Population turnover facilitates cultural selection for efficiency

Michael Chimento^{*1,2}, Gustavo Alarcón-Nieto¹, and Lucy Aplin^{1,2}

¹Cognitive and Cultural Ecology Lab, Max Planck Institute of Animal Behavior, Am Obstberg 1, 78315 Radolfzell, Germany

²Centre for the Advanced Study of Collective Behaviour, Konstanz University; Universitätsstraße 10, 78464 Konstanz, Germany

October 6, 2020

Abstract

Selection for more efficient socially learned behaviors over alternatives is crucial for cumulative cultural evolution, yet our understanding of such cultural selection in animals is limited. We performed a cultural diffusion experiment using 18 populations of wild-caught great tits (*Parus major*) to ask whether more efficient innovations are subsequently selected for, and whether this process is affected by turnover. We show that gradual replacement of individuals greatly increases the probability that a more efficient behavior will invade a population's cultural repertoire, out-competing an established inefficient behavior. Turnover does not increase innovation rates, but instead increases adoption rates, as immigrants are more susceptible to novel, efficient behaviors. An agent based model further supported our results by demonstrating that this effect holds across populations of different types of learners. Altogether, these results provide strong evidence for cultural selection for efficiency in animals, and highlight the importance of population turnover for this process.

1 Introduction

Beyond genes, culture offers a secondary inheritance system that provides an alternative pathway to adaptive plasticity [1, 2]. Culture, here defined as socially transmitted information and behaviors that are shared in groups and persist over time, is increasingly accepted to occur across wide range of taxa and behavioral domains [3]. While persistent, cultural traits are not necessarily static, and can change in frequency and type in response to selective pressures. This has led to the treatment of culture as an evolutionary process, with cultural evolutionary theory arguing that culture exhibits the three fundamental components of Darwinian evolution: variation, competition, and inheritance [4, 5, 6, 7]. Humans have undoubtedly leveraged culture to its fullest extent by accumulating and recombining a wealth of cultural behaviors and artifacts into evermore efficient complexes in a process termed cumulative cultural evolution (CCE) [8]. Despite its potential importance for broader evolutionary theory, cultural evolution is understudied in non-human animals, and especially so in birds, outside of birdsong (however see [9, 10] for recent reviews). It remains an open question to identify and describe processes of cultural evolution in non-human species, including potential precursors to CCE, and the social factors that may affect these evolutionary dynamics.

Recent discussions have proposed that repeated instances of cultural selection for efficiency in a singular cultural trait may represent a potential precursor to more complex forms of CCE unique to humans [11, 12]. While claims of CCE in non-human animals are controversial [13, 14], cases of increases in efficiency have recently been identified in several animals. Reintroduced storks and ungulates refine migration routes over several generations [15, 16], and in homing pigeons, transmission chains refined routes beyond improvement achieved by single individual [11]. There is less evidence for the cultural evolution of efficiency apart from movement, although wild chimpanzees' adoption of moss-sponges over leaf-sponges suggests that they may be able to select tool material based on efficiency [17]. Yet working against any type of selection for efficiency is behavioral conservatism, or a low probability of sampling behavioral variants beyond those already produced. The inability to adopt more profitable alternatives is well documented in studies on chimpanzees [18, 19, 20], but less studied in other taxa (but see [21]).

Culture exists on the substrate of the social network [22]. It follows that cultural evolution should be responsive to network properties such as population size, structure and dynamic turnover in various ways [23, 24, 25, 26, 6, 27, 28]. Notably, larger population sizes have been linked to the generation and maintenance of cultural diversity in humans, with larger populations having an elevated probability of both innovation and faithful cultural transmission [23, 29, 30, 31, 32]. Conversely, population decline and fragmentation can result

in loss of diversity and simplification of traits, as observed in bird song [33, 34, 35]. Population turnover, caused by immigration and emigration or births and deaths, could affect cultural evolution by elevating innovation probabilities as incoming individuals could introduce new innovations or increase 'behavioral noise' [11]. Alternatively, naive individuals might be able to better re-sample the behavioral space, leading to more adaptive cultural outcomes [36]. Yet while population turnover is a fundamental demographic process in populations of any species, its effect on cultural evolution is poorly understood. Linear transmission chains provide an approximation of turnover, but it has not been well explored using populations of socially learning animals, and there are no clear predictions about how turnover might influence the tempo and mode of cultural evolution.

Great tits (*Parus major*), a common passerine bird species, are a useful study system for investigating how population turnover might affect cultural evolution. Tits spontaneously innovate new behavior at high rates [37], and are well known for the social transmission of foraging behavior, such as the spread of piercing milk bottle caps to access cream across Great Britain [38, 39]. Experimental studies have provided strong evidence for the social learning of foraging behavior and establishment of new traditions in tits [40, 41, 21]. In these studies, birds exhibited a conformist bias, in that they disproportionately learned the most commonly demonstrated variants, and maintained stable foraging traditions over multiple generations [41]. Further, turnover within and between generations is a pertinent feature of great tit social systems. Roughly 50% of breeders are newly arrived immigrants each year, and the mean lifespan of breeders is only 2 years [42, 43]. During winter, birds exhibit fission-fusion flocking dynamics in larger social networks [44], with flock membership varying even over the course of a single day depending on density and location of birds [45].

To study the effect of population turnover on cultural selection for efficiency, we conducted a large-scale cultural diffusion and evolution experiment on 18 artificial micro-populations of wild-caught great tits ($n = 181$, over 40 days). A foraging tradition to access food by pushing on a bidirectional door of a puzzle box (Fig. 1A) was established in all populations. Birds could improve their speed at the established foraging tradition, but could also potentially innovate an alternative solution that was on average about .5s faster (Fig. 1C). This solution offered a greater payoff for the solver, as they could receive the same reward for less time investment. In 9 populations, population turnover was simulated by replacing a third of each population with new, wild-caught birds every 7 days (hereafter turnover populations). Another 9 populations maintained stable membership for the duration of the experiment (hereafter static populations). We predicted that turnover populations would be more likely than static populations to innovate and adopt the alternative, more efficient solution. Incoming, naive birds might be more likely 1) to innovate the efficient solution, or 2) to adopt the efficient solution once innovated. Finally, we supported our results by constructing a complementary agent-based-model that replicated the design of the experiment and allowed us to test the effect of turnover while varying relevant individual learning parameters.

2 Results

Puzzle boxes with 2 solutions of differing efficiency were presented to 18 micro-populations of 6 birds each (Methods and Fig. 1A). These populations were either static, or underwent gradual population turnover. In order to establish the inefficient tradition as a starting cultural state in both conditions, all population were provided with a tutor bird trained to perform the inefficient (blue) solution, and the alternative efficient (red) solution was blocked until the first turnover. The tradition for the inefficient solution then spread to naive birds in 17/18 populations. Excluding the 18 tutors, 83 of 163 naive birds learned to solve the puzzle by the end of the experiment. Evidence that solving behavior was socially learned includes 1) that following the first turnover, 87% of birds first solved using the most predominant solution in their population, 2) that the puzzle was difficult to individually learn (tutors took anywhere from 1 to 2 weeks of incremental training to learn to solve) and 3) that the behavior failed to be re-innovated in a population where the solving behavior went extinct.

Overall, populations produced an average of 9,719 total solutions (range: 1,494–17,208), and solvers produced a mean of 1,795 solutions (range: 20–6,612), and a daily average of 98 solutions. Altogether, solvers produced a total of 175,143 recorded solutions, with 89% (156,609) of those with measurable time-to-solve (TTS). Alternatively, some birds adopted a scrounging strategy: waiting for another bird to open the door and taking food before the door closed. In both conditions, there was a bi-modal distribution of solvers and scroungers, with 62 birds scrounging more often than they solved. However, scrounging was excluded from further analysis, as it is not considered a cultural tradition.

Birds that did learn to solve the puzzle became significantly faster at solving with experience, reducing their TTS by an average of 5% with every standard deviation of solutions produced, suggesting that a process

of reinforcement learning was leading to continuous improvements in efficiency (Linear Mixed Model (LMM), solution index (z-scaled): $\exp(\beta) = 0.950$, $t = -31.031$, $P < 0.001$; **SI Table 1**). A comparison of the mean TTS within two solutions, during both lesser and greater stages of solving proficiency, provided further evidence of improvement (Wilcoxon test, efficient: $M = 2.01s$, $M = 1.32s$, $P < 0.001$; inefficient: $M = 2.34s$, $M = 1.97s$, $P < 0.001$). Between solutions, average TTS for the efficient solution was significantly faster than inefficient solution (Wilcoxon test, $M = 1.49s$, $M = 2.06s$, $P < 0.001$). This margin was by several tenths of a second during the first quartile of solving trials, and widened to roughly .7s as birds became more experienced and performance plateaued. (Fig. 1C). Hence, red solutions were more efficient in that they offered a greater payoff on average, as the solver received the same reward for less time investment.

2.1 Turnover facilitates cultural selection for efficiency

Individual populations displayed a variety of dynamics and outcomes, a selection of which are visualized in Fig. 2 (see SI Fig. 1 for an overview of all populations). When aggregated together, turnover populations were clearly more likely to select for the efficient solution as the experiment progressed (Fig. 3A). To quantify selection for efficiency, we used a logistic GLMM to predict the likelihood of producing the efficient solution using a full interaction between condition and experimental day (**SI Table 2**). Once the efficient solution was unblocked at the first turnover, any given bird in either condition was initially unlikely to produce an efficient behavior (GLMM, intercept: $\alpha = -10.852$, $Z = -9.974$, $P < 0.001$; turnover condition: $\beta = -2.169$, $Z = -1.276$, $P = 0.202$). With each passing day, there was an increasing probability of producing efficient solutions in both conditions (GLMM, exp. day: $\beta = 0.291$, $Z = 28.986$, $P < 0.001$). However birds from the turnover condition became significantly more likely to produce efficient solves compared to the static condition (GLMM, exp. day*turnover: $\beta = 0.603$, $Z = 38.71$, $P < 0.001$).

By the end of the experiment, turnover populations had produced 39,216 inefficient solutions and 33,834 efficient solutions, while static populations had produced 81,876 inefficient solutions and only 1,498 efficient solutions. Overall, 7/9 turnover populations selected for and switched to the efficient solution. Interestingly, in one of the two turnover populations that did not select for the efficient solution, the solving behavior went completely extinct due to stochastic events (turnover of all knowledgeable individuals, Fig. 2C). By contrast, only one static population selected for the efficient solution after it was innovated. This population accounted for most of the efficient solutions in the static condition (1,466), while all other static populations produced the efficient solution only a handful of times, if at all (range: 0-10).

2.2 Cultural selection generates differences in solving performance

A GLMM was used to analyze differences in solving performance between conditions in which TTS was predicted by a full interaction between condition and experimental day (**SI Table 3**). While both conditions had a decreasing TTS as the experiment went on, (LMM, exp. day: $\exp(\beta) = -.999$, $t = -1.757$, $P = 0.07$), on average, turnover populations reduced their TTS approximately 0.4% more per day than static populations (LMM, exp. day*turnover condition, $\exp(\beta) = 0.996$, $t = -10.329$, $P < 0.001$; Fig. 3B). Turnover populations were slower than static populations through the first two turnover events, suggesting that the loss of knowledgeable birds by turnover may have initially hindered the accumulation of expertise at the inefficient solution. However, by the final week of the experiment, birds in turnover populations were solving faster than static populations, despite individuals having less experience with either solution (Wilcoxon test, $M = 1.66s$, $M = 1.96s$, $P < 0.001$). This was a result of the increasing frequency of the efficient solution as the dominant solving behavior, which allowed turnover populations to achieve a lower TTS.

2.3 Mechanisms Producing Cultural Outcomes: do innovation rates differ between conditions?

In total, 21 birds were classified as innovators (static: 13, turnover: 8). Innovators were either the first to have produced the efficient solution in a population, or were suspected to have produced the efficient solution without having seen it demonstrated. To test whether turnover increased the probability of innovation, we compared the timing of innovation events between both conditions in several ways: 1) days solving before innovation, 2) exposure time before innovation and 3) experimental day of innovation (**SI Fig. 2**). Innovators in the turnover condition had a significantly shorter period of days solving before innovation (Wilcoxon test: $M_t = 8.8$, $M_s = 13.8$, $p = .0245$), although this difference was driven by only 2 data points (**SI Fig 2B**). Mean days exposure before innovation and mean experimental day of innovation in the turnover condition were not significantly lower than in the static condition (Wilcoxon test: $M_t = 14.1$, $M_s = 15.5$, $p = .168$;

$M_t = 9.5$, $M_s = 11.3$, $p = .112$ respectively). These comparisons reveal that innovators, even in the turnover condition, were typically experienced, knowledgeable solvers. Furthermore, innovators were almost equally split between age (juvenile and adult) and sex (**SI Table 5**). These results, along with the fact that static populations innovated more often during the experiment, suggest that innovation probabilities were comparable between conditions. Surprisingly, only 5 of the 21 innovators permanently adopted the efficient solution after innovation, hinting at some level of behavioral conservatism.

2.4 Mechanisms Producing Cultural Outcomes: do birds exhibit behavioral conservatism?

Most learners held a strong preference for one solution, and relatively few birds were observed producing both solution types after becoming experienced solvers (**SI Fig. 3**). Out of 53 birds that produced both solutions, 36 failed to adopt the more efficient solution. This conservatism could have either been the effect of experience (e.g. habit formation) or conformity. To test their relative importance, we analyzed the probability of failure to adopt the efficient solution in the subset of birds which had produced both solutions using a logistic regression against predictor variables that included experience (days producing the inefficient solution before producing their first efficient solution) and conformity (proportion of socially observed inefficient solutions on day of first efficient solution) summarized in **SI Table 4**. Birds were very likely to adopt the efficient solution if they experienced an efficient solution on the first day of solving (GLMM, intercept: $\alpha = -3.421$, $Z = -1.730$, $P = 0.084$). However, each day of experience with the inefficient solution had a significant positive effect on failure (GLMM, days solving: $\beta = 0.36$, $Z = 2.761$, $P = 0.006$), while conformity had a positive, but non-significant effect on failure (GLMM, socially observed inefficient solutions: $\beta = 0.485$, $Z = 0.235$, $P = 0.814$). Age and sex did not significantly predict failure to adopt. This suggests that experience was the more important factor in conservatism.

2.5 Agent based model

In order to test the robustness of the observed patterns, we reproduced the experimental design in an agent based model. Birds were represented by agents who socially learned to solve the puzzle with an estimated conditional probability of learning derived from the experimental data (**SI Fig. 4**). Following the experimental design, each simulated population was initialized with one agent that was "trained" on the inefficient solution, although not programmed to exclusively produce this solution, and simulations ran for 35 time-steps. Agents made behavioral decisions based on both individual and social information using a modified experience weighted attraction (EWA) model (previously used in similar cultural studies [21, 46, 47, 48]). Three individual-level parameters of interest were systematically varied: 1) sensitivity to social cues, 2) sensitivity to the current payoff versus their memory of payoffs, and 3) level of behavioral conservatism, or sensitivity to differences in payoffs. 100 simulations were run at each point within this parameter space. To quantify selection for efficiency, the proportion of efficient versus inefficient solutions produced by agents was measured at the final time-step.

Our model showed that population turnover increased selection for the efficient solution across all three relevant parameters, resulting in a higher average proportion of efficient solutions in the final time-step (Fig. 4). As agents' sensitivity to the current payoff against previous payoffs increased, so did the proportion of efficient solutions, as agents heavily weighted higher payoffs received from producing the efficient solution in their following decision. As social cue sensitivity increased, the proportion of efficient solves decreased, although at a slower rate in turnover populations. As agent's conservatism (insensitivity to differences in payoffs) increased, the proportion of efficient solutions also decreased. Both conditions responded similarly to the changes in all three parameters, but turnover populations consistently produced more efficient solutions in the final time-step. Finally, while this model used a conditional probability of socially learning solving behavior that was calculated from experimental data, these results were robust even when the probability function was reversed (**SI Fig 5**). These results provide important evidence that the effect of turnover is realized across populations composed of a variety of types of individuals, with minimal assumptions about the learning process. Whatever the true values of these parameters were in the populations of tits in the current study, turnover would have produced the same effect as we observed.

3 Discussion

Our results illustrate a striking positive relationship between population turnover and the cultural evolution of efficiency. The emigration and immigration of individuals lead turnover populations to shift away from an

established foraging tradition towards a more efficient alternative in a process of cultural selection for efficiency. This was not driven by increased innovation rates by immigrants in turnover populations; rather older, knowledgeable residents more often innovated the efficient behavior, which was then adopted by immigrants. This pattern contrasts with the repeated failure of the efficient behavior to invade the cultural repertoire of static populations, despite it having been regularly innovated. The widespread adoption of the efficient solution mitigated the effects of reduced opportunity for development of expertise in turnover populations. They exhibited a faster time to solve at the foraging puzzle by the end of the experiment, despite individuals having less overall contact time with the puzzle. By combining these experimental results along with the agent based model, we propose that turnover provides a mechanism that facilitates cultural selection for efficiency, providing a path to adaptive plasticity that navigates around individual-level behavioral conservatism. Our results point to one further piece of the puzzle of cultural evolution, which is the essential role that inexperienced individuals play not as innovators, but as susceptible adopters of more optimal cultural variants.

This pathway to efficiency contrasts and complements a previous, non-mutually exclusive hypothesis that incoming, naive individuals might provide behavioral variation upon which selection can act [11]. This hypothesis was derived from an experiment that showed the evolution of efficiency in flight paths in diffusion chains of homing pigeons, using diffusion chains. In contrast to the discrete differences in cultural variants presented here, flight paths offer a continuous spectrum of efficiency, and thus innovations might be more easily generated by naive individuals. In populations that experience regular turnover, both processes may be active as immigrants innovate, or successfully resample the behavioral space, and adopt more beneficial, but less frequently produced alternatives. The dominant mechanism through which turnover acts on culture might likely be determined by the ease with which innovations can occur. Future research should seek to identify both processes in their analysis of how populations navigate the payoff landscape of culturally transmitted behaviors.

Overall, birds were behaviorally conservative throughout the experiment; continuing to produce the inefficient solution even after having performed the efficient solution. The role of expertise in suppressing any motivation to adopt alternative behaviors, previously observed in primate studies [19] may explain our finding that behavioral conservatism was best predicted by time spent solving before sampling the efficient solution. Birds which mastered the inefficient solution performed it within the same recorded time as the efficient solution, although ostensibly by applying slightly more force to move the door. The perception of this difference in force may have become negligible with practice. Further, a conformist bias might also account for the resistance to switch in some birds, but not all. This accords with a previous study that produced variable estimates of individual great tits' conformity levels [21].

Behavioral conservatism is cited as a primary reason for the lack of observed CCE in animal species [49]. Such resistance to switching behaviors is one of the key reasons why population turnover is vital to the evolution of efficiency in a cultural system. Turnover alters the memory structure of the population by effectively wiping a portion of the populations' slate clean and allowing for re-sampling of the behavior space. Such memory constraints have been long recognized as necessary for both learning and cultural evolution [50, 51, 52]. Great tits appear to have exceptional memory for socially learned behaviors, remembering preferences years after learning a similar foraging puzzle [41]. If the tits do not quickly forget their preferences, it follows that turnover is required to successfully resample the behavioral space.

As it stands, there is very limited experimental literature directly related to how population turnover might influence the efficiency or adaptiveness of culturally transmitted behaviors. Previous studies on cultural efficiency in chimpanzees used one-on-one demonstrations, or static groups, and largely failed to find evidence for any selection for efficiency [18, 19]. If the social system of the chimpanzees was amenable population turnover, a capacity for selection for efficiency might be more evident. Recent work has suggested strong conformity of immigrant females in the wild after dispersal, who adopted inefficient local nut-cracking techniques [53], presumably because the cost of an inefficient behavior is marginal compared to the benefit of acceptance within a group. In comparison, the immigrant tits were conformist in that they usually first learned the most frequently produced solution. However, they were not conformist enough that it prevented them from sampling and adopting the minority efficient solution. Interestingly, tits appear strongly conformist in the wild [41]. It's possible that the artificial, transient social groups of our experiment might have also diminished the importance of conformity for the tits.

The singular study we have identified that found a similar effect of population turnover (rather than transmission chains as a proxy for turnover) comes from translocation studies of blue-headed wrasse, a fish that maintains spawning sites through cultural inheritance. When all members of a wrasse population were replaced with naive individuals, they resampled the set of potential spawning sites and selected different mating sites from the original population [36]. This particular experiment was designed to test whether spawning sites were culturally inherited, so the entire population was turned over at once. Although, the successful

resampling of sites by naive fish appears analogous to that observed in immigrant tits. Given we find this similar example from a species whose last common ancestor with the great tit was 430 mya, and given that the effect of turnover is replicated across learning parameter values in silico, we propose that population turnover might provide a broad, substrate independent mechanism for resampling socially transmitted information and potentially moving populations towards more adaptive local optima.

The cultural evolution of more refined, efficient, behaviors has been considered as a precursor to more complex, human-like forms of cumulative culture [12]. While we do not argue that this experiment demonstrates CCE in tits, our results add to the growing evidence that cultural evolution of efficiency is possible in animals, and extend this to a new behavior domain (foraging behavior). This small passerine species is capable of moving the cultural ratchet forward by one tick, fulfilling a key description of CCE: a behavior innovated by an individual spreads within a group, and stays in place until future individuals make further modifications that are more efficient, which then spread and supplant the previous behavior [8]. The design of the puzzle box made it such that the birds could not innovate and spread an improvement more than once. However, it is not a far leap to imagine that this process of cultural selection for efficiency might have repeated itself in an extended experiment. In the wild, great tits form fission-fusion foraging flocks, whose membership is constantly changing. Perhaps this demographic churn facilitates rapid cultural evolution of locally adaptive behavior, although additional work on natural populations is needed to test this.

Still, it remains an open question as to why repeated instances of selection for efficiency are so rarely observed in non-human species in the field. Our results suggest this is not due to an inability to select for more efficient variants; rather behavioral conservatism can be mitigated by population turnover. Yet turnover also presents a real cost: memory loss through turnover may be detrimental to the population depending on their knowledge state. One turnover population exhibited a complete extinction of the solving behavior, suggesting that cultural benefits of turnover may also carry a corresponding risk of the potential loss of all knowledgeable individuals and subsequent extinction of cultural traits. This effect could be exacerbated if the underlying resource the cultural behavior is exploiting is ephemeral and inter-generational turnover is particularly high.

This paper illustrates how dynamic populations are more likely to select for efficient cultural behaviors, and details an overlooked mechanism: naive immigrants are more successful at adopting more optimal cultural behaviors that were originally innovated by more experienced residents. This mechanism provides a pathway towards the evolution of efficient, locally adaptive cultural behaviors in populations of behaviorally conservative individuals. If a particular population of animals is capable of maintaining cultural traditions, we might well expect to find an effect of turnover rate on their cultural state.

4 Methods

4.1 Study System

This study was conducted using captive populations of wild-caught great tits, a small passerine bird that is widely distributed across Europe. This experiment attempted to recreate natural foraging flocks by using micro-populations consisting of six individuals. This membership size was based on the average size of foraging flocks in the wild. Sex was well balanced (94 female, 87 male), however age was more difficult to balance (118 juveniles, 63 adults) due to unpredictable variation in the week-to-week catching of wild birds. A total of 181 tits were used in the experiment, with 54 participating in the static condition, and 127 in the turnover condition. The extra bird in the turnover condition was a replacement for a bird which had died during the experiment. Tits were caught using mist nets at 7 different sites, all within a range of 10 km from The Max Planck Institute of Animal Behavior (see SI for map). When caught, birds were fitted with a metal identification ring, as well a passive integrated transponder (PIT) tag (IB Technology). Birds were aged and sexed by plumage. Birds were released back into the wild at the location where they were captured. All work adhered to relevant ethics approval by Regierungspraesidium Freiburg (35-9185.81/G-17/168).

4.2 Experimental Apparatus

The foraging puzzle box consisted of an acrylic box with a bidirectional sliding door. The door (4 × 6-cm) could be slid to the left or right to reveal a 1-cm hole through which the bird could access a reservoir of meal worms. The hole was off-center by .5-cm, such that pushing the red side of the door would be more efficient (1-cm from door-edge to hole) than pushing the blue side (2-cm to hole). This difference provided a significant challenge, given the tit's size. Average solve speed confirms that it was faster to use the red side of the door.

Data collection from the puzzle-boxes was fully automated to minimize disturbance of the birds during the experiment. Puzzle boxes included a micro-computer, a stepper motor to automate the door return mechanism, and an external printed circuit board to handle RFID reading (PriorityOne). An RFID antenna mounted under the perch recorded identities of birds from their PIT tag, as well as their arrival and departure times.

Movement of the door was measured using sensors, which were triggered once the door moved past a position that would allow the bird to access the hole. Once an sensor was triggered and the solving bird left, the door would automatically close either after 3 scrounge attempts, or after 3 seconds, whichever came first. In the 2020 field season (accounting for 12 of 18 total populations), the puzzle also recorded video of the visiting birds, from which identities could be resolved using a bar-code attached to the bird using leg loops [54]. Bar-code data was used 1) to resolve identities of solves without a positive identity from RFID data, 2) to measure TTS.

4.3 Experimental Design

Populations were initially supplied with 1 tutor each that had been trained to produce the inefficient solution. During training, tutors interacted with the puzzle with door open, and then we progressively closed the door in subsequent trials. Once introduced, the populations were given a period of 12 days before the first turnover event to allow the inefficient tradition to become established. During this time, the efficient solution was blocked to promote uptake of only the inefficient solution into the population. Ideally, the efficient solution wouldn't be accessible during this period, in order to prevent birds from receiving null payoffs for trying it while blocked, although this was not possible with the design of the puzzle box. However, this treatment did not appear to affect birds' ability to innovate the efficient solution once unblocked, as evidenced by Results C. Tutors from two populations were re-used as tutors in 2 separate populations, as tutors in these populations had failed to produce the inefficient solution by the first turnover event, and it was not innovated by naive individuals.

To manipulate population turnover, experimental populations were either *static*, where the population consisted of the same set of 6 birds for the duration of the experiment, or *turnover*, where 2 birds within generation were randomly replaced with new, naive birds every 7 days. The first 6 birds in a population were considered Generation 1, with incoming birds considered as Generation 2. This stochastic, within generation turnover was different than a standard diffusion chain, where the oldest participants are always removed, and was meant to replicate conditions similar to turnover of foraging flock composition, rather than generational turnover. The experiment ran for approximately 5 weeks, with a total of 4 turnover events (T1-T4). Experimental day 0 was designated as 7 days prior to T1, and was the day in which the behavior was socially diffused in all populations.

4.4 Statistical Analyses

4.4.1 Criteria for solvers and innovators

Solvers were defined as birds who had ≥ 20 recorded solutions in all of the following analyses.

Innovators were defined in two possible ways. 1) Innovators were birds who were the very first to produce the efficient solution within a population. 2) We calculated the distribution of time between efficient solves (Δt). Innovators were also birds who produced the efficient solution with an outlier Δt (≥ 63 hours), as we deemed it unlikely that the innovator had observed the previous efficient solution.

4.4.2 Time-to-solve

TTS for solutions were calculated similarly to [41], as the time difference between a bird's arrival (recovered from RFID data and QR data) and the time when the door is detected as open (recovered from sensor data). Both RFID and QR reads of birds were noisy and subject to error, so when both RFID and QR data were available for a particular solution, TTS was the average of both measurements.

4.4.3 Individual-level improvement in performance

To make 1C, data was subset to solvers. To assess the difference between inexperienced and experienced solvers, within each bird the data was divided into the first 25% of solutions produced and the remaining 75% of solutions produced. Mean values and 95% confidence intervals were generated by non-parametric

bootstrapping using the `mean.cl.boot` function [55]. These mean values were then compared using a Wilcoxon rank sum test.

To assess differences in solving performance over time within individuals, data was subset to solvers, non-tutors and TTS_{60s}. 59 solutions longer than 60 seconds were excluded, and 18,534 were excluded due to unrecorded solve speeds. We then used a Linear Mixed Effects model in which $\log(TTS + 1)$ was predicted by a Z-scaled within bird trial number, with solution type, population, ID and year controlled for as random effects. TTS was logged to account for non-normality, with 1 added since the shortest TTS value was 0. Exponentiated beta coefficients are presented to make them meaningful after the log transformation.

4.4.4 Results A: Cultural selection for efficiency

To quantify the difference in selection for efficiency between conditions, we first subset the data to solvers only, and experimental day ≥ 6 , as the efficient solution was blocked before this time to ensure the establishment of the inefficient solution before the first turnover event. One static population was excluded, as no social learning had occurred in this population. We then used a logistic GLMM with the probability of producing an efficient solve regressed on age, sex, and an interaction between experimental day and condition. Population and year were included as nested random intercepts. Individual birds were not included as random intercepts, since the intercept value represented experimental day 7. Since many of the birds in the turnover condition were not present on this day, a random intercept would be creating estimates for large proportions of data which didn't exist.

4.4.5 Results B: Condition-level differences in solving performance

To analyze how much each condition had improved by the end of the experiment data was subset to solutions produced by solvers with TTS_{60s}, tutors excluded. We ran a LMM [56, 57] in which $\log(TTS + 1)$ was predicted by sex, age, and experimental day*condition interaction. Population nested within year were included as random effects. Solution type was not included as a random effect, as it is perfectly correlated with W1 of the experiment, and produces a singular model fit. To compare differences in TTS between conditions in Week 5, a Wilcoxon test was used to compare means with the alternative hypothesis that turnover was less than static condition.

4.4.6 Results C: Innovation levels

Comparing a true innovation rates proved difficult, as once an innovation is adopted by a population and produced with any frequency, subsequent true innovation events would be masked due to the assumption that a frequently produced solution would be observed by all birds. Instead, we analyzed differences in innovation timing between conditions using Wilcoxon tests in which the alternative hypothesis was that the mean for the turnover condition was less than the static condition. We compared 1) days solving before innovation, 2) days of exposure to puzzle before innovation, 3) total elapsed experimental days before innovation.

4.4.7 Results D: Behavioral conservatism

To test whether experience or conformity played a more relevant role in the failure of birds to switch from inefficient to efficient we subset to solvers that had produced both solutions. To determine which birds failed to switch, the distribution of the last 10% of their solves were analyzed. If inefficient solutions outnumbered efficient solutions, then that bird was marked as failed to switch. Experience was calculated as days solving before experiencing the efficient solution. To measure conformity, we measured the proportion of socially observed inefficient solutions of all solutions produced the day that bird first produced the efficient solve. There was not enough data from each population to use a GLMM to control for population level differences, so we instead used a GLM in which failure to switch was predicted by age, sex, experience with the inefficient solution prior to sampling the efficient, and the proportion of socially observed inefficient solutions.

4.5 Agent Based Model

Due to the practical limitations of replicating treatments with wild-caught birds, an agent based model was used to 1) test the verbal hypothesis relating demographic features to cultural outcomes, and 2) provide virtual power to the experiment. Simulations were run across parameter space under these conditions to test the hypothesis that population turnover would increase the area of parameter space in which populations of agents would converge on a more efficient solution to solve a foraging puzzle.

The model is comprised of agents capable of learning and producing foraging behaviors. Agents learned both through observation and personal experience, updating their probabilities of producing either solution according to a modified experience-weighted attraction (EWA) learning model [58, 48, 21, 46, 47]. In terms of temporal scale, the time-step approximates one experimental day. The simulation lasts for 35 time-steps, simulating the 5 weeks of the experiment. Each time step, agents interacted with, and observed others produce behaviors. There were two possible behaviors, with agents receiving scalar payoffs of 10 and 20 for each solution respectively. The behavior with the higher reward signal represented the efficient solution in the experiment. To imitate the blocked efficient solution for the establishment period, payoffs were set as efficient: 0, inefficient: 10 for the first 7 time steps. Turnover occurred every 7 time-steps in the same manner as the experiment.

Agents could be in 1 of 2 states: knowledgeable or naive. If knowledgeable, agents produced solution behaviors. If naive, agents simply observed others' solutions, and this information was recorded in their social memory, influencing behavioral decisions after an agent became knowledgeable in proportion to how much agents would weight social information (see parameters below). Possible transition from naive to knowledgeable states occurred once per time-step according to the conditional probability of learning function.

The values for the conditional probability of learning function were determined using latency-to-learn data gathered from the experiment. Data was subset to solvers and non-tutors, and right censored for individuals that never learned. A log-normal parametric survival analyses was performed in R using the rms package Rcore,rms with only the intercept as a predictor, such that a hazard function considering all birds was estimated (log-normal model had a higher log-likelihood than a Weibull model). From this hazard function the conditional probability of learning was estimated. Agents each had a count variable t to keep track of how many time-steps each bird had been exposed to the puzzle, used to return the probability of transitioning from naive to knowledgeable state $R(t)$. This function produces the probability that an individual would acquire the behavior within the interval $[t, t + 1]$ given that it had not acquired it previously. The shape of this probability curve approximates the hazard rate, and was calculated as

$$p(\text{learn}|\text{not learned yet}) = \frac{(R(t) - R(t + 1))}{R(t)} \quad (1)$$

where $R(t)$ is the survival function. To test whether this particular function was an important mechanism in the effect of turnover, we re-ran the sensitivity analysis with the reverse of this function, such that probability of learning increased with time. We found no major differences.

This model allows for variation of the four following learning parameters within each agent:

- g_i : Current payoff sensitivity (weight given to most recently received payoff vs. previous payoffs). Values range from $[0.1, 0.9]$ at increments of 0.2.
- s_i : Social cue sensitivity (weight given to social information vs. personal information). Values ranged from $[0.1, 0.9]$ at increments of 0.2.
- τ : Conservatism (insensitivity to differences in payoffs). Values ranged: mean ranged from $[1, 3]$ at increments of 0.2.
- λ : Conformity (sensitivity to the most frequently observed behavior): Value set at 5.

τ was set above 1, as tits were observed to be behaviorally conservative, rarely alternating between choices. λ was chosen as an arbitrary value greater than 1, to provide the agents with a conformist bias previously demonstrated [21].

For each point in the parameter space described above, data was recorded from 100 simulations per condition. In each simulation, agents were initialized with identical parameter values. One agent was the designated "tutor", and was programmed with a high attraction score for the inefficient solution.

Within one time-step, each knowledgeable individual would produce 100 behaviors, which approximated the real average from the experiment (98 solutions). The order of individuals was not randomized, as social information was not processed until the end of a time step. When a knowledgeable individual produced a solution, they received a payoff. Next, they updated their attraction scores, and finally updated the probabilities for their next choice of solution. At the end of each time-step, the following data was recorded: time-step, condition, mean values of parameters, frequency of solves of each type, and number of solvers. Also, the social cue matrices were updated for all agents using the frequencies of behaviors produced in that time-step. Their choice probabilities were then updated to reflect this new social information. If there were still naive agents in the population, each naive agent may acquire the solving behavior and transition to a knowledgeable state according to $R(t)$. The remaining naive agents update their acquisition probabilities

for the next time-step. If the simulation was in the turnover condition, then every 7 time-steps there was a turnover event.

5 Data Availability

Code and data for statistical analyses and main text figures, as well as code to replicate the agent based model is available at <https://github.com/michaelchimento/DemographicTurnoverEfficiency>.

6 Acknowledgements

Thank you to Anne Kandler, Brendan Barrett, Dora Biro, members of the Cognitive and Cultural Evolution Lab at MPIAB (Barbara Klump, Sonja Wild, Julia Penndorf, Stephen Tyndel, Simeon Smeele), and Elham Nourani for their input and suggestions for the manuscript and assisting in catching birds. Also, thank you to the MPIAB animal caretaker team and veterinarians for their help in running the experiment and keeping the tits happy and healthy.

7 Author Contributions

MC and LA designed the experiment. MC and GAN designed the automated data capture system. MC performed the experiment with support from GAN and LA. MC designed, coded the agent based model. MC analyzed the data with support from LA. MC wrote the first draft of the manuscript, and all authors edited, commented and added to it.

8 Competing Interests

The authors declare no competing interests, financial or otherwise.

9 Materials Correspondence

Correspondence should be addressed to Michael Chimento, e-mail: mchimento@ab.mpg.de

References

- [1] Andrew Whiten et al. "The extension of biology through culture". In: *Proceedings of the National Academy of Sciences* 114.30 (2017), pp. 7775–7781.
- [2] Eva Jablonka and Marion J Lamb. *Evolution in four dimensions, revised edition: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. MIT press, 2014.
- [3] Caroline Schuppli and Carel P van Schaik. "Animal cultures: How we've only seen the tip of the iceberg". In: *Evolutionary Human Sciences* 1 (2019).
- [4] Nicole Creanza, Oren Kolodny, and Marcus W Feldman. "Cultural evolutionary theory: How culture evolves and why it matters". In: *Proceedings of the National Academy of Sciences* 114.30 (2017), pp. 7782–7789.
- [5] Alex Mesoudi, Andrew Whiten, and Kevin N Laland. "Perspective: Is human cultural evolution Darwinian? Evidence reviewed from the perspective of The Origin of Species". In: *Evolution* 58.1 (2004), pp. 1–11.
- [6] Alex Mesoudi. "Pursuing Darwin's curious parallel: Prospects for a science of cultural evolution". In: *Proceedings of the National Academy of Sciences* 114.30 (2017), pp. 7853–7860.
- [7] Robert Boyd and Peter J Richerson. *Culture and the evolutionary process*. University of Chicago press, 1988.
- [8] Michael Tomasello. "The human adaptation for culture". In: *Annual review of anthropology* 28.1 (1999), pp. 509–529.

- [9] Lucy M Aplin. "Culture and cultural evolution in birds: a review of the evidence". In: *Animal Behaviour* 147 (2019), pp. 179–187.
- [10] Andrew Whiten. "A second inheritance system: the extension of biology through culture". In: *Interface Focus* 7.5 (2017), p. 20160142.
- [11] Takao Sasaki and Dora Biro. "Cumulative culture can emerge from collective intelligence in animal groups". In: *Nature communications* 8.1 (2017), pp. 1–6.
- [12] Alex Mesoudi and Alex Thornton. "What is cumulative cultural evolution?" In: *Proceedings of the Royal Society B: Biological Sciences* 285.1880 (2018), p. 20180712.
- [13] Claudio Tennie, Josep Call, and Michael Tomasello. "Ratcheting up the ratchet: on the evolution of cumulative culture". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364.1528 (2009), pp. 2405–2415.
- [14] Robert Boyd, Peter J Richerson, and Joseph Henrich. "The cultural niche: Why social learning is essential for human adaptation". In: *Proceedings of the National Academy of Sciences* 108.Supplement 2 (2011), pp. 10918–10925.
- [15] Brett R Jesmer et al. "Is ungulate migration culturally transmitted? Evidence of social learning from translocated animals". In: *Science* 361.6406 (2018), pp. 1023–1025.
- [16] Thomas Mueller et al. "Social learning of migratory performance". In: *Science* 341.6149 (2013), pp. 999–1002.
- [17] Noemie Lamon et al. "Wild chimpanzees select tool material based on efficiency and knowledge". In: *Proceedings of the Royal Society B* 285.1888 (2018), p. 20181715.
- [18] Sarah Marshall-Pescini and Andrew Whiten. "Chimpanzees (*Pan troglodytes*) and the question of cumulative culture: an experimental approach". In: *Animal cognition* 11.3 (2008), pp. 449–456.
- [19] Christine Hrubesch, Signe Preuschoft, and Carel van Schaik. "Skill mastery inhibits adoption of observed alternative solutions among chimpanzees (*Pan troglodytes*)". In: *Animal cognition* 12.2 (2009), p. 209.
- [20] Lydia M Hopper et al. "Chimpanzees' socially maintained food preferences indicate both conservatism and conformity". In: *Animal Behaviour* 81.6 (2011), pp. 1195–1202.
- [21] Lucy M Aplin, Ben C Sheldon, and Richard McElreath. "Conformity does not perpetuate suboptimal traditions in a wild population of songbirds". In: *Proceedings of the National Academy of Sciences* 114.30 (2017), pp. 7830–7837.
- [22] Hal Whitehead and David Lusseau. "Animal social networks as substrate for cultural behavioural diversity". In: *Journal of theoretical biology* 294 (2012), pp. 19–28.
- [23] Joseph Henrich. "Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case". In: *American antiquity* 69.2 (2004), pp. 197–214.
- [24] Daniel Nettle. "Social scale and structural complexity in human languages". In: *Phil. Trans. R. Soc. B* 367.1597 (2012), pp. 1829–1836.
- [25] Andrea B Migliano et al. "Hunter-gatherer multilevel sociality accelerates cumulative cultural evolution". In: *Science Advances* 6.9 (2020), eaax5913.
- [26] Maxime Derex and Robert Boyd. "Partial connectivity increases cultural accumulation within groups". In: *Proceedings of the National Academy of Sciences* 113.11 (2016), pp. 2982–2987.
- [27] Daizaburo Shizuka and Allison E Johnson. "How demographic processes shape animal social networks". In: *Behavioral Ecology* 31.1 (2020), pp. 1–11.
- [28] Nicolas Fay et al. "Increasing population size can inhibit cumulative cultural evolution". In: *Proceedings of the National Academy of Sciences* 116.14 (2019), pp. 6726–6731.
- [29] Lindell Bromham et al. "Rate of language evolution is affected by population size". In: *Proceedings of the National Academy of Sciences* 112.7 (2015), pp. 2097–2102.
- [30] Maxime Derex et al. "Experimental evidence for the influence of group size on cultural complexity". In: *Nature* 503.7476 (2013), pp. 389–391.
- [31] Marius Kempe and Alex Mesoudi. "An experimental demonstration of the effect of group size on cultural accumulation". In: *Evolution and Human Behavior* 35.4 (2014), pp. 285–290.
- [32] Andrew Foss-Grant, Sharon Bewick, and William F Fagan. "Social transmission of migratory knowledge: quantifying the risk of losing migratory behavior". In: *Theoretical Ecology* 11.3 (2018), pp. 257–270.

- [33] Kristina L Paxton et al. "Loss of cultural song diversity and the convergence of songs in a declining Hawaiian forest bird community". In: *Royal Society open science* 6.8 (2019), p. 190719.
- [34] Kevin A Parker et al. "The effects of translocation-induced isolation and fragmentation on the cultural evolution of bird song". In: *Ecology Letters* 15.8 (2012), pp. 778–785.
- [35] Paola Laiolo and José Luis Tella. "Erosion of animal cultures in fragmented landscapes". In: *Frontiers in Ecology and the Environment* 5.2 (2007), pp. 68–72.
- [36] Robert R Warner. "Traditionality of mating-site preferences in a coral reef fish". In: *Nature* 335.6192 (1988), pp. 719–721.
- [37] Ella F Cole, Dominic L Cram, and John L Quinn. "Individual variation in spontaneous problem-solving performance among wild great tits". In: *Animal Behaviour* 81.2 (2011), pp. 491–498.
- [38] James Fisher and R. A. Hinde. "The opening of milk bottles by birds". In: *British Birds* 42 (1949), pp. 347–357.
- [39] Louis Lefebvre and Lucy M Aplin. "Social learning and innovation". In: *Avian cognition* (2017), p. 93.
- [40] Lucy M Aplin, Ben C Sheldon, and Julie Morand-Ferron. "Milk bottles revisited: social learning and individual variation in the blue tit, *Cyanistes caeruleus*". In: *Animal Behaviour* 85.6 (2013), pp. 1225–1232.
- [41] Lucy M Aplin et al. "Experimentally induced innovations lead to persistent culture via conformity in wild birds". In: *Nature* 518.7540 (2015), p. 538.
- [42] S Bouwhuis et al. "Great tits growing old: selective disappearance and the partitioning of senescence to stages within the breeding cycle". In: *Proceedings of the Royal Society B: Biological Sciences* 276.1668 (2009), pp. 2769–2777.
- [43] RH McCleery et al. "Components of variance underlying fitness in a natural population of the great tit *Parus major*". In: *The American Naturalist* 164.3 (2004), E62–E72.
- [44] Damien R Farine et al. "The role of social and ecological processes in structuring animal populations: a case study from automated tracking of wild birds". In: *Royal Society Open Science* 2.4 (2015), p. 150057.
- [45] Damien R Farine et al. "Collective decision making and social interaction rules in mixed-species flocks of songbirds". In: *Animal behaviour* 95 (2014), pp. 173–182.
- [46] Brendan J Barrett, Richard L McElreath, and Susan E Perry. "Pay-off-biased social learning underlies the diffusion of novel extractive foraging traditions in a wild primate". In: *Proceedings of the Royal Society B: Biological Sciences* 284.1856 (2017), p. 20170358.
- [47] Rachel Kendal et al. "Chimpanzees copy dominant and knowledgeable individuals: implications for cultural diversity". In: *Evolution and Human Behavior* 36.1 (2015), pp. 65–72.
- [48] Richard McElreath et al. "Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 363.1509 (2008), pp. 3515–3528.
- [49] A. Mesoudi. *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. University of Chicago Press, 2011. ISBN: 9780226520445. URL: https://books.google.de/books?id=8KSI0z1%5C_d6sC.
- [50] Shaul Markovitch and Paul D Scott. "The role of forgetting in learning". In: *Machine Learning Proceedings 1988*. Elsevier, 1988, pp. 459–465.
- [51] Hannah Cornish et al. "Sequence Memory Constraints Give Rise to Language-Like Structure through Iterated Learning". In: *PloS one* 12.1 (2017), e0168532.
- [52] Simon Kirby et al. "Compression and communication in the cultural evolution of linguistic structure". In: *Cognition* 141 (2015), pp. 87–102.
- [53] Lydia V Luncz et al. "Costly culture: differences in nut-cracking efficiency between wild chimpanzee groups". In: *Animal Behaviour* 137 (2018), pp. 63–73.
- [54] Gustavo Alarcón-Nieto et al. "An automated barcode tracking system for behavioural studies in birds". In: *Methods in Ecology and Evolution* 9.6 (2018), pp. 1536–1547.
- [55] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN: 978-3-319-24277-4. URL: <https://ggplot2.tidyverse.org>.

- 609 [56] Douglas Bates et al. "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Soft-*
610 *ware* 67.1 (2015), pp. 1–48. DOI: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- 611 [57] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical
612 Computing. Vienna, Austria, 2018. URL: <https://www.R-project.org/>.
- 613 [58] Colin Camerer and Teck Hua Ho. "Experience-weighted attraction learning in normal form games". In:
614 *Econometrica* 67.4 (1999), pp. 827–874.

Figure 1: Diagram of puzzle box alongside a photograph of a great tit at the puzzle, and average time-to-solve for pushing the door one way or the other (efficient: N=44 birds, 34,444 solutions; inefficient: N=50 birds, 78,616 solutions; tutors excluded; confidence intervals computed by non-parametric bootstrapping). Open circles from first quartile of solutions produced by an individual, and closed circles are second-fourth quartile of solutions. **A)** Each population was provided with one puzzle box that had a 4 cm wide door which gave access to a silo of mealworms. The access point was offset from the center of the door by .5 cm, such that pushing from the left (red) side is more efficient. Internal electronics recorded identity and solution type and automatically closed the door after each solve. **B)** The black perch below the door contained an RFID antenna to record identities from the unique transponder (PIT) tags on birds, and a camera above recorded identity from barcodes placed on the backs of birds. Both systems allowed calculation of time-to-solve. **C)** The efficient solution was significantly faster to solve. Solving birds significantly reduced their time-to-solve in both solutions with experience, suggesting a general process of refinement independent of social influence.

Figure 2: Three selected populations' daily average time-to-solve per solution type over experimental day (efficient: red, inefficient: blue; CI computed by non-parametric bootstrapping; turnover events indicated by T1-T4), and below, solutions produced by each bird within these populations over experimental day, with type (color), frequency (size), and innovation events (triangle) ($n = 20$ birds, 29,801 solutions). These populations exemplify selection for efficiency, failure of selection for efficiency, and cultural extinction. **A)** A turnover population that selected for the efficient solution. An experienced bird innovated on day 11, and produced the efficient solution with a low frequency until it was widely adopted by naive immigrants after the third turnover. **B)** A static population that improved upon the established inefficient tradition. Two birds independently innovated the efficient solution, yet the population failed to select for it. **C)** Population 17 was a turnover population in which solving behavior was extinguished after the last remaining solver left the population in the third turnover. **D)** Each point represents the frequency and type of solution produced by an individual bird on a given day, with innovation events marked as triangles. Pop. 6 exemplifies the lengthy latency period between innovation and adoption, in which experienced birds innovate the efficient solution, but it is only adopted by less experienced incoming birds. Pop. 11 is a typical static population, in which the efficient solution is innovated and produced at low frequencies, but fails to invade the cultural repertoire. The gradual removal of knowledgeable birds in pop. 17 leads to behavioral extinction.

Figure 3: Above, the proportion of solution types (efficient: red, inefficient: blue) faceted by condition, and below, average time-to-solve over experimental day for each condition (static: black, turnover: yellow; 95% confidence intervals computed by non-parametric bootstrapping; Y-axis log2 coord. transformation), with turnover events shown by dashed lines. **A)** The proportion of solutions that are inefficient or efficient in all populations on a given experimental day (static: $n = 28$, total solutions = 88,224; turnover: $n = 70$, total solutions = 84,518). All populations started by producing inefficient solutions. Static populations produced mostly inefficient solutions for the majority of the experiment. In contrast, efficient solutions fully invaded the cultural repertoire of nearly all turnover populations by the end of the experiment. **B)** Mean time-to-solve for static and turnover conditions over experimental day. In the last week, turnover populations were solving 0.3s faster than static populations, an increase in efficiency driven by the widespread cultural selection for the efficient solution. (static: $n = 21$, total solutions = 52,458; turnover: $n = 61$, total solutions = 60,735).

Figure 4: Agent based model results showing that population turnover increases the proportion of efficient behaviors in the final time step of the simulation across 3 important parameter values ($N = 54,179$ simulations; excluded 821 simulations where behavior went extinct). **A)** Increased social cue sensitivity decreases selection for efficiency, as the initial tutor agent's preference is more influential. Turnover is more likely to have a higher proportion of efficient solutions in the final time step, even in populations that heavily weight social information. **B)** Current payoff sensitivity generally increases selection for efficiency, as the higher payoffs received for the efficient solution is more important to agents. **C)** Behavioral conservatism, here implemented as sensitivity to differences in attraction score, generally decreases selection for efficiency, as agents are satisfied with producing the established inefficient behavior despite having received better payoffs. Turnover mitigates this effect by introducing new agents who do not already have attraction scores for either solution.