.1437.

Kullberg C, Houston DC, Metcalfe NB. 2002. Impaired flight ability - A cost of reproduction in female blue tits. Behavioral Ecology. 13(4):575–579. doi:10.1093/beheco/13.4.575.

Lande R. 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet Res. 26(3):221–235. doi:10.1017/S0016672300016037. https://www.cambridge.org/core/product/identifier/S0016672300016037/type/journal_article.

Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. Evolution (N Y). 37(6):1210. doi:10.2307/2408842.

Lande R, Engen S, Sæther B-E. 2009. An evolutionary maximum principle for density-dependent population dynamics in a fluctuating environment. Philosophical Transactions of the Royal Society B: Biological Sciences. 364(1523):1511–1518.

doi:10.1098/rstb.2009.0017. https://royalsocietypublishing.org/doi/10.1098/rstb.2009.0017.

Laskowski KL, Moiron M, Niemelä P. 2021. Integrating behavior in life-history theory: allocation versus acquisition? Trends Ecol Evol. 36(2):132–138. doi:10.1016/j.tree.2020.10.017. https://doi.org/10.1016/j.tree.2020.10.017.

Lehtonen TK, McCrary JK, Meyer A. 2010. Territorial aggression can be sensitive to the status of heterospecific intruders. Behavioural Processes. 84(2):598–601. doi:10.1016/j.beproc.2010.02.021. https://linkinghub.elsevier.com/retrieve/pii/S0376635710000859.

Leighton GM, Drury JP, Small J, Miller ET. 2024. Unfamiliarity generates costly aggression in interspecific avian dominance hierarchies. Nat Commun. 15(1):335. doi:10.1038/s41467-023-44613-0. https://www.nature.com/articles/s41467-023-44613-0.

MacArthur RH, Wilson EO. 1967. The theory of island biogeography. Princeton University Press.

Martin PR, Freshwater C, Ghalambor CK. 2017. The outcomes of most aggressive interactions among closely related bird species are asymmetric. PeerJ. 2017(1). doi:10.7717/peerj.2847.

Martin PR, Ghalambor CK. 2014. When David beats Goliath: the advantage of large size in interspecific aggressive contests declines over evolutionary time. Marshall JAR, editor. PLoS One. 9(9):e108741. doi:10.1371/journal.pone.0108741. https://dx.plos.org/10.1371/journal.pone.0108741.

Martin PR, Ghalambor CK. 2023. A case for the “competitive exclusion–tolerance rule” as a general cause of species turnover along environmental gradients. Am Nat. 202(1):1–17. doi:10.1086/724683. https://www.journals.uchicago.edu/doi/10.1086/724683.

Mather K. 1955. Polymorphism as an outcome of disruptive selection. Evolution (N Y). 9(1):52. doi:10.2307/2405357. https://www.jstor.org/stable/2405357?origin=crossref.

Mathot KJ, Wright J, Kempenaers B, Dingemanse NJ. 2012. Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. Oikos. 121(7):1009–1020. doi:10.1111/j.1600-0706.2012.20339.x.

Matthysen E, Dhondt AA, Adriaensen F, Ornis S, Scandinavian S, Apr N. 1991. Sexual Dimorphism in Bill Length and Breeding Success in Great and Blue Tits (Parus major, P. caeruleus). Scandinavian Journal of Ornithology. 22(2):171–173.

Matyjasiak P. 2005. Birds associate species-specific acoustic and visual cues: recognition of heterospecific rivals by male blackcaps. Behavioral Ecology. 16(2):467–471. doi:10.1093/beheco/ari012.

McAdam AG, Webber QMR, Dantzer B, Lane JE, Boutin S. 2022. Social effects on annual fitness in red squirrels. Hughes K, editor. Journal of Heredity. 113(1):69–78. doi:10.1093/jhered/esab051. https://academic.oup.com/jhered/article/113/1/69/6408544.

McNamara J, Houston A. 1996. State-dependent life histories. Nature. 380:215–221. doi:10.1038/380215a0.

