

Modeling how population size drives the evolution of birdsong, a functional cultural trait

Emily J. Hudson^{1,2}  and Nicole Creanza^{1,2,3} 

¹Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37240

²Evolutionary Studies Initiative, Vanderbilt University, Nashville, Tennessee 37240

³E-mail: nicole.creanza@vanderbilt.edu

Received July 15, 2021

Accepted March 9, 2022

Oscine songbirds have been an important study system for social learning, particularly because their learned songs provide an analog for human languages and music. Here, we propose a different analogy: from an evolutionary perspective, could birds' songs change over time more like arrowheads than arias? Small improvements to a bird's song can lead to large fitness differences for its singer, which could make songs more analogous to human tools than languages. We modify a model of human tool evolution to accommodate cultural evolution of birdsong: each song learner chooses the most skilled available tutor to emulate, and each is more likely to produce an inferior copy than a superior one. Similar to human tool evolution, our model suggests that larger populations of birds could foster improvements in song over time, even when learners restrict their pool of tutors to a subset of individuals in their social network. We also demonstrate that song elements could be simplified instead of lost after population bottlenecks if lower quality traits are easier to imitate than higher quality ones. We show that these processes could plausibly generate empirically observed patterns of song evolution for some song traits, and we make predictions about the types of song elements most likely to be lost when populations shrink. More broadly, we aim to connect the modeling approaches used in human and nonhuman systems, moving toward a cohesive theoretical framework that accounts for both cognitive and demographic processes.

Social learning—"learning that is facilitated by observation of, or interaction with, another individual or its products" (Hoppitt and Laland 2013)—is pervasive across the animal kingdom, occurring in diverse taxa including invertebrates, cetaceans, and primates (Reiss and McCowan 1993; Whiten et al. 1999; Kacsoh et al. 2018). Social learning is a special case of inheritance that provides nongenetic pathways for adaptive traits to be passed from one generation to the next. Particularly when learned traits affect reproductive behaviors, social learning is thought to influence genetic divergence and ultimately speciation (Lachlan and Servedio 2004; Yeh and Servedio 2015; Yeh et al. 2018), underscoring its evolutionary importance. The discovery of genes as the primary unit of inheritance meant that Darwin's ideas about natural selection could be synthesized with an underlying mechanism of transmission; a similar synthesis between genetics, evolutionary processes, and broad behavioral patterns is still

needed for learned traits. Mathematical models have already helped generate hypotheses, test assumptions, and, in combination with experiments investigating the mechanisms of learning, advance our understanding of learning processes and how they influence the evolution of behaviors (Lachlan and Servedio 2004; Creanza et al. 2017; Yeh 2019). A better theoretical understanding of the processes underlying the cultural evolution of learned traits could allow us to predict cultural shifts over generations.

The oscine songbirds offer a promising system for modeling the interaction of evolution and culture; like humans, songbirds must undergo a learning process to produce effective vocal communications. However, the evolution of birdsong is hypothesized to be shaped by selection pressures, such as the acoustic transmission properties of the environment (Tobias et al. 2010; Weir et al. 2012; Moseley et al. 2018), more so than human language (although the role of humidity in the evolution of tonal languages

has been debated, e.g., Everett et al. 2015; Collins 2016). Most notably, much research on birdsong has focused on sexual selection, including female preferences for song elaboration or vocal performance (reviewed in Gil and Gahr 2002; Soma and Garamszegi 2011; Wilkins et al. 2013). These preference-shaped song features are, to varying degrees, socially transmitted and constrained by the availability of tutors. For example, syllable-level complexity in songs enhances attractiveness in some bird species such as canaries (Vallet et al. 1998; Nagle et al. 2004), and in house finches there is evidence that complex syllables are preferentially learned (Youngblood and Lahti 2022). However, because songs are learned with very high fidelity in house finches (Youngblood and Lahti 2022), even from heterospecifics (Mann et al. 2021), pupils must have access to tutors with complex syllables to produce these attractive elements themselves. In this scenario, pupils with a limited number of tutors may produce less attractive songs, even if a selection pressure (female preference) still exists. Returning to an analogy with human culture, then, subtle variations in a bird's song can potentially be linked to fitness differences, much like the skill to craft a sharp arrowhead can translate to increased fitness for a human hunter. In both cases, a pupil's learned skill depends on its social environment.

Here, we propose that modeling the evolution of birdsong in the same way that we model historical changes in human functional toolkits could be a novel and complementary approach to analogizing birdsong to human language. The process of learning a language differs from the process of learning to make a tool, and these differences affect the evolutionary dynamics of these cultural traits; we suggest that in the same way, some aspects of birdsong learning might be better modeled by human tool learning. We focus in particular on the skill level that characterizes both singing and toolmaking (not on the physical objects produced by toolmaking), because both toolmaking and birdsongs are learned skills that have fitness consequences, and both may experience cumulative cultural change toward a more effective version. In contrast, the analogy between birdsong and human language could be more limited in its ability to predict evolutionary patterns, because small pronunciation differences in human language are unlikely to lead to large changes in biological fitness for the speaker. Of course, one analogy cannot capture the nuances of how learned song evolves in the thousands of living songbird species. However, by emphasizing the functional aspects of birdsong rather than its potential similarity to learned vocalizations in humans, we gain new methods to understand its evolutionary dynamics; namely, how demographic factors might affect song over time.

In this study, we use a cultural evolutionary approach that has been applied to social learning and human tool evolution (Henrich 2004) to make predictions about song evolution. In essence, we represent song as a difficult skill that all learners in

the population attempt to reproduce, with varying success. This simplified but powerful framework has yielded novel insights into human behavior (e.g., Powell et al. 2009), and it holds promise to do the same for cultural traits in other species. In particular, this approach provides a specific mechanism to explain why attractive elements of song are sometimes lost when bird species expand into new habitat and experience an associated population bottleneck. Unlike cultural drift, which can be dominated by random processes, the model we propose here incorporates cultural selection—the assumption that certain forms of a cultural trait are beneficial or preferable—while integrating population demographics and specific aspects of cognitive ability (Rogers et al. 2009; Steele et al. 2010). It is important to note that in our proposed model, we separate the success of a song from other traits of the male that sings it; a song variant may successfully spread throughout a population via learning without requiring any concurrent biological selection in which certain males have differential reproductive success. After demonstrating the basic features of this model, we apply the tools of social network analysis to show how manipulating population connectivity affects trait evolution (as in Carja and Creanza 2019; Derex et al. 2019). Finally, we model the trajectory of song evolution if the difficulty of learning a trait is not constant but instead varies with the elaboration of the trait; for example, if less attractive songs are easier to learn and transmit than more attractive ones, or if learners instinctively produce songs with certain characteristics even in the absence of tutor input.

Basic Model

In the tradition of anthropologists and cultural evolutionary theorists (Cavalli-Sforza and Feldman 1981; Boyd and Richerson 1988; reviewed in Creanza et al. 2017), Henrich (2004) applied mathematical models to cultural change, showing that the progressive loss of complex skills observed across several thousand years of the Tasmanian archeological record was best explained as an outcome of demographic factors (namely, effective population size). In Henrich's mathematical model, a learned trait (such as a netweaving technique, or the straightness of an arrow shaft; Henrich 2004, p. 200) gradually improves or declines in skill level (z) across generations in a culture, depending on the interaction of three factors: (1) how difficult the trait is to accurately copy (α), (2) how widely learners vary in their attempts to copy the trait (β), and (3) how many individuals exist in the current population (N), a noncognitive factor that nevertheless plays an important role in cultural change. Henrich describes the relationship of these terms in a modified version of the Price equation:

$$\Delta \bar{z} = -\alpha + \beta (\epsilon + \ln(N)). \quad (1)$$

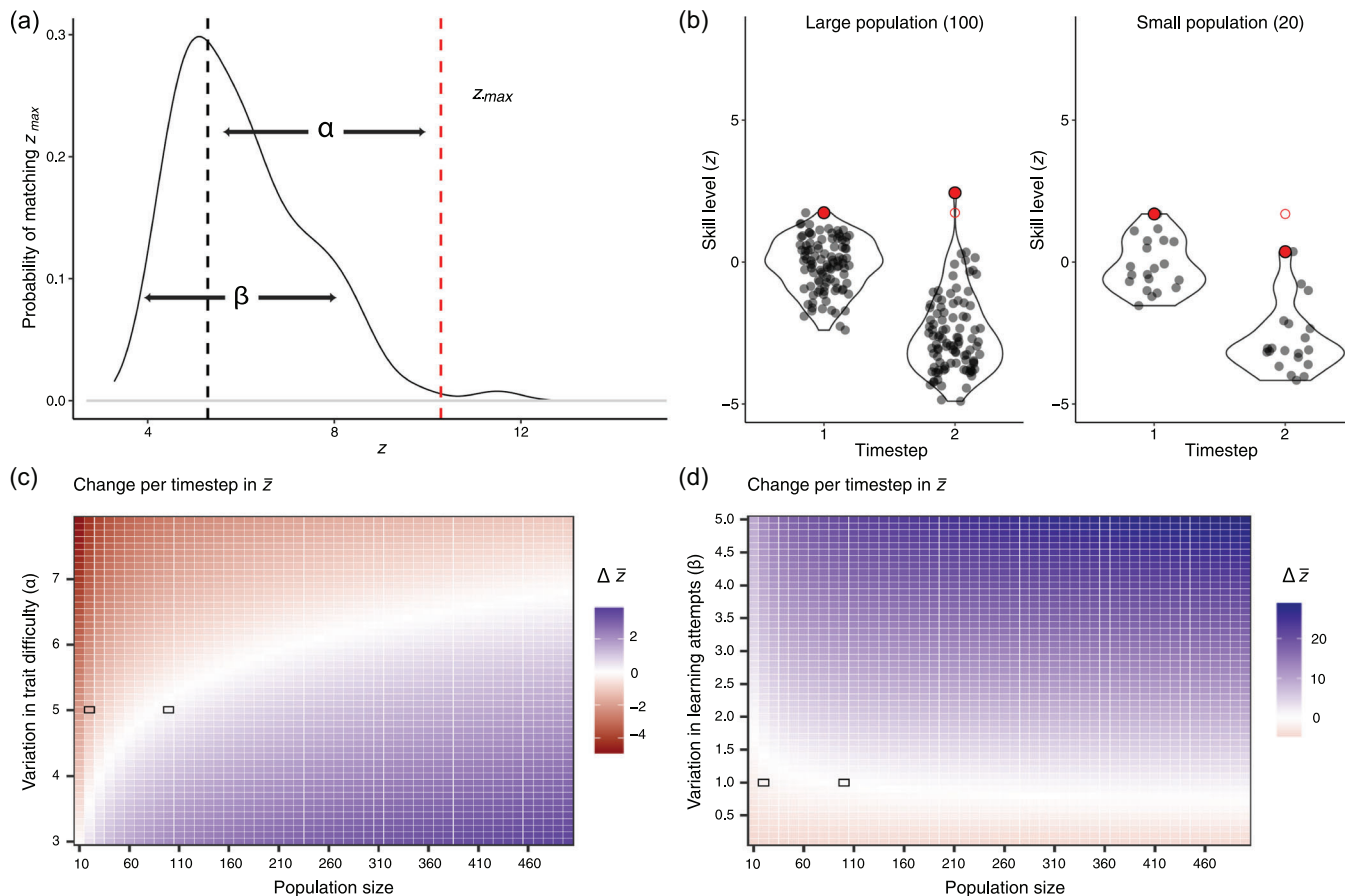


Figure 1. (a) The Gumbel distribution of learning attempts in each generation. Given one individual with the highest skill trait (z_{max} , red dashed line), most individuals attempting to copy this tutor will fall short, such that the mean of the population of learners occurs at $z_{max} - \alpha$. In a large enough population, however, one or more individuals will exceed the tutor’s skill level, increasing z_{max} for the next generation. (b) Here, we show an example with a population of 100 individuals and one with 20 individuals experiencing one learning timestep, or generation, where $\alpha = 5$ and $\beta = 1$ for both (these parameter values are highlighted with black boxes in the heatmaps in panels c and d). The second generation learns from the first generation, attempting to copy the most skilled individual (z_{max} , filled red circle on the left). The open circle shows z_{max} of the previous generation. The value of z_{max} in the larger population increases in generation two (even though most individuals are still lower than the initial z_{max}), whereas the z_{max} of the smaller population decreases. (c) The effect of population size (N) and trait difficulty (α) on the change in mean trait skill level per generation ($\Delta \bar{z}$), where $\beta = 1$. When population size is small, the mean change in learned skill is negative each generation (red), even for lower values of α (easier traits). As population size increases, cultural traits that are increasingly difficult to copy (higher α) can be maintained and improved (blue, positive $\Delta \bar{z}$). (d) In contrast to α (fixed here at 5), increasing β (the variation in learning attempts) leads to overall increases in $\Delta \bar{z}$ over time.

Here, $\Delta \bar{z}$ is the population-wide mean change per timestep in the skill level z of a culturally transmitted behavioral trait, and ϵ is Euler’s constant (≈ 0.577). This equation represents a model of social learning where all individuals attempt to imitate a trait produced by the most competent individual (highest z) in the previous generation, but most ultimately produce an inferior version of the trait. Population size N is important because larger populations have more individuals, increasing the likelihood that at least one individual produces a superior imitation. It is important to note that there is no direct inheritance in this model, genetic or otherwise; each pupil attempts to imitate the best tutor (z_{max}) in the population, and in each generation the resulting z value

for each pupil is drawn from a Gumbel distribution centered at $z_{max} - \alpha$ (Fig. 1a). Additionally, and perhaps counterintuitively, the population variation in skill (z) of available tutors does not influence trait change in the next generation in this model; only the most highly skilled tutor (z_{max}) is chosen by pupils as a model to imitate.

In Henrich’s model, most individuals will produce inferior imitations of their tutor, and thus only sufficiently large populations will contain enough pupils to consistently produce improved imitations by chance, thereby increasing the maximum trait skill level (z_{max}) in the next generation (Fig. 1b). With all other factors being constant, increasing trait difficulty (α) will

tend to reduce $\Delta\bar{z}$ until it is negative, eventually leading to the loss of the trait when $\bar{z} < 0$ (Fig. 1c). In contrast, greater variation in learning attempts (β) will increase $\Delta\bar{z}$ and lead to improved trait skill over time (Fig. 1d). This process differs from drift in that traits that decline in skill level in a population over time are not random, but rather tend to be the most difficult traits to imitate (highest α , Henrich 2004). Although the application of this model to human cultural evolution has been widespread, if contentious (e.g., Read 2011; Henrich et al. 2016; Vaesen et al. 2016), to our knowledge it has never been applied to nonhuman cultures.

Applying the Model to Birdsong: Methods and Results

To apply the above model to birdsong, it is necessary to translate the concept of an individual's skill level along a continuous axis (z) to a measurable feature of song, where a higher value of z is loosely defined as more attractive or otherwise beneficial to the singer. The song traits that appear to influence reproductive success vary across taxa (reviewed in Byers and Kroodsma 2009), and include rapidly trilled notes (Drăgănoiu et al. 2002; Ballentine 2004; Caro et al. 2010; Wilkins et al. 2015), large repertoire size (Lambrechts and Dhondt 1986; Verheyen et al. 1991; Hasselquist et al. 1996), and the production of certain elements (Vallet and Kreuzer 1995; Rehsteiner et al. 1998; Riebel and Slater 1998; Williams et al. 2013). In most oscine birds, this concept can be expressed as a song that is more attractive to the receiver (usually, a prospective female mate), or more effective at displacing competitors (usually other males). Not all components of birdsong are learned, and indeed, many of the examples listed above have strong genetic underpinnings. However, for the predictions of the Price equation to apply, it is only necessary for some degree of social inheritance to take place; song traits with partial genetic control can still be modeled if the distribution of learner performance is affected by the availability of skilled tutors. We use a modified version of the Price equation to explore learned song traits with genetic constraints below. It is also important to note that, although a better song improves a male's chances of mating, the process of song evolution we model here happens completely independently of any selection on male traits outside of song, and no genetic evolution is incorporated in the model. A final important assumption of Henrich's model is that, due to the incentive to produce high-quality imitations of these important traits, learners will choose the best tutor in their population to imitate. This strict assumption may not be true in all species, but there is evidence that learners are selective about what songs to imitate. For example, in swamp sparrows, arti-

cially slowed renditions of natural songs (<40% of the natural speed) are not imitated (Lahti et al. 2011), whereas males given only accelerated tutor songs will attempt to imitate higher trill rates than they are capable of producing, resulting in atypical syntax with a few trilled notes punctuated by longer silences (Podos et al. 1999). These results suggest a directional preference for copying faster trills. In another species of sparrow, artificial tutor songs with a single note removed were much less likely to be imitated than complete tutor songs, suggesting a different type of qualitative criteria for selecting attractive songs to imitate (Soha and Marler 2000). Finally, young male zebra finches whose fathers had small repertoire sizes were more likely to supplement their learning by imitating an additional tutor (Soma et al. 2009). In addition to judging their songs, juvenile birds might also gauge potential tutors using social cues (Clayton 1987; Payne and Payne 1993). This is somewhat analogous to cultural processes of prestige bias and success bias in humans: in addition to judging tool quality (z), humans can also select tutors based on social factors (Henrich and Gil-White 2001; Baldini 2012) without invalidating Henrich's model of cumulative cultural change. Finally, the fact that oblique learning seems to be prevalent in songbirds (e.g., Baptista and Morton 1988; Williams 1990; Liu and Nottebohm 2007) is an important prerequisite for the comparison of multiple possible tutors by learners.

