Leading an urban invasion: risk-sensitive learning is a winning strategy

Alexis J. Breen[1] and Dominik Deffner[2,3]

[1]Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany

[2]Science of Intelligence Excellence Cluster, Technical University Berlin, Berlin 10623, Germany

[3]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin 14195, Germany

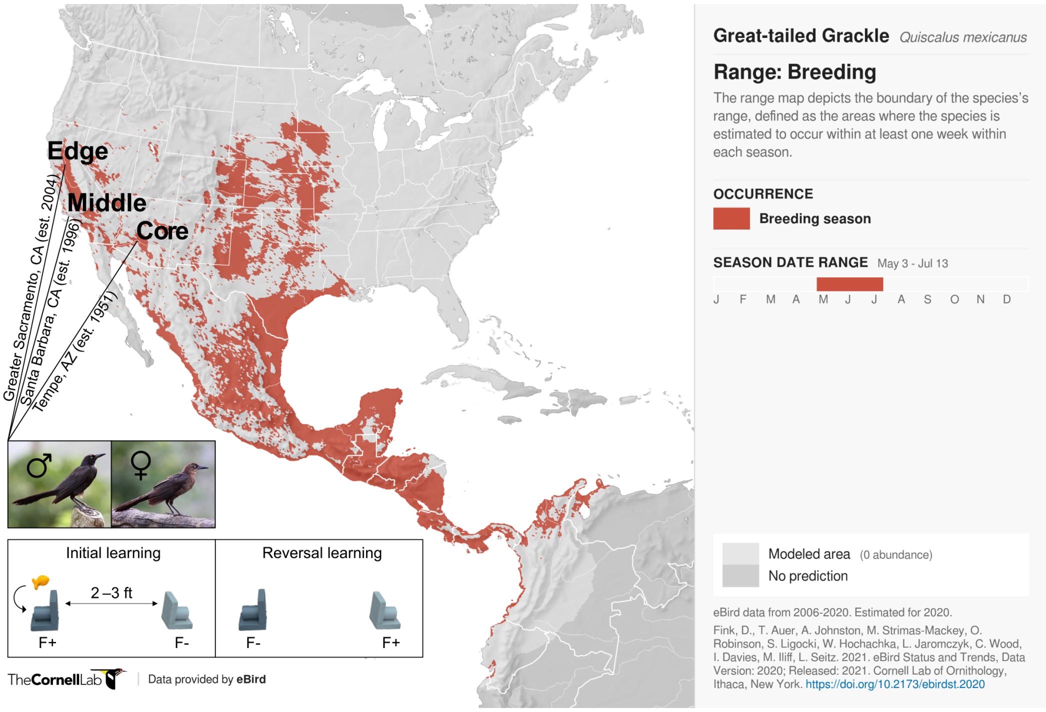
Correspondence: alexis_breen@eva.mpg.de; deffner@mpib-berlin.mpg.de

How do animals successfully invade urban environments? Sex-biased dispersal and learning arguably influence movement ecology, but their joint influence remains unexplored empirically, which might vary by space and time. We assayed reinforcement learning in wild-caught, temporarily-captive core-, middle- or edge-range inhabitants of great-tailed grackles—a bird species undergoing urban-tracking rapid range expansion, led by dispersing males. Across populations, Bayesian models revealed: both sexes initially learn at similar pace, but, when reward contingencies reverse, males—versus females—'relearn' faster via pronounced reward-payoff sensitivity, a risk-sensitive learning strategy. Confirming this mechanism, agent-based forward simulations of reinforcement learning replicate our sex-difference data. Separate evolutionary modelling revealed risk-sensitive learning is favoured by natural selection in stable but stochastic settings—characteristics typical of urban environments. Risk-sensitive learning, then, is a winning strategy for urban-invasion leaders, implying life history (sex-biased dispersal) and cognition (learning) interactively shape invasion success in the unpredictable Anthropocene. Our study sets the scene for future comparative research.

Introduction

Dispersal and range expansion go 'hand in hand'; movement by individuals away from a population's core is a pivotal precondition of witnessed growth in species' geographic limits [1, 2]. Because 'who' disperses—in terms of sex—varies both within and across taxa (for example, male-biased dispersal is dominant among fish and mammals, whereas female-biased dispersal is dominant among birds [3]), skewed sex ratios are apt to arise at expanding range fronts, and, in turn, differentially drive invasion dynamics [4]. Female-biased dispersal, for instance, can 'speed up' staged invertebrate invasions by increasing offspring production [5]. Alongside sex-biased dispersal, learning is also argued to contribute to species' colonisation capacity, as novel environments inevitably present novel (foraging, predation, shelter, and social) challenges that newcomers need to surmount in order to settle successfully [6, 7, 8]. Indeed, a growing number of studies show support for this supposition (recently reviewed in [9]). Carefully controlled choice tests, for example, show urban-dwelling individuals—that is, the invaders—will learn novel stimulus-reward pairings more readily than do rural-dwelling counterparts, supporting the idea that urban invasion selects for learning phenotypes at the dispersal and/or settlement stage(s) [10]. Given the independent influence of sex-biased dispersal and learning on range expansion, it is perhaps surprising, then, that their potential interactive influence on movement ecology remains unexamined empirically (but for recent theoretical work, see [11, 12]), particularly in light of concerns over (in)vertebrates' resilience to ever-increasing urbanisation [13, 14].