Metcalfe NB, Furness RW. 1987. Aggression in shorebirds in relation to flock density and composition. Ibis. 129(s2):553–563. doi:10.1111/j.1474-919X.1987.tb08243.x. https://onlinelibrary.wiley.com/doi/10.1111/j.1474-919X.1987.tb08243.x.

Minot EO. 1981. Effects of interspecific competition for food in breeding blue and great tits. Journal of Animal Ecology. 50(2):375–385. doi:10.2307/4061.

Minot E, Perrins C. 1986. Interspecific interference competition. Nest sites for blue and great tits. Journal of Animal Ecology. 55(1):331–350. doi:10.2307/4712.

Moiron M, Araya-Ajoy YG, Mathot KJ, Mouchet A, Dingemanse NJ. 2019. Functional relations between body mass and risk-taking behavior in wild great tits. Behavioral Ecology. 30(3):617–623. doi:10.1093/beheco/ary199.

Moiron M, Laskowski KL, Niemelä PT. 2020. Individual differences in behaviour explain variation in survival: a meta-analysis. Ecol Lett. 23(2):399–408. doi:10.1111/ele.13438.

Møller AP, Javier B, A DA, Vladimir R, Frank A, Clotilde B, Jordi C, Mariusz C, Blandine D. 2018. Effects of interspecific coexistence on laying date and clutch size in two closely related species of hole- nesting birds. Journal of Animal Ecology. 87(6):1738–1748. doi:10.1111/1365-2656.12896.

Montiglio PO, Dammhahn M, Dubuc Messier G, Réale D. 2018. The pace-of-life syndrome revisited: the role of ecological conditions and natural history on the slow-fast continuum. Behav Ecol Sociobiol. 72(7):1–9. doi:10.1007/s00265-018-2526-2.

Mouchet A, Cole EF, Matthysen E, Nicolaus M, Quinn JL, Roth AM, Tinbergen JM, van Oers K, van Overveld T, Dingemanse NJ. 2021. Heterogeneous selection on exploration behavior within and among West European populations of a passerine bird. Proc Natl Acad Sci U S A. 118(28):1–6. doi:10.1073/pnas.2024994118.

Nakagawa S, Lagisz M, Jennions MD, Koricheva J, Noble DWA, Parker TH, Sánchez-Tójar A, Yang Y, O’Dea RE. 2022. Methods for testing publication bias in ecological and evolutionary meta-analyses. Methods Ecol Evol. 13(1):4–21. doi:10.1111/2041-210X.13724.

Nakagawa S, Parker TH. 2015. Replicating research in ecology and evolution: feasibility, incentives, and the cost-benefit conundrum. BMC Biol. 13(1):1–6. doi:10.1186/s12915-015-0196-3.

Nakagawa S, Santos ESA. 2012. Methodological issues and advances in biological meta-analysis. Evol Ecol. 26(5):1253–1274. doi:10.1007/s10682-012-9555-5.

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

Nicolaus M, Mathot KJ, Araya-Ajoy YG, Mutzel A, Wijmenga JJ, Kempenaers B, Dingemanse NJ. 2015. Does coping style predict optimization? An experimental test in a wild passerine bird. Proceedings of the Royal Society B: Biological Sciences. 282(1799):20142405. doi:10.1098/rspb.2014.2405.

Niemelä PT, Dingemanse NJ. 2018. On the usage of single measurements in behavioural ecology research on individual differences. Anim Behav. 145:99–105. doi:10.1016/j.anbehav.2018.09.012.

Nord A, Nilsson JÅ. 2011. Incubation temperature affects growth and energy metabolism in blue tit nestlings. American Naturalist. 178(5):639–651. doi:10.1086/662172.

Nowakowski JK, Rowiński P. 1996. Wing and body mass measurements in the Great Tit Parus major in Central Poland, errors and methods of standardization. Acta Ornithol. 31(2):107–118.

Nussey DH, Wilson AJ, Brommer JE. 2007. The evolutionary ecology of individual phenotypic plasticity in wild populations. J Evol Biol. 20(3):831–844. doi:10.1111/j.1420-9101.2007.01300.x.