Interpreting the fundamental principles of the Price equation in a birdsong context, where the skill level, z , represents a generic metric of song quality, we find that the mean skill level of the population decreases over time (i.e., $\Delta\bar{z} < 0$) when populations are small (lower N), or when traits are more difficult to imitate (higher α). In this scenario, a decrease in mean skill level would manifest as a song trait degrading over generations and eventually disappearing altogether (Fig. 1c, blue region). On the other hand, the mean song skill level increases over time ($\Delta\bar{z} > 0$) when population sizes are larger or when traits are easier to imitate (Fig. 1c, red region). The trade-off between population size (N) and trait difficulty (α) is shown by the white boundary in the heatmap, where $\Delta\bar{z} \approx 0$ and the skill level remains consistent over generations. Likewise, increasing the learning variance (β) tends to promote an increase in mean skill level (\bar{z}) over time for a given population size (Fig. 1d); this pattern can be thought of as a decrease in conformity, a feature that characterizes some wild songbird populations (e.g., Lachlan et al. 2018). Based on the pattern shown in Figure 1, we identified the following parameters to use in subsequent simulations, unless otherwise stated: a population of size 100, which is realistic and not computationally prohibitive, values of α between 3 and 5, and β set to 1, which together represent a boundary at which minor parameter changes lead to changes in trait skill level (Fig. 1c,d).

APPLICATION TO SPECIFIC BIRDSONG CHARACTERISTICS

So far, z has referred to any generic song trait that is under selection. Next, we apply a variation of this model to specific features of learned song. One aspect of birdsong that lends itself well to the Price equation model of cultural evolution is trill rate. This continuously varying song feature is constrained by morphological limits (Podos 1997), but also can be limited by the quality of the tutor song available. For example, swamp sparrows presented with only artificially slowed songs could not produce songs reaching natural performance levels, although they did improve on their “tutor’s” performance (Lahti et al. 2011), and male splendid fairy wrens inherit trill rates from their social, rather than genetic, fathers (Greig et al. 2012). Moreover, as in the human cultural traits Henrich modeled, trill rate is functionally important; in this case, for both mate attraction (Vallet and Kreutzer 1995; Ballentine 2004; Caro et al. 2010) and territory defense (Illes et al. 2006; de Kort et al. 2009; Phillips and Derryberry 2017). Despite the functional importance of trill rate, however, several bird species show geographic variation in the rate or presence of trills (Shizuka et al. 2016; Wilkins et al. 2018; Searfoss et al. 2020), raising the question of why these seemingly important features of song are present (or more elaborated) in some populations and not others.

In Figure 2a, we show 20 stochastic replicates of possible trait trajectories over 60 generations in a population of $N = 100$ individuals, starting from a uniform arbitrary population trill rate of 4 notes per second ($z = 4$), and assuming that higher trill rates are preferred by receivers. In the majority of replicates, the mean z of the population decreased but remained positive; the mean of all replicates after 60 generations was a trill rate of 2.8 notes/s. Although the precise values are in some sense arbitrary (100 birds may be adequate to maintain most learned vocal behaviors in nature), this result suggests that, at these cognitive parameters ($\alpha = 5$ and $\beta = 1$), a population greater than 100 individuals is required to maintain a trill rate at the starting value of 4 notes/s. It is important to note that for many learned birdsong traits, physical or physiological limits constrain song expression; for example, beak size determines how rapidly a male can trill (Podos 1996, 2001), whereas neurophysiology places limits on auditory sensitivity (Konishi 1970; Prather et al. 2012). Regardless of this upper limit, in our basic model, smaller populations experience a decline in trill rate over successive generations as learners largely fail to exceed the skill level of the best tutor of their generation. Thus, over time a trilled song element may disappear altogether, as the rate becomes so low as to no longer function in mate attraction or territory defense. Indeed, one replicate in this simulation reached a mean $z < 0$ before 60 generations (Fig. 2a, red line), which we suggest represents the loss of the trait.

Even for song characteristics that are more difficult to quantify than trill rate, this model makes useful predictions about the direction of song evolution over time. For example, we can envision a bird’s song composed of multiple syllables that vary in their difficulty to learn. Like trill rate, syllable and song complexity have been shown to play a role in mate attraction for some species (although the strength and ubiquity of this effect is contested; see meta-analyses by Soma and Garamszegi 2011; Robinson and Creanza 2019). For instance, female Bengalese finches prefer more complex songs, but stressful conditions early in life limit a male’s syntactical complexity, indicating that these songs are more difficult to produce (in the terms of our model, large α) (Soma et al. 2006).

Our proposed model offers predictions that differ from those of neutral cultural processes (akin to genetic drift) in that the syllable types that are lost are not random. Rather, we would expect that more difficult-to-imitate syllables would be disproportionately lost if population size is reduced. In the event of demographic changes such as population bottlenecks, our model predicts that syllables with a high trait difficulty (larger α) would be the most likely to disappear from the population, whereas those that are simpler to learn (smaller α) would be maintained even after population size shrinks. Importantly, this prediction holds even if producing the difficult syllables remains beneficial to the singer.

We illustrate an example of selective syllable loss in Figure 2b, with a hypothetical birdsong initially composed of three different syllable types, each with its own imitation difficulty (α). We show the results in a single population over 60 generations, in which each syllable is learned independently following the Price equation. Over time, the terminal syllable (which is the most difficult to imitate) rapidly decreases in z , whereas the middle syllable remains mostly the same and the initial syllables become more attractive, that is, increase in z (Fig. 2b). Note that the exact interpretation of the z value in this context is more complex than in the case of the trill. One way to conceptualize a syllable with a larger z is that the syllable is more salient or attractive to receivers (e.g., one that is faster, more stereotyped, or more complex). In contrast, a syllable with a lower z may be less salient to receivers. Following Henrich (2004), we interpret syllables with skill levels that eventually become negative ($\bar{z} \leq 0$) as disappearing from the population. This example illustrates how repertoire size is another attractive feature of song that could be negatively affected by small population size. Although repertoire size per se may not be a completely culturally transmitted trait, the availability of diverse models to copy necessarily limits the achieved repertoire size of imitative learners. This phenomenon was seen in marsh wrens given a small pool of tutors from which to learn: juveniles presented with small tutor repertoires eventually produced much smaller repertoires than conspecifics

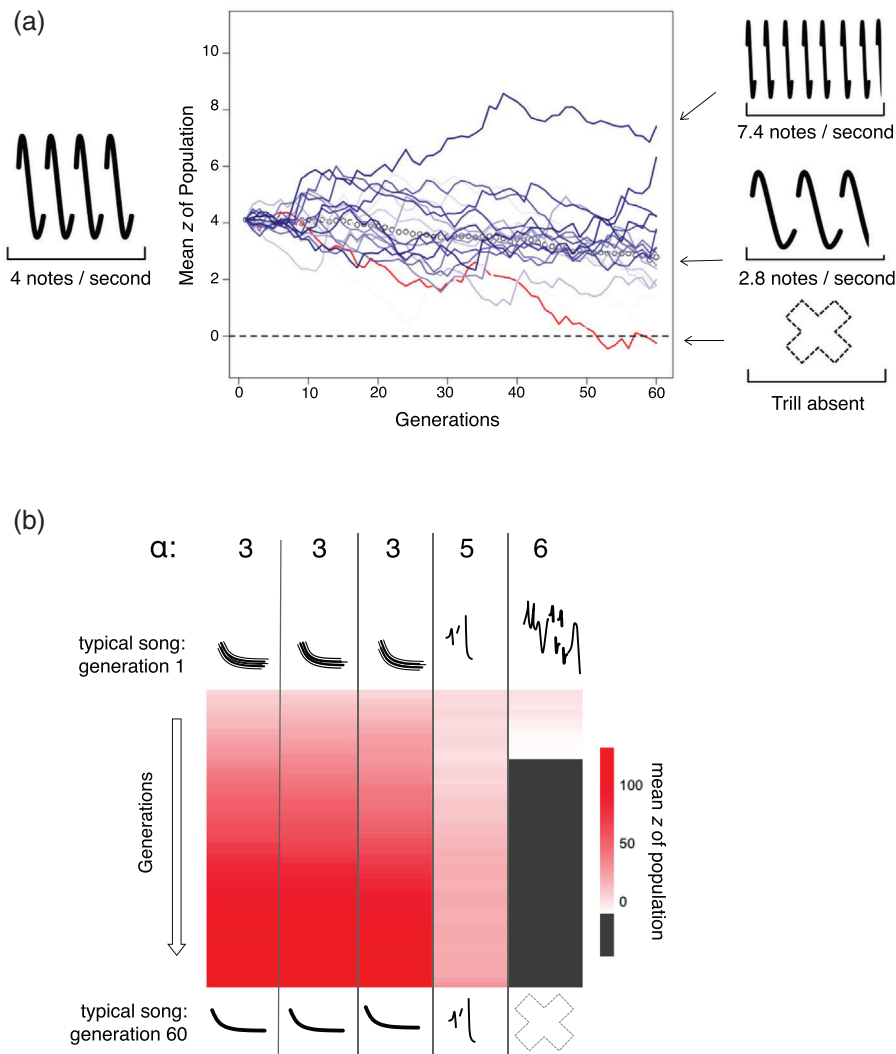


Figure 2. In panel (a), skill level z represents the trill rate (number of notes repeated per second). Lines show the mean \bar{z} of each replicate population over time. Open circles represent the mean \bar{z} value across all 20 replicates for that generation. The range of final z values (maximum, mean, and minimum) after 60 generations are noted on the right y-axis. For all replicates, $\alpha = 5$, $\beta = 1$, and $N = 100$. (b) An ancestral song composed of 5 syllables of varying difficulty (α) is learned by a population of 100 individuals over 60 generations. The mean population value for z for each syllable is indicated by color over time, with red indicating higher values of \bar{z} . In the context of birdsong evolution, $\bar{z} \leq 0$ can be thought of as a syllable disappearing from a population's repertoire. This model shows that a population of 100 is more likely to maintain song components below some threshold degree of difficulty (here, the threshold is $5 < \alpha < 6$).

presented with large tutor repertoires (Brenowitz et al. 1995). In natural populations, individual repertoires could become much smaller or more homogenous gradually over many generations (as shown in Fig. 2b), or more quickly in the case of habitat fragmentation or founder events.