Great-tailed grackles (*Quiscalus mexicanus*; henceforth, grackles) are an excellent model for empirical examination of the interplay between sex-biased dispersal, learning, and ongoing urban-targeted rapid range expansion: over the past $\sim$150 years, they have (seemingly) shifted their historically urban niche to include more variable urban environments (e.g., arid habitat; [15]), moving from their native range in Central America into much of the United States, with several first-sightings spanning as far north as Canada [16, 17, 18]. Notably, the record of this urban invasion is heavily peppered with first-sightings involving a single or multiple male(s) (41 of 63 recorded cases spanning most



eBird data from 2006-2020. Estimated for 2020.
Fink, D., T. Auer, A. Johnston, M. Strimas-Mackey, O. Robinson, S. Ligocki, W. Hochachka, L. Jaromczyk, C. Wood, I. Davies, M. Iliff, L. Seitz. 2021. eBird Status and Trends, Data Version: 2020; Released: 2021. Cornell Lab of Ornithology, Ithaca, New York. https://doi.org/10.2173/ebirdst.2020

Fig. 1. Participants and experimental protocol. Thirty-two male and 17 female wild-caught, temporarily-captive great-tailed grackles either inhabiting a core (17 males, 5 females), middle (4 males, 4 females) or edge (11 males, 8 females) population of their North American breeding range (establishment year: 1951, 1996, and 2004, respectively), are participants in the current study (grackle images: Wikimedia Commons). Each grackle is individually tested on a two-phase reinforcement learning paradigm: *initial learning*, two colour-distinct tubes are presented, but only one coloured tube (e.g., dark grey) contains a food reward (F+ versus F-); *reversal learning*, the stimulus-reward tube-pairings are swapped. The learning criterion is identical in both learning phases: 17 F+ choices out of the last 20 choices, with trial 17 being the earliest a grackle can successfully finish (for details, see Methods and fig. S1).

of the twentieth century; [16]). Moreover, recent genetic data show, when comparing grackles within a population, average relatedness: (*i*) is higher among females than among males; and (*ii*) decreases with increasing geographic distance among females; but (*iii*) is unrelated to geographic distance among males; hence, confirming urban invasion in grackles is male-led via sex-biased dispersal [19]. Considering these life history and genetic data in conjunction with data on grackle wildlife management efforts (e.g., pesticides, pyrotechnics, and sonic booms; [20]), we expect urban invasion to drive differential learning between male and female grackles, potentially resulting in a spatial sorting of the magnitude of this sex difference with respect to population establishment age (i.e., sex-effect: newer population > older population; [21]). In range-expanding western bluebirds (*Sialia mexicana*), for example, more aggressive males disperse towards the invasion front; however, in as little as three years, the sons of these colonisers show considerably reduced aggression, as the invasion front moves on [22].

Here, for the first time (to our knowledge), we examine whether, and, if so, how sex mediates learning across 32 male and 17 female wild-caught, temporarily-captive grackles either inhabiting a core (17 males, 5 females), middle (4 males, 4 females) or edge (11 males, 8 females) population of their North American range (based on year-since-first-breeding: 1951, 1996, and 2004, respectively; details in Methods; Fig. 1). Collating, cleaning, and curating existing reinforcement learning data (see Data provenance)—wherein novel stimulus-reward pairings are presented (i.e., *initial learning*), and, once successfully learned, these reward contingencies are reversed (i.e., *reversal learning*)—we test the hypothesis that sex differences in learning are related to sex differences in dispersal. As range expansion should disfavour slow, error-prone learning strategies, we expect male and female grackles to differ across at least two reinforcement learning behaviours: speed and choice-option switches. Specifically, as documented in our preregistration (see Supplementary Materials), we expect male—versus female—great-tailed grackles: (prediction 1 and 2) to be faster to, firstly, learn a novel colour-reward pairing, and secondly, reverse their colour preference when the colour-reward pairing is swapped; and (prediction 3) to make fewer choice option-switches during their colour-reward learning; if learning and dispersal relate. Finally, we further expect (prediction 4) such sex-mediated differences in learning to be more pronounced in grackles living at the edge, rather than the intermediate and/or core region of their range.

Because (dis)similar behaviour can result from multiple latent processes [23], we next employ computational methods to delimit reinforcement learning mechanisms. Specifically, we adapt a multi-level Bayesian reinforcement learning model (from [24]), which we validate *apriori* via agent-based simulation (see Methods and Supplementary Materials), to estimate the contribution of two core latent learning parameters: the *information-updating rate* ϕ (How rapidly do learners 'revise' knowledge?) and the *risk-sensitivity rate* λ (How strongly do learners 'weight' knowledge?). Both learning parameters capture individual-level internal response to incurred reward-payoffs (full mathematical details in Methods). Specifically, as $\phi_{0\rightarrow1}$, information-updating increases; as $\lambda_{0\rightarrow\infty}$, risk-sensitivity strengthens. In other words, by formulating our scientific model as a statistical model, we can reverse engineer which values of our learning parameters most likely produce grackles' choice-behaviour—an analytical advantage over less mechanistic methods [23].

To determine definitively whether our latent learning parameters are sufficient to generate grackles' observed reinforcement learning, we then conduct agent-based forward simulations; that is, we simulate 'birds' informed by the grackles in our data set. Specifically, whilst maintaining the correlation structure among learning parameters, we randomly assign 5000 'males' and 5000 'females' information-updating rate ϕ and risk-sensitivity rate λ estimates from the full across-population posterior distribution of our reinforcement learning model, and we track synthetic reinforcement learning trajectories. By comparing these synthetic data to our real data, we gain valuable insight into the learning and choice behaviour implied by our reinforcement learning model results. Specifically, a close mapping between the two data sets would indicate our information-updating rate λ and risk-sensitivity rate λ estimates can account for grackles' differential (or not) reinforcement learning; whereas a poor mapping would indicate some important mechanism(s) are missing (e.g., [24]).

Learning mechanisms in grackles obviously did not evolve to be successful in the current study; instead, they likely reflect selection pressures and/or adaptive phenotypic plasticity imposed by urban environments [9, 6, 25, 26, 27]. Applying an evolutionary algorithm model (details in Methods), we conclude by examining how urban environments might favour different information-updating rate ϕ and risk-sensitivity rate λ values, by estimating optimal learning strategies in settings that differ along two key ecological axes: *environmental stability* u (How often does optimal behaviour change?) and *environmental stochasticity* s (How often does optimal behaviour fail to payoff?). Urban environments are generally characterised as both stable (lower u) and stochastic (higher s): more specifically, urbanisation routinely leads to stabilised biotic structure, including predation pressure, thermal habitat, and resource availability, and to enhanced abiotic disruption, such as anthropogenic noise and light pollution (reviews: [28, 29, 30]). Seasonal survey data from (sub)urban British neighborhoods show, for example, 40%-75% of households provide supplemental feeding resources for birds (e.g., seed, bread, and peanuts), the density of which can positively relate to avian abundance within an urban area [31, 32, 33]. But such supplemental feeding opportunities are necessarily traded off against increased vigilance due to unpredictable predator-like anthropogenic disturbances (e.g., automobile and airplane traffic; as outlined in [34]).

