A tale of two birds: cognitive simplicity drives collective route improvements in homing pigeons

Shoubhik Chandan Banerjee,[1,*] Fritz A. Francisco,[1] and Albert B. Kao[1,†]

[1] *Department of Biology, University of Massachusetts Boston, Boston, MA 02125, USA*

Cognitive abilities are central to how animals navigate complex environments. Beyond individual cognition, group living can also enhance navigation by pooling individually acquired information. One way this may be achieved is by following experienced leaders, which requires recognizing expertise within group members. Alternatively, accurate decisions could also emerge without expert opinions, through simpler mechanisms like the 'wisdom of crowds' principle that average out individual biases. Consequently, collective navigation strategies range from cognitively complex to simple, and yet, the prevalence or interplay of different collective strategies in nature remains unexplored. In this study, we asked: what is the navigation mechanism, requiring minimal cognitive demands, that is necessary and sufficient to quantitatively replicate the experimental results of a 2017 study on homing pigeons (*Columba livia*), which showed that sequential chains of bird pairs flying home — similar to a game of telephone — led to shorter homing routes compared to control birds flying individually or in fixed pairs. Our results show that the experimental data aligns closely with the simplest strategy — route averaging. Surprisingly, the complex mechanism of selectively propagating the best flight through social learning offered no additional advantage. We further observed that mixed strategies, although not supported by the experimental data, in theory combined advantages from both averaging and active selection of better routes, resulting in even greater performance. Hence, our results highlight the potential for future research to investigate selective pressures shaping the evolution of cultural learning and trade-offs among different decision mechanisms theoretically available to social animals in nature.

INTRODUCTION

For decades, humans have been intrigued by the remarkable ability of animals to navigate their environment [1–4]. Fascinating examples include the annual pole-to-pole migration of Arctic terns [5, 6] and sooty shearwaters [7], that achieve flight distances spanning thousands of kilometers or the precise navigation of anadromous salmon, swimming from the open ocean to return to their natal spawning grounds [8, 9]. To achieve these marvelous feats, animals utilize a variety of environmental information to guide their journey at different length scales – for example, polarized light [10] or geomagnetic cues [11, 12] can roughly guide animals across long length scales, while olfactory cues [13] and visual landmarks [14] are more useful at shorter length scales.

Animals living in groups can further reduce directional uncertainty through social interactions. One way this can be achieved is by the direct social transfer of information. In such cases, the group leader (or a few members of the group, who are often older or more experienced) guide the rest of the group in their preferred direction via a leader-follower relationship [15–21]. For example, dominant female meerkats have higher influence in leading group movements, especially movement turn and speed, due to their experience and frontmost position in the group

* shoubhik.banerjee001@umb.edu
† albert.kao@umb.edu

[22]. Through these leadership dynamics, expertise can propagate through the group, allowing less knowledgeable individuals to exhibit improved navigational performance.

An alternative socially-mediated mechanism that can improve migration efficiency is the 'wisdom of crowds', where large groups achieve optimum decision by averaging individual opinions [21]. This mechanism benefits from the law of large numbers, whereby individual errors cancel each other out with increasing group size, resulting in greater collective accuracy [23, 24] (assuming that the distribution of opinions is roughly centered around the optimal value [25]). In contrast to the leadership mechanism described above, here the group outcome is emergent. In other words, the collective decision may be highly accurate even if there are no knowledgeable individuals in the group at all. Furthermore, this mechanism has low cognitive requirements, since it does not require group members to recognize differences in knowledge or experience. In both humans and non-human animals, across different behavioral contexts, the wisdom of crowds has been shown to improve outcomes as group sizes increase [26–31].

However, the two mechanisms described above are not mutually exclusive. Weighted average strategies [32, 33], where more experienced, knowledgeable or confident individuals hold greater influence in the collective decision, but other individuals still have some decision weight, may capture the benefits of both existing knowledge and emergent solutions. While sociality can theoretically support a multitude of distinct and non-mutually-exclusive strategies for improving navigation efficiency, the extent to which animals use each strategy in practice remains unclear. In particular, how the optimal strategy depends on the type of decision that animals face, the cognitive ability of the animals, and other details of the species' life history, remains poorly understood.

Learning through time

Many animals usually make repeated trips between two locations in a given period of time. At small spatial scales, these journeys may be between known locations within their home range [18], or at larger spatial scales, they may be seasonal migrations that occur annually [34]. A solitary individual may improve its performance across bouts by sampling multiple routes and preferentially remembering better ones. However, the individual faces a classic exploration-exploitation trade-off [35] between spending time exploring new routes (which may be associated with unknown and possibly severe risks) or utilizing previously known routes that are less costly. Furthermore, an individual can attempt a trip only a finite number of times, ultimately limited by its lifespan, which sets an upper bound on the ability of a single individual to improve its learned route over time.

One way in which animals can overcome the finite time horizon set by an individual lifespan is by genetically encoding migratory information — such as when and where to migrate — so that it can be inherited by future generations [36, 37]. However, genetically encoding information can be slow and susceptible to environmental variations [38–40]. Alternatively, traveling in a group can allow less-informed individuals to incorporate information from others into their own learned route, thereby permitting information to pass across individuals (and time). When such social learning is coupled with the ability to recognize beneficial solutions and prune out costly mistakes, improved solutions can be preferentially spread horizontally and/or vertically. Such 'ratchet'-

like improvements are referred to as cumulative cultural evolution (CCE), in recognition of its functional similarity to genetic evolution (i.e., variation, heritability, and selection). CCE has been observed in humans and some non-human animals [40–44], including primates [45–47], birds [48–51] and cetaceans [52–54]. However, CCE is usually inferred as the underlying process when improvements 'accumulate' over time, even though the social learning mechanism(s) that lead to these improvements have not been directly observed (see [55] for core criteria required for CCE and [48, 56] for examples of such cases). Yet, as previously discussed, alternative mechanisms such as the wisdom of crowds could in principle also cause an increase in task performance, without the need for other components associated with CCE [55].

In this study, we build upon one such example, by Sasaki and Biro [49], in which homing pigeons *(Columba livia)* were shown to improve their navigational performance over time but the mechanism by which they do so remains a mystery. Homing pigeons are a highly social species that have been extensively studied for their ability to develop idiosyncratic routes and recall them across repeated flights, making them an ideal model organism for studying navigational strategies [16, 17, 49, 57–59]. In their experiment, Sasaki and Biro created 'chains' of birds, similar to a game of telephone, and allowed them to fly back home repeatedly from a release site 8.4 km away. In the experimental condition, each chain was composed of five 'generations', with each generation consisting of 12 flights home. In each generation, an experienced bird, having developed knowledge about the homing task from the previous generation, was paired with a naïve bird that lacked this information. At the end of the generation, the experienced bird was removed and a new naïve bird was paired with the remaining (now-experienced) bird (see *Methods* for details). As control conditions, solo and fixed pairs of birds were flown for the same total amount of flights as in the experimental condition (60 flights) (Fig. 1A). To quantify the performance of the birds, the route efficiency, calculated as the ratio of the beeline distance (straight line distance between the release and home sites) to the actual distance traveled by the pair, was measured (a value of 1 indicates a beeline flight, while a value of 0 indicates an infinitely long flight).

The study found that the experimental chains of birds significantly outperformed both the solo and fixed pair controls by the end of the fifth generation (Fig. 1B). The researchers proposed CCE as the mechanism driving this result, where chains of birds improve their routes by exploring different options and selectively choosing better ones. Valentini *et al.* examined the trade-offs between exploration and exploitation of new routes by the naïve and experienced birds [60], showing that both birds had an equally likely chance of initiating collective explorations. Despite these insights, a detailed mechanistic understanding of how route improvements emerge in pigeon chains is still lacking. For instance, it remains unclear what (social and/or environmental) information birds are sensing and using, and what kinds of computations they are performing on that information, to produce the observed increase in route efficiency across generations.

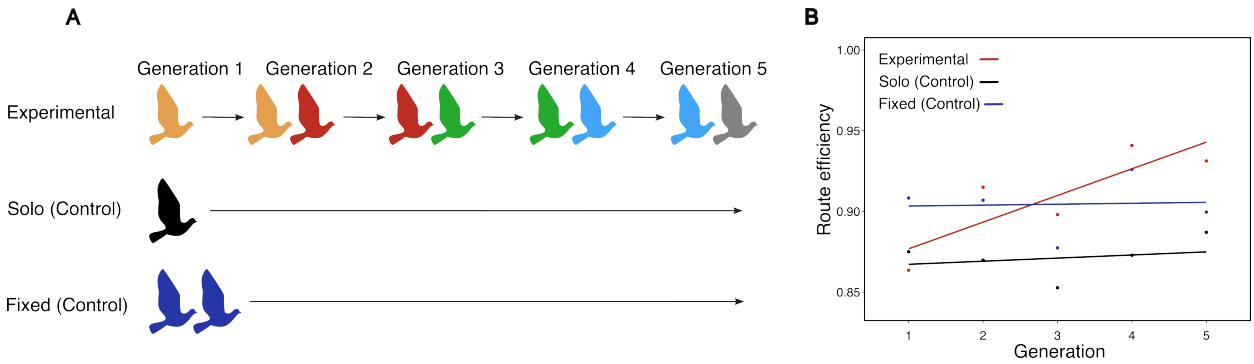

Figure 1: Schematic illustration of the experimental design used in the original study (adapted from Sasaki and Biro, 2017 [49]). (A) In each generation, birds were flown either individually or in pairs for a total of 12 releases. In the experimental condition, partners were replaced in each generation (except the first) whereas in the control conditions, birds were flown either individually or as fixed pairs for a total of 60 releases. (B) Variation in route efficiency in each condition across generations. Route efficiency was calculated from the final flight of each generation (release 12) for the experimental condition, and from releases 12, 24, 36, 48 and 60 for the solo and fixed controls.

Here, we hypothesized several plausible learning mechanisms that bird pairs might employ during homing navigation. We simulated each of these mechanisms using artificial chains of birds composed of trajectories from the solo control dataset and applied a weighted averaging procedure (which depended on the mechanism under study) to mimic the outcome of paired flights in a generation, analogous to the actual experimental design. It is also important to clarify that we use the terms *mechanism* or *strategy* in a more abstract manner rather than in a purely mechanistic way. In other words, we model how two trajectories (representing the 'preferred' path of two birds) map to a new trajectory (representing the joint path of the two birds flying as a pair), rather than the moment-to-moment decisions of each bird. This is because neither the underlying biological basis of the birds' decisions, nor different sources of noise (e.g., perceptual noise, decision noise, environmental noise) can be uniquely inferred using the available geospatial data. Moving forward, we will be using the terms mechanism(s) or strateg(ies) interchangeably to indicate these meta-mechanism(s), while the actual cognitive mechanisms that lead to these meta-mechanisms arising remain an open question. Our learning strategies can be organized into three classes based on the kind of information required to execute the strategy. We quantitatively compared the predictions generated by each strategy with the experimental data to infer which strateg(ies) real birds are most likely to be using. By doing so, we can assess what kind of information and cognition pigeons may be using during paired homing flights. In particular, we ask whether the cognitive requirements of CCE are necessary for the observed improvement in navigation ability.

The navigational model

To simulate the experimental condition, we created artificial chains of bird pairs — similar to the experimental design — using bird trajectories from the solo control condition. We used the 12th flight of each solo control bird, since birds typically established stable idiosyncratic routes by this

point, with minimal improvements in subsequent releases. In our model, we assumed that when two birds are paired together, a shared route is formed from their individual routes, which is the outcome of the learning process that the two birds engage in over the course of 12 releases. We model this learning process using a weighted averaging procedure for each pair (see *Methods* for details). The naïve bird in that pair replaces its idiosyncratic route with this new averaged trajectory and becomes the 'experienced' bird in the next generation, where it is paired with a new naïve bird drawn from the solo control condition to continue the sequence. This iterative process — sampling a bird, generating the averaged route, assigning it as the 'experienced' bird for the next generation, and pairing it with another bird— was repeated for five generations, matching the actual experimental chains. The weighting parameter w_E used in our model quantifies the influence (i.e., leadership) of the 'experienced' bird. As a comparison, for the experimental birds, we measured leadership based on the relative front-back positions of the birds during flights (since leaders have been found to take frontal positions across many species see *Methods* for details). We compared the route efficiencies and social weights of the experienced bird resulting from each simulated strategy with the experimental data to identify which candidate strateg(ies) are compatible with the empirical data.

We simulated seven strategies, which can be categorized into three types based on the kind of information required to execute the strategy. The first type (T_1) represents the simplest process, where birds have no knowledge of their partner's level of experience or performance (Fig. 2A). This includes only the ***averaging*** strategy, where both birds weigh their individual routes equally in each generation, regardless of their respective navigational abilities.

The second type (T_2) assumes that birds are capable of recognizing the more experienced individual within the pair and adjust their social weighting to maximize performance using this knowledge (Fig. 2B). Indeed, empirical evidence suggests that individuals can rely less on social information as they gain more experience [61]. Within type (T_2), the ***experienced bird weighting*** strategy assumes that the experienced bird maintains a constant weight w_E across all generations, where w_E is chosen such that route efficiency, on average, is maximized for the fifth generation. Alternatively, the ***maximize generation*** strategy allows the experienced bird to adjust its weight w_E based on the generation the bird is in, thereby maximizing the route efficiency for each generation.

The third type (T_3), introduces the highest level of cognitive complexity, requiring birds to actively evaluate their individual or paired performance (Fig. 2C). The strategies in T_3 align with the mechanistic criteria required for CCE, where individuals can recognize and select for better-performing routes. Within this type, the first strategy, ***better bird weighting***, assigns the better bird (i.e., the bird in the pair that has a shorter individual route, regardless of its experience level) a weight w_B that maximizes efficiency in the fifth generation. Here, information from the worse-performing bird is still incorporated in the weighting process (with a weight $1-w_B$) before passing to later generations. A more extreme version of this strategy is the ***all-or-nothing*** strategy, wherein the better bird takes full despotic control (i.e., $w_B = 1$), propagating only the best-performing routes across generations, with no contribution from the worse bird. The remaining two strategies focus on adjusting 'local' aspects of a paired flight. In the ***maximize flight*** strategy, birds aim

for finding the best achievable route as a pair. For a particular pair of birds, this is a unique pair of weights w and $1-w$ that maximizes the route efficiency of the resulting trajectory. Finally, the ***maximize time steps*** strategy assumes that birds can judge their joint performance moment-to-moment and make adjustments to their heading at each time step during a flight. In this strategy, the bird with a preferred route closer to the beeline at a given moment, on average, exerts greater influence over the paired flight, pulling the pair closer to the correct homing direction.

Collectively, these strategies range from the minimal cognitive requirement (T_1) to recognizing one's own experience level (T_2) to evaluating the relative quality of the preferred routes of the two birds (T_3). We systematically tested the seven strategies by generating all possible combinations of chains from the solo control trajectories and measured the resulting distribution of route efficiencies and social weights used by the experienced bird for each generation. Using a mixed modeling approach, we fitted two different models, one for the responses of route efficiency and one for the social influence of the experienced bird (see *Methods* for details) and compared the results to the experimental data.

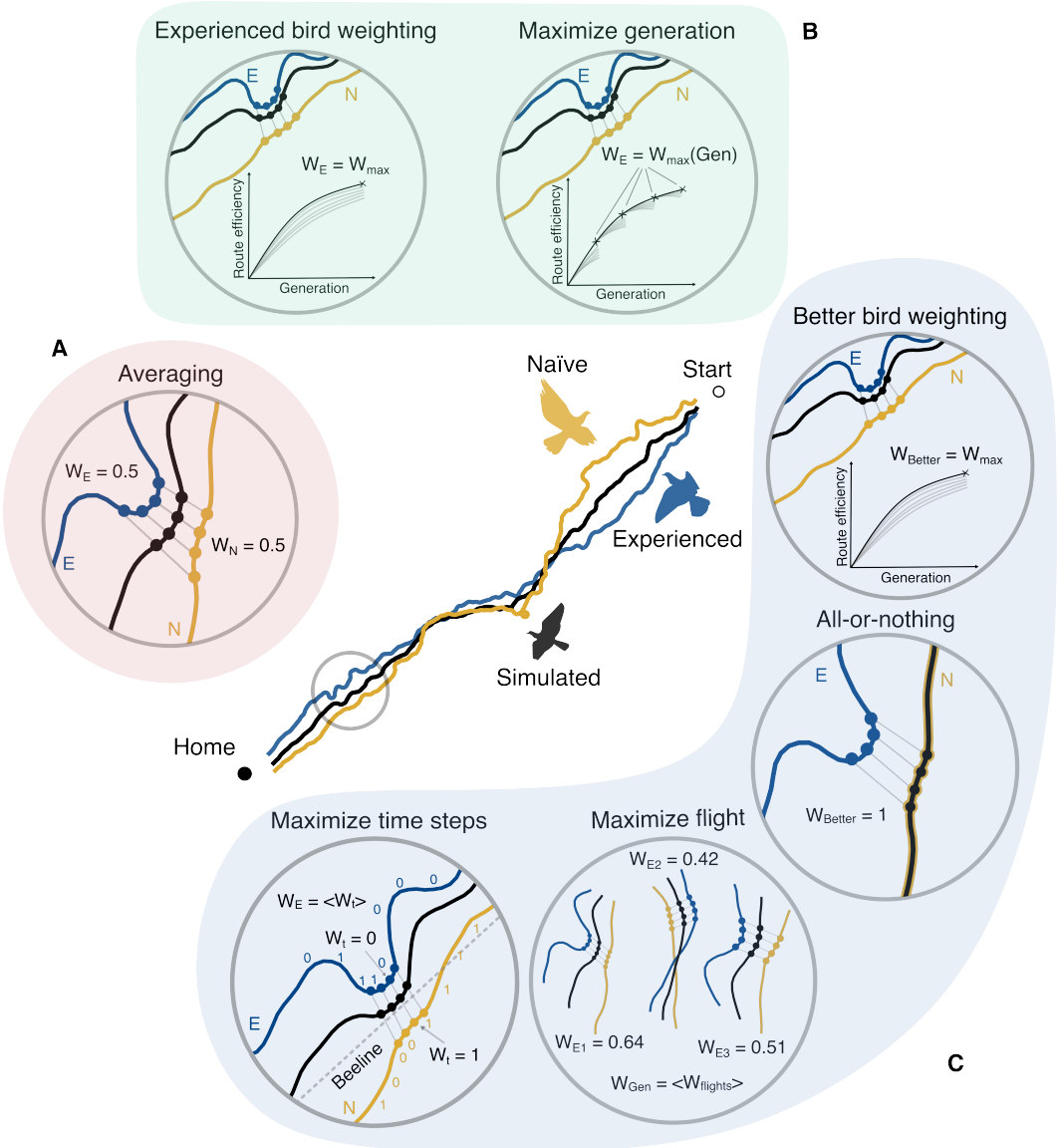

Figure 2: Illustration of the hypothesized social learning strategies. At the center, we depict an example of a simulated averaged bird trajectory (which becomes the experienced bird's route for the next generation) created using an experienced bird (E) and a naïve bird (N) trajectory. The strategies can be organized into three types based on the kind of information that is required to execute the strategy: (A) T_1 (red; requiring no information) — the simple process of equally *averaging* the two bird routes. (B) T_2 (green; requiring individuals to recognize relative experience levels) — strategies that optimize the weight of the experienced bird either to maximize performance in the final generation (*experienced bird weighting*) or in each generation (*maximize generation*). (C) T_3 (blue; requiring individuals to compare the relative path length of flights) — strategies where birds assess paired performance, either by assigning greater social weight to the better-performing bird (*better bird weighting* and *all-or-nothing*) or by fine-tuning local aspects of the flight (*maximize flights* and *maximize time steps*).

RESULTS

Similar route improvements emerge from simple and complex strategies

Across all strategies, we observed an improvement in route efficiency over generations (Fig. 3). This suggests that a wide range of decision-making mechanisms can lead to navigational improvements, not just the ones compatible with the definition of CCE. Interestingly, we found that the experimental birds performed relatively poorly compared to the majority of our hypothesized strategies (Fig. 3B).

After adjusting for generation effects, we found that three strategies significantly outperformed the experimental birds in efficiency gains (Fig. 3A and *SI Appendix*, Table S1). These strategies all belonged to type T_3 (where birds are able to evaluate route performance) and comprised the *better bird weighting* (OR = 1.37, 95% CI = [1.15, 1.63], $p < 0.001$), the *maximize flight* (OR = 1.55, 95% CI = [1.30, 1.86], $p < 0.001$), and the *maximize time steps* (OR = 1.26, 95% CI = [1.06, 1.51], $p < 0.01$) strategies, with the *maximize flight* strategy (where a pair of birds uses the specific set of social weights that optimizes their joint route) producing the highest route efficiency among all strategies tested. These observed effects suggest that the ability to evaluate the performance of routes (type T_3) can result in greater improvements compared to strategies that do not measure route efficiency (types T_1 and T_2).

Furthermore, we found that both strategies within type T_2 (where birds are able to detect which is the more experienced one), the *experienced bird weighting* (OR = 1.13, 95% CI = [0.94, 1.34], $p = 0.34$) and the *maximize generation* (OR = 1.15, 95% CI = [0.96, 1.37], $p = 0.19$) strategies, showed some effect but had non-significant efficiency gains compared to the experimental data.

The remaining two strategies — *averaging* (OR = 1.05, 95% CI = [0.88, 1.25] , $p = 0.93$) (type T_1) and *all-or-nothing* (OR = 0.99, 95% CI = [1.15, 1.63], $p = 1.00$) (type T_3) — aligned closely with the experimental data with little to no difference in improvements compared to the

experimental birds. Interestingly, the *all-or-nothing* strategy, which passes the best routes without integrating any information from the worse-performing bird (i.e., a greedy strategy), performed almost identically (poorly) to the simplest strategy of *averaging* individual routes. This is despite the fact that the *all-or-nothing* strategy has the ability to identify which of the two birds knows the shorter route, while the *averaging* strategy does not.

Examining how the performance of these strategies varied across generations provides further insights into the dynamics of route improvements (Fig. 3B and *SI Appendix*, Table S3). Several strategies showed significant positive effects, but only from generation 3 onwards. These included the *better bird weighting* (Gen 3 : OR = 1.66, 95% CI = [1.16, 2.39], $p < 0.01$; Gen 5: OR = 1.57, 95% CI = [1.05, 2.37], $p = 0.02$), the *maximize flight* (Gen 3 : OR = 1.89, 95% CI = [1.32, 2.72], $p < 0.001$; Gen 4: OR = 1.51, 95% CI = [1.00, 2.27], $p = 0.046$; Gen 5: OR = 1.96, 95% CI = [1.30, 2.94], $p < 0.001$), and the *maximize time steps* (Gen 3 : OR = 1.54, 95% CI = [1.07, 2.22], $p = 0.01$; Gen 5: OR = 1.49, 95% CI = [0.99, 2.24], $p = 0.06$) strategies. However, the increase in effect in generation 3 specifically can be attributed to the dip in efficiency of the experimental birds during this generation, inflating the contrasts with other strategies.

Overall, these observations helped us narrow down our search for candidate mechanisms, excluding three strategies from type T_3 which showed significant deviation from the experimental data. However, the strategies in types T_1 and T_2, as well as the *all-or-nothing* strategy in type T_3, remain statistically indistinguishable from the experimental data, when comparing their route efficiencies.

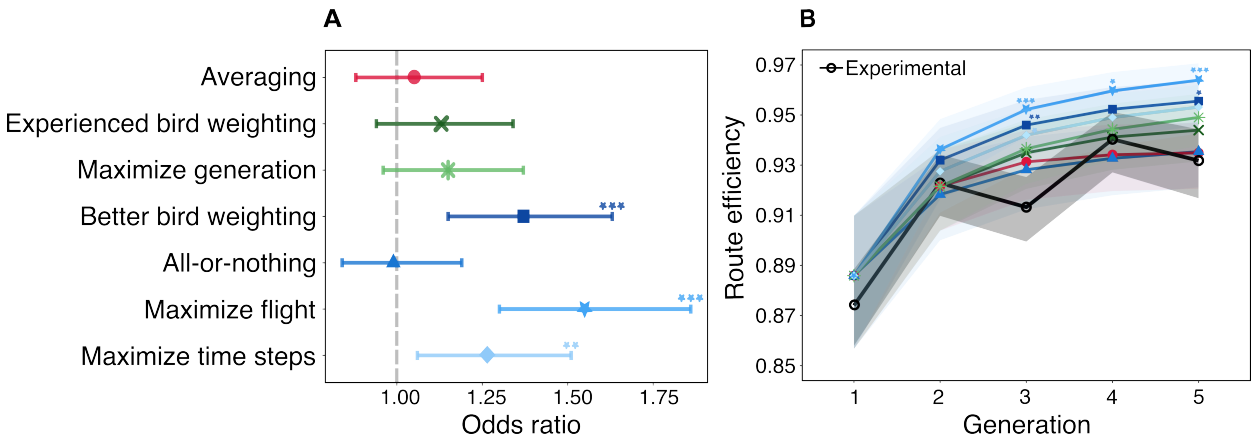

Figure 3: Results of mixed model analysis comparing the route efficiency of the hypothesized strategies to the experimental data. (A) The odds ratio for each strategy relative to the experimental data after adjusting for generation, with the gray dotted line representing the null odds ratio of 1. (B) The route efficiency of each hypothesized strategy and the experimental data (black) across generations. The shaded confidence bands indicate the 95% confidence intervals for the model predictions and reflect the biological uncertainty in the empirical dataset, not simulation noise. Statistical significance levels are denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The general color scheme in the plots match that used in Fig. 2 to highlight different strategy types.

Experimental birds rely on simple route averaging for navigation

Next, we examined how the social weights of the birds varied across different strategies in comparison to the experimental birds (Fig. 4). Although several of our hypothesized strategies assigned social weights to the better/worse bird (type T_3), rather than the experienced/naïve bird (types T_1 and T_2), for all strategies we calculated the mean social weight of the experienced/naïve bird to facilitate comparison with the experimental data. Furthermore, because the social weights of a pair of birds sum to 1, we report just the social weight of the experienced bird in this analysis.

Certain strategies — *averaging*, *experienced bird weighting*, and *maximize generation* — had, by definition, constant social weights assigned to the experienced bird. Specifically, the *averaging* strategy assigned a weight of 0.5, meaning that both birds contributed equally to homing decisions. For the *experienced bird weighting* strategy, the social weights were optimized to achieve maximum efficiency in the last generation, which was estimated to be 0.64 for the experienced bird. For the *maximize generation* strategy, the weights were optimized for each generation, with the experienced bird assuming weights 0.5 (generation 2), 0.67 (generation 3), 0.75 (generation 4), and 0.80 (generation 5) (note that this appears to follow a proportional relation $1/M$ for M generations, which we will return to later). Given their fixed weights, we excluded these strategies from our mixed model and manually contrasted their effects with the experimental birds.

After adjusting for generation effects, we found that the leadership values of the experienced bird in the experiment aligned with only one of the strategies: *averaging* (Mean Difference = 0.003, 95% CI = [-0.08, 0.09], $p = 0.99$) with almost no difference between the two (Fig. 4A and *SI Appendix*, Table S2). In other words, the experienced bird and the naïve bird shared, on average, equal social weight across all generations. All other strategies predicted a significantly higher influence of the experienced bird compared to that observed in the experimental data (*experienced bird weighting* (Mean Difference = 0.14, 95% CI = [0.05, 0.23], $p < 0.001$), *maximize generation* (Mean Difference = 0.18, 95% CI = [0.09, 0.27], $p < 0.001$), *better bird weighting* (Mean Difference = 0.17, 95% CI = [0.06, 0.29], $p < 0.001$), *all-or-nothing* (Mean Difference = 0.23, 95% CI = [0.11, 0.34], $p < 0.001$), *maximize flight* (Mean Difference = 0.23, 95% CI = [0.11, 0.34], $p < 0.001$), and *maximize time steps* (Mean Difference = 0.12, 95% CI = [0.00, 0.23], $p = 0.04$)). Combined with our previous conclusions, only one of our hypothesized strategies quantitatively matched both the route efficiency of the real birds (for all generations) as well as the social weights of the real birds (for all generations): the *averaging* strategy.

We observed that the average influence of the experienced bird monotonically increased across generations, except for the *averaging* and *experienced bird weighting* strategies (Fig. 4B and *SI Appendix*, Table S4). By the end of generation 5, the strategies that showed the strongest positive deviations from the experimental birds included the *better bird weighting* (Mean Difference = 0.30, 95% CI = [0.08, 0.52], $p < 0.01$), *all-or-nothing* (Mean Difference = 0.40, 95% CI = [0.18, 0.62], $p < 0.001$), *maximize generation* (Mean Difference = 0.34, 95% CI = [0.12, 0.56], $p < 0.001$), *maximize flight* (Mean Difference = 0.39, 95% CI = [0.17, 0.61], $p < 0.001$), and *maximize time steps* (Mean Difference = 0.26, 95% CI = [0.04, 0.48], $p = 0.01$).

Interestingly, we found that strategies that had similar average social weights can lead

to very different route efficiency outcomes. For instance, the *all-or-nothing* and *maximize flight* strategies had almost identical trends for the social weights of the experienced bird. However, these strategies had substantially different route efficiency outcomes, with the *all-or-nothing* strategy being the worst performing and the *maximize flight* strategy achieving the highest efficiency gain. Conversely, strategies that exhibited similar route efficiencies could be driven by very different average social weights. For example, the *all-or-nothing* and *averaging* strategies produced similar route efficiency outcomes, yet the former had the highest influence of the experienced bird, while the latter had the lowest influence. These observations demonstrate that the average social weights are not predictive of performance, nor is performance predictive of social weights, at least in their aggregate.

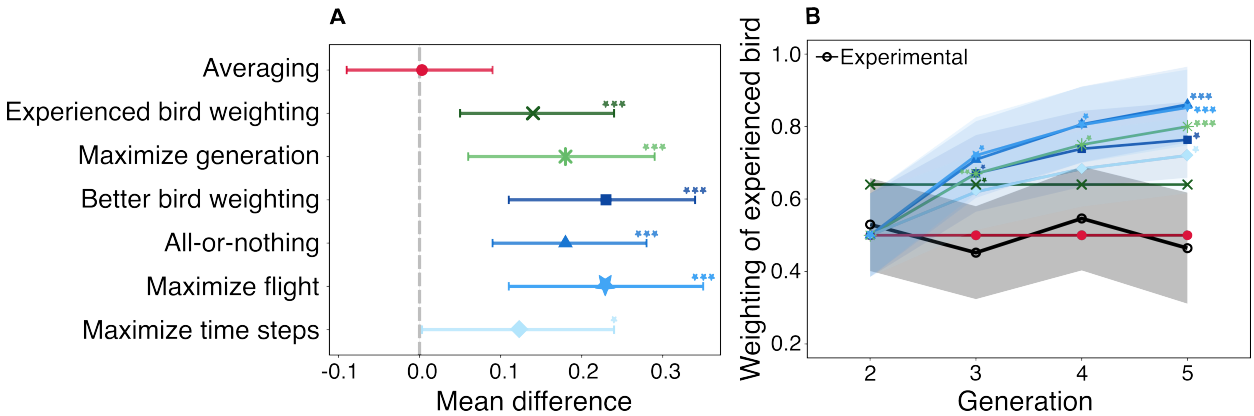

Figure 4: Results of mixed model analysis for the social weighting of the experienced bird. (A) The mean difference for each strategy relative to the experimental data after adjusting for generation, with the gray dotted line representing the mean difference of 0 between the datasets. (B) The variation of the social weighting assigned to the experienced bird for each strategy across generations. The shaded confidence bands indicate the 95% confidence intervals for the model predictions and reflect the biological uncertainty in the empirical dataset, not simulation noise. Statistical significance levels are denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The color scheme in the plots is the same as in Fig. 2.

Chaining with replacement experimental design leads to reduced effective group sizes

Although homing pigeons appear to be simply averaging their flights when flown together, the structure of the 'chaining with replacement' experimental design does not result in the five birds in a chain equally contributing to the final learned flight. Instead, the chain structure produces a 'geometric forgetting' of past information, with the last bird contributing 1/2 of the final flight path, the penultimate bird contributing a 1/4, the third bird contributing an 1/8, etc. Indeed, we predict that the first bird contributes only 1/32 of the final flight path. These unequal contributions of each bird to the final flight trajectory is in contrast to the classic 'wisdom of crowds', for which collective performance is maximized when all individuals are weighed equally (assuming that individuals are unable to perceive differences in competence).

We asked to what extent the chain structure, and correspondingly the unequal weighting of birds, decreases the collective performance of the group. To answer this, we defined a measure that we call 'effective group size', which is the size of a democratic (i.e., equally weighted) group that exhibits equivalent performance to the focal group under study (which is composed of a certain number of individuals, each with a certain social weight). To estimate this, we first mapped the size of democratic flocks to a corresponding route efficiency (Fig. 5, inset) by simulating the outcome of democratic flocks through averaging the routes of independent control flights and measuring the route efficiency of the resulting collective flights (see *Methods* for details). We then re-mapped the route efficiencies resulting from our hypothesized strategies, as well as the experimental data (i.e, Fig. 3B), to an equivalent effective group size.

Most of the hypothesized strategies exhibit an effective group size smaller than five by the end of the fifth generation (i.e., fall below the 1:1 line shown as a dashed orange line in Fig. 5). This demonstrates that the chain structure tends to be an inefficient way to combine information, since for these strategies, the five birds that comprise a chain show a diminished performance compared to a hypothetical flock of five birds flying together. This is even true for all of the strategies of type T_2, where there is knowledge of who is the more experienced bird (knowledge that our hypothetical flock of five birds does not possess). By contrast, two of the strategies of type T_3 exhibit an effective group size greater than five. This is possible because birds using this strategy have direct knowledge of the quality of a route, information that is unavailable to our simulated flocks of five birds.

Our results suggest that the experimental design limits the propagation of information from the past to the future. For a chain of five birds, the average effective group size from the experimental dataset barely exceeds two birds (2.12) (Fig. 5). This suggests that in each generation, contributions from prior individuals are hardly retained and chains of experimental birds fly only marginally better than pairs of birds. In contrast, more complex navigational mechanisms, facilitated by, for example, perceiving route quality, could surpass this limit and even reach effective group sizes exceeding five individuals. Hence, continuous replacement of birds within a chain leads to a substantial loss of information, and, unless compensated by a more cognitively complex mechanism, constrains the chain's ability to improve the efficiency of its route over time.

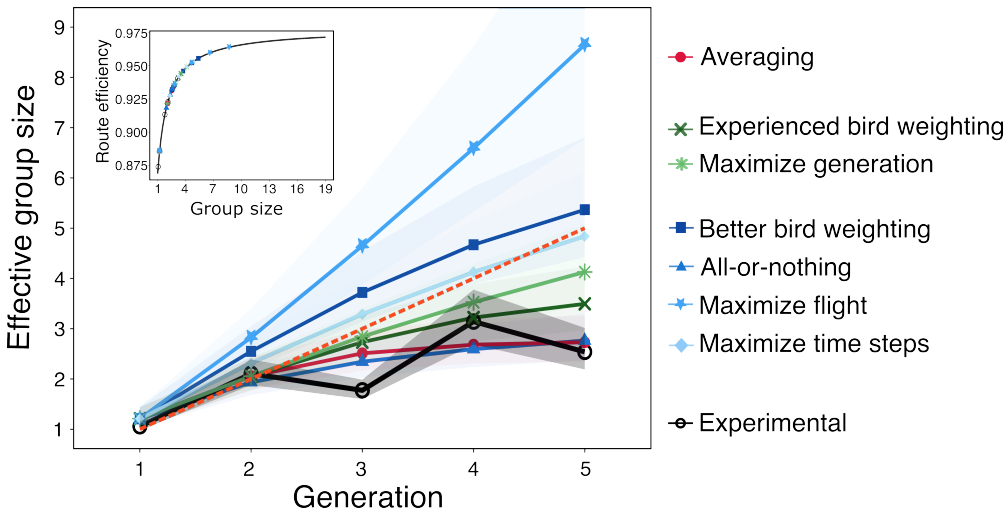

Figure 5: The effective group size achieved by each strategy and the experimental data across generations. The predicted mean route efficiency from each strategy is mapped onto a curve modeling the relationship between the route mean route efficiency and simulated flock sizes. The inset provides the curve, illustrating how route efficiency scales with flock size. The dashed orange line shows the 1:1 between the two variables. Shaded confidence bands denote the 95% confidence intervals for the estimated effective group sizes.

DISCUSSION

In this study, we investigated potential navigational strategies, which varied in the level of cognitive complexity required to execute them, to infer the mechanism of collective route improvements in homing pigeons. Our work builds upon a 2017 experiment which demonstrated that pairs of experienced and naïve pigeons could iteratively improve homing routes over generations. Specifically, our aims were three-fold: 1) provide insights into the mechanism(s) pigeons use to improve route efficiency, 2) investigate how cognitive complexity, as well as experimental design (turnover of birds), relate to performance and 3) investigate if the identified mechanism(s) fall under the criteria required for CCE. To achieve this, we developed seven plausible mechanisms, broadly categorized into three types (T_1, T_2, and T_3), with increasing cognitive complexity. We created a simple navigational model by artificially pairing solo bird trajectories, similar to the experimental design, and tested our strategies to identify candidate strateg(ies) that are quantitatively compatible with the empirical data. All of our strategies resulted in route improvements, regardless of their underlying complexity. However, when combined with the results from the social weight analysis, the experimental data overall aligned best with the simplest strategy: *averaging* individual routes.

The route efficiency achieved generally increased with the cognitive complexity of the strategy — strategies in type T_3 tended to perform the best. However, this was not always the case. Specifically, the *all-or-nothing* strategy (of type T_3), where only the best route was passed to the next generation, performed no better than the simple *averaging* strategy (of type T_1). This can be explained by recalling that there are two sources of information: leadership from experts and emergent wisdom of crowds. The *all-or-nothing* and the *averaging* strategies, respectively, represent these two 'pure' strategies. By contrast, all of the other strategies represent mixed strategies, whereby individuals with better information have greater social weight, but individuals with worse information still contribute to the collective decision. The mixed strategies all outperformed the pure strategies, demonstrating the benefit of exploiting both sources of information.

Furthermore, the strategies of type T_3 are compatible with the definition of CCE [42, 55] but showed a quantitative mismatch with the experimental data. Therefore, an improvement in route efficiency alone is not sufficient evidence for cultural transmission, as the experimental birds did not exhibit evidence of some of the criteria of CCE (evaluation of performance, for instance). This is because of the possibility of the wisdom of crowds improving routes 'for free' without the need for individuals to specifically recognize and maintain improvements. In other words, the wisdom of crowds functions like a statistical ratchet, rather than a deterministic one. Moreover, as evidenced from our social weighting analysis, we did not find direct evidence of social learning

in terms of imitative copying of prior routes (see [62] for similar results). Birds in this experiment, on average, equally influenced each other's routes disregarding any differences in experience. This observation raises broader questions about which social learning mechanisms truly align with the requirements for CCE [63, 64].

We note that the wisdom of crowds is beneficial in some but not all task types [65–67]. In the context of the homing task, the pigeons apparently have a general (noisy) tendency to find their way home. Because of this, averaging the flights of multiple pigeons is an effective way of canceling out noise, resulting in an increasingly efficient route (on average). Hence, averaging routes can lead birds to stumble upon the best routes by chance. By contrast, in human cultural evolution, complex innovations are unlikely to arise *de novo* by averaging existing behaviors [44, 68]. In these types of tasks, selection on novel behaviors at the tails of the distribution — not the average — is probably necessary to 'evolve' towards improved solutions. Future work should investigate the range of tasks that group-living animals typically face, and categorize them into ones that can benefit from the wisdom of crowds and ones that require CCE [69].

This study also reveals that collective strategies must be understood within the context in which they are implemented. Specifically, here the experimental design caused birds to fly in pairs in a particular chain structure. Combined with the *averaging* strategy, this resulted in a geometric dilution of influence over generations. This progressive loss of information due to repeated turnover of birds reflects an inefficient version of the 'wisdom of crowds', where instead of equally incorporating the knowledge of all past individuals, information from early generations is disproportionately lost. This is further supported by our effective group size analysis, which we developed as a reference metric to quantify the effective number of individuals contributing towards route improvements in each generation. We find this to be a potentially valuable metric to quantify such chained experimental design which are usually proposed for studying CCE [70]. Using this method, we showed that despite consisting of five birds, experimental chains of birds exhibited an effective group size only slightly better than two individuals. We expect the behavior of animals to be adaptive in the conditions in which they evolved. The chain structure of this experiment may approximate the dynamics of fission-fusion groups, and additionally forced group sizes to be just two birds. While urban pigeon flocks are highly fluid, they are typically of sizes much larger than two [71]. The *averaging* strategy that the experimental birds exhibited may have performed relatively poorly within this chain structure, but it should exhibit higher performance in larger groups, where there is more scope for the wisdom of crowds. A simple strategy such as the *averaging* strategy may also be beneficial in large temporary groups, where recognizing experienced or more knowledgeable individuals may be difficult.

It is also important to consider what the pigeons themselves are trying to optimize [69]. While our analysis focused on route efficiency, pigeons may not necessarily prioritize minimizing travel distance. Indeed, the average difference in travel distance between the best-performing and worst-performing strategies is only a few hundred meters. Previous studies have shown that pigeons rely on specific landmarks for navigation, suggesting that familiar, slightly longer routes may be preferable to unfamiliar, shorter ones [72–74]. In this context, birds may be prioritizing maximizing their probability of successfully returning home rather than minimizing flight distance

per se (which would require exploring alternative routes, with an associated risk of getting lost). The *averaging* strategy may be best for birds to maximize the probability of returning home.

Another assumption of our model is that the birds execute the strategies perfectly. In other words, the birds are able to accurately identify who is the more experienced bird or the bird with the shorter route, and can apply the optimal social weight for that strategy. Real animals, if they did implement these strategies, would have some error rate (either in differentiating the birds or applying the optimal social weight), resulting in a lower measured performance. If the error rate is high enough, the performance could even be worse than that exhibited by the *averaging* strategy. Therefore, the *averaging* strategy could be a less risky one, guaranteeing (statistically) some measure of improvement.

Our results highlight the need for a more general framework for understanding CCE, calling for a critical re-evaluation of its definition and evaluative criteria, especially in the context of non-human animals. Innovations can be the emergent outcome of collective behavior, rather than necessarily originating from a single individual [75]. In that case, these collectively-generated innovations can transmit within the group, without necessarily relying on social learning or leadership. When individuals lack the ability to recognize beneficial innovations, there is no ratchet in a strict sense, but a ratchet may nonetheless exist statistically, due to the wisdom of crowds. Moreover, CCE and the wisdom of crowds are not necessarily mutually exclusive, and social animals could, in principle, simultaneously exploit both sources of innovation, as illustrated in the mixed strategies we explored. We therefore recommend that future empirical studies of CCE in non-human animals clearly distinguish between these mechanisms and incorporate additional criteria for evaluation. Marrying the fields of collective intelligence and CCE could provide a powerful framework for understanding how social animals can improve their fitness across a range of task types and cognitive abilities.

In conclusion, our study opens avenues for future research into collective decision making across time [76]. Collective performance is the emergent outcome of the social structure, task type, and social strategy. Understanding how these interact with each other could improve our predictions about what collective strategies we should expect to observe in nature given uncertain shifts in timings of events in the future [77]. For this navigation task, simple averaging is sufficient to quantitatively explain the experimental results, but other tasks may be less amenable to the wisdom of crowds. It is crucial to understand collective strategies within the context in which they evolved, including typical group sizes, turnover rate, error rate, and how many times a task is repeated. While more cognitively complex strategies can result in greater performance, they may not be necessary, and, indeed, mixed strategies that draw on both individual and collective sources of intelligence may result in the highest levels of performance.

METHODS

Data pre-processing and experimental procedure

We used GPS trajectory data from a previous study by Valentini *et al.* [60], which was based on a 2017 study conducted by Sasaki and Biro in homing pigeons (*Columba livia*) [49]. The study consisted of three conditions: experimental, solo control, and fixed control. Each condition consisted of birds that were allowed to fly back home from a site 8.4 km away, either individually or as pairs. The experimental trials were structured around sets of 12 bird flights, termed a 'generation'. In generation 1 of the experimental group, a single bird flew 12 times before being paired with a new individual (the naïve bird) in generation 2. This pattern continued in subsequent generations, wherein the prior naïve bird (now experienced) was paired with a new bird, creating a transmission chain. Each chain consisted of five generations, or 60 total flights across the entire chain. The birds in the control conditions completed an equal number of flights (60 releases), with solo control birds flying individually and the fixed control birds flying in constant pairings. In total, each condition had 10 independent replicate chains.

For our analysis, we focused on the final flights (release 12) of each generation for the experimental chains for comparison with our model predictions. Following Valentini *et al.*, we only considered flights where the two birds were less than 250m from each other at all times, as well as excluding flights of birds that never reached their home destination (see *Materials and Methods* in [60]). After accounting for missing flights in the experimental group, we were left with ten flights each for generations 1 to 3, eight flights for generation 4, and seven for generation 5. Each of the trajectories were down-sampled to a frequency of 1 Hz (from 5 Hz) using the R package `trajr` [78]. The trajectory points were converted from their original geographic coordinate system (GPS) to the Universal Transverse Mercator (UTM) system for all of our distance calculations. The joint route flown by a pair of birds was calculated as the center of mass of the two birds at each time point. The route efficiency was then obtained as the ratio of the beeline distance to the actual distance traveled by the birds as a pair. To avoid biases due to circling behavior of the pigeons near release and home sites, we excluded trajectory points within a radius of 800m and 200m from the release and home site, respectively. Our beeline distance was adjusted accordingly to account for these zones.

Weighted averaging

To simulate the experimental chains of birds, we first assumed that each naïve bird has an initial 'preferred' route that it would follow if flown alone. We compiled a set of these preferred routes by taking the 12th release for each of the solo control birds, to match our procedure of analyzing the 12th release of each generation in the experimental condition. Moreover, the route efficiency of the solo control birds reached their maximum, on average, by this flight, and therefore represents the idiosyncratic learned preference of that bird.

We then used a weighted averaging method to simulate the consensus flight that would

emerge when two birds are flown together. Each run in our simulation proceeds as follows: In generation 1, we randomly chose a trajectory from our set, which becomes the 'experienced' bird in generation 2. For generation 2, our experienced bird is paired up with a naïve bird, which is another trajectory sampled with replacement from the representative set. We then computed an averaged route from these paired flights using the following steps: First, we synchronized the number of data points across the two trajectories to match the size of the longer trajectory. This is done by interpolating missing coordinates from two consecutive time points in the shorter trajectory. We used the `TrajResampleTime()` function available from the R package `trajr` to perform our resampling [78]. Next, the x and y coordinates of the paired trajectories were combined using the following weighted averaging equations:

$$X_{\text{avg}}(t) = X_E(t) \cdot w + X_N(t) \cdot (1 - w),$$
$$Y_{\text{avg}}(t) = Y_E(t) \cdot w + Y_N(t) \cdot (1 - w),$$

where $w \in [0, 1]$ and $t \in \{t_1, t_2, \ldots, t_L\}$ corresponds to the time steps along a trajectory of length L. Here, w is the weighting parameter, which quantifies the social influence of the experienced bird (E) relative to the newly sampled naïve (N) bird. Therefore, a weighting of $w = 0$ indicates that the experienced bird has no influence on the averaged route, while $w = 1$ implies full control of the experienced bird on route formation. To account for all possible weighting contributions, w was incremented in steps of 0.01 during simulations of each strategy. We assume that the previously naïve bird adopts this averaged route as its preferred route by the end of the generation, and this averaged route is then propagated as the experienced bird's preferred route for generation 3. A new naïve bird is drawn from our dataset, and its preferred route is again averaged with the experienced bird's. This process of sampling, pairing, averaging, and propagation continued iteratively using all possible combinations of birds in each generation, until generation 5. By systematically varying the weighting parameter w, we investigated seven plausible strategies of navigation and calculated the resulting route efficiencies, similar to the method described for the experimental group.

Leadership calculation

Previous studies have shown that pigeons follow a front-back leadership, wherein the bird at the front of the flock assumes a greater leadership role compared to birds at the back [16, 17]. We calculated a leadership metric for the final release in each generation of the experimental data to compare with our simulation results. To achieve this, we used the method outlined in Valentini *et al.* (see *Appendix 1, Methods* [60]) that is based on calculating the relative position of the birds from each other d_{EN}, at each time step and projecting it onto the direction of motion of the bird pair in the subsequent time step. We used the sign of d_{EN} to determine which bird among the pair was leading at a given time step. The distance d_{EN} is positive if the experienced bird is in front of the pair and negative if the naïve bird is leading the pair. For our analysis, we used the experienced bird as our reference bird in the pair. We calculated the average proportion of time

spent by the experienced bird in the front for each flight and used this for comparing with the simulated social weightings.

Statistical analysis

We utilized a mixed-effects modeling approach for comparing our simulation results with the experimental data. We created two different models to address our two response variables: route efficiency and social weight of the experienced bird across strategies. Since the original experiment provided data from only nine birds in the solo control condition, we were limited by the number of representative trajectories for our simulations. As a result, the same bird trajectories were reused in our simulated chains (i.e., we sampled with replacement), introducing an unwanted dependency in our analysis. This non-independence was evident across generations and simulated chains, effectively reducing the variance within each sample relative to the true variance. To account for this non-independence, we incorporated bird ID (1-9, corresponding to each solo trajectory) and its interaction with each generation as our random effect in our model. For the experimental data, we conditionally excluded the random effect term due to independent bird IDs in the experimental chains. Thus, the modeled bird-specific deviations in the random effect term $\gamma_{\text{Gen:bird}}$ are:

$$\gamma_{\text{Gen:bird}} \sim \mathcal{N}(0, \sigma^2_{\text{dataset}})$$

where the random effect variance $\sigma^2_{\text{dataset}}$ is dataset specific (experimental vs simulated).

To predict route efficiencies, we combined the experimental and simulated datasets and fitted a beta regression model with a logit link using the `glmmTMB()` function from the R package `glmmTMB` (see *SI Appendix* for model details and performance checks) [79]. We used generation, strategy, and their interactions as fixed predictors for our models. Given this, our model equations for the expected route efficiency μ_{ij} for generation i and strategy j are:

$$\text{logit}(\mu_{ij}) = \beta_0 + \beta_1 \text{Gen}_i + \beta_2 \text{Strategy}_j + \beta_3(\text{Gen}_i \times \text{Strategy}_j) + \gamma_{\text{Gen}_i\text{:bird}}$$

Similarly, for the response variable social weight of the experienced bird, we fitted a linear mixed effect model using the `lmer()` function from the R package `lme4` (see *SI Appendix* for model details and performance checks) [80]. We used an identical fixed and random effect structure as the route efficiency model specified above. However, since the estimates for the averaging, better weighting, and maximize generation strategies were constant values, they were removed from the mixed model and contrasted manually with the experimental data. The model equation for the expected social weight ω_{ij} for generation i and strategy j is given by:

$$\omega_{ij} = \beta_0 + \beta_1 \text{Gen}_i + \beta_2 \text{Strategy}_j + \beta_3(\text{Gen}_i \times \text{Strategy}_j) + \gamma_{\text{Gen}_i\text{:bird}} + \epsilon_{ij}$$

From both our models, we computed estimated marginal means (EMMs) adjusting for the random effects, using the `emmeans` package in R [81]. Post hoc analysis tests comparing the simulated

and experimental data was also performed using the `emmeans` package. To validate our model predictions and results, we performed a bootstrapping procedure (see *SI Appendix*, Tables S5 and S6). We generated bootstrap samples for both the experimental and simulated strategies for each generation, using a sample size matching the number of experimental samples in the respective generation, and conducted 10,000 runs. Our model predictions were in accordance with the bootstrap results within reasonable margin of error. We adjusted the calculated p-values and confidence intervals for multiple comparisons within each generation using the Dunnett's method against our experimental control.

Effective group size analysis

We conducted an analysis to map the predicted route efficiencies from our model with those achievable by democratic flocks of varying group sizes (i.e., they make decisions by averaging the preferred routes of all of the individuals in the group). This approach enabled us to better understand the cumulative effects of different mechanisms within the 'chaining with replacement' experimental design used in the study. To achieve this, we first generated groups of different sizes by artificially pairing birds from the final flights of the experimental and solo control conditions in generation 1 (19 trajectories total). Using a weighted averaging approach, similar to our previous analysis (but with equal weights), we calculated an overall mean route from the individual trajectories. This mean route served as the measure of the flock's average route efficiency. We then calculated mean route efficiencies for each group size and observed a non-linear relationship between group size and route efficiency. To characterize this trend, we applied non-linear least squares regression (NLS) in R [82], using various functional forms with group size as the predictor and route efficiency as the response variable. Among these evaluated models, the best fit was obtained using a hyperbolic function (see *SI Appendix*, Fig. S1 for further details regarding model comparisons and goodness of fit). Finally, we used the model coefficients to back-transform our simulated efficiencies and derive corresponding effective group size values.

DATA AVAILABILITY

The dataset and code used for this study have been deposited in figshare with identifier "10.6084/m9.figshare.28950032" and can be accessed through this link: https://doi.org/10.6084/m9.figshare.28950032.v1.

ACKNOWLEDGMENTS

We would like to thank Almond Stöcker, Cooper Kimball-Rhines and Jacob Adamczyk for their valuable feedback on our analysis. S.C.B, F.A.F and A.B.K acknowledge support from the U.S.

National Science Foundation (BRC-BIO DBI-2233416) and a generous anonymous gift fund.

---

[1] Krause J, Ruxton G (2002) *Living in Groups*, Oxford Series in Ecology and Evolution. (Oxford University Press, Oxford, New York).

[2] Wiltschko R, Wiltschko W (2003) Avian navigation: from historical to modern concepts. *Animal Behaviour* 65(2):257–272.

[3] Åkesson S, Boström J, Liedvogel M, Muheim R (2014) Animal navigation in *Animal Movement Across Scales*, eds. Hansson LA, Åkesson S. (Oxford University Press).

[4] Wiltschko R, Wiltschko W (2023) Animal navigation: how animals use environmental factors to find their way. *Eur. Phys. J. Spec. Top.* 232(2):237–252.

[5] Egevang C, et al. (2010) Tracking of Arctic terns Sterna paradisaea reveals longest animal migration. *Proc. Natl. Acad. Sci. U.S.A* 107(5):2078–2081.

[6] Alerstam T, Bäckman J, Grönroos J, Olofsson P, Strandberg R (2019) Hypotheses and tracking results about the longest migration: The case of the arctic tern. *Ecol. Evol.* 9(17):9511–9531.

[7] Shaffer SA, et al. (2006) Migratory shearwaters integrate oceanic resources across the Pacific Ocean in an endless summer. *Proc. Natl. Acad. Sci. U.S.A* 103(34):12799–12802.

[8] Thorpe JE (1988) Salmon migration. *Sci. Prog.* 72(3 (287)):345–370.

[9] Quinn TP (2018) *The behavior and ecology of Pacific salmon and trout.* (University of Washington press).

[10] Muheim R, Phillips JB, Åkesson S (2006) Polarized Light Cues Underlie Compass Calibration in Migratory Songbirds. *Science* 313(5788):837–839.

[11] Brothers JR, Lohmann KJ (2015) Evidence for Geomagnetic Imprinting and Magnetic Navigation in the Natal Homing of Sea Turtles. *Curr. Biol.* 25(3):392–396.

[12] Karwinkel T, et al. (2024) A conceptual framework on the role of magnetic cues in songbird migration ecology. *Biol. Rev.* 99(4):1576–1593.

[13] Bett NN, Hinch SG (2016) Olfactory navigation during spawning migrations: a review and introduction of the Hierarchical Navigation Hypothesis. *Biol. Rev.* 91(3):728–759.

[14] Dreyer D, et al. (2018) The Earth's Magnetic Field and Visual Landmarks Steer Migratory Flight Behavior in the Nocturnal Australian Bogong Moth. *Curr. Biol.* 28(13):2160–2166.e5.

[15] Couzin ID, Krause J, Franks NR, Levin SA (2005) Effective leadership and decision-making in animal groups on the move. *Nature* 433(7025):513–516.

[16] Nagy M, Ákos Z, Biro D, Vicsek T (2010) Hierarchical group dynamics in pigeon flocks. *Nature* 464(7290):890–893.

[17] Nagy M, et al. (2013) Context-dependent hierarchies in pigeons. *Proc. Natl. Acad. Sci. U.S.A* 110(32):13049–13054.

[18] Strandburg-Peshkin A, Farine DR, Couzin ID, Crofoot MC (2015) Shared decision-making drives collective movement in wild baboons. *Science* 348(6241):1358–1361.

[19] Strandburg-Peshkin A, Papageorgiou D, Crofoot MC, Farine DR (2018) Inferring influence and leadership in moving animal groups. *Philos. Trans. R. Soc. B* 373(1746):20170006.

[20] Sampaio E, et al. (2024) Multidimensional social influence drives leadership and composition-dependent success in octopus-fish hunting groups. *Nat. Ecol. Evol.* 8(11):2072–2084.

[21] Berdahl AM, et al. (2018) Collective animal navigation and migratory culture: from theoretical models to empirical evidence. *Philos. Trans. R. Soc. B* 373(1746):20170009.

[22] Averly B, et al. (2022) Disentangling influence over group speed and direction reveals multiple patterns of influence in moving meerkat groups. *Sci. Rep.* 12(1):13844.

[23] Galton F (1907) Vox populi (1907) nature, n. 1949, vol. 75, pp. 450-451 (traduzione di romolo giovanni capuano©). *Nature* 75:450–451.
[24] Simons AM (2004) Many wrongs: the advantage of group navigation. *Trends Ecol. Evol.* 19(9):453–455.
[25] Kao AB, et al. (2018) Counteracting estimation bias and social influence to improve the wisdom of crowds. *J. R. Soc. Interface* 15(141):20180130.
[26] Sumpter DJT, Krause J, James R, Couzin ID, Ward AJW (2008) Consensus decision making by fish. *Curr Biol* 18(22):1773–1777.
[27] Dell'Ariccia G, Dell'Omo G, Wolfer DP, Lipp HP (2008) Flock flying improves pigeons' homing: GPS-track analysis of individual flyers versus small groups. *Animal Behaviour* 76(4):1165–1172.
[28] Faria JJ, Codling EA, Dyer JR, Trillmich F, Krause J (2009) Navigation in human crowds; testing the many-wrongs principle. *Animal Behaviour* 78(3):587–591.
[29] Krause S, James R, Faria JJ, Ruxton GD, Krause J (2011) Swarm intelligence in humans: diversity can trump ability. *Animal Behaviour* 81(5):941–948.
[30] Ward AJW, Herbert-Read JE, Sumpter DJT, Krause J (2011) Fast and accurate decisions through collective vigilance in fish shoals. *Proc. Natl. Acad. Sci. U.S.A* 108(6):2312–2315.
[31] Prelec D, Seung HS, McCoy J (2017) A solution to the single-question crowd wisdom problem. *Nature* 541(7638):532–535.
[32] Bahrami B, et al. (2010) Optimally Interacting Minds. *Science* 329(5995):1081–1085.
[33] Koriat A (2012) When Are Two Heads Better than One and Why? *Science* 336(6079):360–362.
[34] Flack A, Nagy M, Fiedler W, Couzin ID, Wikelski M (2018) From local collective behavior to global migratory patterns in white storks. *Science* 360(6391):911–914.
[35] Mehlhorn K, et al. (2015) Unpacking the exploration-exploitation tradeoff: A synthesis of human and animal literatures. *Decision* 2(3):191–215.
[36] Liedvogel M, Åkesson S, Bensch S (2011) The genetics of migration on the move. *Trends Ecol. Evol.* 26(11):561–569.
[37] Väli Ü, Mirski P, Sellis U, Dagys M, Maciorowski G (2018) Genetic determination of migration strategies in large soaring birds: evidence from hybrid eagles. *Proc. R. Soc. B* 285(1884):20180855.
[38] Cavalli-Sforza LL, Feldman MW (1973) Cultural versus biological inheritance: phenotypic transmission from parents to children. (A theory of the effect of parental phenotypes on children's phenotypes). *Am. J. Hum. Genet.* 25(6):618–637.
[39] Whiten A, Ayala FJ, Feldman MW, Laland KN (2017) The extension of biology through culture. *Proc. Natl. Acad. Sci. U.S.A* 114(30):7775–7781.
[40] Whitehead H, Laland KN, Rendell L, Thorogood R, Whiten A (2019) The reach of gene–culture coevolution in animals. *Nat. Commun.* 10(1):2405.
[41] Boyd R, Richerson PJ (1996) *Why culture is common, but cultural evolution is rare.* (Oxford University Press Inc.) Vol. 88, pp. 77–94.
[42] Tomasello M (2009) *The cultural origins of human cognition.* (Harvard university press).
[43] Dean LG, Vale GL, Laland KN, Flynn E, Kendal RL (2014) Human cumulative culture: a comparative perspective. *Biol. Rev. Camb. Philos. Soc.* 89(2):284–301.
[44] Creanza N, Kolodny O, Feldman MW (2017) Cultural evolutionary theory: How culture evolves and why it matters. *Proc. Natl. Acad. Sci. U.S.A* 114(30):7782–7789.
[45] Yamamoto S, Humle T, Tanaka M (2013) Basis for cumulative cultural evolution in chimpanzees: social learning of a more efficient tool-use technique. *PLOS ONE* 8(1):e55768.
[46] Whiten A, Harrison RA, McGuigan N, Vale GL, Watson SK (2022) Collective knowledge and the dynamics of culture in chimpanzees. *Philos. Trans. R. Soc. B* 377(1843):20200321.
[47] van Leeuwen EJC, DeTroy SE, Haun DBM, Call J (2024) Chimpanzees use social information to acquire a skill they fail to innovate. *Nat. Hum. Behav.* 8(5):891–902.

[48] Hunt GR, Gray RD (2003) Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proc. R. Soc. B* 270(1517):867–874.

[49] Sasaki T, Biro D (2017) Cumulative culture can emerge from collective intelligence in animal groups. *Nat. Commun.* 8:15049.

[50] Aplin LM (2019) Culture and cultural evolution in birds: a review of the evidence. *Animal Behaviour* 147:179–187.

[51] Williams H, Lachlan RF (2022) Evidence for cumulative cultural evolution in bird song. *Philos. Trans. R. Soc. B* 377(1843):20200322.

[52] Krützen M, et al. (2005) Cultural transmission of tool use in bottlenose dolphins. *Proc. Natl. Acad. Sci. U.S.A* 102(25):8939–8943.

[53] Mann J, Stanton MA, Patterson EM, Bienenstock EJ, Singh LO (2012) Social networks reveal cultural behaviour in tool-using dolphins. *Nat. Commun.* 3(1):980.

[54] Garland EC, Garrigue C, Noad MJ (2022) When does cultural evolution become cumulative culture? A case study of humpback whale song. *Philos. Trans. R. Soc. B* 377(1843):20200313.

[55] Mesoudi A, Thornton A (2018) What is cumulative cultural evolution? *Proc. R. Soc. B* 285(1880):20180712.

[56] Claidière N, Smith K, Kirby S, Fagot J (2014) Cultural evolution of systematically structured behaviour in a non-human primate. *Proc. R. Soc. B* 281(1797):20141541.

[57] Matthews GT (1951) The experimental investigation of navigation in homing pigeons. *J. Exp. Biol.* 28(4):508–536.

[58] Wallraff HG (2001) Navigation by homing pigeons: updated perspective. *Ethol. Ecol. Evol.* 13(1):1–48.

[59] Flack A, Biro D (2013) Collective learning in route navigation. *Commun. Integr. Biol.* 6(6):e26521.

[60] Valentini G, et al. (2021) Naïve individuals promote collective exploration in homing pigeons. *eLife* 10:e68653.

[61] Brønnvik H, Nourani E, Fiedler W, Flack A (2024) Experience reduces route selection for conspecifics by the collectively migrating white stork. *Curr. Biol.* 34(9):2030–2037.e3.

[62] Pettit B, Flack A, Freeman R, Guilford T, Biro D (2013) Not just passengers: pigeons, columba livia, can learn homing routes while flying with a more experienced conspecific. *Proc. R. Soc. B* 280(1750):20122160.

[63] Whiten A, Biro D, Bredeche N, Garland EC, Kirby S (2022) The emergence of collective knowledge and cumulative culture in animals, humans and machines. *Philos. Trans. R. Soc. B* 377(1843):20200306.

[64] Dalmaijer ES (2024) Cumulative route improvements spontaneously emerge in artificial navigators even in the absence of sophisticated communication or thought. *PLoS Biol.* 22(6):e3002644.

[65] Anderson C, Franks NR, McShea DW (2001) The complexity and hierarchical structure of tasks in insect societies. *Animal Behaviour* 62(4):643–651.

[66] Sasaki T, Masuda N, Mann RP, Biro D (2022) Empirical test of the many-wrongs hypothesis reveals weighted averaging of individual routes in pigeon flocks. *iScience* 25(10):105076.

[67] Langridge EA, Sendova-Franks AB, Franks NR (2008) How experienced individuals contribute to an improvement in collective performance in ants. *Behav. Ecol. Sociobiol.* 62:447–456.

[68] Migliano AB, Vinicius L (2022) The origins of human cumulative culture: from the foraging niche to collective intelligence. *Philos. Trans. R. Soc. B* 377(1843):20200317.

[69] Gruber T, Chimento M, Aplin LM, Biro D (2022) Efficiency fosters cumulative culture across species. *Philos. Trans. R. Soc. B* 377(1843):20200308.

[70] Caldwell CA, Millen AE (2008) Studying cumulative cultural evolution in the laboratory. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363(1509):3529–3539.

[71] Lefebvre L (1985) Stability of flock composition in urban pigeons. *The Auk* 102(4):886–888.

[72] Meade J, Biro D, Guilford T (2005) Homing pigeons develop local route stereotypy. *Proc. R. Soc. B* 272(1558):17–23.

[73] Guilford T, Biro D (2014) Route following and the pigeon's familiar area map. *J. Exp. Biol.* 217(2):169–179.

[74] Biro D, Meade J, Guilford T (2004) Familiar route loyalty implies visual pilotage in the homing pigeon. *Proc. Natl. Acad. Sci. U.S.A* 101(50):17440–17443.

[75] Berdahl A, Torney CJ, Ioannou CC, Faria JJ, Couzin ID (2013) Emergent Sensing of Complex Environments by Mobile Animal Groups. *Science* 339(6119):574–576.

[76] Biro D, Sasaki T, Portugal SJ (2016) Bringing a time–depth perspective to collective animal behaviour. *Trends Ecol. Evol.* 31(7):550–562.

[77] Kao AB, Banerjee SC, Francisco FA, Berdahl AM (2024) Timing decisions as the next frontier for collective intelligence. *Trends Ecol. Evol.*

[78] McLean DJ, Skowron Volponi MA (2018) trajr: an R package for characterisation of animal trajectories. *Ethology* 124(6):440–448.

[79] Brooks ME, et al. (2017) glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal* 9(2):378–400.

[80] Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 67:1–48.

[81] Searle SR, Speed FM, Milliken GA (1980) Population Marginal Means in the Linear Model: An Alternative to Least Squares Means. *Am. Stat.* 34(4):216–221.

[82] Bates DM, Watts DG (1988) *Nonlinear regression analysis and its applications.* (Wiley New York) Vol. 2.