O’Dea RE, Noble DWA, Nakagawa S. 2022. Unifying individual differences in personality, predictability and plasticity: A practical guide. Methods Ecol Evol. 13(2):278–293. doi:10.1111/2041-210X.13755. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.13755.

Peiman KS, Robinson BW. 2010. Ecology and evolution of resource-related heterospecific aggression. Q Rev Biol. 85(2):133–158. doi:10.1086/652374. https://www.journals.uchicago.edu/doi/10.1086/652374.

Perrier C, Delahaie B, Charmantier A. 2018. Heritability estimates from genomewide relatedness matrices in wild populations: Application to a passerine, using a small sample size. Mol Ecol Resour. 18(4):838–853. doi:10.1111/1755-0998.12886.

Perrins C. 1965. Population fluctuations and clutch-size in the great tit, Parus major L. Journal of Animal Ecology. 34(3):601–647. doi:10.2307/2453.

Pigliucci M. 2001. Phenotypic plasticity. Beyond nature and nurture. Johns Hopkins University Press.

Pigliucci M. 2003. Phenotypic integration: studying the ecology and evolution of complex phenotypes. Ecol Lett. 6(3):265–272. doi:10.1046/j.1461-0248.2003.00428.x.

Pigliucci M, Preston K, editors. 2004. Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes. New York, USA: Oxford University Press.

Planqué B, Vellinga WP, Jongsma J. 2005. Xeno-canto: sharing bird sounds from around the world. www.xeno-canto.org.

R Core Team. 2020. R: A language and environment for statistical computing.

R Core Team. 2023. R: A language and environment for statistical computing.

Réale D, Garant D, Humphries MM, Bergeron P, Careau V, Montiglio PO. 2010. Personality and the emergence of the pace-of-life syndrome concept at the population level. Philosophical Transactions of the Royal Society B: Biological Sciences. 365(1560):4051–4063. doi:10.1098/rstb.2010.0208.

Réale D, Reader SM, Sol D, McDougall PT, Dingemanse NJ. 2007. Integrating animal temperament within ecology and evolution. Biological Reviews. 82(2):291–318. doi:10.1111/j.1469-185X.2007.00010.x.

Reed TM. 1982. Interspecific territoriality in the chaffinch and great tit on islands and the mainland of Scotland: playback and removal experiments. Anim Behav. 30(1):171–181. doi:10.1016/S0003-3472(82)80252-2. https://linkinghub.elsevier.com/retrieve/pii/S0003347282802522.

Ricklefs R, Wikelski M. 2002. The physiology/life-history nexus. Trends Ecol Evol. 17(10):462–468. doi:10.1016/S0169-5347(02)02578-8.

Roff D. 1993. Evolution of life histories: theory and analysis. Roff D, editor. New York, USA: Chapman and Hall.

Roff DA, Fairbairn DJ. 2012. A test of the hypothesis that correlational selection generates genetic correlations. Evolution (N Y). 66(9):2953–2960. doi:10.1111/j.1558-5646.2012.01656.x.

Royauté R, Berdal MA, Garrison CR, Dochtermann NA. 2018. Paceless life? A meta-analysis of the pace-of-life syndrome hypothesis. Behav Ecol Sociobiol. 72(3):1–10. doi:https://doi.org/10.1007/s00265-018-2472-z.

Royauté R, Garrison C, Dalos J, Berdal MA, Dochtermann NA. 2019. Current energy state interacts with the developmental environment to influence behavioural plasticity. Anim Behav. 148:39–51. doi:10.1016/j.anbehav.2018.11.013.

Rueffler C, Van Dooren TJM, Leimar O, Abrams PA. 2006. Disruptive selection and then what? Trends Ecol Evol. 21(5):238–245. doi:10.1016/j.tree.2006.03.003. https://linkinghub.elsevier.com/retrieve/pii/S0169534706000887.

Sæther B. 1988. Pattern of covariation between life-history traits of European birds. Nature. 331:616–617. doi:10.1038/331616a0.

Sæther BE, Visser ME, Grøtan V, Engen S. 2016. Evidence for r-and K-selection in a wild bird population: A reciprocal link between ecology and evolution. Proceedings of the Royal Society B: Biological Sciences. 283(1829). doi:10.1098/rspb.2015.2411.

Sasaki A, Ellner S. 1997. Quantitative genetic variance maintained by fluctuating selection with overlapping generations: variance components and covariances. Evolution (N Y). 51(3):682–696. doi:10.1111/j.1558-5646.1997.tb03652.x. https://academic.oup.com/evolut/article/51/3/682/6757400.

Sasvári L, Orell M. 1992. Breeding success in a north and a central european population of the great tit Parus major. Ornis Scandinavica. 23(1):96. doi:10.2307/3676432.

Sasvári L, Török J, Tóth L. 1987. Density dependent effects between three competitive bird species. Oecologia. 72(1):127–130. doi:10.1007/BF00385056. http://link.springer.com/10.1007/BF00385056.

Scheiner SM. 1993. Genetics and evolution of phenotypic plasticity. Annu Rev Ecol Syst. 24(1):35–68. doi:10.1146/annurev.es.24.110193.000343. https://www.annualreviews.org/doi/10.1146/annurev.es.24.110193.000343.

Scheiner SM. 2002. Selection experiments and the study of phenotypic plasticity. J Evol Biol. 15(6):889–898. doi:10.1046/j.1420-9101.2002.00468.x. http://doi.wiley.com/10.1046/j.1420-9101.2002.00468.x.

Schluter D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution (N Y). 50(5):1766–1774. doi:10.1111/j.1558-5646.1996.tb03563.x.

Schlyter F, Anderbrant O. 1993. Competition and niche separation between two bark beetles: existence and mechanisms. Oikos. 68(3):437–447. doi:https://doi.org/10.2307/3544911.

Searle SR. 1961. Phenotypic, genetic and environmental correlations. Biometrics. 17(3):474. doi:10.2307/2527838.

Sheehy KA, Laskowski KL. 2023. Correlated behavioural plasticities: insights from plasticity evolution, the integrated phenotype and behavioural syndromes. Anim Behav. 200:263–271. doi:10.1016/j.anbehav.2023.04.007. https://linkinghub.elsevier.com/retrieve/pii/S000334722300091X.

Siepielski AM, Dibattista JD, Carlson SM. 2009. It’s about time: the temporal dynamics of phenotypic selection in the wild. Ecol Lett. 12(11):1261–1276. doi:10.1111/j.1461-0248.2009.01381.x.

Sih A, Bell A, Johnson JC. 2004. Behavioral syndromes: an ecological and evolutionary overview. Trends Ecol Evol. 19(7):372–378. doi:10.1016/j.tree.2004.04.009.

Sih A, Bell AM. 2008. Insights for behavioral ecology from behavioral syndromes. In: Advances in the Study of Behavior. Vol. 38. p. 227–281. https://linkinghub.elsevier.com/retrieve/pii/S0065345408000053.

Sih A, Mathot KJ, Moirón M, Montiglio PO, Wolf M, Dingemanse NJ. 2015. Animal personality and state-behaviour feedbacks: a review and guide for empiricists. Trends Ecol Evol. 30(1):50–60. doi:10.1016/j.tree.2014.11.004.

Sinervo B, Svensson E. 2002. Correlational selection and the evolution of genomic architecture. Heredity (Edinb). 89(5):329–338. doi:10.1038/sj.hdy.6800148.

Slarkin M. 1985. Gene flow in natural populations. Annu Rev Ecol Syst. 16(1):393–430. doi:10.1146/annurev.es.16.110185.002141. https://www.annualreviews.org/doi/10.1146/annurev.es.16.110185.002141.

Stamps JA. 2007. Growth-mortality tradeoffs and “personality traits” in animals. Ecol Lett. 10(5):355–363. doi:10.1111/j.1461-0248.2007.01034.x.

Stamps JA. 2016. Individual differences in behavioural plasticities. Biological Reviews. 91(2):534–567. doi:10.1111/brv.12186. https://onlinelibrary.wiley.com/doi/10.1111/brv.12186.

Stearns S. 1992. The evolution of life histories. New York, USA: Oxford University Press.

Stephens PA, Sutherland WJ, Freckleton RP. 1999. What is the Allee Effect? Oikos. 87(1):185. doi:10.2307/3547011.

Stuber EF, Araya-Ajoy YG, Mathot KJ, Mutzel A, Nicolaus M, Wijmenga JJ, Mueller JC, Dingemanse NJ. 2013. Slow explorers take less risk: a problem of sampling bias in ecological studies. Behavioral Ecology. 24(5):1092–1098. doi:10.1093/beheco/art035.

Svensson E, Sheldon BC. 1998. The social context of life history evolution. Oikos. 83(3):466. doi:10.2307/3546674.

Svensson L. 1992. Identification guide to european passerines. 4th ed. Stockholm.

Taborsky B, Oliveira RF. 2012. Social competence: an evolutionary approach. Trends Ecol Evol. 27(12):679–688. doi:10.1016/j.tree.2012.09.003.

Tibbetts EA, Dale J. 2007. Individual recognition: it is good to be different. Trends Ecol Evol. 22(10):529–537. doi:10.1016/j.tree.2007.09.001. https://linkinghub.elsevier.com/retrieve/pii/S0169534707002376.

Tilman D. 1994. Competition and biodiversity in spatially structured habitats. Ecology. 75(1):2–16. doi:https://doi.org/10.2307/1939377.

Tinbergen JM. 1987. Costs of reproduction in the great tit: intraseasonal costs associated with brood size. Ardea. 75(1):111–122.

Tinbergen N. 1963. On aims and methods of ethology. Z Tierpsychol. 20(4):410–433.

Tokeshi M. 2009. Species coexistence: ecological and evolutionary perspectives. John Wiley & Sons.

Török J. 1986. Food segregation in three hole-nesting bird species during the breeding season. Ardea. 74(2):129–136.

Török J, Tóth L. 1999. Asymmetric competition between two tit species: A reciprocal removal experiment. Journal of Animal Ecology. 68(2):338–345. doi:10.1046/j.1365-2656.1999.00283.x.

Urszán TJ, Garamszegi LZ, Nagy G, Hettyey A, Török J, Herczeg G. 2018. Experience during development triggers between-individual variation in behavioural plasticity. Journal of Animal Ecology. 87(5):1264–1273. doi:10.1111/1365-2656.12847.

Verhulst S. 1998. Multiple breeding in the great tit, II. The costs of rearing a second clutch. Funct Ecol. 12(1):132–140. doi:10.1046/j.1365-2435.1998.00165.x.

Viechtbauer W. 2010. Conducting meta-analyses in R with the metafor. J Stat Softw. 36(3):1–48.

West-Eberhard MJ. 1989. Phenotypic plasticity and the origins of diversity. Annu Rev Ecol Syst. 20(1):249–278. doi:10.1146/annurev.es.20.110189.001341. https://linkinghub.elsevier.com/retrieve/pii/B9780128179963000177.

Westneat DF, Wright J, Dingemanse NJ. 2015. The biology hidden inside residual within-individual phenotypic variation. Biological Reviews. 90(3):729–743. doi:10.1111/brv.12131.

Wilson D. 1998a. Adaptive individual differences within single populations. Magurran AE, May RM, editors. Philos Trans R Soc Lond B Biol Sci. 353(1366):199–205. doi:10.1098/rstb.1998.0202. https://royalsocietypublishing.org/doi/10.1098/rstb.1998.0202.

Wilson E. 1998b. Consilience: the unity of knowledge. London, UK: Little, Brown & Company.

Wolf M, Van Doorn GS, Leimar O, Weissing FJ. 2007. Life-history trade-offs favour the evolution of animal personalities. Nature. 447(7144):581–584. doi:10.1038/nature05835.

Wolf M, van Doorn GS, Weissing FJ. 2008. Evolutionary emergence of responsive and unresponsive personalities. Proceedings of the National Academy of Sciences. 105(41):15825–15830. doi:10.1073/pnas.0805473105. https://pnas.org/doi/full/10.1073/pnas.0805473105.

Wolf M, Weissing FJ. 2010. An explanatory framework for adaptive personality differences. Philosophical Transactions of the Royal Society B: Biological Sciences. 365(1560):3959–3968. doi:10.1098/rstb.2010.0215.

Wright J, Bolstad GH, Araya-Ajoy YG, Dingemanse NJ. 2019. Life-history evolution under fluctuating density-dependent selection and the adaptive alignment of pace-of-life syndromes. Biological Reviews. 94(1):230–247. doi:10.1111/brv.12451.

Yang Y, Sánchez-Tójar A, O’Dea RE, Noble D, Koricheva J, Jennions MD, Parker TH, Lagisz M, Nakagawa S. 2023. Publication bias impacts on effect size, statistical power, and magnitude (Type M) and sign (Type S) errors in ecology and evolutionary biology. BMC Bio.:1–20. doi:10.1186/s12915-022-01485-y.

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(3):515–521. doi:10.1016/j.anbehav.2012.05.024. https://linkinghub.elsevier.com/retrieve/pii/S0003347212002448.

8. Acknowledgments

This journey began several years ago in Mexico when I reached out to Niels about the possibility of pursuing a PhD with him. After he agreed, the pursuit of a scholarship started. Some months later, I finally received a positive response from the DAAD, stating that my proposal was accepted. Then, I moved to Germany to start this new chapter.

I am deeply grateful to Niels. Thank you for sharing your time, expertise, and passion for science. Your infinite patience allowed me to achieve all my goals. Your ability to find a solution to whatever setback we faced, especially during COVID, made this research possible. For all the trust in every step I made and for being an excellent supervisor and professor.

I extend sincere thanks to Benedikt Holtmann for his tireless efforts during the extensive fieldwork. All his feedback and ideas contributed to the success of this project.

A special thank you needs to go to Joanna Bialas. I could not have asked for a better colleague with whom to share so many hours of fieldwork. Joanna, your willingness to go above and beyond, whether it was checking boxes in Plot A or supervising students, has been truly appreciated. Your contributions have been instrumental in facilitating the progress of my research.

I also want to express my gratitude to all the bachelor and master students that contributed to the data collection. This type of research would have been impossible without all your help.

I would also like to thank the former and current members of the Behavioral Ecology group. It was a pleasure to have shared time and space with you. Especially, to Alexia for helping with the database, to Alex sorting and providing the song records, to Magdalena and Michelle for their willingness to help with all the administrative tasks, and to Corné and Rori for always having time to provide feedback about my research.

My sincere thanks go to the Deutsche Akademische Austauschdienst (DAAD) for funding my thesis project.

My heartfelt appreciation goes to my family. The support you gave me when I decided to leave everything I knew and start this adventure highly motivated me to finish this chapter successfully. For sharing your experiences and words of encouragement. Though some are no longer with us, their belief in my dreams continues to inspire me every day.

I am deeply grateful for the enriching experiences and collaborations with all the people I have had the privilege to work with over the years. Each of you has contributed to shaping me into the scientist I am today.

A special recognition to my husband, José Enrique, for standing by me and believing in my professional endeavors. Thank you for your support and continuous encouragement throughout this journey.

IRENE GAONA GORDILLO

Munich, Germany | +4901748732732 | gaona-gordillo@bio.lmu.de |

Behavioral Ecology, Department II, Faculty of Biology, Ludwig-Maximilians-Universität (*LMU)*

Großhaderner Str. 2, 82152 Planegg-Martinsried

EDUCATION

PhD	• 2019 – 2024; PhD in Behavioral Ecology, Ludwig-Maximilians-Universität (LMU) • 2020 – today; PhD, Max Planck Research School (IMPRS) for Organismal Biology (now Biological Intelligence)
Research project:	“Social interactions: intra- and inter-specific competition between blue and great tits”
Advisor:	Dr. Niels Dingemanse
Masters	• 2015 – 2018; Masters in Biological Science, Tlaxcala Center for the Study of Behavior (CTBC), Autonomous University of Tlaxcala (UAT), Tlaxcala, México
Research project:	“Thermal stress during early development: trade-offs among growth components”
Advisors:	Dr. Roxana Torres and Dr. Sergio Ancona
Bs	• 2008 – 2014; Biology, Science Faculty, Autonomous National University of Mexico (UNAM), Mexico City, Mexico
Research project:	“Natural history of the Western Bluebird (*Sialia mexicana*) at La Malinche National Park (PNLM): mating and breeding system”
Advisor:	Dr. Constantino Macías-García

PEER-REVIEWED PAPERS

- Gaona-Gordillo I, Holtmann B, Mouchet A, Hutfluss A, Sánchez-Tójar A, Dingemanse NJ. 2023. Are animal personality, body condition, physiology and structural size integrated? A comparison of species, populations and sexes, and the value of study replication. Journal of Animal Ecology. 92(9):1707–1718. doi:10.1111/1365-2656.13966. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2656.13966.

SUPERVISED THESIS

- Liedl, S (2022) Effect of breeding density on the reproductive fitness of great tits (*Parus major)* and blue tits (*Cyanistes caeruleus*). BSc-thesis (LMU).
- *Mollenhauer, M (2022) Comparison of intra- and interspecific aggression between male blue tits (Cyanistes caeruleus) and great tits (Parus major).* BSc-thesis (LMU).
- *Braun, D (2022)* Trade-offs between size and condition of the offspring in great tis (*Parus major*) and blue tits (*Cyanistes caeruleus*). BSc-thesis (LMU).
- Fröhlich, F (2022) Influence of parental exploratory behaviour on the body condition of their chicks in great tits (*Parus major)* and blue tits (*Cyanistes caeruleus*). BSc-thesis (LMU).
- Glombik S (2022) Hello neighbor: the effects of neighbors on blue tits and great tits. BSc-thesis (LMU).
- Riemens, L (2022) The effect of intraspecific and interspecific competition on reproduction success and behaviour in a sympatric population of Great (*Parus major*) and Blue Tits (*Cyanistes caeruleus*). MSc-thesis (University of Groningen).

EMPLOYMENT HISTORY

Student assistant:	• December 2023 – February 2024; Prof. Dr. Niels Dingemanse, LMU, Munich
Research assistant:	• August 2017 – August 2018; Dra. Margarita Martinez Gomez, National Council of Science and Technology (CONACyT) Number of affiliation: 15277
Electoral assistant trainer:	• February 2009 – July 2009; Federal Electoral Institute (IFE, now INE); capacitation of people for their participation in Mexican regional elections

TEACHING

2022-2023:	• Basic statistics in R; undergraduate students at the LMU; 2 hours
2018-2019:	• Popular science; talks about birds and conservation; audience: general public from 5 to 80 years old and students from 8 to 15 years (Morelia, Mexico)
2015-2018:	• Popular science; talks about birds and conservation “Malinche treasures”; audience: students from 5 to 18 years old and adults from 25 to 70 years old (Tlaxcala, Mexico)
2018:	• XXVII “National Biology Olympiad”; lectures of genetics and biostatistics; 20 hours; high school students
2017:	• XXVI “National Biology Olympiad”; lectures of genetics; 10 hours; high school students

TECHNICAL EXPERIENCE

Social service:
- Protein purification, western blot at National Institute of Cardiology “Ignacio Chávez”, Mexico

Bs:
- Bird manipulation at La Malinche National Park. DNA purification, PCR, DNA analysis at Ecology Institute, UNAM, Mexico

Masters:
- Bird manipulation at La Malinche National Park, DNA purification, qPCR, statistical analysis at CTBC, UAT, Mexico

PhD:
- Bird manipulation at the *Forstenrieder Park*, statistical analysis at LMU, Germany

PARTICIPATION IN PROJECTS

2018:
- Bird watching at the “3th swamps festival”; audience: people from 5 to 70 years old

2017:
- Bird watching at the “2nd swamps festival”; audience: people from5 to 70 years old
- Talk about “Nocturnal birds of prey of La Malinche National Park”; INFOCAB PB201116; audience: students from high school

2011:
- "La Malinche Scientific Field Station: Popular science actions for natural resources conservation of template forest"; audience: high school students

ORAL PRESENTATIONS

2023:
- “Are animal personality, body condition, physiology, and structural size integrated? A comparison of species, populations, and sexes, and the value of study replication”; Talk; audience: international postgraduate students and international researchers at Behaviour 2023
- “Are animal personality, body condition, physiology and structural size integrated? A comparison of species, populations and sexes, and the value of study replication”; Talk; audience: international postgraduate students at 1st joint IMPRS Seminar

2022:
- “Phenotypic integration between morphology and behavior at different hierarchical levels”; Talk; audience: international postgraduate students and international researchers at the Second (and hopefully las) virtual conference of the Ethologische Gesellshaft e. V.
- “Phenotypic integration between morphology and behavior at different hierarchical levels”; Poster; audience: international postgraduate students and international researchers at International Society for Behavioral Ecology Congress 2022

2021:
- “Differences in the pace-of-life syndromes between and within species”, at the First Virtual Conference of the Ethologische Gesellschaft e.V.

2018:

- Nicolaitas at the park; Talk; audience: general public from 5 to 80 years

2017:

- "Why do we get old?"; Conference; audience: undergraduate psychology students
- "Thermal stress during early development: trade-offs among growth components"; Poster; audience: undergraduate and postgraduate students, international researchers at the Congreso Internacional Bases Biológicas de la Conducta

2016:

- "Thermal stress during early development: physiological implication and of life-history"; Poster; audience: undergraduate and postgraduate students, international researchers at the Congreso Internacional Bases Biológicas de la Conducta
- "Growing ardent or glacial affects your size"; Poster; audience: undergraduate and postgraduate students, international researchers
- "Thermal stress during early development: physiological implication and of life-history"; Talk; audience: undergraduate and postgraduate students of the CTBC

2015:

- "Effect of variation of the ambient temperature over the allocation of resources towards body maintenance or growth of *Troglodytes aedon* nestlings"; Poster; audience: undergraduate and postgraduate students, international researchers at the Congreso Internacional Bases Biológicas de la Conducta

AWARDS

2016:

- Science in on minute; award for oral presentation

GRANTS

2022:

- One-month grant for an academic research at the Uniwersytet Przyrodniczy w Poznaniu, Poznan, Poland given by PROM Programme – International scholarship exchange of doctoral students and academic staff.

2019-2023:

- Grant for PhD studies at LMU, Munich, Germany; Research Grants - Doctoral Programmes in Germany given by Deutscher Akademischer Austauschdienst (DAAD). Grant number: 91731707

2015-2017:

- Grant for Masters Studies at CTBC, UAT, Tlaxcala given by CONACyT. Grant number: 590304

CERTIFICATIONS

2020:

- "Theoretical qualification pursuant to the ordinance on animal protection in experiments (*TierSchVersV*)", Germany

2015-2017:

- Handling and manipulation of laboratory animals, Mexico

LANGUAGE SKILLS

- Spanish (native language)
- English: TOEFL iBT Test 103 points; April 2018
- Certification in reading and translation in English given by Teaching Center of Foreign Languages (CELE), UNAM; 2012-2017
- German: A2.1 given by the International University Club Munich (IUCM); Germany; February 2023

SKILLS

- R
    - LMM and GLMM analyses and interpretation
    - Multivariate analysis
    - Path analyses
    - Data visualization
- Organization of database in Excel and basic Access
- Office (Word, Excel, PowerPoint)