In summary, here, we map a full pathway from reinforcement learning behaviour to underlying reinforcement learning mechanisms to their selection and adaptive value in urban-like environments, to comprehensively examine links between life history (male-biased dispersal) and cognition (learning) in an active urban-invader—grackles.

Results

Reinforcement learning behaviour

We observe robust reinforcement learning dynamics across populations (full model outputs and 89% highest posterior density intervals (HPDI) for all three study populations in table S1 and S2). As such, we compare male and female grackles' reinforcement learning across populations. Both sexes start out as similar learners, finishing initial learning in comparable trial numbers, and with comparable counts of choice-option switches (median number of trials and switches in initial learning: males, 32 and 10.5; females, 35 and 15; respectively; Fig. 2, A to C and table S1 and S2). Once reward contingencies reverse, however, male—versus female—grackles overwhelmingly finish this 'relearning' in fewer trials and with fewer choice-option switches (median number of trials and switches in reversal learning: males, 64 and 25; females, 81 and 35; respectively; Fig. 2, D to F and table S1 and S2). Environmental unpredictability, then, dependably directs disparate reinforcement learning trajectories between male and female

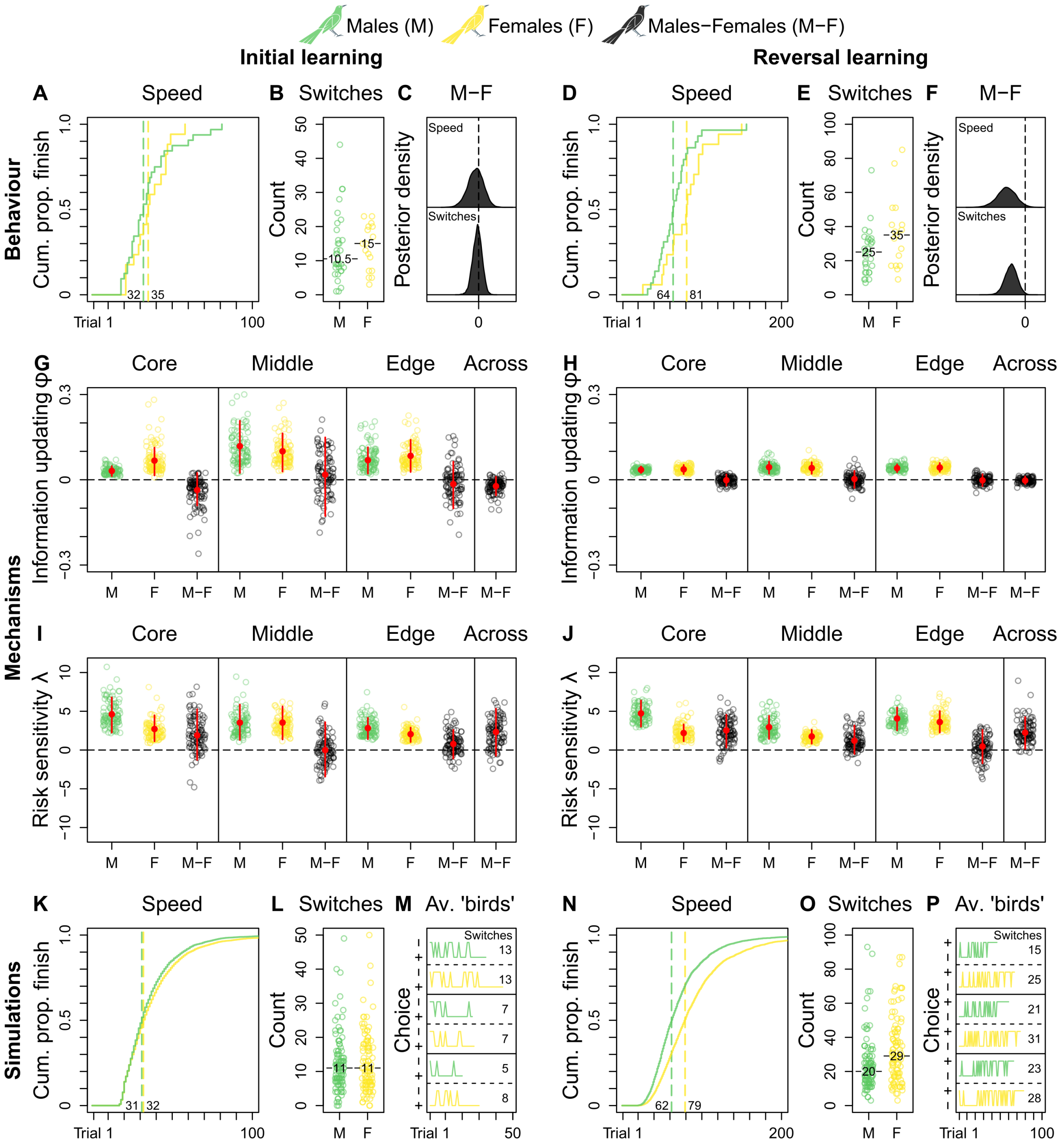


Fig. 2. Grackle reinforcement learning: behaviour, mechanisms, and simulations. (**A** to **F**) *Reinforcement learning behaviour*. Across-population comparison of grackles' reinforcement learning speed and choice-option switches in (**A** and **B**) initial learning (males, 32; females, 17) and (**D** and **E**) reversal learning (males, 29; females, 17), with (**C** and **D**) respective posterior estimates and contrasts. In (**A** and **D**) dashed vertical lines show labelled median trials-to-finish; in (**B** and **E**) floating numbers show median choice-option switch counts. (**G** to **J**) *Reinforcement learning mechanisms*. Within- and across-population estimates and contrasts of grackles' information-updating rate ϕ and risk-sensitivity rate λ in (**G** and **I**) initial and (**H** and **J**) reversal learning. In (**G** to **J**) open circles show 100 random posterior draws; red filled circles and red vertical lines show posterior means and 89% highest posterior density intervals, respectively. (**K** to **P**). *Reinforcement learning agent-based forward simulations*. Comparison of reinforcement learning speed and choice-option switches: by 10,000 full posterior-informed 'birds' (n = 5000 per sex) in (**K** and **L**) initial learning and (**N** and **O**) reversal learning; and by six average posterior-informed 'birds' (n = 3 per sex) in (**M**) initial learning and (**P**) reversal learning. In (**K** and **N**) the full simulation sample is plotted with labelled median trials-to-finish (dashed vertical lines); in (**L** and **O**) open circles correspond to 100 random simulant draws but the floating switch-count medians are calculated from our full simulation; in (**M** and **P**) simulant switch counts are also labelled. Note (**K** and **N**) x-axes are truncated to match (**A** and **D**) x-axes. For all M-F contrasts, the further the distribution falls from zero the stronger the sex difference. Plots (**E** to **L**) are generated via model estimates using our full sample size: 32 males and 17 females.

grackles (faster versus slower finishers, respectively), supporting our overall expectation of sex-mediated differential learning in urban-invading grackles.

Reinforcement learning mechanisms

Looking at our reinforcement learning model's estimates between populations to determine replicability (full model output and 89% HPDI in tables S3 and S4), we observe: in initial learning, the information-updating rate ϕ of core- and edge-inhabiting male grackles is largely lower than that of female counterparts, with smaller sample size likely explaining the middle population's more uncertain estimates (Fig. 2G and table S3); while in reversal learning, the information-updating rate ϕ of both sexes is nearly identical irrespective of population membership, with females dropping to the reduced level of males (Fig. 2H and table S3). Therefore, the information-updating rate ϕ across male and female grackles is initially different (males $<$ females), but converges downwards over reinforcement learning phases (see across-population contrasts in Fig. 2, G and H and table S3).

These primary mechanistic findings are, at first glance, perplexing: if male grackles generally outperform females grackles in reversal learning (Fig. 2, D to F), why do all grackles ultimately update information at matched, dampened pace? This apparent conundrum, however, in fact highlights the potential for multiple latent processes to direct behaviour. Case in point: the risk-sensitivity rate λ is distinctly higher in male grackles, compared to female counterparts, regardless of population membership and learning phase, outwith the middle population in initial learning likely due to sample size (Fig. 2, I and J and table S4). In other words, choice behaviour in male grackles is more strongly governed by the relative differences in predicted reward-payoffs (see across-population contrasts in Fig. 2, I and J and table S4). Thus, these combined mechanistic data reveal, when reward contingencies reverse, male—versus female—grackles 'relearn' faster via pronounced reward-payoff sensitivity—a persistence-based risk-sensitive learning strategy.

Agent-based replication of reinforcement learning

Ten thousand synthetic reinforcement learning trajectories, together, compellingly show our 'birds' behave just like our grackles: 'males' overwhelming outpace 'females' in reversal but not in initial learning (median number of trials in initial and reversal learning: 'males', 31 and 62; females, 32 and 79; respectively; Fig. 2, K and N); and 'males' make markedly fewer choice-option switches in initial but not in reversal learning, compared to 'females' (median number of switches in initial and reversal learning: 'males', 11 and 20; 'females', 11 and 29; respectively; Fig. 2, L and O). Fig. 2, M and P show, respectively, synthetic initial learning and reversal learning trajectories by three average 'males' and three average 'females' (i.e., simulants informed via learning parameter estimates that average over our posterior distribution), for the reader interested in representative individual-level reinforcement learning dynamics.

Selection and adaptive value of reinforcement learning mechanisms under urban-like environments

Strikingly, under characteristically urban-like (i.e., stable but stochastic) conditions, our evolutionary model shows the learning parameter constellation robustly exhibited by males grackles in our study—that is, low information-updating rate ϕ and high risk-sensitivity rate λ—is favoured by natural selection (Fig. 3, A and B). These results imply, in urban and other statistically similar environments, learners benefit by averaging over prior experience (i.e., gradually updating 'beliefs'), and by informing behaviour based on this experiential history (i.e., proceeding with 'caution'), highlighting the adaptive value of strategising risk-sensitive learning in urban-like environments.

Discussion

In the unpredictable Anthropocene, one particularly pressing open question is how specific species invade urban environments. Contemporary evolutionary theory predicts life history and cognition combine to scaffold successful 'city living' [6, 8, 11, 12, 9]—we provide the first empirical evidence to support this idea. Assaying available data on stimulus-reward initial and reversal reinforcement learning in wild-caught, temporarily-captive male and female grackles either inhabiting a core, middle, or edge population of their range, we show: irrespective of population

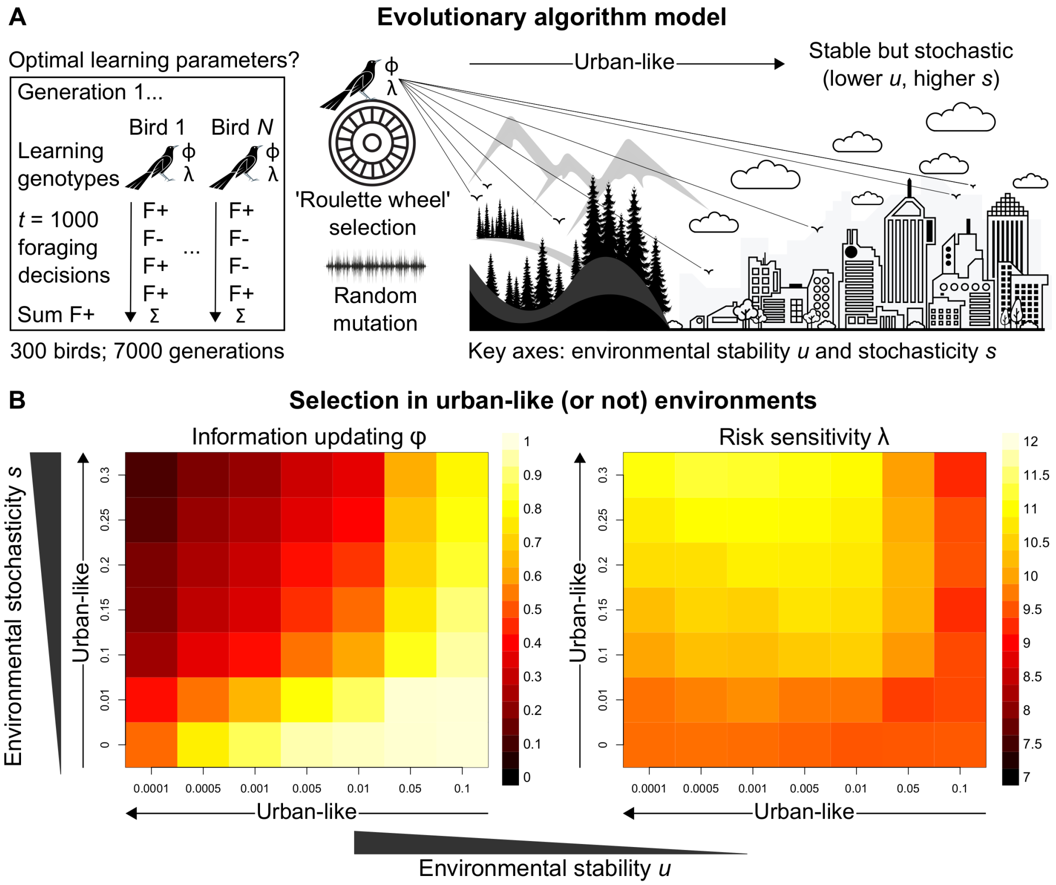


Fig. 3. Evolutionary optimality of strategising risk-sensitive learning. (**A**) Illustration of our evolutionary algorithm model to estimate optimal learning parameters that evolve under systematically varied pairings of two key (urban) ecology axes: environmental stability u and environmental stochasticity s. Specifically, 300-member populations run for 10 independent 7000-generation simulations per pairing, using 'roulette wheel' selection (parents are chosen for reproduction with a probability proportional to collected F+ rewards out of 1000 choices) and random mutation (offspring inherit learning genotypes with a small deviation in random direction). (**B**) Mean optimal learning parameter values discovered by our evolutionary model (averaged over the last 5000 generations). As the statistical environment becomes more urban-like (lower u and higher s values), selection favours lower information-updating rate ϕ and higher risk-sensitivity rate λ (darker and lighter squares, respectively). We note arrows are intended as illustrative aids and do not correspond to a linear scale of 'urbanness'.

membership, both sexes initially learn stimulus-reward pairings at comparable pace, but, when reward contingencies reverse, male grackles overwhelming finish this 'relearning' faster, compared to female counterparts. Under environmental unpredictability, then, male grackles dependably outperform female grackles. How do male grackles do this? Using a mechanistic Bayesian reinforcement learning model, we estimated how (hidden) learning parameters generate grackles' (observed) choice behaviour. We show choice-behaviour in male—versus female—grackles is governed more strongly by the 'weight' of relative differences in recent foraging returns—in other words, male grackles show pronounced reward-payoff sensitivity. Indeed, subsequent agent-based forward simulations of reinforcement learning—where we simulate 'new' birds based on the empirical estimates of grackles' reinforcement learning in our study—show 'males' more readily refocus disrupted foraging behaviour relative to 'females', making fewer choice-option switches and finishing stimulus-reward 'relearning' faster. Such robust quantitative replication thus substantiates our assertion that strategising risk-sensitive learning underpins our behavioural sex-difference data. Finally, using evolutionary modelling to move out from our study into more general types of environments, we show natural selection favours risk-sensitive learning in characteristically urban-like environments: that is, stable but stochastic settings. Together, these data imply risk-sensitive learning is a viable adaptive strategy to help explain how male great-tailed grackles—the dispersing sex—currently lead their species' remarkable North American urban invasion.

The term 'behavioural flexibility'—broadly defined as some 'attribute', 'cognition', 'characteristic', 'feature', 'trait' and/or 'quality' that enables animals to adapt behaviour to changing circumstances [35, 8, 36, 37]—has been previously hypothesised to explain invasion success [7], including that of grackles [15]. But as eloquently argued elsewhere [37], this term is conceptually uninformative, given the many ways in which it is applied and assessed. Of these approaches, reversal learning and serial—multiple back-to-back—reversal learning tasks are the most common experimental assays of behavioural flexibility (non-exhaustive examples of each assays in: bees [38, 39]; birds [40, 41]; fish [42, 43]; frogs [44, 45]; reptiles [10, 46]; primates [47, 48]; and rodents [49, 50]). We have shown, however, at least for our grackles, faster reversal learning is governed primarily by pronounced reward-payoff sensitivity, so: firstly, these go-to experimental assays do not necessarily measure the unit they claim to measure (a point similarly highlighted in [51]); and secondly, formal models based on the false premise that variation in learning speed relates to variation in behavioural flexibility require reassessment [35, 52]. Heeding previous calls [53, 54, 55], our study provides an analytical solution to facilitate productive research on proximate and ultimate explanations of seemingly flexible (or not) behaviour: because we publicly provide step-by-step code to map a full pathway from behaviour to mechanisms through to selection and adaptation (see Data and materials availability), which can be tailored to specific research questions. The reinforcement learning model, for example, generalises to, in theory, a variety of choice-option paradigms [56], and these learning models can be extended to estimate asocial *and* social influence on individual decision making (e.g., [57, 24, 58, 59, 60]), facilitating useful insight into the multi-faceted feedback process between individual cognition and social systems [61]. Our open-access analytical resource thus allows researchers to move beyond the umbrella term behavioural flexibility, and to biologically inform and interpret their science—only then can we begin to meaningfully compare behavioural variation across taxa and/or contexts.

Related to this final point, it is useful to outline how additional drivers outwith sex-biased risk-sensitive learning might contribute towards urban invasion success in grackles, too. Grackles exhibit a polygynous mating system, with territorial males attracting multiple female nesters [62]. Recent learning 'style' simulations show the sex with high reproductive skew approaches pure individual learning, while the other sex approaches pure social learning [63]. During population establishment, then, later-arriving female grackles could rely heavily on vetted information provided by male grackles on 'what to do' [7], as both sexes ultimately face the same urban-related challenges. Moreover, risk-sensitive learning in male grackles should help reduce the elevated risk associated with any skew towards acquiring knowledge through individual learning. And as the dispersing sex this process would operate independently of their proximity to a range front—a pattern suggestively supported by our mechanistic data (i.e., risk-sensitivity: males $>$ females; Fig. 2, G and H). As such, future research on potential sex differences in social learning propensity in grackles seem particularly prudent, alongside systematic surveying of population-level environmental and fitness components, across spatially (dis)similar populations; for this, our annotated and readily available analytical approach should prove useful, as highlighted above.

The lack of spatial replicates in the existing data set used herein inherently poses limitations on inference. But it is worth noting that phenotypic filtering by invasion stage is not a compulsory signature of successful (urban) invasion; instead, phenotypic plasticity and/or inherent species trait(s) may be facilitators [9, 6, 25, 26, 27]. For urban-invading grackles, both of these biological explanations seem strongly plausible, given: firstly, grackles' highly plastic foraging and nesting ecology [64, 65, 17]; secondly, grackles' apparent historic and current realised niche being—albeit in present day more variable—urban environments, a consistent habit preference that cannot be explained by changes in habitat availability or habitat connectivity [15]; and finally, our combined reinforcement learning and evolutionary modelling results showing environments approaching grackles' general species niche—urban environments—select for particular traits that exist across grackle populations (here, sex-biased risk-sensitive learning; Fig. 3B). Admittedly, our evolutionary model is not a complete representation of urban ecology dynamics. Relevant factors—e.g., spatial dynamics and realistic life histories—are missed out. These omissions are tactical ones. Our evolutionary model solely focuses on the response of reinforcement learning parameters to two core urban-like (or not) environmental statistics, providing a baseline for future study to build on; for example, it would be interesting to investigate such selection on learning parameters of 'true' invaders and not their descendants, a logistically tricky but nonetheless feasible research possibility (e.g., [22]).

In conclusion, by revealing robust interactive links between the dispersing sex and risk-sensitive learning in an urban invader (grackles), these fully replicable insights, coupled with our finding that urban-like environments favour pronounced risk-sensitivity, underscore the potential for life history *and* cognition to scaffold invasion success in the unpredictable Anthropocene. Our modelling methods, which we document in-depth and make freely available, can now be comparatively applied, establishing a biologically meaningful analytical approach for much-needed study on (shared or divergent) drivers of geographic and/or phenotypic distributions [9, 66, 67, 68, 69, 70].

Methods

Data provenance

The current study uses data from two types of sources: publicly archived data at the Knowledge Network for Biocomplexity [71, 72]; or privately permissed access to A.J.B. of unpublished data by Grackle Project principal investigator Corina Logan, who declined participation on this study.

Data contents

The data sets used herein encompass colour-reward reinforcement learning data from 32 male and 17 female wild-caught, temporarily-captive grackles across three study sites that differ in their range-expansion demographics; that is, defined as a core, middle or edge population (based on time-since-settlement population growth dynamics, as outlined in [1]). Specifically, the colour-reward reinforcement learning data originated from the following three populations: (*i*) Tempe, Arizona (17 males and five females)—herein, the core population (estimated to be breeding since 1951, by adding the average time between first sighting and first breeding to the year first sighted [17, 73]); (*ii*) Santa Barbara, California (four males and four females)—herein, the middle population (known to be breeding since 1996 [74]); and (*iii*) Greater Sacramento, California (eleven males and eight females)—herein, the edge population (known to be breeding since 2004 [75]).

Experimental protocol

A step-by-step description of the experimental protocol carried out by the original experiments is reported elsewhere [52]. As such, below we detail only the protocol for the colour-reward reinforcement learning test that we analysed herein.

Reinforcement learning test

The reinforcement learning test consists of two experimental phases (Fig. 1): (*i*) stimulus-reward initial learning (hereafter, initial learning) and (*ii*) stimulus-reward reversal learning (hereafter, reversal learning). In both experimental phases, two different coloured tubes are used: for Santa Barbara grackles, gold and grey; for all other grackles, light and dark grey. Each tube consists of an outer and inner diameter of 26 mm and 19 mm, respectively; and each is mounted to two pieces of plywood attached at a right angle (entire apparatus: 50 mm wide × 50 mm tall × 67 mm deep); thus resulting in only one end of each coloured tube being accessible (Fig. 1).

In initial learning, grackles are required to learn that only one of the two coloured tubes contains a food reward (e.g., dark grey); this colour-reward pairing is counterbalanced across grackles within each study site. Specifically, the rewarded and unrewarded coloured tubes are placed—either on a table or on the floor—in the centre of the aviary run (distance apart: table, 2 feet; floor, 3 feet), with the open tube-ends facing, and perpendicular to, their respective aviary side-wall. Which coloured tube is placed on which side of the aviary run (left or right) is pseudorandomised across trials. A trial begins at tube-placement, and ends when a grackle has either made a tube-choice or the maximum trial time has elapsed (eight minutes). A tube-choice is defined as a grackle bending down to examine the contents (or lack thereof) of a tube. If the chosen tube contains food, the grackle is allowed to retrieve and eat the food, before both tubes are removed and the rewarded coloured tube is rebaited out of sight (for the grackle). If a chosen tube does not contain food, both tubes are immediately removed. Each grackle is given, first, up to three minutes to make a tube-choice, after which a piece of food is placed equidistant between the tubes to entice participation; and then, if no choice has been made, an additional five minutes maximum, before both tubes are removed. All trials are recorded as either correct (choosing the rewarded coloured tube), incorrect (choosing the unrewarded coloured tube), or incomplete (no choice made). To successfully finish initial learning, a grackle must meet the learning criterion, detailed below.

In reversal learning, grackles are required to learn that the colour-reward pairing has been swapped; that is, the previously unrewarded coloured tube (e.g., light grey) now contains a food reward. The protocol for this second and final experimental phase is identical to that, described above, of initial learning.

Learning criterion

For all grackles in the current study, we apply the following learning criterion: to successfully finish their respective learning phase, grackles must make a correct choice in 17 of the most recent 20 trials. Therefore, the earliest a grackle can successfully finish initial or reversal learning in the current study is at trail 17. This applied learning criterion is the most compatible with respect to previous learning criteria used by the original experimenters. Specifically, Logan [72] and Logan et al. [72] used a fixed-window learning criterion for core- and middle-population grackles, in which grackles were required to make 17 out of the last 20 choices correctly, with a minimum of eight and nine correct choices across the last two sets of 10 trials, assessed at the end of each set. If a core- or middle-population grackle successfully satisfied the fixed-window learning criterion, the grackle was assigned by Logan or colleagues the end-trial number for that set (e.g., 20, 30, 40), which did not always coincide with the true passing trial (by a maximum of two additional trials). For edge-population grackles, Logan and colleagues used a sliding-window learning criterion (determined via inspection of the original data file Woodland_CA_data.csv available at our GitHub repository), in which grackles were required to again make 17 out of the last 20 choices correctly, with the same minimum correct-choice counts for the previous two 10-trial sets, except that this criterion was assessed at every trial (from 18 onward) rather than at the end of discrete sets. This second method is problematic because a grackle can successfully reach criterion via a shift in the sliding window *before* making a choice. For example, a grackle could make three wrong choices followed by 17 correct choices (i.e., 7/10 correct and 10/10 correct in the last two sets of 10 trials), and at the start of the next trial, the grackle will reach criterion because the summed choices now consist of 8/10 correct and at least 9/10 correct in the last two sets of 10 trials no matter their subsequent choice—see initial learning performance by bird 'Kel' for a real example. Moreover, the use of different learning criteria (fixed- and sliding-window) by Logan and colleagues in different populations represents a confound when populations are compared. Thus, our applied learning criterion ensures our assessment of grackles' reinforcement learning is informative, straightforward, and consistent.

We note, as a consequence of applying our learning criterion, grackles can remain in initial and/or reversal learning beyond reaching criterion. These extra learning trials, however, are already inherent to core- and middle-population grackles originally assessed via the fixed-window learning criterion (see above). And our cleaning of the original data detected additional cases where edge-population grackles already remained in-test after meeting the sliding-window learning criterion (see our Data_Processing.R script at our GitHub repository), presumably due to experimenter oversight. In any case, to ensure our analyses capture grackles' reinforcement learning up until our applied learning criterion, we exclude extra learning trials, where applicable. Prior to doing so, however, we verified: (*i*) a similar proportion of male and female grackles experience extra initial learning trials (females, 15/17; males, 30/32); and (*ii*) our learning parameter estimations during initial learning remain relatively unchanged irrespective of whether we exclude or include extra initial learning trails (fig. S1). In short, we are confident that any carry-over effect of extra initial learning trials on grackles' reversal learning is negligible if nonexistent.

Statistical analyses

We analysed, processed, and visually present our data using, respectively, the 'rstan' [76], 'rethinking' [23], and 'tidyverse' [77] packages in R [78]. We note our reproducible code is available at our GitHub repository. We further note our reinforcement learning model, defined below, does not exclude cases—two males in the core, and one male in the middle population—where a grackle was dropped (due to time constraints) early on from reversal learning by experimenters: because individual-level ϕ and λ estimates can still be estimated irrespective of trial number; the certainty around the estimates will simply be wider [23]. Our Poisson models, however, do exclude these three cases for our modelling of reversal learning, to conserve estimation. The full output from each of our models, which use weakly informative and conservative priors, is available in tables S1 to S4, including 89% highest-posterior density intervals [23].

Poisson models

For our behavioural assay of reinforcement learning finishing trajectories, we used a multi-level Bayesian Poisson regression to quantify the effect(s) of sex and learning phase (initial versus reversal) on grackles' recorded number of trials to successfully finish each phase. This model was performed at both the population and across-population level, and accounted for individual differences among birds through the inclusion of individual-specific varying (i.e., random) effects.

For our behavioural assay of reinforcement learning choice-option switching, we used an identical Poisson model to that described above, to predict the total number of switches between the rewarded and unrewarded coloured tubes.

Reinforcement learning model

We employed an adapted (from [24]) multi-level Bayesian reinforcement learning model, to examine the influence of sex on grackles' initial and reversal learning. Our reinforcement learning model, defined below, allows us to link observed coloured tube-choices to latent individual-level attraction updating, and to translate the influence of latent attractions (i.e., expected payoffs) into individual tube-choice probabilities. As introduced above, we can reverse engineer which values of our two latent learning parameters—the information-updating rate ϕ and the risk-sensitivity rate λ—most likely produce grackles' choice-behaviour, by formulating our scientific model as a statistical model. Therefore, this computational method facilitates mechanistic insight into how multiple latent learning parameters simultaneously guide grackles' reinforcement learning [23].

Our reinforcement learning model consists of two equations:

$$A_{i,j,t+1} = (1 - \phi_{k,l})A_{i,j,t} + \phi_{k,l}\pi_{i,j,t} \tag{1}$$

$$P(i)_{t+1} = \frac{\exp(\lambda_{k,l}A_{i,j,t})}{\sum\limits_{m=1}^{2} \exp(\lambda_{k,l}A_{m,j,t})} \tag{2}$$

Equation 1 expresses how attraction A to choice-option i changes for an individual j across time $(t + 1)$ based on their prior attraction to that choice-option $(A_{i,j,t})$ plus their recently experienced choice reward-payoffs $(\pi_{i,j,t})$, whilst accounting for the relative influence of recent reward-payoffs $(\phi_{k,l})$. As $\phi_{k,l}$ increases in value, so, too, does the rate of individual-level attraction updating based on reward-payoffs. Here, then, $\phi_{k,l}$ represents the information-updating rate. We highlight that the k, l indexing (here and elsewhere) denotes that we estimate separate ϕ parameters for each population (k = 1 for core; k = 2 for middle; k = 3 for edge) and for each learning phase (l = 1 for females/initial, l = 2 for females/reversal; l = 3 for males/initial; l = 4 for males/reversal).

Equation 2 is a *softmax* function that expresses the probability P that choice-option i is selected in the next choice-round $(t+1)$ as a function of the attractions A and the parameter $\lambda_{k,l}$, which governs how much relative differences in attraction scores guide individual choice-behaviour. In the reinforcement learning literature, the λ parameter is known by several names—for example, 'inverse temperature', 'exploration' or 'risk-appetite' [79, 58]—since the higher its value the more deterministic the choice-behaviour of an individual becomes (note $\lambda = 0$ generates random choice). In line with foraging theory [80], we call λ the risk-sensitivity rate, where higher values of λ imply foragers are more sensitive to risk, seeking higher expected payoffs based on their prior experience, instead of randomly sampling alternative options.

From the above reinforcement learning model, then, we generate inferences about the effect of sex on $\phi_{k,l}$ and $\lambda_{k,l}$ from at least 1000 effective samples of the posterior distribution, at both the population- and across-population-level. We note our reinforcement learning model also includes bird as a random effect (to account for repeated measures within individuals); however, for clarity, this parameter is omitted from our equations (but not our code—see our GitHub repository). Our reinforcement learning model does not, on the other hand, include trials where a grackle did not make a tube-choice, as this measure cannot clearly speak to individual learning—for example, satiation rather than any learning of 'appropriate' colour tube-choice could be invoked as an explanation in such cases. Indeed, there are, admittedly, a number of intrinsic and extrinsic factors (e.g., temperament and temperature, respectively) that might bias grackles' tube-choice behaviour, and, in turn, the output from our reinforcement learning model [81]. But the aim of such models is not to replicate the entire study system. Finally, we further note, while we exclude extra learning trials from all of our analyses (see above), our reinforcement learning model initiates estimation of ϕ and λ during reversal learning, based on individual-level attractions encompassing all previous choices. This parameterisation ensures we precisely capture grackles' attraction scores up to the point of stimulus-reward reversal (for details, see our RL_Execution.R script at our GitHub repository).

Agent-based simulations: pre- and post-study

Prior to analysing our data, we used agent-based simulations to validate our reinforcement learning model (full details in our preregistration–see Supplementary Materials). In brief, the tube-choice behaviour of simulants is governed by a set of rules identical to those defined by equations 1 and 2, and we apply the exact same learning criterion for successfully finishing both learning phases. Crucially, this *apriori* model vetting verifies our reinforcement learning model can (*i*) detect simulated sex effects and (*ii*) accurately recover simulated parameter values in both extreme and more realistic scenarios.

After model fitting, we used the same agent-based approach to forward simulate—that is, simulate via the posterior distribution—synthetic learning trajectories by 'new' birds via individual-level parameter estimates generated from our across-population reinforcement learning model. Specifically, maintaining the correlation structure among sex- and phase-specific learning parameters, we draw samples from the full or averaged random-effects multivariate normal distribution describing the inferred population of grackles. We use these post-study forward simulations to gain a better understanding of the implied consequences of the estimated sex differences in grackles' learning parameters (see Fig. 2 and associated main text; for an example of this approach in a different context, see [24]).

Evolutionary model

To investigate the evolutionary significance of strategising risk-sensitive learning, we used algorithmic optimisation techniques [82, 83]. Specifically, we construct an evolutionary model of grackle learning, to estimate how our learning parameters—the information-updating rate ϕ and the risk-sensitivity rate λ—evolve in environments that systematically vary across two ecologically relevant (see main text) statistical properties: the rate of environmental stability u and the rate of environmental stochasticity s. The environmental stability parameter u represents the probability that behaviour leading to a food reward changes from one choice to the next. If u is small, individuals encounter a world where they can expect the same behaviour to be adaptive for a relatively long time. As u becomes larger, optimal behaviour can change multiple times within an individual's lifetime. The environmental stochasticity parameter s describes the probability that, on any given day, optimal behaviour may not result in a food reward due to external causes specific to this day. If s is small, optimal behaviour reliably produces rewards. As s becomes larger, there is more and more daily 'noise' regarding which behaviour is rewarded.

We consider a population of fixed size with $N = 300$ individuals. Each generation, individual agents are born naïve and make $t = 1000$ binary foraging decisions resulting in a food reward (or not). Agents decide and learn about the world through reinforcement learning governed by their individual learning parameters, ϕ and λ (see equations 1 and 2). Both learning parameters can vary continuously, corresponding to the infinite-alleles model from population genetics [82]. Over the course of their lifetime, agents collect food rewards, and the sum of rewards collected over the last 800 foraging decisions (or 'days') determines their individual fitness. We ignore the first 200 choices because selection should respond to the steady state of the environment, independently of initial conditions [82].

To generate the next generation, we assume asexual, haploid reproduction, and use fitness-proportionate (or 'roulette wheel') selection to choose individuals for reproduction [82, 83]. