

The demographic drivers of cultural evolution in bird song: a multilevel study

Nilo Merino Recalde  ^{1,*}, Andrea Estandía  ¹, Sara C. Keen  ^{1,2}, Ella F. Cole  ¹, and Ben C. Sheldon  ¹

¹ Edward Grey Institute, Department of Biology, University of Oxford, Oxford, UK

² Earth Species Project, 1536 Oxford St. Berkeley CA 94709, US

*Corresponding author: nilo.recalde@biology.ox.ac.uk

Summary

Social learning within communities sometimes leads to behavioural patterns that persist over time, which we know as culture. Examples of culture include learned bird and whale songs, cetacean feeding techniques, and avian and mammalian migratory routes. Shaped by neutral and selective forces, animal cultures evolve dynamically and lead to cultural traditions that differ greatly in their diversity and stability. These cultural traits can influence individual and group survival, population structure, and even inform conservation efforts, underscoring the importance of understanding how other population processes interact with social learning to shape culture. Although the impact of social learning mechanisms and biases has been extensively explored, the role of demographic factors—such as population turnover, immigration, and age structure—on cultural evolution has received theoretical attention but rarely been subject to empirical investigation in natural populations. Doing so requires very complete trait sampling and detailed individual life history data, which are hard to acquire in combination. To this end, we built a multi-generational dataset containing over 109,000 songs from >400 individuals from a population of Great Tits (*Parus major*), which we study using a deep metric learning model to re-identify individuals and quantify song similarity. We show that demographic variation at the small spatial scales at which learning takes place has the potential to strongly impact the pace and outcome of animal cultural evolution. For example, age distributions skewed towards older individuals are associated with slower cultural change and increased diversity, while higher local population turnover leads to elevated rates of cultural change. Our analyses support theoretical expectations for a key role of demographic processes resulting from individual behaviour in determining cultural evolution, and emphasize that these processes interact with species-specific factors such as the timing of song acquisition. Implications extend to large-scale cultural dynamics and the formation of dialects or traditions.

Keywords: animal culture; bird song; demography, cultural evolution

RESULTS AND DISCUSSION

Some behavioural traits are shared and persist within communities due to social learning (Vicinaña, 2021). We refer to these behaviours as ‘animal culture’, exemplified by tool use in capuchin monkeys (Falótico et al., 2019), the learned songs of oscine birds, migration routes (Berdahl et al., 2018; Byholm et al., 2022; Jesmer et al., 2018), and the feeding techniques of some cetaceans (Allen et al., 2013; Rendell & Whitehead, 2001). Animal cultures are not static: neutral and selective mechanisms influence the frequency of cultural traits (Potvin & Clegg, 2015; Williams & Lachlan, 2021), leading to a process of cultural evolution. The resulting cultural traditions differ considerably in their diversity and stability (Tchernichovski et al., 2017), determined by both learning biases and mechanisms and the demographic structure of populations (Deffner et al., 2022; Kandler et al., 2017).

While the role of social learning strategies and

biases—frequency dependence, tutor biases, etc.—has been extensively studied (Aplin et al., 2017; Kendal et al., 2015; Lachlan et al., 2018; Pike & Laland, 2010; Tchernichovski et al., 2021), there exists a substantial gap in our understanding of how demography contributes to the emergence and persistence of distinct cultural traits within wild populations. Processes such as the recruitment of juveniles and immigration, emigration, mortality, and variation in age structure are likely to strongly affect an individual’s opportunities for learning and exposure to different cultural variants, which has been amply emphasised by theoretical work (Barta et al., 2023; Deffner & McElreath, 2020, 2022; Deffner et al., 2022; Derex & Boyd, 2016; Fogarty et al., 2019; Kandler et al., 2023; Kirby & Tamariz, 2021; Nunn et al., 2009). Despite this, translating theoretical expectations into empirical evidence remains a challenge (see Chimento et al., 2021; Fayet et al., 2014 for exceptions).

Culture is increasingly recognised as both a funda-

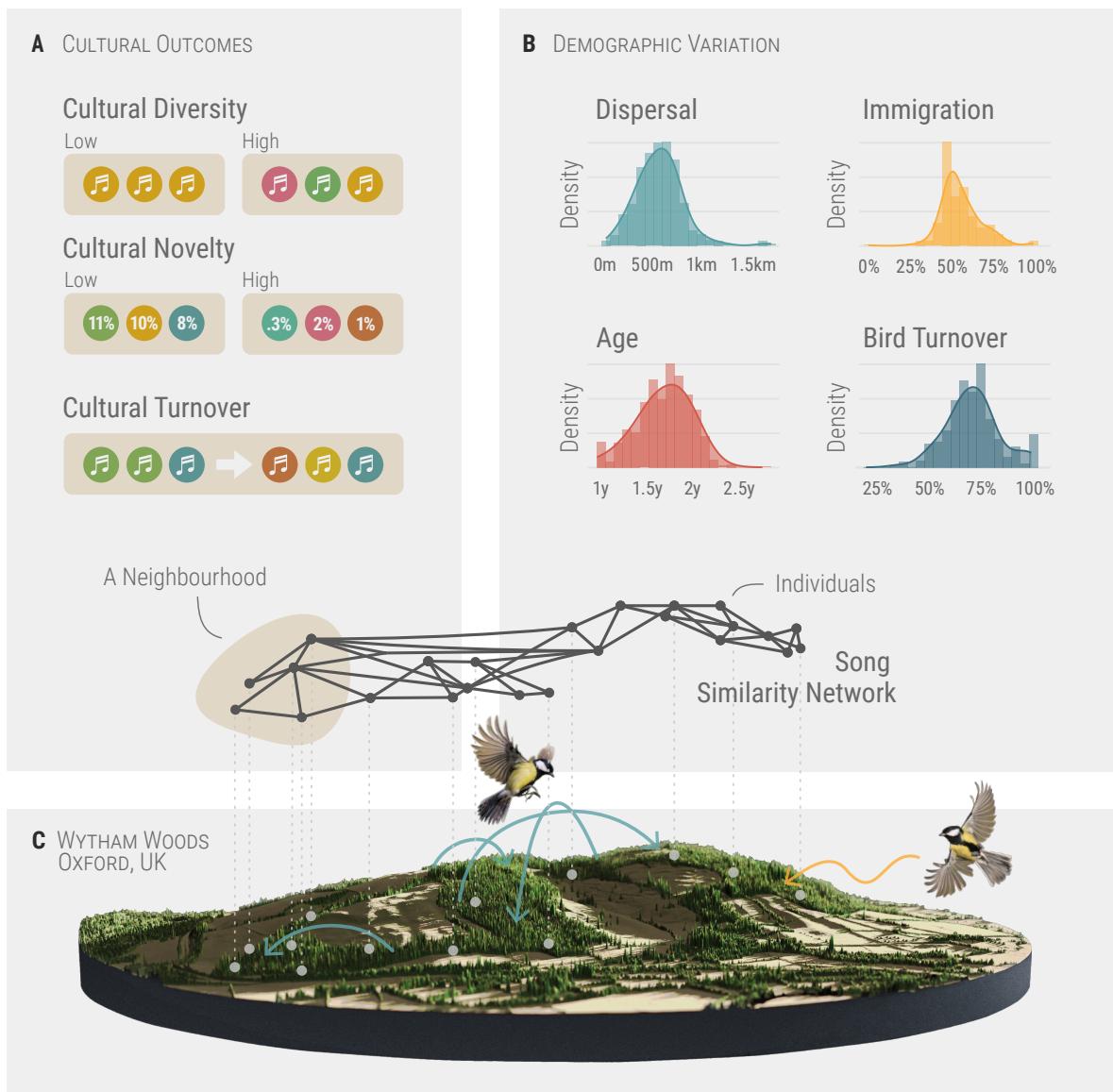


Figure 1. Study system and main variables in our analysis.

(A) Cultural variables measured at the neighbourhood level. See methods for definitions. (B) Variation in the properties and composition of neighbourhoods across the population. See methods for definitions. (C) 3D render of our study site, Wytham Woods, based on first return LiDAR data (for Environment Food & Affairs, 2020) and made with rayshader (Morgan-Wall, 2023). Elevation is exaggerated. The network represents pairwise repertoire similarity between individuals with known spatial locations, used in the models reported in Fig. 2. Aquamarine (darker) arrows represent natal dispersal within the population, and the yellow (lighter) arrow represents immigration into the population, two of the variables used in this work. Age and bird (individual) turnover are not depicted.

mental aspect of many animals' lives and a valuable tool in monitoring and conservation efforts (Brakes et al., 2019, 2021). Cultural traits play a role in the survival and reproduction of individuals and social groups; they reflect or even shape the structure of the population (Brakes et al., 2019), and can be lost when habitat fragmentation and population decline lead to reduced learning opportunities (Crates et al., 2021; Paxton et al., 2019). A comprehensive understanding of cultural change and loss, then, requires that we have the ability to detect and study not just intrinsic factors social, cultural, cognitive but also extrinsic, ecological and demographic processes. This entails identifying the relevant spatial and temporal scales at which these processes manifest within natural populations, as well as their relative importance.

To contribute to this goal, we built a comprehensive dataset that spans three years and documents the dawn songs produced by male great tit birds during 454 breeding attempts in a single population located in Wytham Woods, UK. This population has marked variation in individual turnover, postnatal dispersal distances, age structure, and immigration across space (Figure 1), which allowed us to estimate their effects on song cultural repertoires at both individual and group levels within the long-term population study of this species (Lack, 1964). First, we assign more than 109,000 songs in 330 song repertoires to 242 individual birds through a combination of direct physical capture, radio frequency identification microchips, and a novel song-based reidentification method using a deep metric learning model.

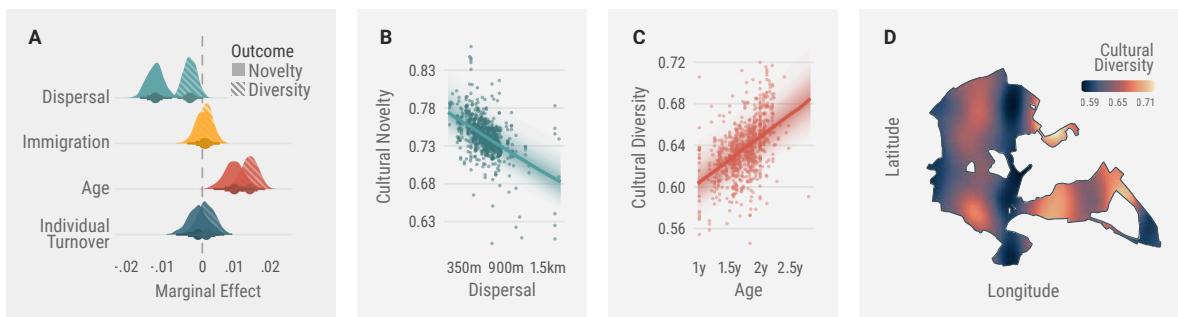


Figure 2. Influence of demographic variables on cultural diversity and novelty.

(A) Marginal effects at the mean of neighbourhood characteristics including mean dispersal distance, proportion of immigrant birds, average age, and individual turnover. (B) Adjusted predictions and partial residuals of the effect of mean neighbourhood dispersal distance on cultural novelty. Low-dispersal neighbourhoods are those in which birds were born in the same area. (C) Adjusted predictions and partial residuals of the effect of the mean age of the neighbourhood on cultural diversity. A neighbourhood with a mean age of 1 would be one where all birds are breeding for the first time. (D) The average distribution of cultural diversity in the population across space during the study period (2020-2022). This map captures the residual variation in cultural diversity after taking into account the demographic variables in (A). The map is based on a Gaussian process model with an exponentiated-quadratic kernel covariance function, which allows us to interpolate between the locations where we have data.

Then we quantified individual and group-level traits and analysed variation in song cultural similarity, diversity, and turnover using network and spatially explicit Bayesian multilevel regression models.

Our results reveal an interplay of demographics and song cultural dynamics that, albeit complex, largely matches theoretical expectations, as discussed below. This work also demonstrates that bird song, which already provides what is perhaps the largest body of evidence for cultural change in animals (Laland & Janik, 2006), also has the potential to help us shed light on the impact of other population processes on animal cultures, owing to the fact that we can sample song cultural repertoires with relative ease.

Reduced dispersal, increased immigration and an aged population are associated with higher cultural diversity

Population genetics provides robust evidence supporting the notion that high dispersal rates facilitate gene flow, which, in turn, reduces the efficacy of selection and diversification. Conversely, low dispersal facilitates genetic differentiation through mechanisms such as mutation and drift, leading to allopatric population divergence (Claramunt et al., 2011; Papadopoulou et al., 2009; Suárez et al., 2022). Drawing an analogy from genetics to culture, we anticipate that reduced dispersal rates will decelerate the diffusion of cultural traits (Nunn et al., 2009). This, in turn, should result in the maintenance of distinct behavioural patterns within populations (Planqué et al., 2014; Whitehead & Lusseau, 2012), leading to a greater abundance of cultural variants unique to a specific area. Our analysis indeed indicates that neighbourhoods where more birds have remained in proximity to their natal areas harbour greater cultural diversity and novelty (*diversity*: $P(\beta_{disp(\bar{m})} < 0|D) = 1$, $mem = -0.018$, $CI_{95\%} [-0.023, -0.012]$; *novelty*: $P(\beta_{disp(\bar{m})} < 0|D) = 0.96$, $mem =$

-0.005 , $CI_{95\%} [-0.01, 0]$; Figure 2A&B, Table S2), in line with prior research at a much coarser grain (Fayet et al., 2014).

However, the analogy breaks down when we examine our individual-level models more closely. Although the outcome at the group level resembles the homogenization of populations resulting from gene flow, the underlying mechanisms differ significantly, due to complex species-specific interactions between the timing of dispersal and learning mechanisms. In the case of great tits, these mechanisms are believed to involve selective retention or modification of songs encountered during early life and the establishment of territories following dispersal, a process that results in crystallised song repertoires that resemble those of their new neighbours at breeding sites (Marler & Peters, 1982; Nelson, 1992; Peters & Nowicki, 2017). Birds that dispersed over longer distances tend to have repertoires composed of songs that are common within the population (*novelty*: $P(\beta_{disp(m)} < 0|D) = 1$, $mem = -0.2$, $CI_{95\%} [-0.3, -0.09]$; Figure 3B, Table S2), and possibly smaller repertoires as well (*rep.size*: $P(\beta_{disp(m)} < 0|D) = 0.91$, $mem = -0.2$, $CI_{95\%} [-0.44, 0.05]$; Figure 3B; Table S2). We speculate that birds with more extensive movements are more likely to sample a larger proportion of common cultural variants, simply because they are exposed to more songs while dispersing. In contrast, birds with a more restricted and stable neighbour pool tend to be equally exposed to common and globally rarer songs, and this is sufficient to give rise to the differences that we detect at the group level (see repository for a simulation demonstrating this).

Building on our understanding of cultural dynamics in relation to dispersal we expect that, when song learning is relatively precise and dispersal is limited, cultural differences will accumulate, and immigration will introduce cultural novelty to the recipient

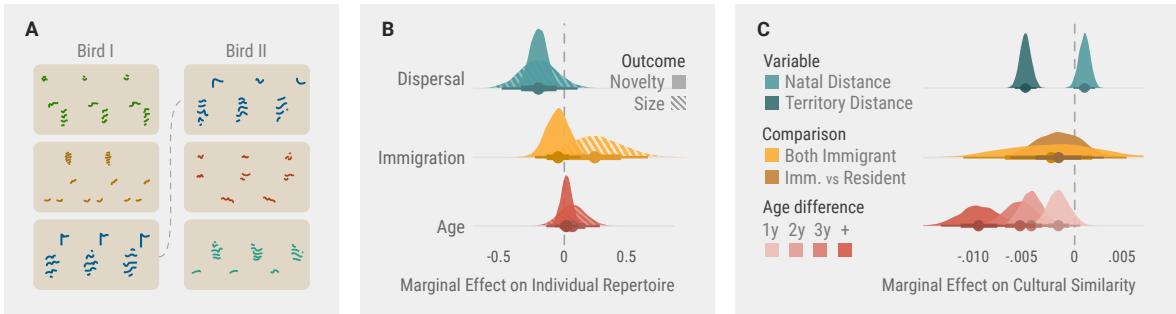


Figure 3. Individual and dyadic analysis of cultural diversity and similarity.

(A) Illustrative example showing the repertoires of two different birds in the population, with three songs in each, one of which is shared. Each sub-panel shows a stylised spectrogram, with time in the horizontal axis and frequency in the vertical. Units not shown. (B) Marginal effects of dispersal, immigration status, and age on individual song repertoires, described in terms of their absolute size (the total number of distinct song variants sang by that bird) and their relative novelty (how frequent, on average, those songs are in the entire population within that year). (C) Marginal effects of dispersal, immigration status, and age on song repertoire similarity between individuals. Dispersal: Birds that are close neighbours are more culturally similar, regardless of where they were born, whereas natal distance may have a weak positive effect on cultural similarity. Immigration: There is no strong evidence that birds born outside the population are dissimilar from resident birds. Age difference: Birds are less culturally similar the greater the difference in their birth years.

population. However, the extent to which immigration introduces new cultural variants into the population also hinges on an interplay between the species' learning programme, the timing of dispersal, and the spatial movements of individuals. Animals that learn first and then disperse, for example, may bring cultural novelty with them. But this is not the case for great tits, whose young disperse in late summer and autumn, shortly after achieving independence; learn their songs until the end of their first winter (Rivera-Gutierrez et al., 2011), and become chiefly sedentary as adults (Dhondt, 1979; Dingemanse et al., 2003; Greenwood et al., 1979). In this species, then, we anticipate that immigrant birds will learn or retain songs they encounter upon arrival, either before or during the establishment of their territories (Graham et al., 2018; Keen, 2020).

Indeed, in our population, we find no evidence that the repertoires of birds originating from outside the population significantly differ from those of resident birds ($mem = -0.002$, $CI_{95\%} [-0.006, 0.002]$; Figure 3C). This, in conjunction with the observation that cultural similarity between individuals is predicted by the distance between breeding territories ($mem = -0.005$, $CI_{95\%} [-0.006, -0.004]$; Figure 3C; Table S2), supports the hypothesis that great tits are predominantly closed-end learners that learn primarily from territorial neighbours after dispersal (Graham et al., 2017; McGregor & Krebs, 1982b; Rivera-Gutierrez et al., 2011).

This leads to a somewhat contradictory scenario, however: immigrant birds, while not acoustically distinct, tend to exhibit larger repertoires compared to their resident counterparts (Figure 3B; $P(\beta_{imm.} > 0 | D) = 0.87$, $mem = 0.24$, $CI_{95\%} [-0.098, 0.593]$; Table S2). At the group level, this small and uncertain effect amplifies, such that neighbourhoods with a higher proportion of immigrant birds do not exhibit increased cultural diversity relative to the total

number of songs ($mem = 0.002$, $CI_{95\%} [-0.004, 0.007]$; Figure 2A); but they do have a higher absolute cultural diversity—above what would be expected based solely on the number of birds ($P(\beta_{imm.} > 0 | D) = 0.98$, $mem = 0.47$, $CI_{95\%} [0.1, 0.84]$; Figure S5, Table S2).

Previous research (Verhulst et al., 1997) has revealed that most birds arriving from outside the population disperse over two kilometres, significantly farther than the typical distances observed within the population (median for males = 558 metres Greenwood et al., 1979). This extended dispersal may have qualitative consequences for cultural diversity, through a combination of factors: first, an initial exposure to songs from the source population; then, a heightened pressure to adopt vocalisations similar to those of territorial neighbours to avoid any social or reproductive costs associated with non-local signals (Baker et al., 1981; Beecher, 2008; Lachlan et al., 2014; Mortega et al., 2014; Payne, 1983).

Finally, we find that individual turnover does not significantly affect cultural diversity or novelty, and we uncover an association between age structure and cultural diversity and novelty (Figure 2B). Individuals of the same generation share the most similar song repertoires and, while age itself doesn't directly relate to changes in the repertoires of individual birds (Figure 3B), the acoustic similarity between pairs of individuals decreases as the age gap between them widens (Figure 3C; Table S2). This is expected when a species ceases to learn new songs as they age, and has detectable consequences for neighbourhoods: those with a higher proportion of older individuals have heightened levels of cultural diversity and novelty. Conversely, in areas where the majority of the population comprises active learners surrounded by their peers, birds tend to produce fewer unique songs that are simultaneously more common within the population (Figure 2A; Fig-

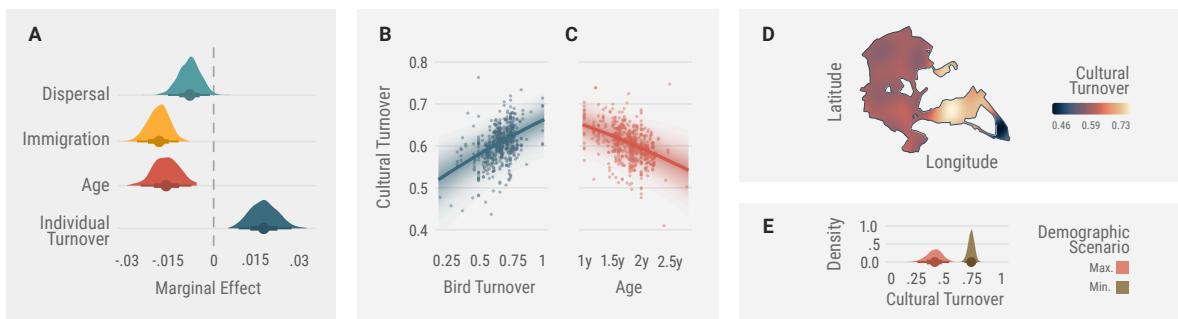


Figure 4. Influence of demographic variables on the rate of local cultural change.

(A) Marginal effect at the means for mean dispersal distance, proportion of immigrant birds, average age, and individual turnover on the rate of cultural turnover. (B and C) Adjusted predictions and partial residuals of the effects of the proportion of individual turnover on cultural turnover (B) and the effect of mean neighbourhood age on cultural turnover (C). (D) The population's average distribution of cultural turnover across space during the study period (2020–2022). (E) Posterior counterfactual predictions for two scenarios: all variables at their maximum (Max.) and minimum (Min.) observed values, adjusting for individual turnover. Cultural turnover is expected to be over two times higher when neighbourhood dispersal, immigration and age are low.

ure S5; diversity: $P(\beta_{\text{age}} < 0|D) = 1$, mem = 0.021, CI_{95%} [0.014, 0.027]; novelty: $P(\beta_{\text{age}} < 0|D) = 0.99$, mem = 0.012, CI_{95%} [0.005, 0.019]).

Demographic processes strongly moderate the rate of cultural change at small spatio-temporal scales

We now shift our focus from static measures of cultural diversity to cultural turnover, examining how quickly song variants disappear from neighbourhoods and the consequences this has for their cultural makeup. The primary driver of cultural turnover is individual turnover (total effect mem = 0.072 CI_{95%} [0.051, 0.093]): as birds leave or die, many song variants disappear with them. Accounting for this, we also assess the direct impact of mean natal dispersal distance, the proportion of immigrant birds, and mean neighbourhood age: higher levels of each of these factors correlate with slower cultural change in the neighbourhood (Figure 4A; Table S2). When there is substantial dispersal, a high influx of immigrants, and an age distribution skewed towards older individuals, the model predicts slower cultural change, at less than half the rate compared to the converse scenario (0.28 CI_{95%} [0.23, 0.34] vs. 0.61 CI_{95%} [0.49, 0.76], as illustrated in Figure 4E). Modelling work suggests that learning from older individuals should slow down cultural change (Kirby & Tamariz, 2021), aligning with our observations ($P(\beta_{\text{age}} < 0|D) = 1$, mem = -0.044, CI_{95%} [-0.063, -0.026]; Figure 4C). Age may serve as a brake on change, potentially increasing the relative cultural diversity and novelty within neighbourhoods by maintaining song types now less frequent in the population, as supported by the individual-level analysis where birds become more dissimilar as they are further in time. Across the three-year study period, now considering the entire population, cultural turnover between consecutive years hovers around 45% (0.47, 0.44). If all variants faced an equal chance of disappearing, this high

turnover rate would lead to complete cultural replacement within a short time span. However, with a two-year gap, turnover only slightly increases to 0.59. We expect this rate to taper over longer periods, as rare variants encounter greater stochasticity while common songs endure (Figure S4A). This is exemplified by some common song types documented over four decades ago that persist within the same population (Keen, 2020; McGregor & Krebs, 1982b), either through accurate learning or, more likely, strong convergent biases (Claidière & Sperber, 2007; James & Sakata, 2017; Lachlan et al., 2018; Tchernichovski et al., 2021).

Consequences for cultural structure, stability and diversity

Cultural traits, learnt bird song in this case, are shaped by many factors: some external, such as those discussed here, others intrinsic to learning and culture, and yet others that arise from selective processes driven by preference and function. Even within the confines of a relatively small population—Wytham Wood spans a mere four kilometres—we have recovered associations between heterogeneity in the demographic composition of neighbourhoods and cultural outcomes. This emphasises the need for both empirical studies and modelling efforts on cultural change to account for the population's demographic characteristics and their inherent heterogeneity across time and space, which shape individuals' exposure to cultural variants and opportunities for learning and, therefore, emergent group-level cultural dynamics.

METHODS

Resource availability

The complete Wytham great tit song dataset is available in osf.io/n8ac9/, and documented [here](#). The main repository with code and data to reproduce the analysis and figures in this article can be found at [birdsong-demography](#).

Data collection

Study system and fieldwork

Great tits are small, short-lived birds—average lifespan: 1.9 years—that sing acoustically simple yet highly diverse songs. Each male great tit has a repertoire of one to over 10 song variants, referred to as ‘song types,’ which are repeated multiple times in short bursts separated by longer periods of silence. During the breeding season, from March to June, great tit pairs are socially monogamous and defend territories around their nests (Hinde, 1952). In Wytham Woods, Oxfordshire, UK ($51^{\circ}46'N$, $1^{\circ}20'W$), a population of these birds has been the focus of a long-term study since 1947 (Lack, 1964). Wytham Woods is a semi-natural predominantly deciduous woodland that spans an area of approximately 385 hectares and is surrounded by farmland. Most great tits in this population breed in nest boxes with known locations, and the majority of individuals are marked with a unique British Trust for Ornithology (BTO) metal leg ring as either nestlings or adults.

We collected data from late March to mid-May during the 2020, 2021, and 2022 breeding seasons. Every year, fieldworkers checked each of the 1018 nest boxes at least once a week before and during the egg-laying period, which typically lasts from one to 14 days (Perrins, 1965), and recorded the identities of breeding males and females, the dates of clutch initiation and egg hatching, clutch size, and fledgling number and condition under standardised protocols. We found the first egg date by assuming that one egg is laid every day and counting back from the day of observation. In cases where we did not observe the chicks on their day of hatching, the actual hatching date was determined by assessing the weight of the heaviest chicks and extrapolating their age from established growth curves (Cresswell & McCleery, 2003; Gibb, 1950).

To record the vocalisations of male great tits we took advantage of their behaviour during the reproductive period, when they engage in continuous singing near their nests at dawn before and during egg laying (Mace, 1987). Collectively, this vocal display is referred to as the dawn chorus, and has been demonstrated to yield a reliable estimation of the song repertoire of individuals when recorded in full (Rivera-Gutierrez et al., 2012; Van Duyse et al., 2005). As soon as we suspected that a pair of great tits were using a nest box—based on nest lining materials, egg size if present, or other signs of activity—we deployed an autonomous sound recorder nearby. Our goal was to maintain a consistent position and orientation for the recorder. The microphone faced upwards and slightly away from the nest box, aligning with the nest box entrance hole’s direction. The birds sang near the recorder, and although we did

not gather data on this aspect, our anecdotal observations were in line with a different population where the average distance to the nest box was 10 metres (Halfwerk et al., 2012). The birds also changed perches and moved around during our recording. Although variation in sound amplitude due to changes in distance and direction could affect song selection, we did not observe systematic bias, ruling out potential issues like consistently low signal-to-noise ratios causing exclusion of entire song types.

All work involving birds was subject to review by the University of Oxford, Department of Zoology, Animal Welfare and Ethical Review Board (approval number: APA/1/5/ZOO/NASPA/ Sheldon/TitBreedingEcology). Data collection adhered to local guidelines for the use of animals in research and all birds were caught, tagged, and ringed by BTO licence holders (NMR’s licence: C/6904).

Recording equipment and schedule

We used 60 (30 in 2020) AudioMoth recorders (Hill et al., 2019), which were housed in custom-built waterproof enclosures. Recording began approximately one hour before sunrise (05:36 – 04:00 UTC during the recording period) and consisted of seven consecutive 60-minute-long recordings with a sample rate of 48 kHz and a depth of 16-bit. To sample as many birds as possible, we left each recorder at the same location for at least three consecutive days before moving it to a different nest box. We relocated 20 recorders (10 in 2020) every day throughout the recording period.

Data processing and annotation

We processed and annotated the song recordings, 109,963 in total, using custom software and scripts written in Python 3 (van Rossum, 1995) and the open source package `pykanto` (Merino Recalde, 2023a). These are available from github.com/nilomr/great-tit-hits-setup (Merino Recalde, 2023b). Our annotated dataset and a detailed description of the process can be found in Merino Recalde et al. (2023).

Identifying individuals and their traits

We further augmented our dataset by training a deep metric learning model (see (Merino Recalde et al., 2023) for details) to recognize individual songs, which we then used to assign individual IDs to a subset of birds that we failed to physically capture or identify using PIT (Passive Integrated Transponder) tags. This increased the number of identified breeding attempts for which we also had songs from 299 to 330, belonging to 242 unique birds. Briefly, we calculated pairwise song distances using the feature vectors obtained from the trained model. Then we assigned unknown song repertoires to known birds if they met two conservative criteria: that at least two songs had a Euclidean distance below 0.9, and that

the unknown singer was recorded less than 100 metres apart from the known individual (see Figure S3 for a graphic explanation). Natal dispersal distance was calculated as the straight line distance from the natal site to the breeding site. The dispersal distances of birds classed as immigrants (not ringed as chicks in the population) are not known, but most are thought to come from other populations at least 1 km, and likely more than 2.5 km, away (Quinn et al., 2011; Verhulst et al., 1997). We determined age based on the year of hatching for birds born in the population; and plumage characteristics for immigrants, which are most often caught as yearlings (76%)—allowing us to age them accurately (Woodman et al., 2023).

Characterising repertoire similarity

Our analyses require i) a measure of the acoustic similarity between any two birds, and ii) a way to identify song cultural variants. The underlying assumption is that song repertoires will be more similar if one bird has learned it at least in part from a second, or if they have both learnt from other individuals who are themselves similar due to intergenerational cultural descent. There is no single optimal solution for this problem, both due to technical challenges and because we do not know enough about song perception and learning mechanisms in this species. There are three main possible approaches, each with its own advantages and disadvantages.

Continuous similarity

Traditional methods used to compare bird vocalisations include visual inspection of spectrograms and measurement of hand-picked acoustic features. However, these approaches have limitations in dealing with noise and variations in performance and can be extremely time-consuming. They also fail to capture complex features such as the syntactic relationships between notes. So, instead, we adopted a data-driven approach by training a Vision Transformer (ViT) model for feature extraction in a metric learning task. Our goal was to create a similarity space based on inherent variation in the data, using categorical labels of song types sung by individual birds, which we know to be perceptually and behaviourally significant (Lind et al., 1996). Further details and code are available at (Merino Recalde, 2023a) and (Merino Recalde et al., 2023). We used the resulting model to calculate feature vectors for each song in the dataset (109,963 samples x 384 dimensions), which serve as compressed representations that can be used to compare them.

Great tits have variable repertoire sizes and there is no evidence that they ever learn them en bloc (McGregor & Krebs, 1982b; Rivera-Gutierrez et al., 2010). Therefore, the simplest continuous measure (an average pairwise Euclidean distance between all songs) would mask any signatures of learning if the aver-

age repertoire similarity is similar across the population, and does not take into account the asymmetry in total repertoire size. To improve on this, we define repertoire similarity as the average minimum Euclidean distance (AMED), given by

$$\text{AMED} = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} |a - b|_2 \quad (1)$$

where we compare each song feature vector a in set A with all song feature vectors b in set B and compute their Euclidean distance $|a - b|_2$. We then retain the minimum distance for each element in set A and obtain the AMED by averaging these minimum distances over all elements in set A . The main advantage of this approach is that it allows us to avoid imposing discrete population-wide song categories. On the other hand, if song learning is categorical and not very precise in terms of fine song structure, this method could underestimate it or fail to detect it. We used this approach for all individual-level analyses in this paper.

Automated clustering

Instead of calculating a continuous measure of repertoire similarity, we can build a pairwise distance matrix for all songs, assign them to discrete clusters using a clustering algorithm, and then calculate the intersection between repertoires by using the Jaccard coefficient or modelling it as a binomial process, with n = the combined repertoire size and s = the number of songs in the same cluster. Here we used UPGMA hierarchical clustering and dynamic tree-cut techniques to classify the syllables into distinct types, allowing a minimum cluster size of 1 to ensure the representation of rare song types. The usefulness of this method relies on the global properties of the embedding space derived from section ???. In a low-dimensional space where linear distances effectively capture meaningful variation, creating clusters by cutting the hierarchical tree at different heights yields varying cluster counts while maintaining meaningful groupings. However, in a high-dimensional space where global distances are not meaningful, only relatively small clusters of nearby points remain interpretable. This is the case with our dataset and embedding space: we find that the method reliably groups song renditions by the same bird across different years, alone or together with other birds with highly similar songs, yet consistently splits songs that are similar by human (and perhaps great tit (Falls et al., 1982)) standards, ultimately leading to a very large number of clusters (the most stable clustering solutions were close to the total number of different individual song types, >1000). Due to these issues, we did not use song types defined in this way.

Manual categorization

To date, all research on great tit song has relied on a visual classification of songs into population-level types (Baker et al., 1987; Falls et al., 1982; Fayet et al., 2014; Hutfluss et al., 2022; McGregor & Krebs, 1982a; McGregor & Krebs, 1982b; McGregor et al., 1981). This process is both inevitable and very subjective. However, despite its clear problems, human perceptual judgments might be our best available substitute for those of the birds (but see recent work by Morfi et al., 2021; Zandberg et al., 2022) for some tasks. Indeed, across fields, advanced classification algorithms are often evaluated against ground truth created by humans, and this is also the case in bird song research.

Our neighbourhood-level analyses require that we define discrete cultural units, so, given the difficulties with the alternatives described above, we adopted a variant of this approach and used the criteria followed by McGregor and Krebs (1982b) and most subsequent work. With over 100,000 songs, our dataset is much larger than is common in the field and would have been impossible to label entirely manually. Instead, we used the output of the process described above, consisting of labelled song repertoires (birdID x song type). This made the problem 57 times smaller: 1920 song variants that were already assigned to small clusters of highly similar songs, which we reviewed manually.

Following common practice in the field, we validated our manually assigned labels statistically, although we note that i) the ability of a statistical method to differentiate between manually defined clusters does not mean that these are perceptually meaningful, only that they can be distinguished in a manner that aligns with human classification, and ii) a large range of clustering solutions will be compatible with the data. To do this, we retrained the ResNet50-based classifier described in Merino Recalde (2023a) using a random subset of the data and obtained an accuracy of 0.87 on the validation set (see other metrics in the repository). With the caveats already mentioned, this means that our manual classification following McGregor and Krebs (1982b) is successful at finding a stable solution that reduces intraclass variation. A comparable process by Fayet et al. (2014) was able to reach 0.71 accuracy for 374 songs. We further explored the result by building a dendrogram based on the confusion matrix during test time and reviewing the classes that were not well supported, which led us to combine seven classes into two. There is an inverse relationship between how densely occupied a region of the song space is and the ease with which we can find categorical divisions: the more examples the more graded the variation and, in consequence, what may have seemed like clear-cut categories if we had fewer data blend

into one another without an obvious transition.

In practical terms, because most of the great tits in our population sing some variation of the well-known 'tea-cher, tea-cher' song, these are much harder to categorize than the many rare songs with complex structures only sung by one or a few birds. This was our impression when manually labelling the songs, and it was also the case when applying the supervised classification algorithm. As mentioned in the main text, the direct consequence of this for our analysis is that the absolute estimates of cultural turnover depend on the granularity of this process: when we lump all similar 'tea-cher' songs, as McGregor and Krebs (1982b) do, the estimates of turnover are necessarily lower—but, crucially, any relative differences remain the same. The code used to perform the song type validation process, along with the figures generated during it, can be found in [the main narrative notebook](#) and [a dedicated repository](#).

Quantification and statistical analysis

Pairwise similarity and individual repertoire models

It is common for analysis of song similarity to fit simple linear regression models using all pairwise comparisons in a population. However, this leads to very strong pseudo-replication and, therefore, an increased chance of Type I errors. To avoid this, we treat our song similarity data as a fully connected network and build Bayesian multilevel models with a multi-membership structure and the pairwise AMED described above as the response variable. The full model specifications can be found in the [main repository](#) for this project; also see a summary in [Table S1](#).

Individual repertoires We first modelled individual repertoire size using Poisson and negative binomial models, but this led to poor performance as assessed through posterior predictive checks (both underestimation of mean values and either under or overestimation of very low repertoire sizes). Instead, we built continuation ratio models, a type of sequential ordinal model where reaching a particular level (number of song types in the repertoire) requires first reaching all lower levels (Chambers & Drovandi, 2023; Warti et al., 2020). rep_{m_1} , $rep_{m_{1,1}}$, and $rep_{m_{1,2}}$ estimate the association between immigrant status, distance dispersed, age, and repertoire size. Three further log-normal models, $repnov_{m_1}$, $repnov_{m_{1,1}}$, $repnov_{m_{1,2}}$, do the same for the average cultural diversity of individual repertoires

Pairwise similarity Our first model ($disp_{m_1}$) explores the interaction between natal distance, that is, the distance between the nests where two resident birds were born, and the distance between the centre of their breeding territories, adjusting for year and absolute age difference. We do not have direct information on how long birds have spent around one

another, so instead we estimate the effect of the interaction of the distance at which they were born and the distance at which they subsequently breed: If both are small, they will have had more opportunities for interaction and learning. We extract predictions for the interaction and calculate marginal effects at minimum distances, to answer the questions 'How does cultural similarity change with distance for birds that were born nearby' and 'Does how close a bird was born matter for birds that hold territories nearby'. We use a similar model structure (age_{m_1}) to estimate the marginal effects of the absolute age difference, this time adjusting for the natal and territorial distance between birds. Then, to study the effect of immigration, we fit a model (imm_{m_1}) with the possible combinations of immigration status (both immigrant, both residents, one of each) and adjust for age difference and territorial distance.

Group-level properties

Defining neighbourhoods and their demographic properties Song turnover, diversity, and novelty are group-level properties. However, our study lacks naturally occurring distinct subpopulations that we can use as units for analysis. Rather than partitioning the population using a discrete polygonal grid or non-overlapping areas, we opted to model neighbourhoods continuously across space, with a radius of 200 m around each nest box occupied at least once during the study (Fayet et al., 2014) which we sampled across the duration of the study. This radius is necessarily arbitrary but strikes a good compromise between describing the relevant spatial scale at which vocal interactions occur, which extends up to around 180 metres (Bircher et al., 2021; Blumenrath & Dabelsteen, 2004), and maintaining an adequate sample size in areas of low density. Neighbourhoods defined in this way are highly non-independent, so we model both this methodological spatial dependence and other sources of spatial autocorrelation intrinsic to the study site by including a 2D Gaussian process (GP), which estimates a length-scale parameter defining a variance-covariance matrix for the spatial locations based on their distance (Dearmon & Smith, 2016; Gelfand & Schliep, 2016; Wright et al., 2021). We confirmed that this eliminated the residual spatial autocorrelation via Moran's I tests. Note that we fit a separate GP for each year, as treating the spatial dependence as fixed across the study duration, as is often done, risks further underestimating uncertainty.

We define our predictor variables in the following way: Individual turnover is the proportion of birds that were not already in a neighbourhood in the preceding year ($Ind. Turn. = 1 - \frac{|N_{current} \cap N_{previous}|}{|N_{current}|}$). Dispersal is the mean of the distances that birds in the neighbourhood travelled to get from their natal territories to their breeding territories if they were born within

the Wytham population. Immigration is the proportion of birds that were not ringed as nestlings in the population, and neighbourhood age is the mean age of the birds within it. [Figure S1](#) illustrates that our sampling process did not introduce bias into any of these predictor variables: the birds from which we recorded song repertoires were, on average, representative of the true neighbourhood composition.

Operational definitions of cultural diversity, novelty, and turnover We calculated a simple diversity index by dividing the number of different song types by the total number of songs in a neighbourhood. To calculate the novelty index, we computed the relative frequency of each class label in the current year in the entire population. We then took the mean of these relative frequencies for each song type in the neighbourhood, took the logarithm of the inverse of this proportion and scaled it between 0 and 1. In this way, 'diversity' describes the proportion of unique songs in a neighbourhood, and 'novelty' refers to how uncommon, on average, the songs of the birds in a neighbourhood are. These two ways of characterising cultural diversity are (as expected) anti-correlated in our study site due to the effect of sampling: more frequent songs are sampled more readily, causing larger sample sizes (neighbourhoods with more density and therefore songs) to yield lower average estimates of diversity and higher average estimates of novelty, in a nonlinear manner. Once this is adjusted for, diversity and novelty are positively correlated, as expected (see [Figure S2](#); models nov_{m_2} and $nov_{m_{2,1}}$). All of our models adjust for this sampling effect.

Models To study the effect of dispersal and immigration on local cultural diversity and novelty, we built log-normal models (div_{m_1} , nov_{m_1}) and estimated the marginal effects of the proportion of immigrants, mean dispersal distance, and mean neighbourhood age, while also adjusting for individual turnover, year, and spatial dependence. Lastly, to examine whether the effects of immigration and dispersal on cultural diversity were related to individual differences in repertoire size and novelty, we fit two further models predicting the absolute number of unique songs in a neighbourhood while also adjusting for the number of birds (div_{m_2}) and the number of songs ($div_{m_{2,1}}$).

We calculated the rate of song cultural turnover as the proportion of unique song types in a given year that were not already present in the same neighbourhood the preceding year, and this was the response variable in two models: one ($turn_{m_1}$) trying to estimate the total effect of turnover and a second ($turn_{m_2}$) estimating the marginal effects of the proportion of immigrants, mean dispersal distance, and mean age while also adjusting for individual turnover, year, and spatial dependence. In both cases,

we modelled the response distribution as a truncated log-normal with a hurdle (logistic) part to account for the zeroes.

Model estimates and reporting

We build the models and approximate the posterior distributions of the parameters of interest using brms (Bürkner, 2017), an interface to the Hamiltonian Monte Carlo engine Stan (Stan Development Team, 2023). We then processed the posterior distributions with the help of the marginal effects package. We checked model convergence via the effective number of samples, visual inspection of the chain trace plots, and the Gelman-Rubin diagnostic. Estimation in a Bayesian framework returns a posterior distribution of possible values instead of point estimates. By convention, we report posterior central estimates (means or medians) and their 95% credible intervals, but also include plots with full posteriors. Note that categorical predictors are dummy-coded and continuous predictions z-score transformed.

For each parameter of interest, we calculate predictions or marginal effects at the means or other relevant values. Regression plots show predicted values of the mean and their credible intervals, as well as partial residuals adjusted to the means or other relevant values of the explanatory terms included in the model (Fox & Weisberg, 2018; Larsen & McCleary, 1972). We have tried to build reasonable models, but even then our estimates should not be interpreted causally. See the software section at the end for a complete list of libraries used in the various analyses and the code repository for full model specifications.

ACKNOWLEDGEMENTS

We thank all those who have contributed to the long-term nest box study in Wytham Woods and the collection of associated data. This work was supported by a Clarendon-Mary Frances Wagley Graduate Scholarship and an EGI scholarship to Nilo Merino Recalde, and made use of the University of Oxford Advanced Research Computing facility (Richards, 2015).

AUTHOR CONTRIBUTIONS

Nilo Merino Recalde: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Andrea Estandía:** Investigation, Data Curation, Writing - Review & Editing. **Sara C. Keen:** Writing – Review & Editing. **Ella F. Cole:** Supervision, Project Administration. **Ben C. Sheldon:** Supervision, Project Administration, Writing – Review & Editing, Funding Acquisition.

REFERENCES

- Allen, J., Weinrich, M., Hoppitt, W., & Rendell, L. (2013). Network-Based Diffusion Analysis Reveals Cultural Transmission of Lobtail Feeding in Humpback Whales. *Science*, 340(6131), 485–488. <https://doi.org/10.1126/science.1231976>
- Aplin, L. M., Sheldon, B. C., & McElreath, R. (2017). Conformity does not perpetuate suboptimal traditions in a wild population of songbirds. *Proceedings of the National Academy of Sciences of the United States of America*, 114(30), 7830–7837. <https://doi.org/10.1073/pnas.1621067114>
- Baker, M. C., McGregor, P. K., & Krebs, J. R. (1987). Sexual Response of Female Great Tits to Local and Distant Songs. *Ornis Scandinavica*, 18(3), 186. <https://doi.org/10.2307/3676765>
- Baker, M. C., Thompson, D. B., & Sherman, G. L. (1981). Neighbor/Stranger Song Discrimination in White-Crowned Sparrows. *The Condor*, 83(3), 265–267. <https://doi.org/10.2307/1367322>
- Barta, K. A., Garamszegi, L. Z., Scheuring, I., & Zsebk, S. (2023). Effects of positive frequency-dependent learning, learning mistakes, and immigration on complex culturesValidation on the song of collared flycatcher (*Ficedula albicollis*) by individual-based modeling. *Frontiers in Ecology and Evolution*, 11.
- Beecher, M. D. (2008, January). Chapter 4 Function and Mechanisms of Song Learning in Song Sparrows*. In *Advances in the Study of Behavior* (pp. 167–225, Vol. 38). Academic Press. [https://doi.org/10.1016/S0065-3454\(08\)0004-1](https://doi.org/10.1016/S0065-3454(08)0004-1)
- Berdahl, A. M., Kao, A. B., Flack, A., Westley, P. A., Codling, E. A., Couzin, I. D., Dell, A. I., & Biro, D. (2018). Collective animal navigation and migratory culture: From theoretical models to empirical evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1746), 20170009. <https://doi.org/10.1098/rstb.2017.0009>
- Bircher, N., van Oers, K., & Naguib, M. (2021). Anyone listening? No evidence for eavesdropping on male singing interactions in the great tit, *Parus major*. *Animal Behaviour*, 176, 67–76. <https://doi.org/10.1016/j.anbehav.2021.03.009>
- Blumenrath, S. H., & Dabelsteen, T. (2004). Degradation of Great Tit (*Parus major*) Song before and after Foliation: Implications for Vocal Communication in a Deciduous Forest. *Behaviour*, 141(8), 935–958.
- Brakes, P., Carroll, E. L., Dall, S. R. X., Keith, S. A., McGregor, P. K., Mesnick, S. L., Noad, M. J., Rendell, L., Robbins, M. M., Rutz, C., Thornton, A., Whiten, A., Whiting, M. J., Aplin, L. M., Bearhop, S., Ciucci, P., Fishlock, V., Ford, J. K. B., Notarbartolo di Sciara, G., ... Garland, E. C. (2021). A deepening understanding of animal culture suggests lessons for conservation. *Proceedings of the Royal Society B: Biological Sciences*, 288(1949), 20202718. <https://doi.org/10.1098/rspb.2020.2718>
- Brakes, P., Dall, S. R., Aplin, L. M., Bearhop, S., Carroll, E. L., Ciucci, P., Fishlock, V., Ford, J. K., Garland, E. C., Keith, S. A., McGregor, P. K., Mesnick, S. L., Noad, M. J., Notarbartolo di Sciara, G., Robbins, M. M., Simmonds, M. P., Spina, F., Thornton, A., Wade, P. R., ... Rutz, C. (2019). Animal cultures matter for conservation. *Science*, 363(6431), 1032–1034. <https://doi.org/10.1126/science.aaw3557>
- Bürkner, P.-C. (2017). Brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01>

- Byholm, P., Beal, M., Isaksson, N., Lötberg, U., & Åkesson, S. (2022). Paternal transmission of migration knowledge in a long-distance bird migrant. *Nat Commun*, 13, 1566. <https://doi.org/10.1038/s41467-022-29300-w>
- Chambers, M., & Drovandi, C. (2023, May). Many-levelled continuation ratio models for frequency of alcohol and drug use data.
- Chimento, M., Alarcón-Nieto, G., & Aplin, L. M. (2021). Population turnover facilitates cultural selection for efficiency in birds. *Current Biology*, 31(11), 2477–2483.e3. <https://doi.org/10.1016/j.cub.2021.03.057>
- Claidière, N., & Sperber, D. (2007). Commentary: The role of attraction in cultural evolution. *Journal of Cognition and Culture*, 7(1-2), 89–111. <https://doi.org/10.1163/156853707X171829>
- Claramunt, S., Derryberry, E. P., Remsen, J. V., & Brumfield, R. T. (2011). High dispersal ability inhibits speciation in a continental radiation of passerine birds. *Proceedings of the Royal Society B: Biological Sciences*, 279(1733), 1567–1574. <https://doi.org/10.1098/rspb.2011.1922>
- Crates, R., Langmore, N., Ranjard, L., Stojanovic, D., Rayner, L., Ingwersen, D., & Heinsohn, R. (2021). Loss of vocal culture and fitness costs in a critically endangered songbird. *Proceedings of the Royal Society B: Biological Sciences*, 288(1947), 20210225. <https://doi.org/10.1098/rspb.2021.0225>
- Cresswell, W., & McCleery, R. (2003). How great tits maintain synchronization of their hatch date with food supply in response to long-term variability in temperature. *Journal of Animal Ecology*, 72(2), 356–366. <https://doi.org/10.1046/j.1365-2656.2003.00701.x>
- Dearmon, J., & Smith, T. E. (2016). Gaussian Process Regression and Bayesian Model Averaging: An Alternative Approach to Modeling Spatial Phenomena: Gaussian Process Regression and BMA. *Geogr Anal*, 48(1), 82–111. <https://doi.org/10.1111/gean.12083>
- Deffner, D., Kandler, A., & Fogarty, L. (2022). Effective population size for culturally evolving traits. *PLOS Computational Biology*, 18(4), e1009430. <https://doi.org/10.1371/journal.pcbi.1009430>
- Deffner, D., & McElreath, R. (2020). The importance of life history and population regulation for the evolution of social learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1803), 20190492. <https://doi.org/10.1098/rstb.2019.0492>
- Deffner, D., & McElreath, R. (2022). When does selection favor learning from the old? Social learning in age-structured populations. *PLOS ONE*, 17(4), e0267204. <https://doi.org/10.1371/journal.pone.0267204>
- Derex, M., & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences*, 113(11), 2982–2987. <https://doi.org/10.1073/pnas.1518798113>
- Dhondt, A. A. (1979). Summer dispersal and survival of juvenile Great Tits in southern Sweden. *Oecologia*, 42(2), 139–157. <https://doi.org/10.1007/BF00344854>
- Dingemanse, N. J., Both, C., van Noordwijk, A. J., Rutten, A. L., & Drent, P. J. (2003). Natal dispersal and personalities in great tits (*Parus major*). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1516), 741–747. <https://doi.org/10.1098/rspb.2002.2300>
- Falls, J. B., Krebs, J. R., & McGregor, P. K. (1982). Song matching in the great tit (*Parus major*): The effect of similarity and familiarity. *Animal Behaviour*, 30(4), 997–1009. [https://doi.org/10.1016/S0003-3472\(82\)80188-7](https://doi.org/10.1016/S0003-3472(82)80188-7)
- Falótico, T., Proffitt, T., Ottoni, E. B., Staff, R. A., & Haslam, M. (2019). Three thousand years of wild capuchin stone tool use. *Nature Ecology and Evolution*. <https://doi.org/10.1038/s41559-019-0904-4>
- Fayet, A. L., Tobias, J. A., Hintzen, R. E., & Seddon, N. (2014). Immigration and dispersal are key determinants of cultural diversity in a songbird population. *Behavioral Ecology*, 25(4), 744–753. <https://doi.org/10.1093/beheco/aru047>
- Fogarty, L., Creanza, N., & Feldman, M. W. (2019). The life history of learning: Demographic structure changes cultural outcomes. *PLOS Computational Biology*, 15(4), e1006821. <https://doi.org/10.1371/journal.pcbi.1006821>
- for Environment Food, D., & Affairs, R. (2020). Defra data services platform.
- Fox, J., & Weisberg, S. (2018). Visualizing Fit and Lack of Fit in Complex Regression Models with Predictor Effect Plots and Partial Residuals. *Journal of Statistical Software*, 87, 1–27. <https://doi.org/10.18637/jss.v087.i09>
- Gelfand, A. E., & Schliep, E. M. (2016). Spatial statistics and Gaussian processes: A beautiful marriage. *Spatial Statistics*, 18, 86–104. <https://doi.org/10.1016/j.spasta.2016.03.006>
- Gibb, J. (1950). The Breeding Biology of the Great and Blue Titmice. *Ibis*, 92(4), 507–539. <https://doi.org/10.1111/j.1474-919X.1950.tb01759.x>
- Graham, B. A., Heath, D. D., & Mennill, D. J. (2017). Dispersal influences genetic and acoustic spatial structure for both males and females in a tropical songbird. *Ecology and Evolution*, 7(23), 10089–10102. <https://doi.org/10.1002/ece3.3456>
- Graham, B. A., Heath, D. D., Walter, R. P., & Mennill, D. J. (2018). Immigrant song: Males and females learn songs after dispersal in a tropical bird. *Behavioral Ecology*, 29(3), 711–723. <https://doi.org/10.1093/beheco/ary029>
- Greenwood, P. J., Harvey, P. H., & Perrins, C. M. (1979). The Role of Dispersal in the Great Tit (*Parus major*): The Causes, Consequences and Heritability of Natal Dispersal. *Journal of Animal Ecology*, 48(1), 123–142. <https://doi.org/10.2307/4105>
- Halfwerk, W., Bot, S., & Slabbekoorn, H. (2012). Male great tit song perch selection in response to noise-dependent female feedback. *Functional Ecology*, 26(6), 1339–1347. <https://doi.org/10.1111/j.1365-2435.2012.02018.x>
- Hill, A. P., Prince, P., Snaddon, J. L., Doncaster, C. P., & Rogers, A. (2019). AudioMoth: A low-cost acoustic device for monitoring biodiversity and the environment. *HardwareX*, 6, e00073. <https://doi.org/10.1016/j.hwx.2019.e00073>
- Hinde, R. A. (1952). The Behaviour of the Great Tit (*Parus Major*) and Some Other Related Species. *Behaviour. Supplement*, (2), III–201.
- Hutfluss, A., Bermúdez-Cuamatzin, E., Mouchet, A., Briffa, M., Slabbekoorn, H., & Dingemanse, N. J. (2022). Male song stability shows cross-year repeatability but does not affect reproductive success in a wild passerine bird. *Journal of Animal Ecology*, 91(7), 1507–1520. <https://doi.org/10.1111/1365-2656.13736>
- James, L. S., & Sakata, J. T. (2017). Learning Biases Underlie “Universals” in Avian Vocal Sequencing. *Current Biology*,

- 27(23), 3676–3682. <https://doi.org/10.1016/j.cub.2017.10.019>
- Jesmer, B. R., Merkle, J. A., Goheen, J. R., Aikens, E. O., Beck, J. L., Courtemanch, A. B., Hurley, M. A., McWhirter, D. E., Miyasaki, H. M., Monteith, K. L., & Kauffman, M. J. (2018). Is ungulate migration culturally transmitted? Evidence of social learning from translocated animals. *Science*, 361(6406), 1023–1025. <https://doi.org/10.1126/science.aat0985>
- Kandler, A., Fogarty, L., & Karsdorp, F. (2023). The interplay between age structure and cultural transmission. *PLOS Computational Biology*, 19(7), e1011297. <https://doi.org/10.1371/journal.pcbi.1011297>
- Kandler, A., Wilder, B., & Fortunato, L. (2017). Inferring individual-level processes from population-level patterns in cultural evolution. *Royal Society Open Science*, 4(9), 170949. <https://doi.org/10.1098/rsos.170949>
- Keen, S. C. (2020). The role of social environment in shaping vocal communication systems in wild songbirds. *Cornell University, PhD Thesis*. <https://doi.org/10.7298/qz29-2v92>
- Kendal, R., Hopper, L. M., Whiten, A., Brosnan, S. F., Lambeth, S. P., Schapiro, S. J., & Hoppitt, W. (2015). Chimpanzees copy dominant and knowledgeable individuals: Implications for cultural diversity. *Evolution and Human Behavior*, 36(1), 65–72. <https://doi.org/10.1016/j.evolhumbehav.2014.09.002>
- Kirby, S., & Tamariz, M. (2021). Cumulative cultural evolution, population structure and the origin of combinatoriality in human language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1843), 20200319. <https://doi.org/10.1098/rstb.2020.0319>
- Lachlan, R. F., Anderson, R. C., Peters, S., Searcy, W. A., & Nowicki, S. (2014). Typical versions of learned swamp sparrow song types are more effective signals than are less typical versions. *Proceedings of the Royal Society B: Biological Sciences*, 281(1785). <https://doi.org/10.1098/rspb.2014.0252>
- Lachlan, R. F., Ratmann, O., & Nowicki, S. (2018). Cultural conformity generates extremely stable traditions in bird song. *Nature Communications*, 9(1), 2417. <https://doi.org/10.1038/s41467-018-04728-1>
- Lack, D. (1964). A Long-Term Study of the Great Tit (*Parus major*). *Journal of Animal Ecology*, 33, 159–173. <https://doi.org/10.2307/2437>
- Laland, K. N., & Janik, V. M. (2006). The animal cultures debate. *Trends in Ecology & Evolution*, 21(10), 542–547. <https://doi.org/10.1016/j.tree.2006.06.005>
- Larsen, W. A., & McCleary, S. J. (1972). The Use of Partial Residual Plots in Regression Analysis. *Technometrics*, 14(3), 781–790. <https://doi.org/10.1080/00401706.1972.10488966>
- Lind, H., Dabelsteen, T., & McGregor, P. K. (1996). Female great tits can identify mates by song. *Animal Behaviour*, 52(4), 667–671. <https://doi.org/10.1006/anbe.1996.0211>
- Mace, R. (1987). The dawn chorus in the great tit *Parus major* is directly related to female fertility. *Nature*, 330(6150), 745–746. <https://doi.org/10.1038/330745a0>
- Marler, P., & Peters, S. (1982). Developmental overproduction and selective attrition: New processes in the epigenesis of birdsong. *Dev Psychobiol*, 15(4), 369–378. <https://doi.org/10.1002/dev.420150409>
- McGregor, P. K., & Krebs, J. R. (1982a). Mating and song sharing in the great tit. *Nature*, 297(5861), 60–61. <https://doi.org/10.1038/297060a0>
- McGregor, P. K., & Krebs, J. R. (1982b). Song Types in a Population of Great Tits (*Parus major*): Their Distribution, Abundance and Acquisition by Individuals. *Behaviour*, 79(2/4), 126–152.
- McGregor, P. K., Krebs, J. R., & Perrins, C. M. (1981). Song Repertoires and Lifetime Reproductive Success in the Great Tit (*Parus major*). *The American Naturalist*, 118(2), 149–159. <https://doi.org/918>
- Merino Recalde, N. (2023a). Pykanto: A python library to accelerate research on wild bird song. *Methods in Ecology and Evolution*, 00, 1–9. <https://doi.org/10.1111/2041-210X.14155>
- Merino Recalde, N. (2023b, June). Nilomr/ggeat-tit-hits-setup: V0.1.0. <https://doi.org/10.5281/zenodo.8092397>
- Merino Recalde, N., Estandía, A., Pichot, L., Vansse, A., Cole, E. F., & Sheldon, B. C. (2023). A densely sampled and richly annotated acoustic dataset from a wild bird population. *bioRxiv*. <https://doi.org/10.1101/2023.07.03.547484>
- Morfi, V., Lachlan, R. F., & Stowell, D. (2021). Deep perceptual embeddings for unlabelled animal sound events. *The Journal of the Acoustical Society of America*, 150(1), 2–11. <https://doi.org/10.1121/10.0005475>
- Morgan-Wall, T. (2023). *Rayshader: Create maps and visualize data in 2d and 3d* [<https://www.rayshader.com>, <https://github.com/tylernmorganwall/rayshader>].
- Mortega, K. G., Flinks, H., & Helm, B. (2014). Behavioural response of a migratory songbird to geographic variation in song and morphology. *Frontiers in Zoology*, 11(1), 85. <https://doi.org/10.1186/s12983-014-0085-6>
- Nelson, D. A. (1992). Song overproduction and selective attrition lead to song sharing in the field sparrow (*Spizella pusilla*). *Behav Ecol Sociobiol*, 30(6), 415–424. <https://doi.org/10.1007/BF00176177>
- Nunn, C. L., Thrall, P. H., Bartz, K., Dasgupta, T., & Boesch, C. (2009). Do transmission mechanisms or social systems drive cultural dynamics in socially structured populations? *Animal Behaviour*, 77(6), 1515–1524. <https://doi.org/10.1016/j.anbehav.2009.02.023>
- Papadopoulou, A., Anastasiou, I., Keskin, B., & Vogler, A. P. (2009). Comparative phylogeography of tenebrionid beetles in the Aegean archipelago: The effect of dispersal ability and habitat preference. *Molecular Ecology*, 18(11), 2503–2517. <https://doi.org/10.1111/j.1365-294X.2009.04207.x>
- Paxton, K. L., Sebastián-González, E., Hite, J. M., Crampston, L. H., Kuhn, D., & Hart, P. J. (2019). Loss of cultural song diversity and the convergence of songs in a declining Hawaiian forest bird community. *Royal Society Open Science*, 6(8). <https://doi.org/10.1098/rsos.190719>
- Payne, R. B. (1983). The social context of song mimicry: Song-matching dialects in indigo buntings (*Passerina cyanea*). *Animal Behaviour*, 31(3), 788–805. [https://doi.org/10.1016/S0003-3472\(83\)80236-X](https://doi.org/10.1016/S0003-3472(83)80236-X)
- Perrins, C. M. (1965). Population Fluctuations and Clutch-Size in the Great Tit, *Parus major* L. *Journal of Animal Ecology*, 34(3), 601–647.
- Peters, S., & Nowicki, S. (2017). Overproduction and attrition: The fates of songs memorized during song learning in songbirds. *Animal Behaviour*, 124, 255–261. <https://doi.org/10.1016/j.anbehav.2016.09.019>

- Pike, T. W., & Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biology Letters*, 6(4), 466–468. <https://doi.org/10.1098/rsbl.2009.1014>
- Planqué, R., Britton, N. F., & Slabbekoorn, H. (2014). On the maintenance of bird song dialects. *Journal of Mathematical Biology*, 68(1-2), 505–531. <https://doi.org/10.1007/s00285-012-0632-8>
- Potvin, D. A., & Clegg, S. M. (2015). The relative roles of cultural drift and acoustic adaptation in shaping syllable repertoires of island bird populations change with time since colonization. *Evolution*, 69(2), 368–380. <https://doi.org/10.1111/evo.12573>
- Quinn, J. L., Cole, E. F., Patrick, S. C., & Sheldon, B. C. (2011). Scale and state dependence of the relationship between personality and dispersal in a great tit population. *Journal of Animal Ecology*, 80(5), 918–928.
- Rendell, L., & Whitehead, H. (2001). Culture in whales and dolphins. *Behavioral and Brain Sciences*, 24(2), 309–324. <https://doi.org/10.1017/S0140525X0100396X>
- Richards, A. (2015, August). *University of oxford advanced research computing*. Manual. <https://doi.org/10.5281/zenodo.22558>
- Rivera-Gutierrez, H. F., Matthysen, E., Adriaensen, F., & Slabbekoorn, H. (2010). Repertoire sharing and song similarity between great tit males decline with distance between forest fragments. *Ethology*, 116(10), 951–960. <https://doi.org/10.1111/j.1439-0310.2010.01813.x>
- Rivera-Gutierrez, H. F., Pinxten, R., & Eens, M. (2011). Difficulties when Assessing Birdsong Learning Programmes under Field Conditions: A Re-Evaluation of Song Repertoire Flexibility in the Great Tit (M. Giurfa, Ed.). *PLoS ONE*, 6(1), e16003. <https://doi.org/10.1371/journal.pone.0016003>
- Rivera-Gutierrez, H. F., Pinxten, R., & Eens, M. (2012). Tuning and fading voices in songbirds: Age-dependent changes in two acoustic traits across the life span. *Animal Behaviour*, 83(5), 1279–1283. <https://doi.org/10.1016/j.anbehav.2012.03.001>
- Stan Development Team. (2023). Stan modeling language users guide and reference manual, version 2.18.0. <http://mc-stan.org/>
- Suárez, D., Arribas, P., Jiménez-García, E., & Emerson, B. C. (2022). Dispersal ability and its consequences for population genetic differentiation and diversification. *Proceedings of the Royal Society B: Biological Sciences*, 289(1975), 20220489. <https://doi.org/10.1098/rspb.2022.0489>
- Tchernichovski, O., Eisenberg-Edidin, S., & Jarvis, E. D. (2021). Balanced imitation sustains song culture in zebra finches. *Nat Commun*, 12(1), 2562. <https://doi.org/10.1038/s41467-021-22852-3>
- Tchernichovski, O., Feher, O., Fimiarz, D., & Conley, D. (2017). How social learning adds up to a culture: From birdsong to human public opinion. *Journal of Experimental Biology*, 220(1), 124–132. <https://doi.org/10.1242/jeb.142786>
- Van Duyse, E., Pinxten, R., Snoeijns, T., & Eens, M. (2005). Simultaneous treatment with an aromatase inhibitor and an anti-androgen decreases the likelihood of dawn song in free-living male great tits, *Parus major*. *Hormones and Behavior*, 48(2), 243–251. <https://doi.org/10.1016/j.yhbeh.2005.02.013>
- van Rossum, G. (1995, January). Python reference manual.
- Verhulst, S., Perrins, C. M., & Riddington, R. (1997). Natal Dispersal of Great Tits in a Patchy Environment. *Ecology*, 78(3), 864–872. [https://doi.org/10.1890/0012-9658\(1997\)078\[0864:NDOGTI\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1997)078[0864:NDOGTI]2.0.CO;2)
- Viciano, H. (2021). Animal culture: But of which kind? *Studies in History and Philosophy of Science Part A*, 90, 208–218. <https://doi.org/10.1016/j.shpsa.2021.10.012>
- Warti, R., Kurnia, A., & Sadik, K. (2020). Evaluation of Proportional Odds and Continuation Ratio Models for Smoker in Indonesia. *Proceedings of the 1st International Conference on Statistics and Analytics, ICSA 2019*, 2-3 August 2019, Bogor, Indonesia.
- Whitehead, H., & Lusseau, D. (2012). Animal social networks as substrate for cultural behavioural diversity. *Journal of Theoretical Biology*, 294, 19–28. <https://doi.org/10.1016/j.jtbi.2011.10.025>
- Williams, H., & Lachlan, R. F. (2021). Evidence for cumulative cultural evolution in bird song. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1843), 20200322. <https://doi.org/10.1098/rstb.2020.0322>
- Woodman, J. P., Cole, E. F., Firth, J. A., Perrins, C. M., & Sheldon, B. C. (2023). Disentangling the causes of age-assortative mating in bird populations with contrasting life-history strategies. *Journal of Animal Ecology*, 92(5), 979–990. <https://doi.org/10.1111/1365-2656.13851>
- Wright, W. J., Irvine, K. M., Rodhouse, T. J., & Litt, A. R. (2021). Spatial Gaussian processes improve multi-species occupancy models when range boundaries are uncertain and nonoverlapping. *Ecology and Evolution*, 11(13), 8516–8527. <https://doi.org/10.1002/ece3.7629>
- Zandberg, L., Morfi, V., George, J., Clayton, D. F., Stowell, D., & Lachlan, R. F. (2022, December). Bird song comparison using deep learning trained from avian perceptual judgments. <https://doi.org/10.1101/2022.12.23.521425>

SUPPLEMENTARY INFORMATION**Table S1.** Model information

Model	Formula	Family	N	Groups
Individual Repertoires				
rep_m_1	repertoire_size ~ 1 + immigrant + s(sampling_effort) + year + (1 father)	cratio	301	242
rep_m_1.1	repertoire_size ~ 1 + dispersal_distance + s(sampling_effort) + year + (1 father)	cratio	133	105
rep_m_1.2	repertoire_size ~ 1 + age + s(sampling_effort) + year + (1 father)	cratio	256	205
repnov_m_1	average_frequency ~ 1 + immigrant + s(sampling_effort) + year + (1 father)	lognormal	300	242
repnov_m_1.1	average_frequency ~ 1 + dispersal_distance + s(sampling_effort) + year + (1 father)	lognormal	133	105
repnov_m_1.2	average_frequency ~ 1 + age + s(sampling_effort) + year + (1 father)	lognormal	256	205
Cultural Similarity				
disp_m_1	mean_dist1 ~ 0 + natal_distance + nest_distance + year_born_diff + year + (1 mm(father, father2))	Gaussian	8745	105
imm_m_1	mean_dist1 ~ 0 + resident_status + year_born_diff + nest_distance + year + (1 mm(father, father2))	Gaussian	11029	205
age_m_1	mean_dist1 ~ 0 + natal_distance + nest_distance + year_born_diff + year + (1 mm(father, father2))	Gaussian	8745	105
Cultural Novelty and Diversity				
nov_m_1	novelty ~ 0 + prop_immigrant + mean_dispersal_distance + prop_same_birds + mean_age + s(n_current_songs) + year + gp(x, y, by = year, k = 25, c = 5/4, scale = F)	lognormal	791	GP
div_m_1	diversity ~ 0 + prop_immigrant + mean_dispersal_distance + prop_same_birds + mean_age + s(n_current_songs) + year + gp(x, y, by = year, k = 25, c = 5/4, scale = F)	lognormal	791	GP
nov_m_2	novelty ~ 0 + diversity + s(n_current_songs) + year + gp(x, y, by = year, k = 25, c = 5/4, scale = F)	lognormal	791	GP
nov_m_2.1	novelty ~ 0 + diversity + year + gp(x, y, by = year, k = 25, c = 5/4, scale = F)	lognormal	791	GP
div_m_2	n_unique_current_songs ~ 0 + prop_immigrant + mean_dispersal_distance + prop_same_birds + mean_age + s(recorded) + year + gp(x, y, by = year, k = 25, c = 5/4, scale = F)	Gaussian	791	GP
div_m_2.1	n_unique_current_songs ~ 0 + prop_immigrant + mean_dispersal_distance + prop_same_birds + mean_age + s(n_current_songs) + year + gp(x, y, by = year, k = 25, c = 5/4, scale = F)	Gaussian	791	GP
Cultural Turnover				
turn_m_1	prop_shared ~ 0 + prop_same_birds + year + gp(x, y, by = year, k = 25, c = 5/4)	hurdle lognormal	544	GP
turn_m_2	prop_shared ~ 0 + prop_immigrant + mean_dispersal_distance + mean_age + prop_same_birds + year + gp(x, y, by = year, k = 25, c = 5/4)	hurdle lognormal	544	GP

Table S2. Model estimates

Model	Hypothesis	Estimate ^a	Evid. Ratio	Post. Prob
Individual Repertoires				
rep_m_1	immigrant > 0	0.239 [-0.098, 0.593]	6.963	0.874
rep_m_1.1	dispersal distance < 0	-0.201 [-0.443, 0.045]	10.111	0.910
rep_m_1.2	age > 0	0.064 [-0.108, 0.241]	2.701	0.730
repnov_m_1	non-immigrant < 0	-0.049 [-0.2, 0.1]	2.401	0.706
repnov_m_1.1	dispersal distance > 0	0.203 [0.088, 0.316]	741.857	0.999
repnov_m_1.2	age > 0	-0.017 [-0.093, 0.058]	0.540	0.351
Cultural Similarity				
disp_m_1	natal distance > 0	0.001 [0, 0.002]	22.529	0.958
disp_m_1	nest distance > 0	-0.005 [-0.006, -0.004]	Inf	1
imm_m_1	both resident-both immigrant < 0	0.002 [-0.005, 0.01]	0.438	0.304
imm_m_1	both resident-one immigrant < 0	0.002 [-0.002, 0.006]	0.331	0.248
age_m_1	age difference 0-1 > 0	0.002 [0, 0.003]	12.289	0.925
age_m_1	age difference 0-2 > 0	0.004 [0.002, 0.006]	Inf	1
age_m_1	age difference 0-3 > 0	0.006 [0.003, 0.008]	1999	1
age_m_1	age difference 0-4+ > 0	0.01 [0.006, 0.013]	Inf	1
Cultural Novelty and Diversity				
nov_m_1	mean dispersal distance < 0	-0.018 [-0.023, -0.012]	Inf	1
nov_m_1	proportion immigrant < 0	0.001 [-0.005, 0.006]	0.752	0.429
nov_m_1	mean age < 0	0.012 [0.005, 0.019]	399	0.998
nov_m_1	individual turnover < 0	0.001 [-0.005, 0.008]	1.733	0.634
div_m_1	mean dispersal distance < 0	-0.005 [-0.01, 0]	26.65	0.964
div_m_1	proportion immigrant < 0	0.002 [-0.004, 0.007]	0.442	0.306
div_m_1	mean age < 0	0.021 [0.014, 0.027]	Inf	1
div_m_1	individual turnover < 0	-0.002 [-0.008, 0.005]	0.446	0.309
nov_m_2	diversity > 0	0.713 [0.629, 0.795]	Inf	1
nov_m_2.1	diversity > 0	-0.099 [-0.216, 0.018]	0.086	0.080
div_m_2	mean dispersal distance < 0	-0.658 [-0.999, -0.315]	570.429	0.998
div_m_2	proportion immigrant < 0	0.469 [0.1, 0.837]	62.492	0.984
div_m_2	mean age < 0	1.045 [0.576, 1.495]	Inf	1
div_m_2	individual turnover < 0	0.204 [-0.291, 0.683]	3.077	0.755
div_m_2.1	mean dispersal distance < 0	0.019 [-0.213, 0.249]	0.824	0.452
div_m_2.1	proportion immigrant < 0	0.072 [-0.168, 0.312]	2.259	0.693
div_m_2.1	mean age < 0	0.928 [0.599, 1.245]	Inf	1
div_m_2.1	individual turnover < 0	-0.031 [-0.349, 0.279]	0.726	0.421
Cultural Turnover				
turn_m_1	individual turnover > 0	0.072 [0.054, 0.09]	Inf	1
turn_m_2	mean dispersal distance < 0	-0.022 [-0.039, -0.006]	79	0.988
turn_m_2	proportion immigrant < 0	-0.051 [-0.065, -0.037]	Inf	1
turn_m_2	mean age < 0	-0.044 [-0.063, -0.026]	3999	1
turn_m_2	individual turnover < 0	0.047 [0.028, 0.066]	Inf	1

^a Estimates are Medians and 95% Credible Intervals

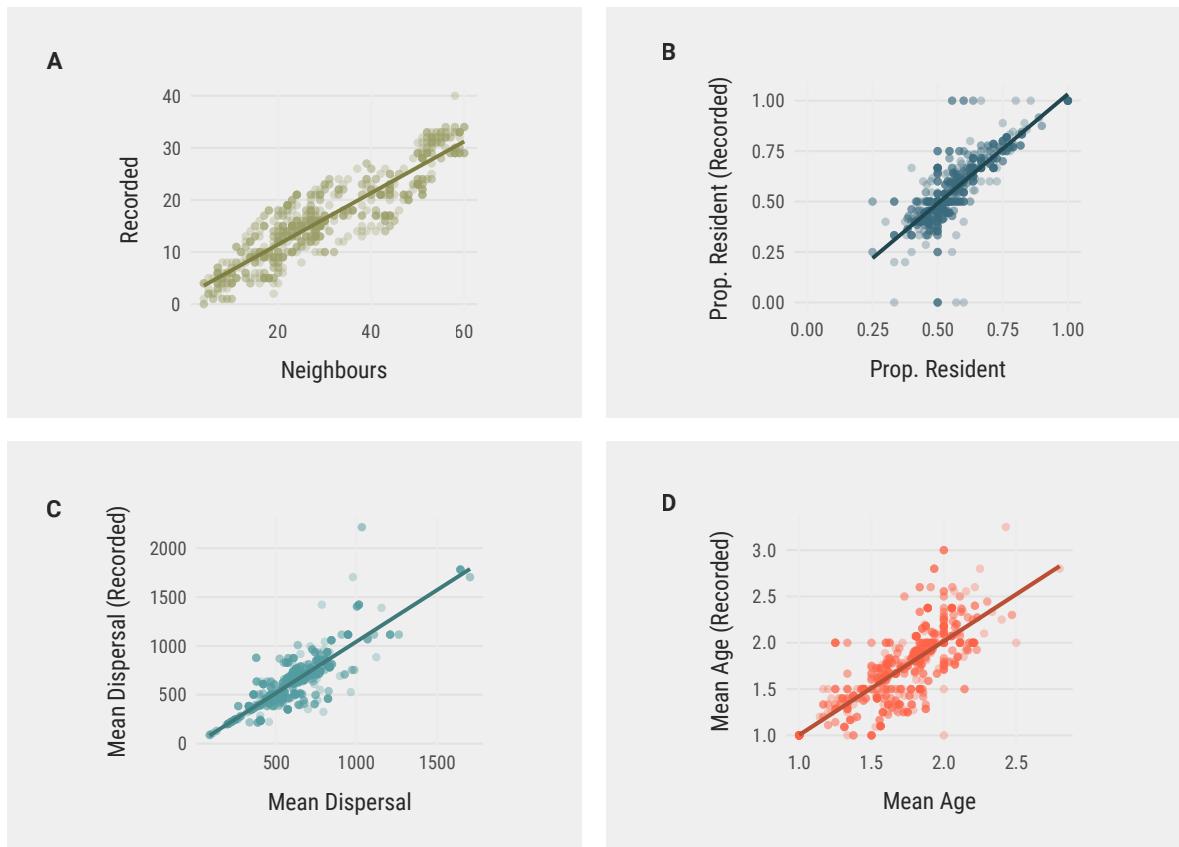


Figure S1. Absence of bias in the sampling of neighbourhood properties.

Correlation between the actual neighbourhood properties and neighbourhood properties estimated from birds for which we have song recordings. (A) Neighbourhood size (number of individuals) and number of individuals with song recordings. (B) Proportion of resident birds from monitoring data and only those birds with song recordings. (C) Mean dispersal distance calculated from birds born in the study site and only those birds born in the study site with song recordings. (D) Mean age of birds in the study site and only those birds with song recordings.

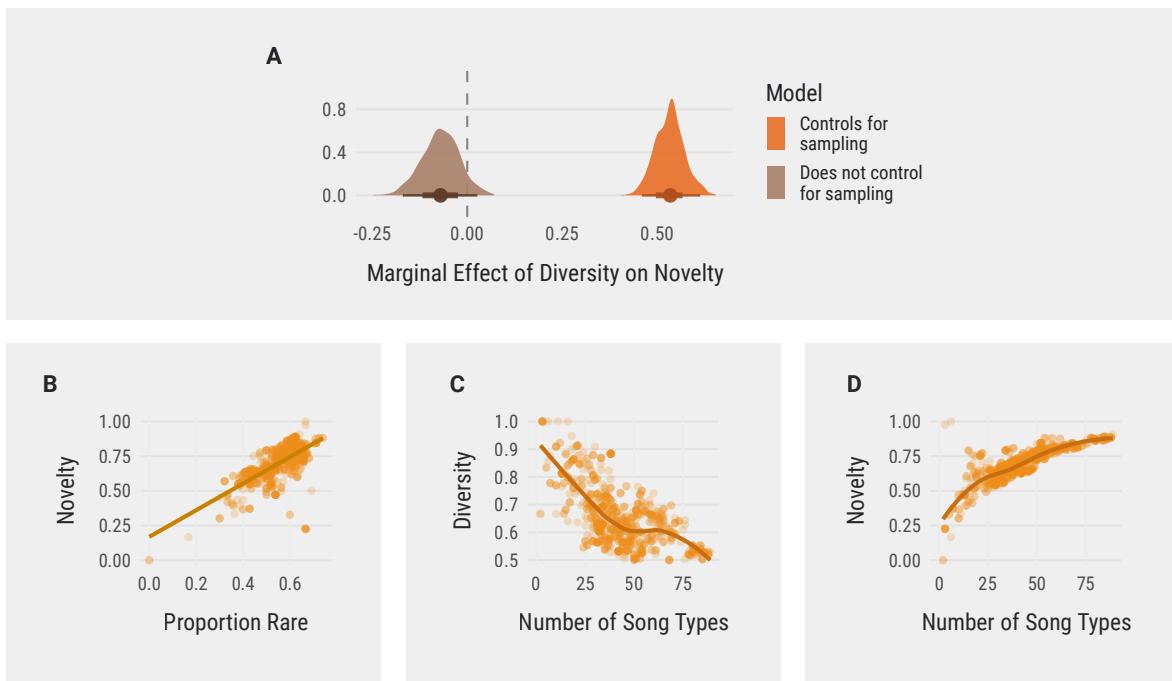


Figure S2. Relationships among outcome variables and sampling effects.

(A) Marginal effect of diversity—which describes the proportion of unique songs in a neighbourhood—on novelty, that is, how uncommon, on average, the songs of the birds in a neighbourhood are. These two ways of characterising cultural diversity are (as expected) anti-correlated in our study site due to the effect of sampling: more frequent songs are sampled more readily, causing larger sample sizes (neighbourhoods with more birds and therefore songs) to yield lower average estimates of diversity (C) and higher average estimates of novelty (D), in a nonlinear manner. Once this is adjusted for, diversity and novelty are positively correlated, as expected. (B) Our measure of cultural novelty (y-axis) has the advantages of being continuous and not using an arbitrary cutoff, but is nonetheless correlated with definitions traditionally used in the literature, such as ‘songs shared by fewer than 4 birds’ (McGregor & Krebs, 1982b)

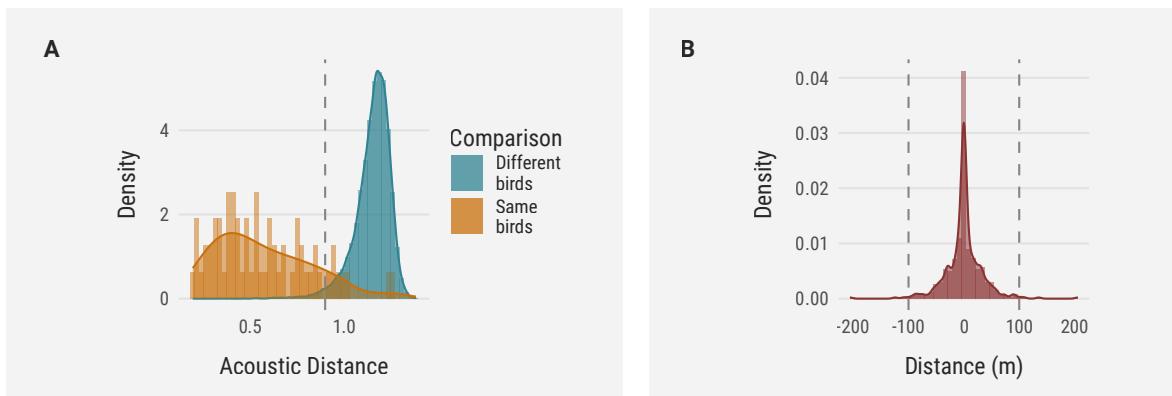


Figure S3. Thresholds used during the process of reidentifying individual birds based on their songs.

(A) Shows the distribution of the acoustic distances between the same song type sung by the same known bird in different years, in orange, and the minimum pairwise distance between different birds and years. The x-intercept of the vertical line = 0.9. (B) Shows the distribution of the change in distance from the natal nest box to the breeding site in different years for birds that bred more than once. Adult birds have high nest site fidelity, which we used as a further constraint when reidentifying them from their songs.

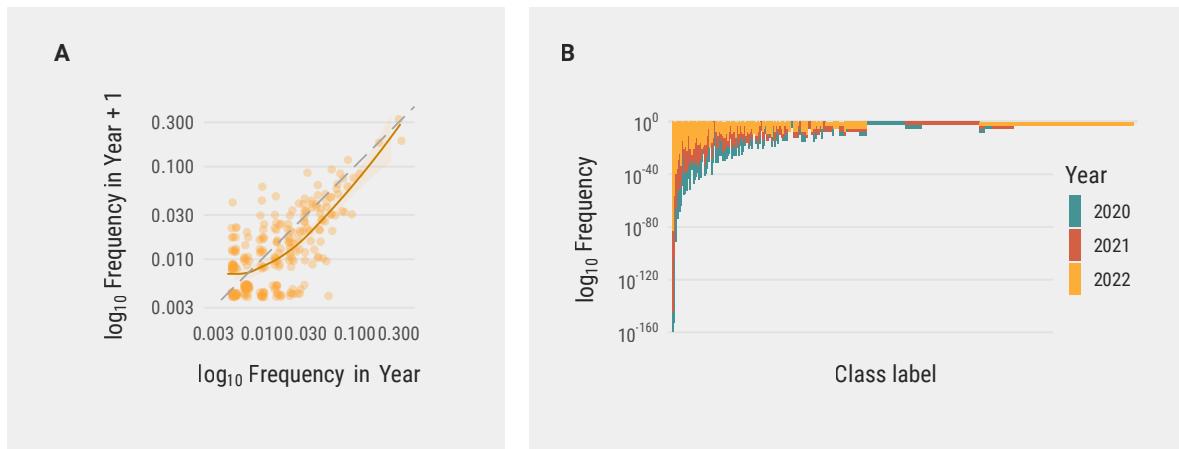


Figure S4. Song frequencies and their relationship with abundance in the following year.

(A) The abundance of a song type in a given year predicts its abundance in the following year, with higher variance around rare songs.
 (B) Histogram showing the frequency of individual song types in the study.

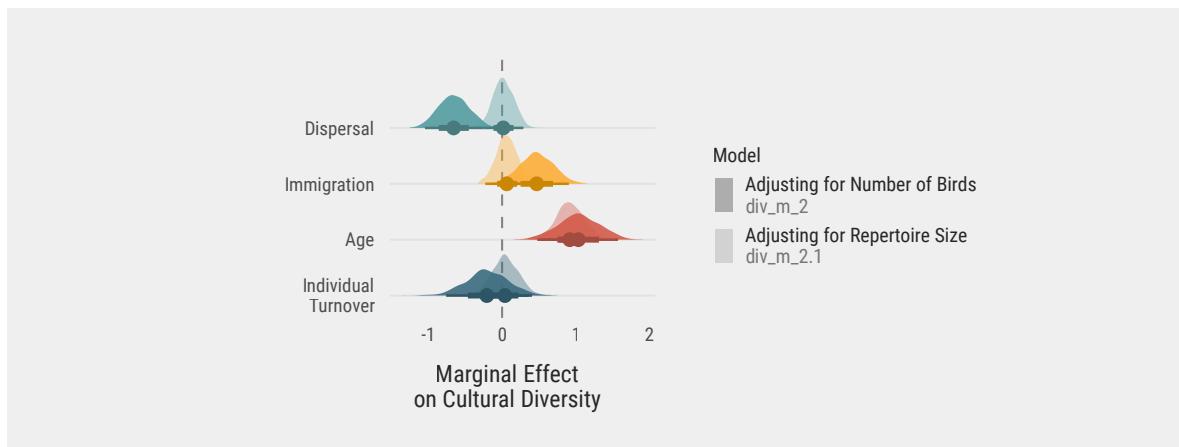


Figure S5. Marginal effects of demographic variables on absolute cultural diversity.

Marginal effects of our predictor variables on absolute cultural diversity (the number of different song types sampled in a neighbourhood), while adjusting for the effect of either number of individuals (higher opacity fill, corresponding to model div_m_2) or number of song variants, including repeated variants (lower opacity fill, $div_m_2.1$).

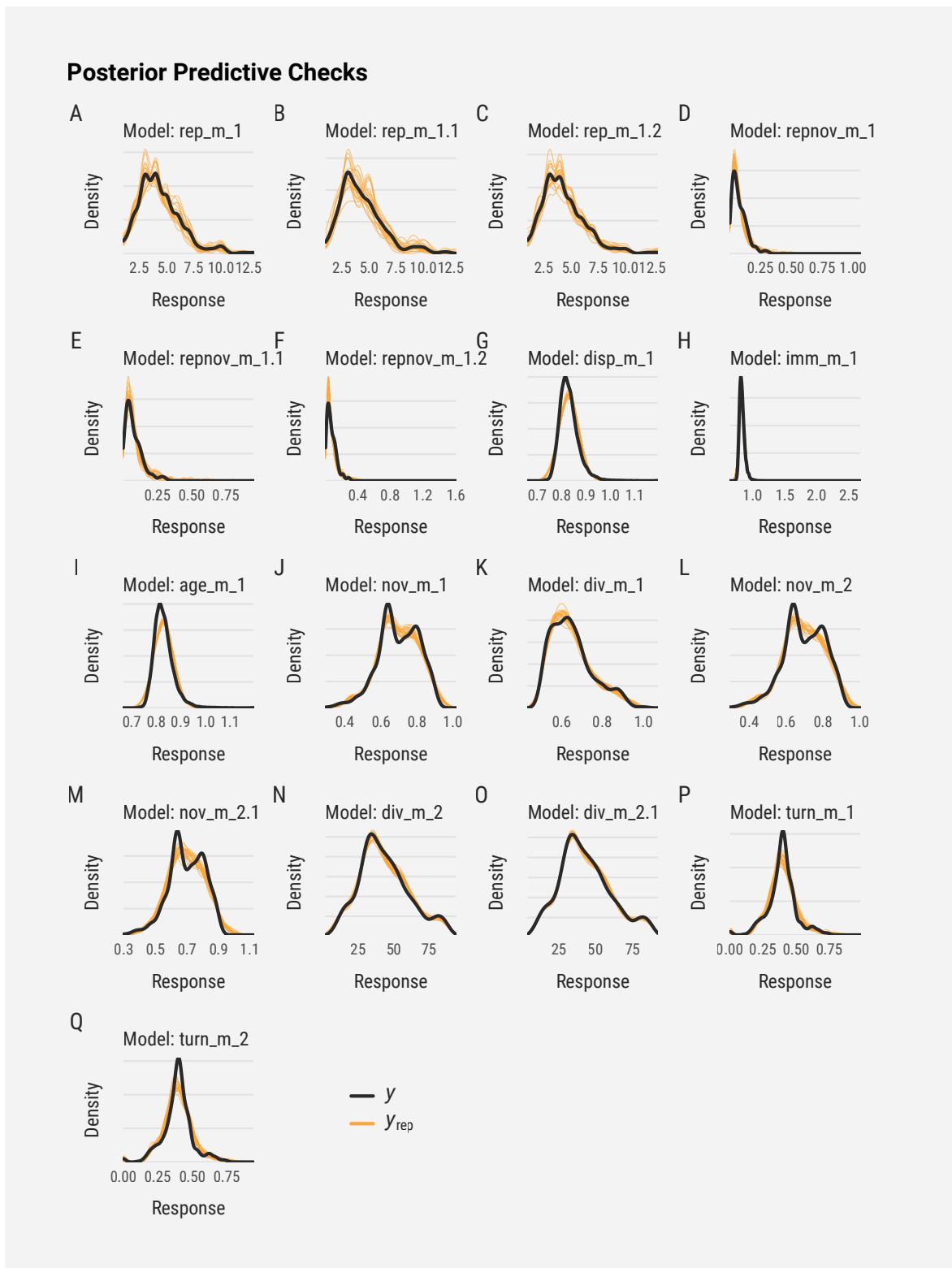


Figure S6. Posterior predictive checks for the main models in the study.

Comparing simulations from the posterior predictive distribution y^{rep} (thin orange lines) with the outcome y (black lines) using Kernel density estimates. The posterior predictive distribution is a distribution of possible outcomes of the model given the data and the model parameters, here used to check the fit of the model to the data.