ORIGINAL ARTICLE

Male aggressiveness and risk-taking during reproduction are repeatable but not correlated in a wild bird population

Stephen M. Salazar[1,2] · Julia I. Camacho García[3] · Miriam Kuspiel[4] · Rienk W. Fokkema[1] · Jan Komdeur[2] · Peter Korsten[1]

Abstract

The existence of among-individual variation in behaviour within populations is poorly understood. Recent theory suggests that fine-scale individual differences in investment into current versus future reproduction may lead to a 'slow-fast'-pace-of-life continuum, also referred to as the 'pace-of-life-syndrome' (POLS) hypothesis. According to this idea, individuals are predicted to differ in their level of risk-taking, which may drive among-individual variation and covariation of behaviours. Consistent individual differences in aggression, an ecologically relevant and potentially risky behaviour, have been reported across the animal kingdom. Here we test whether such individual differences in aggression are a manifestation of underlying differences in risk-taking. In a wild blue tit (*Cyanistes caeruleus*) population, we used standard behavioural tests to investigate if male territorial aggressiveness and risk-taking during breeding are positively related. At the start of breeding, we simulated conspecific territorial intrusions to obtain repeated measures of male aggressiveness. Subsequently, we measured male risk-taking as their latency to resume brood provisioning after presenting two different predators at their nest: human and sparrowhawk, a common predator of adult songbirds. First, we found substantial repeatability for male aggressiveness ($R = 0.56 \pm 0.08$ SE). Second, while males took longer to resume provisioning after presentation of a sparrowhawk mount as compared to a human observer, risk-taking was repeatable across these two predator contexts ($R = 0.51 \pm 0.13$ SE). Finally, we found no evidence for a correlation between male aggressiveness and risk-taking, thereby providing little support to a main prediction of the POLS hypothesis.

Significance statement

Consistent, and often correlated, individual differences in basal behaviours, such as aggression, exploration and sociability, are found across the animal kingdom. Why individuals consistently differ in their behaviour is poorly understood, as behavioural traits would seem inherently flexible. The 'pace-of-life syndrome' (POLS) hypothesis proposes observed behavioural variation to reflect differences in risk-taking associated with individual reproductive strategies. We tested this idea in a wild blue tit population by investigating whether individual males that were more aggressive toward territorial intruders also took more risk when provisioning their nestlings under a threat of predation. While we found consistent individual differences in both aggressiveness and risk-taking, these behaviours were not significantly correlated. Therefore, our study demonstrates among-individual variation in ecologically relevant behaviours in wild blue tits but provides little support for the POLS hypothesis.

Introduction

Aggression is a basal behaviour expressed in the acquisition or defence of fitness-enhancing resources such as territory, food and mates, but which can also have energetic and other costs such as injury or death. The evolution of variation in aggressive behaviour has since long intrigued

behavioural biologists (e.g. Lorenz 1966; Riechert 1978; Maynard Smith et al. 1988). More recent theoretical work suggests that life history trade-offs between current and future reproduction may give rise to within-population variation where individuals differ consistently in aggressiveness as well as risk-taking behaviours (Wolf et al. 2007; see also Baldauf et al. 2014), but empirical studies of this remain scarce.

The hypothesized covariation of life history, behaviour and also physiological traits (Ricklefs and Wikelski 2002) is gaining traction as a framework to explain phenotypic variation within populations (Réale et al. 2010; Dammhahn et al. 2018). According to this pace-of-life-syndrome hypothesis (hereafter POLS hypothesis), individuals broadly vary along a slow-fast continuum over suites of correlated behavioural, physiological and life history traits. For example, individuals on the 'fast' end may be more aggressive, take more risks (to invest in current reproduction), reproduce earlier in life and consequently have shorter lifespans. In contrast, individuals on the 'slow' end may be characterized by lower aggressiveness, take fewer risks, reproduce later in life and invest more in survival (and thus future reproduction) (Wolf et al. 2007; Réale et al. 2010; Dammhahn et al. 2018; Wright et al. 2019).

Despite growing appreciation for this idea and much research into behavioural correlations (reviewed in, e.g. Brommer and Class 2017) — often referred to as behavioural syndromes (e.g. Sih et al. 2004) — broad empirical support for POLS as a driver of behavioural variation and co-variation is lacking (e.g. Royauté et al. 2018; Moiron et al. 2020). Heterogeneity among species and populations, likely due to differences in selection regimes, may play a role in this (see, e.g. Dingemanse et al. 2007; Montiglio et al. 2018). Another possible reason for the lack of overall support for the POLS hypothesis may be that until now a large proportion of studies aimed at understanding among-individual variation in behaviour is based on behavioural assays conducted in artificial or laboratory conditions which may weaken their biological interpretation (as discussed in Carter et al. 2013; David and Dall 2016; Royauté et al. 2018; see also Moiron et al. 2020).

In this study, conducted on a wild population of blue tits (*Cyanistes caeruleus*), we investigated if individual levels of male territorial aggressiveness, a biologically relevant behaviour, are positively related to risk-taking when provisioning young in the face of predation. Covariation in the expression of such behaviours would hint at their non-independent evolution and existence of a latent axis of variation (Sih et al. 2004; Sih and Bell 2008; Dingemanse et al. 2012a; Korsten et al. 2013). Based on the POLS hypothesis, we predict that male blue tits that are highly aggressive in territorial conflict should favour fitness gains from current reproduction and, therefore, in general, take greater risks.

Conversely, less aggressive males should favour fitness gains from future reproduction and thus, in general, be risk averse.

While the repeatability of 'aggressiveness' in various contexts has been documented in a wide range of taxa (reviewed by Bell et al. 2009; Garamszegi et al. 2013; Briffa et al. 2015), among-individual variation in territorial aggressiveness is increasingly reported in wild songbirds (e.g. western bluebird (*Sialia mexicana*), Duckworth 2006; dark-eyed junco (*Junco hyemalis*), Cain et al. 2011; eastern bluebird (*Sialia sialis*), Harris and Siefferman 2014; collared flycatcher (*Ficedula albicollis*), Garamszegi et al. 2015; song sparrow (*Melospiza melodia*), Davies and Sewall 2016; great tit (*Parus major*), Araya-Ajoy and Dingemanse 2017; wood warbler (*Phylloscopus sibilatrix*), Szymkowiak and Kuczyński 2017; thorn-tailed rayadito (*Aphrastura spinicauda*), Botero-Delgadillo et al. 2020). Males of our study species, the blue tit, are known to vigorously defend their territories against conspecific intruders (Colquhoun 1942), particularly during early stages of breeding (Alonso-Alvarez et al. 2004; Korsten et al. 2007a; Mutzel et al. 2013b), but the repeatability of this aggressive behaviour among individual males has not yet been established. We simulated territorial intrusions to obtain repeated measures of territorial aggressiveness of males at the beginning of breeding, around their female's fertile period, and investigated the repeatability of this behaviour. Thereafter, to quantify among-male variation in risk-taking behaviour, we simulated predation-risk at the nest during brood provisioning. We measured the latencies of males to resume brood provisioning after we presented a predation threat. This is a well-established paradigm to study risk-taking in the context of parental care (e.g. Dale et al. 1996; Lambrechts et al. 2000; Ghalambor and Martin 2001; Mahr et al. 2015; Mutzel et al. 2019; Vincze et al. 2019). We investigated each of the males' responses to different levels of predation risk by presenting two types of predators at the nest: a sparrowhawk (*Accipiter nisus*), which is a predator of adult blue tits (Vedder et al. 2014) and poses a severe threat, and a human (a potential nest-predator; Müller et al 2006; Hollander et al. 2008), which poses a relatively low threat to adult blue tits.

Specifically, our study had the following three aims: (1) to establish to what extent levels of territorial aggressiveness are repeatable in male blue tits; (2) to establish to what extent risk-taking during offspring provisioning in the face of predation is repeatable across two predator types (i.e. contexts) posing potentially different levels of risk to adult birds; and (3) to investigate whether more aggressive individuals take more risk in the face of predation, thereby testing a key prediction of the POLS hypothesis. We focused on variation in male aggressiveness, as the interpretation of female responses to simulated territorial intrusions may be ambiguous (e.g. females may also approach male intruders with a non-aggressive motivation).

Methods

General field methods

We conducted this study on a nestbox-breeding population of blue tits (*Cyanistes caeruleus*) that has been monitored since 2001 (Korsten et al. 2006). The study site is located on the 'De Vosbergen' estate south of Groningen in the Netherlands (53° 08′ N, 06° 35′ E). Around 210 nestboxes designed specifically for blue tits (entrance hole: 26 mm diameter) are distributed over 54 hectares of mixed, deciduous and coniferous forest (Amininasab et al. 2016), with intermittent open farm and grasslands. Nestboxes are mounted onto tree trunks at about 2 m from the ground and are accessed by ladders. Every year around 80–120 nestboxes are taken up by breeding blue tit pairs.

Prior to the breeding season, within a single evening (19th February 2018), we caught all birds roosting in nestboxes in the study area and ringed them (if not ringed) with a standard aluminium ring provided by the Dutch ringing station along with a unique combination of colour rings. This allowed us to identify individual birds in the simulated territorial intrusions which we ran during the breeding season later on (see 'Measuring territorial aggressiveness' below).

From mid-March (2018) onwards, we conducted checks on every nestbox at least once a week to monitor nest-building activity. As soon as we found a nest that was complete (i.e. having a cup-like shape lined with feathers and/or hairs), daily checks were carried out until the first egg was laid (hereon, laying date), after which we resumed checking only once a week. Blue tit females lay an egg a day and produce clutches of 11 eggs on average at this study site (Amininasab et al. 2016). Incubation starts towards the end of egg laying, during which we halted checking nests to minimize disturbance. Eggs hatch about 2 weeks from the start of incubation, and so 10 days after the last egg was laid, we resumed daily nest checks to obtain the exact date of hatching of the first nestling (hereon, hatching date).

When broods were 7–9 days old (day 0 = day of hatching of the first nestling), we caught the adult breeders inside their nestbox during nestling provisioning using a 'spring-trap'. We sexed (presence or absence of a brood patch; present in females only) and aged the birds (first year breeder or older; following Svensson 1992), took standard morphological measures (mass, tarsus and third primary feather length) as part of the long-term data collection and provided each individual with an RFID ring (2.3 mm EM4102 PIT tag, Eccel Technology Ltd., Great Britain) allowing us to automatically record their provisioning visits (see 'Measuring risk-taking behaviour

during brood provisioning' below). To be able to fit the RFID rings, any colour rings present were removed using a small pair of scissors before applying the RFID rings. Finally, a small blood sample (ca. 10 µL) was taken by brachial venepuncture for paternity analyses (for detailed methods, see de Jong et al. 2017), which we used to identify polygynous males (following the criteria described in Vedder et al. 2011). Thereafter, we released the birds at the site of capture.

Measuring territorial aggressiveness

We obtained repeated measures of territorial aggressiveness for males during the presumed fertile period of their female (i.e. towards completion of nest-building and during the early egg-laying phase; following Korsten et al. 2007a) by recording territory owners' behaviour in response to simulated territorial intrusion (STI) tests. At nestboxes where we found complete nests (see above), we conducted tests opportunistically (up to a maximum of three tests) until the first egg was laid. Thereafter, we aimed to conduct two tests during egg-laying: one on the day after the first egg was laid and an additional one within 8 days after the first egg was laid. We thus obtained a maximum of 5 repeated measures of aggressiveness for the males. See electronic supplementary material Fig. S1 for the frequency distribution of the number of aggressiveness tests we carried out per day relative to the laying date at each of the nestboxes where the tests were conducted. Consecutive tests at the same nestbox were conducted with at least one day in between on which no test was run. All tests were conducted from 1st–30th April 2018 between 08:00 and 12:00.

We adopted an STI test setup similar to that used by Korsten et al. (2007a), wherein a male taxidermic blue tit mount along with playback of male song was used as a stimulus. We caged the mounts using green wire-mesh to protect them from actual attack. The mount atop a 1.5-m-long wooden pole was placed within 2 m of the nestbox facing its entrance. The upper end of the pole was provided with two 'arms' (1 m each) at the centre of which the taxidermic mount was placed. This T-shaped setup was designed to even out the opportunity for focal blue tits to approach the stimulus despite variation in the vegetation immediately surrounding different nestboxes. See Fig. S2 for a photograph of the STI test setup.

An mp3 player (iPod Nano, Apple Inc., USA) connected to a speaker (JBL Go, Harman International Industries, Inc., USA) was placed at the base of the pole for playback of male song. Six male blue tit songs (recorded from Dutch blue tit populations geographically distant from our study site) acquired from the Xeno-Canto online database (www.xeno-canto.org; Catalogue numbers: XC95324, XC97938, XC130384, XC235424, XC292031, XC293631), and 6 male

taxidermic mounts (sex verified by molecular sex determination; Griffiths et al. 1998) were used. We looped each male song with added breaks that mimicked natural song, into a 20-min track using the Audacity software (v2.2.0, Audacity Team 2018). The first 30 s of the track were left silent, allowing the observers to move away from the stimulus before the start of the test (see below for further details). All playback was set at the same volume mimicking natural singing amplitude of blue tits. The taxidermic mounts and male songs were independently and randomly assigned to each test, under the condition that each of the songs and mounts was only used once at a given nestbox.

The sexes of blue tits are difficult to distinguish from a distance. Therefore, we conducted the tests as pairs of observers and recorded responses of both members of the focal breeding pair. Each observer tracked the behaviour of one individual from the pair using binoculars and standing at a distance of around 15 m from the stimulus mount. During the entire STI, each observer recorded the behaviour of their focal individual in the form of a voice recording (using a Voice Tracer LFH0648, Koninklijke Philips N.V., the Netherlands). We identified individual birds by the combination of their colour rings (fitted in previous years or the preceding winter; see 'General field methods' above) which we verified upon catching them later in the season. These were then replaced by RFID rings (see 'General field methods' above). As female behaviour in response to the STIs cannot unambiguously be interpreted as aggression (as females may also advance towards the male mounts with a non-aggressive motivation), we excluded behavioural recordings from females from further analyses. In the large majority of aggressiveness tests, the female was present during the trial (95.5%, $N = 132$), thereby generating little additional heterogeneity among the male responses.

A test began at the start of song playback. A focal individual was considered to have responded to the stimulus if it arrived within a 5 m radius from the mount, within 15 min of the start of song playback. A test was terminated if no birds responded within 15 min. Along with the time taken to respond, the following data were recorded for each focal bird during a 5-min observation period starting from the moment of its response: (1) occasions when it landed on either 'arm' of the setup, (2) occasions when it landed on the cage of the mount (considered as 'attack'; following Korsten et al. 2007a), (3) and occasions when it was elsewhere. If a blue tit was already heard to alarm call or seen around the nestbox prior to the start of a test, the attempt was postponed until later that same morning. Immediately after the test, the stage of nest-building or number of eggs laid was recorded. It was not possible to record data blind because our study involved focal animals in the field.

In all, we conducted 247 aggressiveness tests at 88 nestboxes. From these, a number of tests did not yield usable data ($N = 115$ at 36 nestboxes) for one or more of the following reasons: (1) nests were abandoned or depredated and hence breeding birds could not be caught during brood provisioning for verification of individual identification ($N = 42$ tests at 14 nest boxes), (2) birds of a pair could not be individually distinguished during the test for example due to absence of rings (i.e. colour or aluminium; $N = 11$ tests at 4 nestboxes), (3) the birds' colour ring combinations recorded during the test were different from those of the individuals caught at the nestbox later on ($N = 9$ tests at 6 nestboxes), (4) attempts to catch the breeding males were unsuccessful ($N = 27$ tests at 10 nestboxes), and (5) nests turned out to belong to great tits ($N = 8$ tests at 2 nestboxes). From the remaining 150 tests at 56 nestboxes, we had to exclude another 18 tests at 13 nestboxes, because males did not respond. Thus, further analyses were based on 132 responses from 52 males. The number of responses obtained from each male varied over 5 repeats ($N = 1$), 4 repeats ($N = 5$), 3 repeats ($N = 21$), 2 repeats, ($N = 19$) and a single measure ($N = 6$).

To analyse the voice recordings taken during the aggressiveness tests, we used the BORIS v 6.2.1 software (Friard and Gamba 2016). We extracted the following variables from the recordings (in seconds): (1) the latency to respond to the stimulus (i.e., entering the 5 m radius from the mount) from the start of the test, (2) the latency to first land on either arm of the setup from the time of first response, (3) the latency to 'attack' from the time of response, (4) the total time spent on either arm of the STI setup, and (5) the total time spent on the cage of the taxidermic mount, during the 5-min observation period (hereon labelled as the 'time spent attacking'). In our further analyses, we used the total time spent attacking as a proxy for aggressiveness (see Text S1 for justification). When birds responded but did not attack the caged taxidermic mount ($N = 34$ tests, $N = 21$ males), the time spent attacking was taken as 0 s, and we included these cases in the analysis thereby avoiding potential sampling bias (following Stuber et al. 2013).

Measuring risk-taking during brood provisioning

As a measure of risk-taking, we recorded the males' latencies to resume brood provisioning after 10 min of simulated predator presence at their nestbox. We attempted to obtain risk-taking measures at those nestboxes where we ran territorial aggressiveness tests. To automatically record the birds' provisioning visits, we installed RFID logging equipment at nestboxes (an antenna at the entrance connected to a data logger) when broods were 9 days old. This allowed us to detect every entry and exit (undistinguished) between 06:00 and 20:00 of birds fitted with an RFID ring (see 'General field methods' above') at brood ages of 10, 11 and 12 days. One human predator (hereon HP) trial and

one sparrowhawk predator (hereon SP) trial were randomly assigned to be conducted on one of these days, one predator trial per day. On the remaining — non-treatment — day no trial was conducted to record the brood provisioning rate in the absence of a simulated predation threat. Predator trials were conducted between 08:00 and 12:00.

For an HP trial, an observer ascended a ladder to access the nestbox. The nestbox was quickly checked, and if an adult bird was found to be present inside, it was gently let out of the nestbox. The observer subsequently 'logged' the start of the 10-min presentation by running a unique RFID transponder through the entrance of the nestbox. Thus, the RFID data logger recorded the exact time at which a trial began. The observer remained ascended on the ladder until the end of the 10-min predator presentation. Thereafter, the observer 'logged' the end time of the presentation using the RFID transponder again, descended the ladder and quickly left the area. A total of 6 observers conducted the HP trials. In 3 cases we found a bird inside the nestbox at the start of an HP trial. All of these turned out to be females (verified from last entry of RFID data).

An SP trial was conducted in a similar manner as described above. However, after 'logging' the start time, the observer quickly placed a taxidermic sparrowhawk mount on the nestbox, immediately removed the ladder and left the area for 10 min before returning to collect the mount and 'logging' the end of the predator presentation. Except when placed on the nestbox, the sparrowhawk mount was always completely covered by an opaque plastic sheet (e.g. during transportation). A total of 6 observers used 3 different sparrowhawk mounts to conduct SP trials. In 7 cases we found a bird inside the nestbox at the start of an SP trial. Only one of these turned out to be a male. The mounts were assigned to a trial at random.

Thereafter, on day 13 we collected the RFID antenna and logger from the nestbox and read out the recorded nest visits. From these data, we obtained the males' latencies to resume brood provisioning from the end of the predator presentations ('latency to resume provisioning'). Furthermore, from the 2 h immediately preceding the predator trials, we estimated the mean time between consecutive provisioning events (i.e. the 'pre-trial inter-visit interval', hereon pre-trial IVI; see Text S2 for RFID data processing and calculation of pre-trial IVI). We used these pre-trial IVI estimates to control in our analyses for potential differences in latency due to differences in individual provisioning rates immediately prior to the trials (see statistical analyses below). In addition, we also estimated the inter-visit interval from the day on which no predator trial was conducted (hereon IVI; see Text S2). Mean IVIs were estimated over the same time period in which the predator trials were conducted (08:00–12:00). The observer (SMS) was not strictly blinded to the identity of the birds or the predator type during the data extraction

process. However, since measurement of the latencies and IVI was automated, it is improbable that non-blinding could introduce an observer bias in the data.

In total, we measured IVIs and latencies after HP and SP presentations at $N=40$, $N=38$, and $N=38$ nestboxes, respectively. Of these, we excluded trials in which males and/or their female partners were not provisioning their broods in the morning prior to the predator trials ($N=3$ SP trials). Furthermore, we excluded three males that turned out to be polygynous (following criteria in Vedder et al. 2011; IVI estimates; $N=3$; HP trials: $N=3$; SP trials: $N=3$), as their nest attendance may be deviating compared to that of monogamously breeding males (Kempenaers 1995; Schlicht and Kempenaers 2021). Finally, we excluded one suspected polygynous male (IVI estimate; $N=1$; HP trials: $N=1$; SP trials: $N=1$), which showed extreme outlier values for estimates of both its pre-trial IVI and latency (see Fig. S3). We thus retained $N=36$ IVIs and latencies after $N=34$ HP and $N=31$ SP presentations along with their corresponding pre-trial IVIs ($N=65$) from a total of 38 males for further analysis (see Fig. S4 for more details on sample sizes). For $N=33$ of these males we obtained responses in the aggressiveness tests.

Statistical analyses

In summary, we conducted our analyses in three main steps matching our specific research aims (for details, see further below). First, we estimated (a) the repeatability of aggressiveness within males ($N=52$ males). Second, we estimated (b) the repeatability of their risk-taking (i.e., the latency to resume brood provisioning after the predator presentations) across the two predator trials (HP and SP) ($N=36$ males). Finally, we investigated (c) the relationship between the males' level of territorial aggressiveness and their risk-taking responses toward the two predator types. For this last step, we ran a bivariate analysis allowing for the use of the full dataset ($N=57$ males) with no need to calculate average values across repeated measures of aggressiveness or risk-taking (following, for example, Jacobs et al. 2014; Niemelä and Dingemanse 2017).

(a) Repeatability of aggressiveness. When estimating the repeatability of male territorial aggressiveness (i.e., the time spent attacking during a 5-min observation period), we controlled for a number of fixed effects. These included two key biological variables that are often found to relate to behaviour and reproductive success: individual age and timing of breeding (e.g. Amininasab et al. 2017; Araya-Ajoy and Dingemanse 2017). Hence, we included (1) age of the male (first year breeder or older) and (2) the laying date at its nest. Additionally, we controlled for variables associated with our study design which potentially influenced the level of aggressiveness measured (see Wilson 2018): (3) the

number of days relative to the laying date (= day 0) on which the test was conducted, taking days before the laying date as negative integers, (4) the time between 08:00 and 12:00 (measured in minutes) at which the test was conducted, and finally (5) the sequence in which the tests were conducted (as a factor with five levels) to control for potential habituation effects (see e.g. Dingemanse et al. 2012b). Along with male ID, we also included taxidermic mount ID and song ID as random effects. The 'adjusted' repeatability estimate for aggressiveness was thus calculated as the proportion of the total observed variance explained by male ID, conditioned on the above fixed effects (Nakagawa and Schielzeth 2010; Wilson 2018). The aggressiveness measures showed a non-Gaussian distribution with zero inflation due to tests in which birds responded but did not land on the cage of the mount. Residuals of the models, however, showed approximate Gaussian distributions. Generally, mixed-effects models are robust to minor violations of their distributional assumptions (Schielzeth et al. 2020).

(b) Repeatability of risk-taking. In two models separately fitting risk-taking in HP and SP trials, we first assessed the amount of variance in latency explained by the random effects of human observer ID and sparrowhawk mount ID, respectively. In each of these two analyses, we controlled for a number of fixed effects again consisting of biologically relevant and study design variables. Biologically relevant variables: (1) age of the bird (first year breeder or older), (2) the size of its brood (Curio et al. 1985; Wetzel and Westneat 2014), and (3) the hatching date of its brood (Hollander et al. 2008; Wetzel and Westneat 2014). Study design-associated variables: (4) the age of the brood (in days) and (5) the time between 08:00 and 12:00 (measured in minutes) at which the trial was conducted. Additionally, we controlled for the potential effect of variation in individual provisioning rates prior to the predator trials by including (6) the pre-trial IVI as a fixed effect. The distribution of latency measures was positively skewed but approximated a Gaussian distribution after a natural log transformation. As both human observer ID and sparrowhawk ID explained relatively little variance in the latency to resume provisioning (see Results; see Tables S1, S2, S3, S4, S5, S6, S7 and S8), we excluded these random effects from subsequent analyses. This allowed us to estimate the repeatability of risk-taking behaviour (i.e., latency to resume provisioning) by including the risk-taking measures in response to the two predators in the same model. In this model, we included individual male ID as a random variable. In addition to the fixed effects described above, we also included the type of predator presented. The adjusted repeatability estimate for risk-taking was thus calculated as the proportion of the total variance explained by male ID, conditioned on the fixed effects. Additionally, to validate that our predator presentations were perceived as an actual threat, we also compared the time taken by male blue tits to resume provisioning after presentation of a predator (HP or SP) with their regular IVIs in the absence of a predator on the non-treatment day (including male ID as a random effect). IVI from the non-treatment day were natural-log transformed like the latency data to make them directly comparable on the same scale.

(c) Aggression and risk-taking. We fitted a bivariate model with the measures of aggressiveness and risk-taking as response variables. To reduce model complexity, in this model we only included predator type and hatching date as fixed effects on risk-taking as only these predictors were found to explain significant variation in risk-taking in the univariate models, while for aggressiveness, none of the fixed effects were significant (see Results). The random effect of male ID was included to estimate the among-individual covariance between aggressiveness and risk-taking (Dingemanse and Dochtermann 2013). Since the (repeated) measures of aggressiveness and risk-taking were taken at two separate time-points during breeding and are therefore not paired in time, it was not possible to estimate the residual (within-individual) covariance (as explained in Dingemanse and Dochtermann 2013). We calculated the among-individual correlation between aggressiveness and risk-taking as: among-individual covariance / √(among-individual variance for aggressiveness × among-individual variance for risk-taking).

The statistical analyses were carried out in R version 4.0.3 (R core Team 2020). We fitted the univariate linear mixed-effects models with a Gaussian error distribution using the 'lme4' package (Bates et al. 2014). We constructed 'full' models providing estimates of all fixed and random variables, when fitted simultaneously (see Tables 1 and 2). For each full model, we also obtained the corresponding reduced model by stepwise backward elimination of the fixed predictors with a P value greater than 0.05, starting with the predictors with the highest P values. This alternative modelling approach did not yield qualitatively different conclusions (see electronic supplementary material). All estimates of variance components and other parameters presented are based on models fit with the restricted maximum likelihood method. P values of individual random and fixed effects were inferred from log-likelihood ratio tests comparing nested models refitted with the maximum likelihood method. We estimated the adjusted repeatabilities and their respective standard errors from the full models using the 'rptR' package (Stoffel et al. 2017).

We fitted the bivariate mixed-effects model with a Gaussian error distribution using the 'asreml' package (ASReml-R version 4.1.0; Butler 2020). P values of the fixed effects were inferred from conditional Wald F tests as implemented in ASReml-R. To obtain the P value for the among-individual covariance estimate, we used a log-likelihood ratio test comparing the full model with a model where we constrained

Table 1 Summary of the linear mixed-effects model investigating sources of variation in aggressiveness of male blue tits measured as the total time spent attacking (in seconds) a caged conspecific male taxidermic mount in simulated territorial intrusion tests ($N = 132$ observations from 52 males)

Fixed effects	Coefficient (SE)	χ^2	df	P
Intercept (sec)	96.10 (18.37)			
Age (ref: 1st year breeder)	−27.81 (21.41)	1.737	1	0.19
Laying date	0.76 (4.45)	0.029	1	0.86
Day relative to laying date	−0.99 (1.80)	0.280	1	0.60
Time (min)	-0.001 (0.10)	0.001	1	0.99
Test sequence (ref: 1st)	2nd: 4.61 (13.98)	1.929	4	0.75
	3rd: 1.92 (23.28)			
	4th: −17.45 (41.32)			
	5th: 63.79 (75.00)			
Random effects	Variance			
Male ID	4248.82	33.772	1	<0.001
Mount ID	451.47	4.897	1	0.027
Song ID	47.86	0.121	1	0.73
Residual	2853.38			
	Repeatability (SE)			
Male ID	0.56 (0.08)			<0.001

Predictor variables included were (1) age of the male (factor with two levels; first year breeder or older), (2) the laying date at its nest (continuous variable), (3) the number of days relative to the laying date ($= 0$) on which the test was conducted (continuous variable), (4) the time at which the test was conducted (in min; continuous variable), and (5) the sequence in which the tests were conducted (factor with five levels; 1st to 5th). Blue tit male ID ($N = 52$), taxidermic mount ID ($N = 6$) and playback song ID ($N = 6$) were included as random variables. Reported are estimates of the fixed and random variables along with the adjusted repeatability of aggressiveness within males. P values of individual random and fixed effects were inferred from log-likelihood ratio tests comparing nested models refitted with the maximum likelihood method. SE and P value of the repeatability estimate are respectively based on parametric bootstrapping and likelihood ratio tests through permutation of residuals both obtained using the 'rptR' package

Table 2 Summary of the linear mixed-effects model investigating sources of variation in risk-taking by male blue tits measured as the latency to resume provisioning their nestlings in seconds (natural log transformed) after exposure to a predator ($N = 65$ observations from 36 males)

Fixed effects	Coefficient (SE)	χ^2	df	P
Intercept (log sec)	6.261 (0.256)			
Predator (ref: human)	0.531 (0.099)	20.561	1	<0.001
Age (ref: 1st year breeder)	0.116 (0.173)	0.516	1	0.47
Brood size	−0.034 (0.045)	0.680	1	0.41
Hatching date	−0.074 (0.031)	6.051	1	0.014
Brood age (ref: 10 days)	11 days: −0.096 (0.135)	3.187	2	0.20
	12 days: 0.139 (0.126)			
Time (min)	0.001 (0.001)	0.004	1	0.95
Pre-trial IVI (sec)	−0.001	0.001	1	0.67
Random effects	Variance			
Male ID	0.156	6.725	1	<0.01
Residual	0.147			
	Repeatability (SE)			
Male ID	0.51 (0.13)			<0.01

Predictor variables tested were (1) the type of predator (human or sparrowhawk), (2) age of the male blue tit (factor with two levels; first year breeder or older), (3) the number of nestlings in its nest, i.e. its brood size (continuous variable), (4) the hatching date of its brood (continuous variable) (5) the age of its brood on the day of predator trial (factor with 3 levels; 10, 11 and 12 days old), (6) the time at which the trial was conducted (in min; continuous variable), (7) pre-trial inter-visit interval (IVI; in seconds; continuous variable). Blue tit male ID ($N = 36$) was included as a random variable. For further details, see Table 1

the among-individual covariance to 0. In both these models, we constrained the within-individual covariance to 0 since it was non-estimable by design (Dingemanse and Dochtermann 2013; Niemelä and Dingemanse 2017).

For all models we mean-centred the continuous fixed variables to aid in interpretation of the model coefficients (Schielzeth 2010). We assessed homogeneity of variances and normality of residuals by visually inspecting plots of the residual versus fitted values and quantile–quantile ('QQ') plots, respectively. For testing significance of mean differences in pairwise comparisons between the IVIs and the latency to resume provisioning in the HP and SP trials, we used the Tukey's adjustment method as implemented in the

'lsmeans' package (Lenth 2016). Alpha was set to 0.05 for all analyses. All graphs were plotted using the 'ggplot2' package (Wickham 2016).

Results

Sources of variation in aggressiveness

We found the level of territorial aggressiveness of male blue tits to be substantially repeatable ($R = 0.56 \pm 0.08$ SE, $P < 0.001$; Table 1, see also Table S9). Taxidermic mount ID and song ID further explained little of the total variance (mount ID: 5.9%, $P = 0.027$; song ID: 0.6%, $P = 0.73$; Table 1). The fixed variables included as predictors explained no significant variation in aggressiveness (Table 1). A model reduction approach essentially yielded the same results (see Table S9). In this estimation of the repeatability of individual aggressiveness, we included the

males' responses to STIs both when they attacked and did not attack the mounts. When we excluded the non-attack responses (i.e., where time spent attacking = 0 s), focusing on the individual variation in aggressiveness for the responses with attacks only ($N = 98$ observations from 40 males), the time spent attacking remained repeatable within individuals ($R = 0.37 \pm 0.12$ SE, $P = 0.02$; Table S10).

Sources of variation in risk-taking

First, in separate analyses of the HP and SP trials, we found that the identities of the human observers and sparrowhawk mounts presented in the predator trials explained only small parts of the variance in the level of risk-taking, measured as males' latencies to resume brood provisioning after predator exposure (all models: < 9.4% and < 0.2%, respectively, all $P \geq 0.50$; see Tables S1, S2, S3, S4, S5 and S6). There was an indication that latencies in the SP trial were shorter for larger broods (see brood size in Tables S5 and S6), but this effect did not hold in the across-treatment analysis (see below and brood size effect in Table 2). Second, when combining risk-taking measures from HP and SP trials in a single model, we found the level of risk-taking by male blue tits to be repeatable across the two predator types ($R = 0.51 \pm 0.13$ SE, $P < 0.01$; Table 2; see also Table S11). Additionally, birds took longer to resume brood provisioning when confronted with a sparrowhawk mount (median = 732 s) as compared to a human (median = 522 s; Fig. 1a; Table 2). The latency to resume provisioning decreased with progressing hatching

dates at males' nests (Table 2). We found no evidence for relationships between the latency to resume provisioning and the age of the males, the size and age of their broods and the time at which the trials were conducted (Table 2). Also, the observed latencies did not relate to the pre-trial IVI (Table 2). Finally, the latencies to resume provisioning after the predator presentations (both HP and SP) were substantially longer than the inter-visit intervals measured on the non-treatment day (Fig. 1a; median IVI = 204 s; pairwise comparisons after natural-log transformation; HP vs IVI: $\beta \pm SE = 0.852 \pm 0.090$, t ratio = 9.407, $P < 0.001$; SP vs IVI: $\beta \pm SE = 1.389 \pm 0.095$, t ratio = 14.663, $P < 0.001$).

Aggressiveness and risk-taking

We found no significant relationship (among-individual correlation) between male territorial aggressiveness and risk-taking (i.e. the latency to resume brood provisioning after the predator presentations) (see Table 3; Fig. 1b).

Discussion

In this study we found that breeding male blue tits show repeatable variation in the expression of two ecologically relevant behaviours: territorial aggressiveness and risk-taking in the context of parental care. Additionally, in the risk-taking trials, birds adjusted their response to different levels of predator threat at their nest. However, contrary to
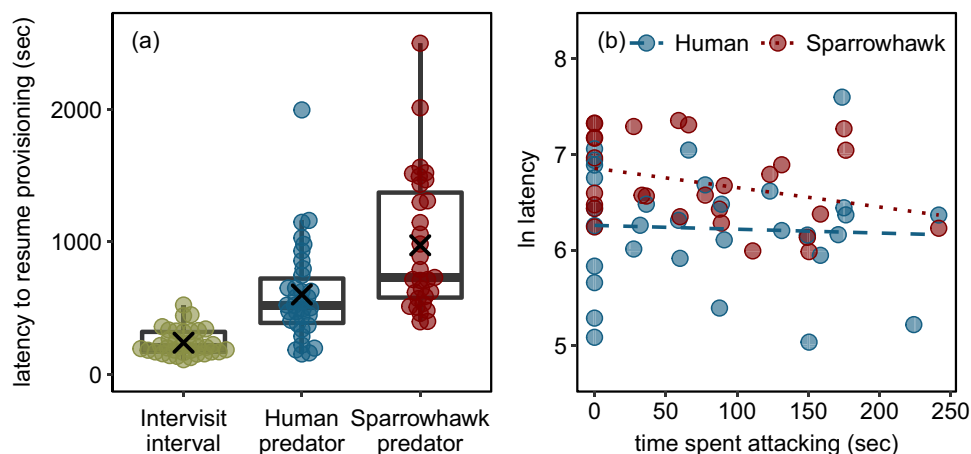

Fig. 1 Risk-taking of male blue tits measured as their latency (in seconds) to resume nestling provisioning after 10-min presentations of human (HP; $N = 34$) and sparrowhawk (SP; $N = 31$) predators at their nests. **a** The time birds took to resume provisioning after presentation of both types of predator (i.e. latency) was much longer than the regular inter-visit interval (IVI; $N = 36$) in the absence of a predator. The latencies were substantially longer after presentation of a taxidermic sparrowhawk mount compared to a human observer. Boxplots show the median, 1st and 3rd quartiles; whiskers indicate values within 1.5

times the interquartile range; cross points indicate group means. **b** Overall, males' latencies to resume provisioning did not relate to their level of aggressiveness (Table 3). For illustration, shown is the average time spent attacking a caged taxidermic male blue tit model (in seconds) across repeated simulated territorial intrusion tests against the latency measures in the human and sparrowhawk predator trials. Note that in panel (**b**) latencies (in seconds) are natural log transformed

Table 3 Summary of the bivariate linear mixed-effects model estimating the covariance for aggressiveness and risk-taking in male blue tits

	Aggressiveness	Risk-taking			
		Response variables			
Fixed effects	Coefficient (SE)	Coefficient (SE)	Wald F	df	P
Intercept	84.78 (10.20)	6.238 (0.091)			
Predator type	-	0.529 (0.099)	29.900	1, 30.3	< 0.001
Hatching date (days)	-	−0.064 (0.030)	4.700	1, 36.2	0.036
Random effects	Variance (SE)	Variance (SE)			
Male ID	4015.02 (1093.92)	0.14 (0.06)			
Residual	3214.93 (506.97)	0.15 (0.04)			
	Covariance (SE)		χ^2	df	P
Male ID	−4.39 (6.55)		0.496	1	0.50
Residual	NA				
	Among-individual correlation (SE)				
	−0.19 (0.27)				

Aggressiveness was measured as the total time spent attacking (in seconds) a conspecific male taxidermic mount during a 5-min simulated territorial intrusion test. Risk-taking was measured as the latency to resume provisioning nestlings (in seconds; natural log transformed) after presentation of a predator. Predator type (factor with two levels; human and sparrowhawk) and hatching date (in days; continuous variable) were included as fixed effects for risk-taking as these were found to significantly explain variation in risk-taking in univariate models (see Tables 2, Table S11). Residual covariance in this model is inestimable due to the different time-points at which aggressiveness and risk-taking are measured

our expectation, we found no evidence that more aggressive males took greater risks to provision their broods in the face of a predation threat. We discuss the relevance and implications of these findings below.

Aggressiveness

We found strong evidence that individual male blue tits consistently differ in their level of territorial aggressiveness. Our estimate of the within-individual repeatability of aggressiveness ($R = 0.56 \pm 0.08$ SE) falls well within the range of estimates reported in lab and field studies across different taxa, focusing specifically on territorial aggressiveness (see Table S12; repeatability estimates varying between $R = 0.07$–0.91). Two meta-analyses of the repeatability of different behaviours also found the expression of aggressive behaviour measured in many different contexts to be overall repeatable within individuals and to be among the most repeatable classes of behaviour (Bell et al. 2009; Garamszegi et al. 2013: $R = 0.45$, 95% CI, 0.21–0.67). Hence, by establishing the repeatability of aggressiveness in male blue tits, our findings add to accumulating evidence that individuals show consistent differences along this basal axis of behavioural variation.

Araya-Ajoy and Dingemanse (2017) assessed the repeatability of aggressiveness in male great tits (*P. major*), a closely related and often sympatric species to the blue tit, which shares a similar breeding ecology (Hinde 1952; Perrins et al. 1979) and competes with blue tits for the same

nesting sites (Fokkema et al. 2018; Samplonius 2019). Yet, using a similar STI setup, we found that blue tits are much more likely to engage in physical 'attack' of a taxidermic mount (i.e., jumping onto the mount's cage; 74.2%, $N = 132$ STI tests of 52 males) as compared to great tits (12.2%, $N = 1285$ STI tests of 596 males, also conducted during the egg-laying stage; Araya-Ajoy et al. 2016a) underlining the importance of considering species-specific behaviour despite ecological or phylogenetic similarity when designing and interpreting behavioural assays. The cross-year repeatability for aggressiveness found in great tits ($R = 0.57$, 95% CI, 0.37–0.77; Araya-Ajoy and Dingemanse 2017) was similar to our within-year estimate. Furthermore, aggressiveness in great tits was found to be moderately heritable ($h^2 = 0.26$, 95% CI, 0.005–0.55), though this estimate was accompanied by relatively large uncertainty (Araya-Ajoy and Dingemanse 2017). It remains to be investigated whether the observed among-individual variation in aggressiveness in our blue tit population also has a heritable component.

Our estimate of the repeatability of male aggressiveness within a single breeding season is not independent of the potential effects of characteristics of the males' territories (as pointed out by Niemelä and Dingemanse 2017; see also Wilson 2018 for further discussion). For example, the aggressive responses we obtained could (partly) represent the value of the males' territory, since each male was assayed for aggressiveness only at a single breeding site. Notably, in a population of dunnocks (*Prunella modularis*) behavioural phenotype-environment matching has been found (i.e. the

non-random distribution of individuals among territories; Holtmann et al. 2017), which implies that territory/environment characteristics may be seen as an integral part of the individual phenotype (Nicolaus and Edelaar 2018; Fokkema et al. 2021). In any case, partitioning of territory effects from among-individual variation in aggressiveness will be challenging in our study species, as blue tits are short-lived birds (less than half of the adult breeders in our study population breed more than once; Korsten 2006) that show considerable site fidelity across years (Colquhoun 1942; also see Korsten et al. 2007b).

Risk-taking

By means of our predator trials conducted at the nest, we aimed to measure the propensity of individuals to invest in their current brood at the cost of their own survival. In response to simulated predation threat at their nests, males interrupted provisioning their nestlings for much longer than their regular intervals between provisioning events. This indicates that birds likely perceived both human observers and sparrowhawk mounts as a threat of predation. They also differentiated between the two predators, with a longer latency to resume provisioning after the presence of a sparrowhawk (*A. nisus*) than of a human observer. Moreover, we found that males that quickly resumed provisioning their nestlings in the sparrowhawk predator trial also resumed provisioning quicker in the human predator trial, as indicated by the substantial within-male repeatability of latency measures ($R = 0.51 \pm 0.13$ SE). This among-male variation in risk-taking behaviour may reflect differences in individuals' investment into current versus future reproduction.

Vincze et al. (2019) examined risk-taking behaviour of breeding great tits (*P. major*) in urban versus forest habitats using a setup similar to ours, i.e. measuring the latencies of individuals to resume provisioning after presentation of human and sparrowhawk stimuli. In this study, the human and sparrowhawk stimuli were presented at a short distance from the nestbox, instead of directly next to/on top of it as in our study, and latency measures of both males and females were included. As in our study, Vincze et al. also found risk-taking to be correlated across the two predator treatments (Spearman's $r = 0.23$), although their finding was no longer significant when controlling for nestbox ID effects and baseline provisioning rates (comparable to our pre-trial IVIs; note that we found risk-taking responses to be independent of the provisioning rate prior to the predator trials; discussed below). An important methodological difference between the study of Vincze et al. and ours was that they limited the measurement of the latencies after the predator presentations to a maximum of 10 min. This may have reduced the study's ability to accurately quantify risk-taking across the entire range of individual responses, potentially reducing the power

to detect within-individual repeatability in risk-taking. For comparison, in our study in 38% and 71% of the human and sparrowhawk trials, respectively, males took longer than 10 min before resuming brood provisioning. In conclusion, the results of Vincze et al. (2019) and ours appear largely consistent, although differences in methodology may have reduced their repeatability estimate for risk-taking.

Other studies also using latencies to resume provisioning to estimate the repeatability of risk-taking, but which used different predator stimuli, found similar results. Mutzel et al. (2013a) who presented a great spotted woodpecker (*Dendrocopos major*) and a novel object (red rubber ball) on a pole at a distance of 2 m from blue tit nests, found parental provisioning latencies in response to the two stimuli to be repeatable ($R = 0.37 \pm 0.12$ SE) within a single breeding season. In another study, Mutzel et al. (2019) found the cross-year repeatability of latency responses by great tits towards a great spotted woodpecker to be somewhat lower ($R = 0.16$, 95% CI, 0.14, 0.20). Overall, within-individual repeatability of risk-taking measured as the latency to resume provisioning nestlings after a simulated predator threat appears a consistent finding both within and across years.

We found a clear difference in the males' risk-taking responses to the different predator treatments (see results, Fig. 1a, Table 2). Specifically, males responded to the human and sparrowhawk predators in a threat-sensitive manner. Plastic responses to prevailing levels of predation threat are a well-established finding in birds (Mutzel et al. 2013a, 2019; Mahr et al. 2015; Carlson et al. 2017) and other taxa (Kavaliers and Choleris 2001; Owings et al. 2001; Ferrari et al. 2010). Yet, few studies have investigated the existence of among-individual variation in such plastic responses. Building on the current study, we suggest further investigation into individual variation in plasticity of risk-taking. By obtaining repeated risk-taking measures both within and between different predator stimuli, it would be possible to quantify among-individual variation not only in average expression but also plasticity of risk-taking behaviour (Dingemanse et al. 2010; Westneat et al. 2015; Houslay and Wilson 2017). The latter is also considered a potentially important component of POLS theory (Wright et al. 2019).

Individuals with a high provisioning frequency (i.e., a low inter-visit interval) might be pre-disposed to resume provisioning more quickly after predator presence (Wetzel and Westneat 2014). We found that males in general provisioned nestlings at a high frequency (median frequency ca. once every 3.5 min) with little variation in provisioning frequencies across individuals (median: 204 s ± 99 SD, see Fig. 1a). In contrast, the variation in the latencies after predator presentation was considerably larger (HP: median: 522 s ± SD = 367 s; SP: 732 s ± SD = 520 s). Thus, the variation in provisioning rates is unlikely to contribute substantially to the observed variation in latency responses. Indeed,

we found latencies to be independent of the rates at which males were provisioning prior to predator presentations. Furthermore, the longest IVI observed was 523 s, which means that almost all males will likely have visited the nest at least once within the 10-min interval of the predator presentations (the 10-min interval was chosen based on our prior knowledge of provisioning rates in blue tits). This implies that it is highly probable that in all predator trials males (and their female partners) will have detected the (simulated) presence of the predator. The latencies to resume provisioning are therefore likely a function of the males' actual risk-taking responses in the face of a predation threat rather than a function of the combined effect of risk-taking and the provisioning rate prior to the predator trial. Therefore, we interpret the latency measures as an important individual feature of risk-taking in the context of parental care.

We found that males breeding later in the season (later hatching date at the nest), took greater risks to provision their young. This result concurs with that of Hollander et al. (2008) who found that the intensity of nest defence (alarm calling) by breeding great tits (*P. major*) towards a human intruder increased with progression of the season (for a similar recent result see de Jong et al. 2020). It is well documented that fitness pay-offs to the parents are often lower for later broods (reviewed by Verhulst and Nilsson 2008). The finding of apparent increased parental investment in later broods may therefore appear counter intuitive. Yet, an increase in parental risk-taking could be expected if the opportunity to raise a replacement brood after brood failure decreases over the season due to a decline in food availability (Curio et al. 1984; Verhulst et al. 1995).

Aggressiveness and risk-taking syndrome

The POLS hypothesis posits that individuals within populations broadly differ along a general slow-fast pace-of-life continuum (Réale et al. 2010; Dammhahn et al 2018). In line with the POLS hypothesis, we expected that male territorial aggressiveness and risk-taking in the context of parental care are correlated (at the among-individual level), in a manner reflecting fine-scale differences in investment into current versus future reproduction. However, although we found substantial repeatability in these two ecologically relevant behaviours, we found no evidence that more aggressive males take more risks to provision their young.

Our results are in agreement with an earlier study on great tits wherein risk-taking during a great spotted woodpecker (*D. major*) presentation were not correlated to levels of male aggressiveness (Mutzel et al. 2019). A meta-analysis studying the relationships between broadly different behaviours has found that 'aggressiveness' is generally weakly correlated with other behaviours such as risk-taking, activity and novel environment exploration but not novel object exploration (Garamszegi et al. 2013). Furthermore, the strength and sign of behavioural correlations are suggested to deviate from expectations of the POLS hypothesis depending on prevailing ecological conditions and selection pressures and consequentially across populations (Dingemanse et al. 2007; Réale et al 2010; Morinay et al. 2019).

By virtue of a slow-fast continuum across biological traits, individuals investing more into current reproduction should have higher within-year fitness. From our limited dataset, we found that the number of young reared to fledging at a male's nest was independent of its average aggressiveness (Pearson's $r = -0.15$, $N = 47$, $P = 0.31$) and also of its risk-taking levels (human predator: $r = -0.16$, $N = 34$, $P = 0.36$; sparrowhawk predator: $r = -0.06$, $N = 31$, $P = 0.75$). The number of young reared to fledging is certainly also a function of the female partner's fecundity and parental care investment as shown by Mutzel et al. (2013b). Additionally, the above-mentioned relationships could be influenced by alternative siring routes (i.e., differences in within and extra-pair paternities; as shown in Duckworth 2006; Araya-Ajoy et al. 2016b) that we have not accounted for in the above estimations. In forthcoming work on our study species, we intend to investigate these questions in detail using data from multiple years.

The observed lack of a correlation between territorial aggressiveness and risk-taking, despite both behaviours being substantially repeatable, is inconsistent with straightforward predictions from the POLS hypothesis. As an alternative hypothesis to explain repeatable variation in individual competitiveness (which may relate to aggressiveness), Baldauf et al. (2014) suggest through a modelling approach that investment into resource-acquiring behaviours enhancing competitiveness may trade off with an individual's ability to efficiently exploit the acquired resource (such as a high-quality territory), for example, by contributing to parental care. This may explain the presence of individual strategies within populations differing in their competitiveness but sharing equal fitness. While such a mechanism may drive the maintenance of individual variation in competition-related behaviours (such as aggressiveness), variation in resource allocation behaviours (such as risk-taking) may have other underpinnings.

Our study admittedly possesses relatively low power as that ideally required for the detection of among-individual correlations (Dingemanse and Dochtermann 2013). We were also unable to partition among-individual covariance from residual covariance due to the nature of our dataset (see Dingemanse and Dochtermann 2013), which renders our estimate of the among-individual correlation essentially nothing more than a phenotypic correlation. Obtaining data from individuals across years would allow for partitioning within and among-individual covariances, but obtaining repeated measures for a substantial number

of individuals across breeding seasons is challenging due to low survival. In any case, from a compilation of estimates of phenotypic and individual-specific correlations in the literature, Brommer and Class (2017) show that phenotypic correlations are often an adequate representation of among-individual correlations of behaviour.

Conclusions

Using field behavioural assays of ecologically relevant traits conducted on a population of wild birds, we found little evidence for a cross context behavioural correlation as predicted by the POLS hypothesis (Réale et al. 2010; Dammhahn et al. 2018). Males defending their territories more aggressively did not take greater risk to provision their broods in the face of predation, i.e. at the cost of uncertainty of their survival. We nonetheless encourage further investigation into the consistency and covariation of behaviours in light of theoretical predictions. With the help of larger sample sizes, investigation of how lifetime reproductive success and survival relates to the behaviours in question may shed light on the broader ecological implications of among-individual differences in behaviour.

Declarations

References

Alonso-Alvarez C, Doutrelant C, Sorci G (2004) Ultraviolet reflectance affects male-male interactions in the blue tit (*Parus caeruleus ultramarinus*). Behav Ecol 15:805–809

Amininasab SM, Hammers M, Vedder O, Komdeur J, Korsten P (2017) No effect of partner age and lifespan on female age-specific reproductive performance in blue tits. J Avian Biol 48:544–551

Amininasab SM, Vedder O, Schut E, de Jong B, Magrath MJ, Korsten P, Komdeur J (2016) Influence of fine-scale habitat structure on nest-site occupancy, laying date and clutch size in blue tits *Cyanistes caeruleus*. Acta Oecol 70:37–44

Araya-Ajoy YG, Dingemanse NJ (2014) Characterizing behavioural "characters": an evolutionary framework. Proc R Soc B 281:20132645

Araya Ajoy YG, Dingemanse NJ (2016) Data from: Repeatability, heritability, and age-dependence in the aggressiveness reaction norms of a wild passerine bird. Dryad, Dataset. https://doi.org/10.5061/dryad.7p160

Araya-Ajoy YG, Dingemanse NJ (2017) Repeatability, heritability, and age-dependence of seasonal plasticity in aggressiveness in a wild passerine bird. J Anim Ecol 86:227–238

Araya-Ajoy YG, Kuhn S, Mathot KJ, Mouchet A, Mutzel A, Nicolaus M, Wijmenga J, Kempenaers B, Dingemanse N (2016b) Sources of (co)variation in alternative siring routes available to male great tits (*Parus major*). Evolution 70:2308–2321

Baldauf SA, Engqvist L, Weissing FJ (2014) Diversifying Evolution of Competitiveness Nat Commun 5:5233

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783

Botero-Delgadillo E, Quirici V, Poblete Y, Poulin E, Kempenaers B, Vásquez RA (2020) Exploratory behavior, but not aggressiveness, is correlated with breeding dispersal propensity in the highly philopatric thorn-tailed rayadito. J Avian Biol 2020:e02262

Briffa M, Sneddon LU, Wilson AJ (2015) Animal personality as a cause and consequence of contest behaviour. Biol Lett 11:20141007

Brommer JE, Class B (2017) Phenotypic correlations capture between-individual correlations underlying behavioral syndromes. Behav Ecol Sociobiol 71:50

Butler D (2020) asreml: fits the linear mixed model. R package version 4.1.0.143, https://www.vsni.co.uk

Cain KE, Rich MS, Ainsworth K, Ketterson ED (2011) Two sides of the same coin? Consistency in aggression to conspecifics and predators in a female songbird. Ethology 117:786–795

Carlson NV, Pargeter HM, Templeton CN (2017) Sparrowhawk movement, calling, and presence of dead conspecifics differentially impact blue tit (*Cyanistes caeruleus*) vocal and behavioral mobbing responses. Behav Ecol Sociobiol 71:133

Carter AJ, Feeney WE, Marshall HH, Cowlishaw G, Heinsohn R (2013) Animal personality: what are behavioural ecologists measuring? Biol Rev 88:465–475

Colquhoun MK (1942) Notes on the social behaviour of blue tits. British Birds 35:234–240

Curio E, Regelmann K, Zimmermann U (1984) The defence of first and second broods by great tit (*Parus major*) parents: a test of predictive sociobiology. Z Tierpsychol 66:101–127

Curio E, Regelmann K, Zimmermann U (1985) Brood defence in the great tit (*Parus major*): the influence of life-history and habitat. Behav Ecol Sociobiol 16:273–283

Dale S, Gustaven R, Slagsvold T (1996) Risk taking during parental care: a test of three hypotheses to the pied flycatcher Applied. Behav Ecol Sociobiol 39:31–42

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 72:62

David M, Dall SRX (2016) Unravelling the philosophies underlying 'animal personality' studies: a brief re-appraisal of the field. Ethology 122:1–9

Davies S, Sewall KB (2016) Agonistic urban birds: elevated territorial aggression of urban song sparrows is individually consistent within a breeding period. Biol Lett 12:20160315

de Jong B, Lens L, van der Velde M, Korsten P, Groothuis T, Komdeur J (2017) Testosterone reduces promiscuity of female blue tits (*Cyanistes caeruleus*): an experimental study. Ethology 123:69–82

de Jong ME, Nicolaus M, Fokkema RW, Loonen MJJE (2020) State-dependence explains individual variation in nest defence behaviour in a long-lived bird. J Anim Ecol 90:809–819

Dingemanse NJ, Bouwman KM, van de Pol M, van Overveld T, Patrick SC, Matthysen E, Quinn JL (2012a) Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. J Anim Ecol 81:116–126

Dingemanse NJ, Dochtermann NA (2013) Quantifying individual variation in behaviour: Mixed-effect modelling approaches. J Anim Ecol 82:39–54

Dingemanse NJ, Dochtermann NA, Nakagawa S (2012b) Defining behavioural syndromes and the role of 'syndrome deviation' in understanding their evolution. Behav Ecol Sociobiol 66:1543–1548

Dingemanse NJ, Kazem AJN, Réale D, Wright J (2010) Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol Evol 25:81–89

Dingemanse NJ, Wright J, Kazem AJN, Daem DK, Hickling R, Dawnay N (2007) Behavioural syndromes differ predictably between 12 populations of three-spined stickleback. J Anim Ecol 76:1128–1138

Duckworth RA (2006) Behavioral correlations across breeding contexts provide a mechanism for a cost of aggression. Behav Ecol 17:1011–1019

Ferrari MCO, Wisenden BD, Chivers DP (2010) Chemical ecology of predator–prey interactions in aquatic ecosystems: a review and prospectus. Can J Zool 88:698–724

Fokkema RW, Ubels R, Tinbergen JM (2018) Experimentally quantifying the effect of nest-site depth on the predation risk and breeding success of blue tits. Auk 135:919–932

Fokkema RW, Korsten P, Schmoll T, Wilson AW (2021) Social competition as a driver of phenotype-environment correlations: implications for ecology and evolution. Biol Rev (published online, https://doi.org/10.1111/brv.12768)

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

Garamszegi LZ, Markó G, Herczeg G (2013) A meta-analysis of correlated behaviors with implications for behavioral syndromes: relationships between particular behavioral traits. Behav Ecol 24:1068–1080

Garamszegi LZ, Markó G, Szász E, Zsebőket S, Azcárate M, Herczeg G, Török J (2015) Among-year variation in the repeatability, within- and between-individual, and phenotypic correlations of behaviors in a natural population. Behav Ecol Sociobiol 69:2005–2017

Ghalambor CK, Martin TE (2001) Fecundity-survival trade-offs and parental risk-taking in birds. Science 292:494–497

Griffiths R, Double MC, Orr K, Dawson RJ (1998) A DNA test to sex most birds. Mol Ecol 7:1071–1075

Harris MR, Siefferman L (2014) Interspecific competition influences fitness benefits of assortative mating for territorial aggression in eastern bluebirds (*Sialia sialis*). PLoS ONE 9:e88668

Hinde RA (1952) The behaviour of the great tit (*Parus major*) and some other related species. Behav Suppl 2:1–201

Hollander FA, Van Overveld T, Tokka I, Matthysen E (2008) Personality and nest defence in the great tit (*Parus major*). Ethology 114:405–412

Holtmann B, Santos ESA, Lara CE, Nakagawa S (2017) Personality-matching habitat choice, rather than behavioural plasticity, is a likely driver of a phenotype – environment covariance. Proc R Soc B 284:20170943

Houslay TM, Wilson AJ (2017) Avoiding the misuse of BLUP in behavioural ecology. Behav Ecol 28:948–952

Jacobs CGC, van Overveld T, Careau V, Matthysen E, Adriaensen F, Slabbekoorn H (2014) Personality-dependent response to field playback in great tits: slow explorers can be strong responders. Anim Behav 90:65–71

Kavaliers M, Choleris E (2001) Antipredator responses and defensive behavior: ecological and ethological approaches for the neurosciences. Neurosci Biobehav Rev 25:577–586

Kempenaers B (1995) Polygyny in the blue tit: intra- and inter-sexual conflicts. Anim Behav 49:1047–1064

Korsten P (2006) Avian sex allocation and ornamental coloration: a study on blue tits. PhD thesis, University of Groningen, Groningen, https://pure.rug.nl/ws/portalfiles/portal/2880527/thesis.pdf

Korsten P, Dijkstra T, Komdeur J (2007a) Is UV signalling involved in male-male territorial conflict in the blue tit (*Cyanistes caeruleus*)? A new experimental approach. Behaviour 144:447–470

Korsten P, Lessells CK, Mateman AC, van der Velde M, Komdeur J (2006) Primary sex ratio adjustment to experimentally reduced male UV attractiveness in blue tits. Behav Ecol 17:539–546

Korsten P, van Overveld T, Adriaensen F, Matthysen E (2013) Genetic integration of local dispersal and exploratory behaviour in a wild bird. Nat Commun 4:2362

Korsten P, Vedder O, Szentirmai I, Komdeur J (2007b) Absence of status signalling by structurally based ultraviolet plumage in wintering blue tits (*Cyanistes caeruleus*). Behav Ecol Sociobiol 61:1933–1943

Lambrechts MM, Prieur B, Caizergues A, Dehorter O, Galan MJ, Perret P (2000) Risk-taking restraints in a bird with reduced egg–hatching success. Proc R Soc Lond B 267:333–338

Lenth RV (2016) Least-squares means: the R package lsmeans. J Stat Softw 69:1–33

Lorenz K (1966) On aggression. Bantam, New York

Mahr K, Riegler G, Hoi H (2015) Parental risk management in relation to offspring defence: bad news for kids. Proc R Soc Lond B 282:20141670

Maynard Smith J, Harper DG, Brookfield JFY (1988) The evolution of aggression: can selection generate variability? Phil Trans R Soc B 319:557–570

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

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 72:116

Morinay J, Daniel G, Gustafsson L, Doligez B (2019) No evidence for behavioural syndrome and genetic basis for three personality traits in a wild bird population. Anim Behav 153:69–82

Müller C, Jenni-Eiermann S, Blondel J, Perret P, Caro SP, Lambrechts M, Jenni L (2006) Effect of human presence and handling on circulating corticosterone levels in breeding blue tits (*Parus caeruleus*). Gen Comp Endocr 148:163–171

Mutzel A, Blom MPK, Spagopoulou F, Wright J, Dingemanse NJ, Kempenaers B (2013a) Temporal trade-offs between nestling provisioning and defence against nest predators in blue tits. Anim Behav 85:1459–1469

Mutzel A, Dingemanse NJ, Araya-Ajoy YG, Kempenaers B (2013b) Parental provisioning behaviour plays a key role in linking personality with reproductive success. Proc R Soc B 280:20131019

Mutzel A, Olsen A-L, Mathot KJ, Araya-Ajoy YG, Nicolaus M, Wijmenga JJ, Wright J, Kempenaers B, Dingemanse NJ (2019) Effects of manipulated levels of predation threat on parental provisioning and nestling begging. Behav Ecol 30:1123–1135

Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. Biol Rev 85:935–956

Nicolaus M, Edelaar P (2018) Comparing the consequences of natural selection, adaptive phenotypic plasticity, and matching habitat choice for phenotype–environment matching, population genetic structure, and reproductive isolation in meta-populations. Ecol Evol 8:3815–3827

Niemelä PT, Dingemanse NJ (2017) Individual versus pseudo-repeatability in behaviour: lessons from translocation experiments in a wild insect. J Anim Ecol 86:1033–1043

Owings DH, Coss RG, Mckernon D, Rowe MP, Arrowood PC (2001) Snake-directed antipredator behavior of rock squirrels (*Spermophilus variegatus*): population differences and snake-species discrimination. Behaviour 138:575–595

Perrins CM (1979) British tits. Harper Collins, London

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/

Réale D, Garant D, Humphries MM, Bergeron P, Careau V, Montiglio P-O (2010) Personality and the emergence of the pace-of-life syndrome concept at the population level. Phil Trans R Soc B 265:4051–4063

Ricklefs RE, Wikelski M (2002) The physiology/life-history nexus. Trends Ecol Evol 17:462–468

Riechert SE (1978) Games spiders play: behavioral variability in territorial disputes. Behav Ecol Sociobiol 3:135–162

Royauté R, Berdal MA, Garrison CR, Dochtermann NA (2018) Paceless life? A meta-analysis of the pace-of-life syndrome hypothesis. Behav Ecol Sociobiol 72:64

Samplonius JM (2019) Interspecific aggression declines seasonally in breeding great tits *Parus major*. Ethology 125:159–163

Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. Methods Ecol Evol 1:103–113

Schielzeth H, Dingemanse NJ, Nakagawa S et al (2020) Robustness of linear mixed-effects models to violations of distributional assumptions. Methods Ecol Evol 11:1141–1152

Schlicht E, Kempenaers B (2021) Origin and outcome of social polygyny in the blue tit. Ardea 109:91–118

Sih A, Bell AM (2008) Insights for behavioral ecology from behavioral syndromes. Adv Stud Behav 38:227–281

Sih A, Bell A, Johnson JC (2004) Behavioral syndromes: an ecological and evolutionary overview. Trends Ecol Evol 19:372–378

Stoffel MA, Nakagawa S, Schielzeth H (2017) rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods Ecol Evol 8:1639–1644

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. Behav Ecol 24:1092–1098

Svensson L (1992) Identification guide to European passerines. Fingraf, Stockholm

Szymkowiak J, Kuczyński L (2017) Song rate as a signal of male aggressiveness during territorial contests in the wood warbler. J Avian Biol 48:275–283

Vedder O, Bouwhuis S, Sheldon BC (2014) The contribution of an avian top predator to selection in prey species. J Anim Ecol 83:99–106

Vedder O, Komdeur J, van der Velde M, Schut E, Magrath MJL (2011) Polygyny and extra-pair paternity enhance the opportunity for sexual selection in blue tits. Behav Ecol Sociobiol 65:741–752

Verhulst S, Nilsson JÅ (2008) The timing of birds' breeding seasons: a review of experiments that manipulated timing of breeding. Phil Trans R Soc B 363:399–410

Verhulst S, van Balen JH, Tinbergen JM (1995) Seasonal decline in reproductive success of the great tit: variation in time or quality? Ecology 76:2392–2403

Vincze E, Pipoly I, Seress G, Preszner B, Papp S, Nemeth B, Liker A, Bokony V (2019) Great tits take greater risk toward humans and sparrowhawks in urban habitats than in forests. Ethology 125:686–701

Westneat DF, Wright J, Dingemanse NJ (2015) The biology hidden inside residual within-individual phenotypic variation. Biol Rev 90:729–743

Wetzel DP, Westneat DF (2014) Parental care syndromes in house sparrows: Positive covariance between provisioning and defense linked to parent identity. Ethology 120:249–257

Wickham H (2016) *ggplot2:* Elegant graphics for data analysis. Springer-Verlag, New York

Wilson AJ (2018) How should we interpret estimates of individual repeatability? Evol Lett 2:4–8

Wolf M, McNamara JM (2012) On the evolution of personalities via frequency-dependent selection. Am Nat 179:679–692

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

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. Biol Rev 94:230–247

Authors and Affiliations

Stephen M. Salazar[1,2] ⓘ · **Julia I. Camacho García**[3] · **Miriam Kuspiel**[4] · **Rienk W. Fokkema**[1] ⓘ · **Jan Komdeur**[2] ⓘ · **Peter Korsten**[1] ⓘ

1 Department of Animal Behaviour, Bielefeld University, Morgenbreede 45, 33615 Bielefeld, Germany

2 Behavioural & Physiological Ecology, Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, the Netherlands

3 Department of Biology, University of Antwerp, Antwerp, Belgium

4 Department of Behavioural Ecology, Wageningen University, Wageningen, the Netherlands