Song Learning in a Social Network

Henrich's application of the Price equation necessarily simplifies the nature of social interaction in a population. A more sophisticated model would include an estimation of the size of both the overall population and the number of individuals accessible

for copying by a pupil in each generation. Kobayashi and Aoki (2012) devised such a model, in which learners are randomly assigned k individuals out of the total population as potential cultural tutors, from which they choose the best tutor (highest z) among those k individuals. Below, we extend this model in two ways to apply social network methodology to the question of how birdsong might evolve in structured populations.

DISENTANGLING POPULATION SIZE (N) FROM CONNECTIVITY OF A NETWORK (DEGREE)

A key characteristic of social networks is degree, or the number of connections possessed by each member (node) of the network,

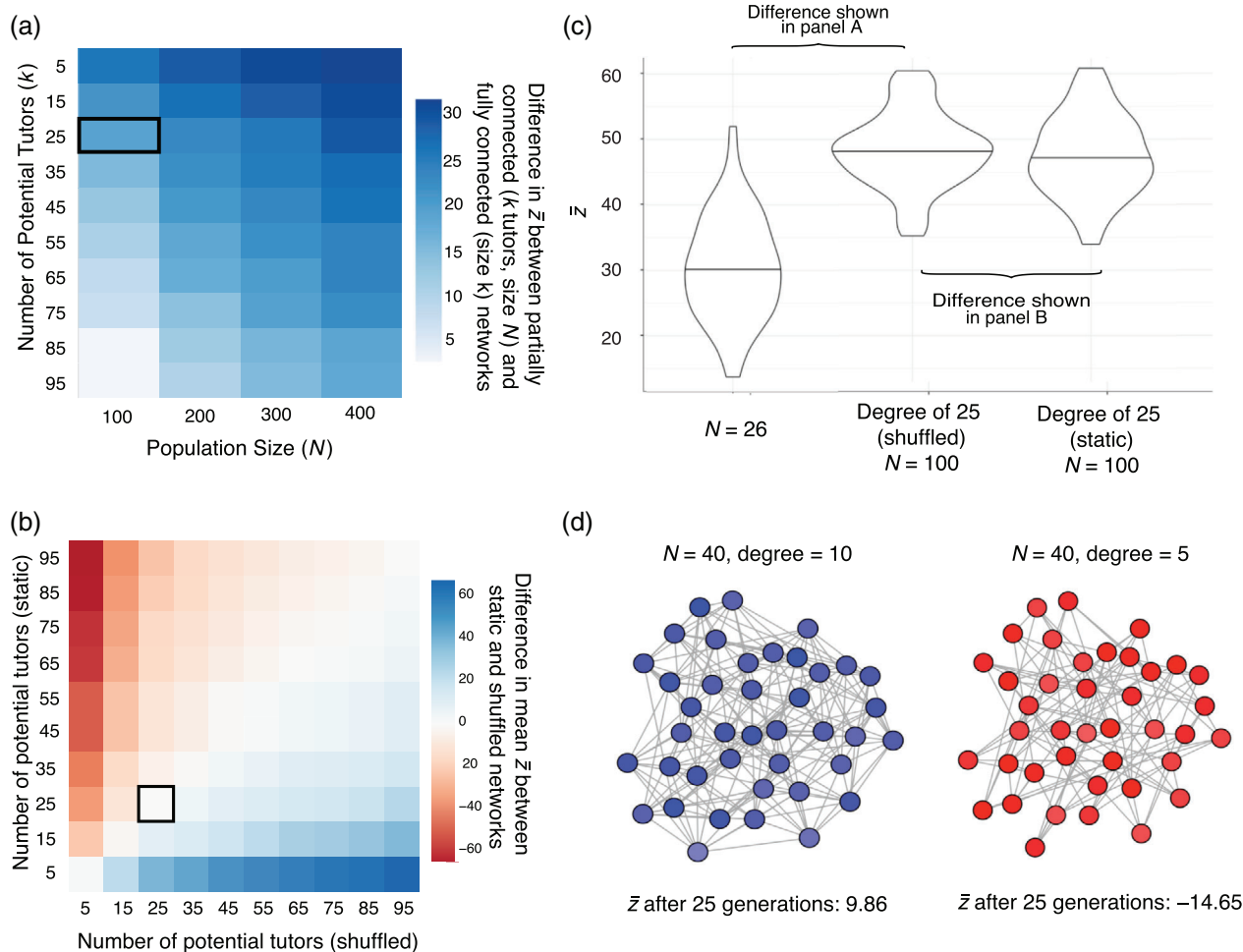


Figure 3. (a) The color scale shows the difference in skill level when the tutor pool is sampled from a larger population: that is, the final mean \bar{z} of a population after 50 generations with the total N and k on the x- and y-axes, respectively, minus the final mean \bar{z} of a population with the same parameters but a total N equal to k . Values shown are the mean of 20 replicates, the black box highlights the comparison in panel c, and $\alpha = 3$ and $\beta = 1$. (b) The color scale shows the difference in skill level when the tutor pool is shuffled versus static: that is, the mean \bar{z} of networks with a k between 5 and 95 where k potential tutors are sampled randomly from the network each generation, minus the mean \bar{z} where the same k connections are maintained across generations, for populations of $N = 100$ after 50 generations. Values shown are the mean of 20 replicates, $\alpha = 3$ and $\beta = 1$. The black box highlights the comparison shown in panel c (right). (c) Violin plots show the mean \bar{z} (horizontal bar) in networks based on 100 simulations run for 50 generations each where $\alpha = 3$ and $\beta = 1$. From left to right, the plots show the means of a fully connected 26-member network with a uniform degree of 25; a 100-member network with a uniform degree of 25, where tutors are shuffled every generation; and a 100-member network with 25 static connections. (d) Example networks of 40 with a uniform degree of 10 (left) and 5 (right), $\alpha = 3$ and $\beta = 1$. Networks with intermediate steps are illustrated in Figure S2.

as illustrated in Figures 3d and S2. To tease apart the impact of local social network dynamics and overall population size on song evolution, we modeled two different network scenarios. First, we simulated a network with $N = 100$ in which every individual was connected to 25 others in the population (degree of 25). Initially, all individuals are assigned $z = 1$. Thus, in the first timestep of our model, individuals randomly select an identical tutor from the 25 connected individuals. In all subsequent timesteps, each learner selects the individual with the highest skill level (z) of its 25 connections to be its tutor. The likelihood that a pupil exceeds

the tutor's z value is defined by α and β , as before. Although our model is not spatially explicit, this scenario might approximate, for example, territorial bird species in which a son inherits a territory from his father, and neighboring males from their fathers (Woolfenden and Fitzpatrick 1978; Komdeur and Edelaar 2001; Suh et al. 2020), such that the connections to potential tutors in neighboring territories are maintained over time.

Second, we compared the above network to a fully connected network in a smaller population with $N = 26$, such that each individual still has 25 connections, but they are to all other

individuals in the population. We found that the larger network with a population of 100 and degree of 25 consistently achieves a higher mean z value after 50 generations than a network of 26 fully connected individuals (Fig. 3a, black box, 3c), showing that it is indeed easier to maintain traits in a larger population than a smaller one, even if individuals in both populations have the same number of potential tutors. This “benefit” increases when the subset k is small, which might occur in territorial birds that sample only a few neighboring tutors (dark-colored region, Fig. 3a); the benefit of being a member of a larger population decreases as the number of connections to potential tutors (k) approaches the size of the whole population, which might occur in colony-breeding species (light-colored region, Fig. 3a). In addition to population and subset size, the outcome of the simulation depends on parameters α and β ; traits are more likely to be lost at larger α and smaller β values (Fig. S1), as well as at smaller degrees of connectivity (k).

THE EFFECT OF RANDOMIZED VERSUS SPATIALLY CONSISTENT TUTORS

Next, we investigated two different scenarios in which only a subset of individuals in the network were available as potential tutors. In the first, we created a 100-member network and ran simulations exactly as in the previous section, with each individual connected to a subset (k) of between 5 and 95 others. Again as before, these individual connections were retained for the entire run (here, 50 generations). At each generation, each learner chooses the highest z among its connected nodes and, based on that maximum value, generates its new z value according to the Price equation. We replicated this simulation 20 times for 50 generations each.

To elucidate the effects of more dynamic network connections, we also simulated 100-member networks where each learner’s k potential tutors were randomly drawn from the population in each generation. In contrast to the previous scenario, the k individuals (between 5 and 95) were assigned randomly for each individual during each generation. From this random pool of k tutors, learners still choose the individual with the highest z to copy, as before. We hypothesized that redrawing connections between individuals every generation might enhance trait improvement relative to the static connection method above, for example, by preventing pockets of lower quality song from developing around certain nodes. However, we found that networks with static versus resampled tutor connections yield a very similar mean z after 100 generations, as illustrated in Figure 3b (light region along diagonal). This suggests that whether the connections between nodes are static (e.g., if sons inherit their fathers’ territories), or shuffled each generation, does not meaningfully affect the population skill level.

The Relationship between Trait Quality and Difficulty: Dynamic α as a Function of z_{\max}

Following Henrich (2004), we have so far assumed that difficulty (α) for a given trait remains constant over time, regardless of the tutor’s skill level (z_{\max}) or average skill level (\bar{z}) of that trait in the population. However, it is plausible that if the population-wide skill level of a trait decreases (e.g., trill rate decreases or syllables become less structurally complex), the most skilled tutor’s trait (z_{\max}) will also decrease, and learners will find that trait easier to successfully imitate. Conversely, as traits increase in skill level, they may become increasingly difficult to imitate. Thus, we propose a modification of our model to reflect a positive association between α and z : as the skill level of the trait increases in tutors (higher z_{\max} values), the trait difficulty should also increase (higher α), whereas decreasing z_{\max} reduces α .

A related aspect of songbird behavior is the fact that some song traits are controlled in part by innate mechanisms, such that even in the absence of skilled (or any) tutors, the traits still persist in the population. For example, in certain species, individuals still produce a form of species-typical vocalizations when raised in the lab without tutors (Marler 1970a), or improve their performance above artificially poor tutors (Brenowitz et al. 1995). Similarly, in zebra finch populations founded by an isolated bird (who had no tutors and thus sang an aberrant song), pupils in each subsequent generation imitated tutors, but song variations tended to accumulate in the direction of species-typical songs (Fehér et al. 2009). In our revised model, this situation could be approximated by not only varying α contingent on z_{\max} , but also assigning α a negative value, meaning that most learners will improve on the tutor’s song, when the maximum z in a population is below a certain threshold. Specifically, in this section, we model a scenario where α changes according to the maximum z value in the population (z_{\max}). After each generation, trait difficulty (α) is set to one of three predetermined values, with α_0 at very low values of z_{\max} , α_1 at intermediate (species-typical) values of z_{\max} , and α_2 at very high values of z_{\max} . In the first version of this variable α model, all values of α are positive ($0 < \alpha_0 < \alpha_1 < \alpha_2$). In the second version, the lowest value of α is negative, such that $\alpha_0 < 0 < \alpha_1 < \alpha_2$. A negative α means that the average attempt by learners *exceeds* the tutor’s skill level (z_{\max}), the reverse of the usual situation modeled by the Price equation; this modification reflects the empirical observation that birds can innately perform some important vocalizations even in the absence of a skilled tutor.

Incorporating this plausible association between z and α produced two culturally and biologically relevant patterns. First, when α increases with z , mean trait performance improves until reaching a plateau, rather than an unrealistic linear increase ad infinitum as seen in Henrich’s original model (Fig. S3a). In other

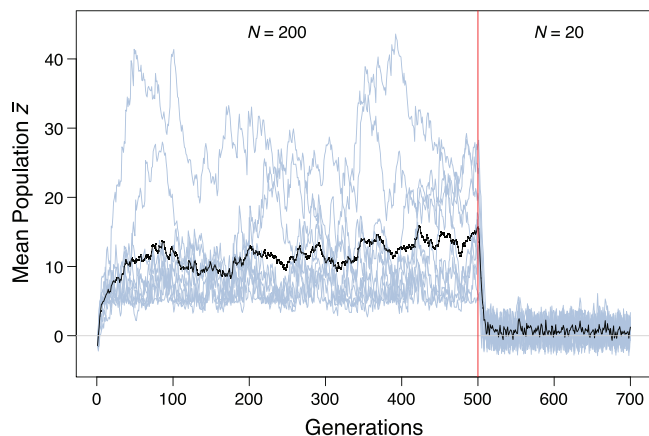


Figure 4. In these 10 simulations, α varies between -1 and 6 , changing each generation depending on the maximum value of z . When $z_{\max} > 10$, $\alpha = 6$; when $2 > z_{\max} > 10$, $\alpha = 5$; and when $z_{\max} < 2$, $\alpha = -1$. For all generations, $\beta = 1$, and the population shrinks from $N = 200$ to $N = 20$ at generation 500 (red line). Blue lines represent the population \bar{z} at each generation for each simulation; the black line represents the mean z across the 10 simulations.

words, as the maximum skill level of a trait improves over generations in a population, α also increases and the trait becomes more difficult to learn, acting as a check on the unbounded increase in trait quality seen in the original model. Second, simplified traits may be maintained in small populations, including after drastic decreases in population size, rather than being lost altogether (Fig. 4, after red line). Thus, as trait quality (z) declines in a small population below a threshold skill value, learners will find it easier to match or exceed tutors, and the trait can be rescued from loss. This concords with the relatively stable song traditions observed in many oscines, even in small or isolated populations. We found qualitatively similar results when varying α linearly with z instead of limiting α to three possible values (Fig. S4).

Discussion

Analogies between research fields can be useful in sparking the application of novel techniques to longstanding problems and encouraging interdisciplinary thinking. It is thus worth periodically revisiting the scientific analogies we use most and considering new and complementary ways of thinking about well-studied traits. Birdsong has often been compared to human language (Doupe and Kuhl 1999; Marler 1970b) and music (Oikkonen et al. 2016), both symbolic traits; these analogies have no doubt encouraged and shaped the many productive decades of birdsong research in animal behavior, neuroscience, and psychology. We argue here that framing birdsong as a functional tool is likely to spur further fruitful research. We draw an analogy between birdsong and human technologies—both of which

are fitness-enhancing, culturally transmitted traits—and leverage models from cultural evolutionary theory to shed new light on the evolution of animal behavior. We have shown, using a relatively simple model based on the Price equation, that the size and social network structure of a population can influence the retention of song traits over time. Population size and network structure in turn interact with cognitive parameters, such as the difficulty of copying a certain song feature and the variance of copying attempts, to shape song evolution. The interaction of demographic and cognitive parameters, we argue, is likely to affect the evolution of signals in oscine songbirds (and potentially other bird clades that experience vocal learning, such as parrots and hummingbirds) in similar ways to their effects on the evolution of human functional traits; specifically, populations that are small or sparsely connected should tend to lose their most difficult-to-copy traits over time, even if these traits continue to be advantageous to their bearers.

In general, the modified Price equation model we describe here is most applicable to relatively simple traits that vary along a single continuous axis, but there are many other types of learned signals that influence reproductive success in birds. For example, repertoire size is an important and multidimensional trait posited in some species to influence both female mate choice (Hasselquist et al. 1996; Gil and Gahr 2002) and intrasexual competition in the form of neighbor matching (Payne 1982). Although in some ways, a large repertoire size would be expected to behave as a difficult-to-learn trait according to our proposed model, and thus to correlate positively with population size, empirical support for this correlation is mixed. Importantly, evidence for the prevalence of socially learned repertoire size is lacking, although there is evidence that repertoire size is more innately influenced in some species (e.g., catbirds [Kroodsma et al. 1997], swamp sparrows and song sparrows [Marler and Sherman 1985]) than others (e.g., marsh wrens [Brenowitz et al. 1995], Java sparrows [Lewis et al. 2021], and great tits [McGregor et al. 1981; Johannessen et al. 2006]). Species in which individuals have larger repertoires may also tend to be species with open-ended learning periods (Robinson et al. 2019); this complicates predictions based on the Price equation, because learning events would likely occur repeatedly and the effective population size of tutors must be calculated across a learner's entire lifetime. In addition, repertoires of the same size might not require equal skill levels to learn; the syllables that comprise the repertoire could be more or less elaborate or complex. Another complicating factor in the relationship between population size and repertoire size (as with all song traits discussed here) is the potential disparity between total population size and the number of immediately available tutors; for example, song sparrows in Québec sang larger syllable repertoires in denser populations where presumably more tutors were available (Harris and Lemon 1972). Finally, birds that engage in neighbor

matching might choose which vocalizations to produce at a given time based on their own skill relative to competitors, rather than always performing the song that requires the highest skill (Logue and Forstmeier 2008). Thus, the population distribution of song skill levels heard by a pupil may be influenced by the dynamics of one-on-one competitions in ways that do not strictly match the predictions we lay out here.

Only in species where there is both a preference for large repertoires and, critically, where learning has strong effects on repertoire size (i.e., not in species where repertoire size is mostly innate, or improvisation is common) would we predict individual repertoire size to behave according to the Price equation. Moreover, these species would likely have to undergo a drastic population bottleneck to see an effect on repertoire size, because in most cases a few tutors with nonoverlapping repertoires would be sufficient to provide a large repertoire for pupils to learn. Accordingly, it is not surprising that the empirical evidence for an association between population size and repertoire size is mixed. For example, a comparison of 49 island and mainland species pairs did not find a consistent decrease in complexity in island species (Morinay et al. 2013). By contrast, in species where population bottlenecks are associated with decreases in repertoire size (Laiolo and Tella 2007) or number of unique syllables as part of a broader metric of song complexity (Crates et al. 2021), repertoire size may be a more learning-dependent feature of song. In other words, according to the predictions we outline here, the extent to which average repertoire size in island populations behaves according to the Price equation likely depends not only on experiencing drastic fluctuations in population size but on the importance of social learning in the species in question. Whether such scenarios exist and how common they are is, at this point, speculative. In any case, we might only expect to see an effect of population size or connectedness on repertoire size in particular under certain circumstances, which, although rare, are still evolutionarily important (e.g., habitat fragmentation, founder events, and/or populations where tutors have largely overlapping repertoires).

Studies finding associations between song features and population bottlenecks are hard to replicate, as they rely on infrequent, large-scale phylogeographic events, which can in turn be accompanied by morphological and other confounding changes (Podos 2001); however, studies conducted on island birds have shown rapid changes in particular song elements (Baker et al. 2003; Parker et al. 2012) and in the associated recognition behavior (Parker et al. 2010), shedding light on the types of song traits that are most likely to be affected by demography. Some common patterns emerge; for example, in the comparison of 49 pairs of mainland and island species mentioned previously, island species were less likely to sing rattles or buzzes (Morinay et al. 2013), which are fast, complex sounds and thus potentially difficult to imitate (although, interestingly, trills

showed no differences in occurrence). Another comparison of island and mainland birds found no overall difference in song structure, but found that birds occupying smaller islands sang fewer syllables (Reudink et al. 2021). After successive translocations of the North Island saddleback (*tīke*), high-pitched calls became lower pitched before eventually disappearing (Parker et al. 2012; K. Parker, pers. commun.); it remains unknown whether these lower pitched calls are easier to produce. There are also hints that difficult-to-produce elements (such as trilled notes) may be vulnerable to loss in small wild populations not restricted to islands, as predicted by our model; for example, golden-crowned sparrows in the furthest northwest breeding population lack terminal trills in their songs, which are found in more central populations (Shizuka et al. 2016). Similarly, in the barn swallow, the populations that have the shortest and least complex songs also have the slowest trills (Wilkins et al. 2018); there is molecular evidence that these populations went through a bottleneck and subsequently remained small (Zink et al. 2006). Our model predicts that degradation of hard-to-learn elements should occur in small, isolated mainland populations, in addition to those on islands; one complication is that mainland birds in small breeding populations may in fact be in contact with a larger number of potential tutors throughout their lifetime, for example, during migration or wintering, depending on the species' seasonal singing behavior (Otter et al. 2020). Determining the demographic history, as well as the current effective number of tutors, of populations that have slow or no trills versus those with rapid trills could provide a test of the hypothesis that trills are lost when the number of tutors falls below a certain threshold.

A key assumption of Henrich's model is that trait difficulty (α) stays constant for a given trait over time; in reality, a trait may become easier to learn as it simplifies. We modified the model to account for varying difficulty (α), and found an intuitive and important result: skill level (z) approaches a minimum value as a song feature becomes easier to imitate, which may enable that trait to be stably maintained, although in a simpler form, in small populations that might have lost the trait under a regime where α was constant. That simpler renditions of a species-typical song could be copied and maintained is consistent with empirical work in swamp sparrows (Podos et al. 1999), which shows that "poor" copies of song (with broken syntax, to accommodate trill rates beyond performance limits) are accepted as models by young sparrows and retained in future generations. However, even when trait difficulty varies with trait quality, the song trait can be lost from the population under some parameter combinations (i.e., when z approaches a stable, but negative, value, as in Fig. S3b). By modifying the Price equation to change α according to z_{\max} , we are able to accommodate an important facet of songbird learning: that song traits, such as aspects of song structure, syntax, and repertoire, can have both genetic and cultural components. Our

modified model shows that even song traits whose social learning has intrinsic limits can behave according to the predictions of the Price equation, where population size affects the realized elaboration of those traits.

Although population size has often been invoked in human cultural evolutionary theory as a factor in the maintenance or loss of learned skills, the existence of a straightforward relationship between population size and cultural complexity in humans has been debated (Henrich 2004; Carlino et al. 2007; Powell et al. 2009; Collard et al. 2013; Henrich et al. 2016; Vaesen et al. 2016; Fogarty and Creanza 2017). Empirical research in humans has provided evidence both for and against such a relationship (see Strassberg and Creanza 2021 for a review); similarly, there is also debate surrounding the relationship between population size and maintenance of a large inventory of sounds in language (Hay and Bauer 2007; Donohue and Nichols 2011). These seemingly conflicting patterns may be complicated by other properties of a population, such as migration rates and connectivity. For example, an experimental study showed that multiple, partially connected human groups can maintain a higher diversity of solutions to a complex problem than fully connected groups of the same size (Dereux and Boyd 2016). Similarly, our model shows the importance of network factors, such as degree, in maintaining simulated song features. This framework offers a promising perspective for investigating the origin and maintenance of dialects in many bird species, where spatial heterogeneity in song types can persist in the absence of obvious physical barriers.

We cannot rule out that there is selection directly on elements of song, as in some aspects of human language (Newberry et al. 2017), although there is no reason to suppose this will tend toward simplifying songs. Very rapid changes in birdsong cultures do occur (e.g., Otter et al. 2020), and such changes may be tied to cultural selection. Anthropogenic changes to the soundscape may also be a major selective pressure on song elements, including elements that are transmitted socially (Moseley et al. 2018), producing changes that can persist even after anthropogenic noise is removed (Reichard et al. 2020). Although our model does not address the effects of cultural selection or biased transmission on birdsong, clearly these mechanisms are extremely important for understanding evolution of learned vocalizations.

Bird species vary in the relative importance of song for mate attraction. Determining the relationship between song elaboration and female choice in the field, which would provide a key empirical test of the model we propose in any given species, is an important unresolved problem. There is an interesting discrepancy between different types of studies of repertoire size, with laboratory studies often finding a positive association between repertoire size and female preference, and field studies reporting such an association only in a subset of species (Byers and

Kroodsma 2009). One explanation that these authors put forward is the existence of “preferences that do not translate to choices”—that females in the field, although they may perceive differences in song repertoire and even prefer larger repertoires, are also influenced by many other interacting factors, such as territory quality, in making their ultimate choice of mate (Byers and Kroodsma 2009). We believe this principle likely applies to many aspects of female choice in the field, and potentially complicates the task of determining the fitness benefits for given song traits. The strength of selection on song, and whether it arises from mate choice, intrasexual competition, or other mechanisms, likely affects how closely the assumptions of the model presented here would apply: species with strong directional selection on song via mate choice by females seem most likely to conform to the predictions of the model. In other words, the stronger the link between song traits and fitness outcome, the more we would expect song to behave in a tool-like way, and in turn, the more its evolution might respond to demographic changes. We encourage researchers interested in vocal learning and other forms of nonhuman cultural evolution to consider what novel predictions could be made in their system by embracing this subtle but important change of perspective.

AUTHOR CONTRIBUTIONS

EJH and NC conceived the study; EJH performed analyses and simulations; EJH and NC interpreted results, generated figures, and wrote the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank three anonymous reviewers whose incisive reading and constructive feedback greatly improved the manuscript. We also wish to thank members of the Creanza lab for valuable feedback at all stages of writing. This work was supported by Vanderbilt's Evolutionary Studies Initiative. NC is supported by the John Templeton Foundation (62187), and the National Science Foundation (NSF BCS-1918824).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA ARCHIVING

Code to reproduce the simulations underlying all figures is available at github.com/CreanzaLab/PopulationSizeSongEvo.

REFERENCES

- Baker, M.C., Baker, M.S.A. & Baker, E.M. (2003) Rapid evolution of a novel song and an increase in repertoire size in an island population of an Australian songbird. *Ibis*, 145, 465–471.
- Baldini, R. (2012) Success-biased social learning: cultural and evolutionary dynamics. *Theoretical Population Biology*, 82, 222–228.
- Ballentine, B. (2004) Vocal performance influences female response to male bird song: an experimental test. *Behav. Ecol.*, 15, 163–168.
- Baptista, L.F. & Morton, M.L. (1988) Song learning in montane white-crowned sparrows: from whom and when. *Animal Behaviour*, 36, 1753–1764.

- Boyd, R. & Richerson, P.J. (1988) Culture and the Evolutionary Process. Univ. of Chicago Press, Chicago.
- Brenowitz, E.A., Lent, K. & Kroodsma, D.E. (1995) Brain space for learned song in birds develops independently of song learning. *Journal of Neuroscience*, 15, 6281–6286.
- Byers, B.E. & Kroodsma, D.E. (2009) Female mate choice and songbird song repertoires. *Animal Behaviour*, 77, 13–22.
- Carja, O. & Creanza, N. (2019) The evolutionary advantage of cultural memory on heterogeneous contact networks. *Theoretical Population Biology*, 129, 118–125.
- Carlino, G.A., Chatterjee, S. & Hunt, R.M. (2007) Urban density and the rate of invention. *J. Urban Econ.*, 61, 389–419.
- Caro, S.P., Sewall, K.B., Salvante, K.G. & Sockman, K.W. (2010) Female Lincoln's sparrows modulate their behavior in response to variation in male song quality. *Behav. Ecol.*, 21, 562–569.
- Cavalli-Sforza, L.L. & Feldman, M.W. (1981) Cultural transmission and evolution: a quantitative approach. Princeton Univ. Press, Princeton, NJ.
- Clayton, N.S. (1987) Song tutor choice in zebra finches. *Animal Behaviour*, 35, 714–721.
- Collard, M., Ruttle, A., Buchanan, B. & O'Brien, M.J. (2013) Population size and cultural evolution in nonindustrial food-producing societies. *Plos One*, 8, e72628.
- Collins, J. (2016) Commentary: the role of language contact in creating correlations between humidity and tone. *Journal of Language Evolution*, 1, 46–52.
- 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. *Proc. R. Soc. B Biol. Sci.*, 288, 20210225.
- Creanza, N., Kolodny, O. & Feldman, M.W. (2017) Cultural evolutionary theory: how culture evolves and why it matters. *Proceedings National Academy of Science USA*, 114, 7782–7789.
- deKort, S.R., Eldermire, E.R.B., Cramer, E.R.A. & Vehrencamp, S.L. (2009) The deterrent effect of bird song in territory defense. *Behavioral Ecology*, 20, 200–206.
- Derech, M., Bonnefon, J.-F., Boyd, R. & Mesoudi, A. (2019) Causal understanding is not necessary for the improvement of culturally evolving technology. *Nature human behaviour*, 3, 446–452.
- Derech, M. & Boyd, R. (2016) Partial connectivity increases cultural accumulation within groups. *PNAS*, 113, 2982–2987.
- Donohue, M. & Nichols, J. (2011) Does phoneme inventory size correlate with population size? *Linguist. Typol.*, 15, 161–170.
- Doupe, A.J. & Kuhl, P.K. (1999) Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567–631.
- Drăgănoiu, T.I., Nagle, L. & Kreuzer, M. (2002) Directional female preference for an exaggerated male trait in canary (*Serinus canaria*) song. *Proc. Biol. Sci.*, 269, 2525–2531.
- Everett, C., Blasi, D.E. & Roberts, S.G. (2015) Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. *PNAS*, 112, 1322–1327.
- Fehér, O., Wang, H., Saar, S., Mitra, P.P. & Tchernichovski, O. (2009) De novo establishment of wild-type song culture in the zebra finch. *Nature*, 459, 564–568.
- Fogarty, L. & Creanza, N. (2017) The niche construction of cultural complexity: interactions between innovations, population size and the environment. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 372, 20160428.
- Gil, D. & Gahr, M. (2002) The honesty of bird song: multiple constraints for multiple traits. *Trends in Ecology & Evolution*, 17, 133–141.
- Greig, E.I., Taft, B.N. & Pruett-Jones, S. (2012) Sons learn songs from their social fathers in a cooperatively breeding bird. *Proceedings of the Royal Society B: Biological Sciences*, 279, 3154–3160.
- Harris, M.A. & Lemon, R.E. (1972) Songs of song sparrows (*Melospiza melodia*): individual variation and dialects. *Canadian Journal of Zoology*, 50, 301–309.
- Hasselquist, D., Bensch, S. & vonSchantz, T. (1996) Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature*, 381, 229–232.
- Hay, J. & Bauer, L. (2007) Phoneme inventory size and population size. *Language*, 83, 388–400.
- Henrich, J. (2004) Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *American Antiquity*, 69, 197–214.
- Henrich, J., Boyd, R., Derech, M., Kline, M.A., Mesoudi, A., Muthukrishna, M., Powell, A.T., Shennan, S.J. & Thomas, M.G. (2016) Understanding cumulative cultural evolution. *Proceedings National Academy of Science USA*, 113, E6724–E6725.
- Henrich, J. & Gil-White, F.J. (2001) The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.*, 22, 165–196.
- Hoppitt, W. & Laland, K.N. (2013) Social learning: an introduction to mechanisms, methods, and models. Princeton Univ. Press, Princeton, NJ.
- Illes, A.E., Hall, M.L. & Vehrencamp, S.L. (2006) Vocal performance influences male receiver response in the banded wren. *Proc. R. Soc. B Biol. Sci.*, 273, 1907–1912.
- Johannessen, L.E., Slagsvold, T. & Hansen, B.T. (2006) Effects of social rearing conditions on song structure and repertoire size: experimental evidence from the field. *Animal Behaviour*, 72, 83–95.
- Kacsoh, B.Z., Bozler, J. & Bosco, G. (2018) *Drosophila* species learn dialects through communal living. *PLoS Genet*, 14, e1007430.
- Kobayashi, Y. & Aoki, K. (2012) Innovativeness, population size and cumulative cultural evolution. *Theoretical Population Biology*, 82, 38–47.
- Komdeur, J. & Edelaar, P. (2001) Evidence that helping at the nest does not result in territory inheritance in the Seychelles warbler. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268, 2007–2012.
- Konishi, M. (1970) Comparative neurophysiological studies of hearing and vocalizations in songbirds. *Z. Vergl. Physiol.*, 66, 257–272.
- Kroodsma, D.E., Houlihan, P.W., Fallon, P.A. & Wells, J.A. (1997) Song development by grey catbirds. *Animal Behaviour*, 54, 457–464.
- Lachlan, R.F., Ratmann, O. & Nowicki, S. (2018) Cultural conformity generates extremely stable traditions in bird song. *Nature Communication*, 9, 2417.
- Lachlan, R.F. & Servedio, M.R. (2004) Song learning accelerates allopatric speciation. *Evolution*, 58, 2049–2063.
- Lahti, D.C., Moseley, D.L. & Podos, J. (2011) A tradeoff between performance and accuracy in bird song learning. *Ethology*, 117, 802–811.
- Laiolo, P. & Tella, J.L. (2007) Erosion of animal cultures in fragmented landscapes. *Frontiers in Ecology and the Environment*, 5, 68–72.
- Lambrechts, M. & Dhondt, A.A. (1986) Male quality, reproduction, and survival in the great tit (*Parus major*). *Behavioral Ecology and Sociobiology*, 19, 57–63.
- Lewis, R.N., Soma, M., deKort, S.R. & Tucker Gilman, R. (2021) Like father like son: cultural and genetic contributions to song inheritance in an estrildid finch. *Frontiers in Psychology*, 12, 654198.
- Liu, W.-C. & Nottebohm, F. (2007) A learning program that ensures prompt and versatile vocal imitation. *PNAS*, 104, 20398–20403.
- Logue, D.M. & Forstmeier, W. (2008) Constrained performance in a communication network: implications for the function of song-type matching

- and for the evolution of multiple ornaments. *The American Naturalist*, 172, 34–41.
- Mann, D.C., Lahti, D.C., Waddick, L. & Munding, P.C. (2021) House finches learn canary trills. *Bioacoustics*, 30, 215–231.
- Marler, P. (1970a) A comparative approach to vocal learning: song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology*, 71, 1–25.
- . (1970b) Birdsong and speech development: could there be parallels? *Am. Sci.*, 58, 669–673.
- Marler, P. & Sherman, V. (1985) Innate differences in singing behaviour of sparrows reared in isolation from adult conspecific song. *Animal Behaviour*, 33, 57–71.
- 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, 149–159.
- Morinay, J., Cardoso, G.C., Doutrelant, C. & Covas, R. (2013) The evolution of birdsong on islands. *Ecol. Evol.*, 3, 5127–5140.
- Moseley, D.L., Derryberry, G.E., Phillips, J.N., Danner, J.E., Danner, R.M., Luther, D.A. & Derryberry, E.P. (2018) Acoustic adaptation to city noise through vocal learning by a songbird. *Proc. R. Soc. B Biol. Sci.*, 285, 20181356.
- Nagle, L., Pasteau, M. & Kreuzer, M. (2004) Preferences and predispositions for intra-syllabic diversity in female canaries (*Serinus canaria*). *Behaviour*, 141, 571–583.
- Newberry, M.G., Ahern, C.A., Clark, R. & Plotkin, J.B. (2017) Detecting evolutionary forces in language change. *Nature*, 551, 223–226.
- Oikkonen, J., Onkamo, P., Järvelä, I. & Kanduri, C. (2016) Convergent evidence for the molecular basis of musical traits. *Science Reports*, 6, 39707.
- Otter, K.A., Mckenna, A., LaZerte, S.E. & Ramsay, S.M. (2020) Continent-wide shifts in song dialects of white-throated sparrows. *Current Biology*, 30, 3231.e3–3235.e3.
- Parker, K.A., Anderson, M.J., Jenkins, P.F. & Brunton, D.H. (2012) The effects of translocation-induced isolation and fragmentation on the cultural evolution of bird song. *Ecol. Lett.*, 15, 778–785.
- Parker, K.A., Hauber, M.E. & Brunton, D.H. (2010) Contemporary cultural evolution of a conspecific recognition signal following serial translocations. *Evolution*, 64, 2431–2441.
- Payne, R.B. (1982) Ecological consequences of song matching: breeding success and intraspecific song mimicry in indigo buntings. *Ecology*, 63, 401–411.
- Payne, R.B. & Payne, L.L. (1993) Song copying and cultural transmission in indigo buntings. *Animal Behaviour*, 46, 1045–1065.
- Phillips, J.N. & Derryberry, E.P. (2017) Vocal performance is a salient signal for male–male competition in White-crowned Sparrows. *The Auk*, 134, 564–574.
- Podos, J. (1997) A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). *Evolution*, 51, 537.
- Podos, J. (2001) Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature*, 409, 185–188.
- Podos, J. (1996) Motor constraints on vocal development in a songbird. *Animal Behaviour*, 51, 1061–1070.
- Podos, J., Nowicki, S. & Peters, S. (1999) Permissiveness in the learning and development of song syntax in swamp sparrows. *Animal Behaviour*, 58, 93–103.
- Powell, A., Shennan, S. & Thomas, M.G. (2009) Late Pleistocene demography and the appearance of modern human behavior. *Science*, 324, 1298–1301.
- Prather, J., Peters, S., Mooney, R. & Nowicki, S. (2012) Sensory constraints on birdsong syntax: neural responses to swamp sparrow songs with accelerated trill rates. *Animal Behaviour*, 83, 1411–1420.
- Read, D. (2011) The misuse of a mathematical model: the Tasmanian case. *SSRN Electron. J.*, <https://doi.org/10.2139/ssrn.2805822>.
- Rehsteiner, U., Geisser, H. & Reyer, H.-U. (1998) Singing and mating success in water pipits: one specific song element makes all the difference. *Animal Behaviour*, 55, 1471–1481.
- Reichard, D.G., Atwell, J.W., Pandit, M.M., Cardoso, G.C., Price, T.D. & Ketterson, E.D. (2020) Urban birdsongs: higher minimum song frequency of an urban colonist persists in a common garden experiment. *Animal Behaviour*, 170, 33–41.
- Reiss, D. & McCowan, B. (1993) Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): evidence for vocal learning. *Journal of Comparative Psychology*, 107, 301–312.
- Reudink, M.W., Pageau, C., Fisher, M., Mount, N., Buckler, N., Otter, K.A. & Briskie, J.V. (2021) Evolution of song and color in island birds. *Wilson Journal of Ornithology*, 133, 1–10.
- Riebel, K. & Slater, P.J.B. (1998) Testing female chaffinch song preferences by operant conditioning. *Animal Behaviour*, 56, 1443–1453.
- Robinson, C.M. & Creanza, N. (2019) Species-level repertoire size predicts a correlation between individual song elaboration and reproductive success. *Ecol. Evol.*, 9, 8362–8377.
- Robinson, C.M., Snyder, K.T. & Creanza, N. (2019) Correlated evolution between repertoire size and song plasticity predicts that sexual selection on song promotes open-ended learning. *eLife*, 8, e44454.
- Rogers, D.S., Feldman, M.W. & Ehrlich, P.R. (2009) Inferring population histories using cultural data. *Proc. R. Soc. B Biol. Sci.*, 276, 3835–3843.
- Searfoss, A.M., Liu, W.-C. & Creanza, N. (2020) Geographically well-distributed citizen science data reveals range-wide variation in the chipping sparrow's simple song. *Animal Behaviour*, 161, 63–76.
- Shizuka, D., Lein, M. & Chilton, G. (2016) Range-wide patterns of geographic variation in songs of Golden-crowned Sparrows (*Zonotrichia atricapilla*). *Auk*, 133, 520–529.
- Soha, J.A. & Marler, P. (2000) A species-specific acoustic cue for selective song learning in the white-crowned sparrow. *Animal Behaviour*, 60, 297–306.
- Soma, M. & Garamszegi, L.Z. (2011) Rethinking birdsong evolution: meta-analysis of the relationship between song complexity and reproductive success. *Behavioral Ecology*, 22, 363–371.
- Soma, M., Hiraiwa-Hasegawa, M. & Okanoya, K. (2009) Song-learning strategies in the Bengalese finch: do chicks choose tutors based on song complexity? *Animal Behaviour*, 78, 1107–1113.
- Soma, M., Takahasi, M., Ikebuchi, M., Yamada, H., Suzuki, M., Hasegawa, T. & Okanoya, K. (2006) Early rearing conditions affect the development of body size and song in Bengalese finches. *Ethology*, 112, 1071–1078.
- Steele, J., Glatz, C. & Kandler, A. (2010) Ceramic diversity, random copying, and tests for selectivity in ceramic production. *Journal of Archaeological Science*, 37, 1348–1358.
- Strassberg, S.S. & Creanza, N. (2021) Cultural evolution and prehistoric demography. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 376, 20190713.
- Suh, Y.H., Pesendorfer, M.B., Tringali, A., Bowman, R. & Fitzpatrick, J.W. (2020) Investigating social and environmental predictors of natal dispersal in a cooperative breeding bird. *Behavioral Ecology*, 31, 692–701.
- Tobias, J.A., Aben, J., Brumfield, R.T., Derryberry, E.P., Halfwerk, W., Slabbekoorn, H. & Seddon, N. (2010) Song divergence by sensory drive in Amazonian birds. *Evolution*, 64, 2820–2839.

- Vaesen, K., Collard, M., Cosgrove, R. & Roebroeks, W. (2016) Population size does not explain past changes in cultural complexity. *PNAS*, 113, E2241–E2247.
- Vallet, E., Beme, I.I. & Kreutzer, M. (1998) Two-note syllables in canary songs elicit high levels of sexual display. *Anim. Behav.*, 55, 291–297.
- Vallet, E. & Kreutzer, M. (1995) Female canaries are sexually responsive to special song phrases. *Animal Behaviour*, 49, 1603–1610.
- Verheyen, R.F., Eens, M. & Pinxten, R. (1991) Male song as a cue for mate choice in the European starling. *Behaviour*, 116, 210–238.
- Weir, J.T., Wheatcroft, D.J. & Price, T.D. (2012) The role of ecological constraint in driving the evolution of avian song frequency across a latitudinal gradient. *Evolution*, 66, 2773–2783.
- Whiten, A., Goodall, J., McGrew, W.C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C.E.G., Wrangham, R.W. & Boesch, C. (1999) Cultures in chimpanzees. *Nature*, 399, 682–685.
- Wilkins, M.R., Scordato, E.S.C., Semenov, G.A., Karaardıç, H., Shizuka, D., Rubtsov, A., Pap, P.L., Shen, S.-F. & Safran, R.J. (2018) Global song divergence in barn swallows (*Hirundo rustica*): exploring the roles of genetic, geographical and climatic distance in sympatry and allopatry. *Biological Journal of the Linnean Society*, 123, 825–849.
- Wilkins, M.R., Seddon, N. & Safran, R.J. (2013) Evolutionary divergence in acoustic signals: causes and consequences. *Trends in Ecology & Evolution*, 28, 156–166.
- Wilkins, M.R., Shizuka, D., Joseph, M.B., Hubbard, J.K. & Safran, R.J. (2015) Multimodal signalling in the North American barn swallow: a phenotype network approach. *Proc. R. Soc. B Biol. Sci.*, 282, 20151574.
- Williams, H. (1990) Models for song learning in the zebra finch: fathers or others? *Animal Behaviour*, 39, 745–757.
- Williams, H., Levin, I.I., Norris, D., Newman, A.E.M. & Wheelwright, N.T. (2013) Three decades of cultural evolution in Savannah sparrow songs. *Animal Behaviour*, 85, 213–223.
- Woolfenden, G.E. & Fitzpatrick, J.W. (1978) The inheritance of territory in group-breeding birds. *Bioscience*, 28, 104–108.
- Yeh, D.J. (2019) Assortative mating by an obliquely transmitted local cultural trait promotes genetic divergence: a model. *The American Naturalist*, 193, 81–92.
- Yeh, D.J., Boughman, J.W., Sætre, G. & Servedio, M.R. (2018) The evolution of sexual imprinting through reinforcement. *Evolution*, 72, 1336–1349.
- Yeh, D.J. & Servedio, M.R. (2015) Reproductive isolation with a learned trait in a structured population. *Evolution*, 69, 1938–1947.
- Youngblood, M. & Lahti, D.C. (2022) Content bias in the cultural evolution of house finch song. *Animal Behaviour*, 185, 37–48.
- Zink, R.M., Pavlova, A., Rohwer, S. & Drovetski, S.V. (2006) Barn swallows before barns: population histories and intercontinental colonization. *Proc. Biol. Sci.*, 273, 1245–1251.

Associate Editor: J. A. Tobias
 Handling Editor: A. G. McAdam

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. The population (N=100) mean skill level (z) values after 50 generations (average of 20 replicates), with degree of network (k , shown on the y axis), α (panel A), and β (panel B) varying systematically.

Figure S2. The simulated change in z in a population of 24 individuals where each member has a degree of 10 (top) or 5 (bottom).

Figure S3. The skill level of a population before and after a bottleneck.

Figure S4. In contrast to Figure 4, here α is drawn from a predetermined range of 100 evenly spaced values from -1 to 6 .

Figure S5. In these ten simulations, α varies between 1 and 6, changing each generation depending on the maximum value of z . When $z_{\max} > 10$, $\alpha = 6$; when $2 > z_{\max} > 10$, $\alpha = 5$; when $z_{\max} < 2$, $\alpha = 1$.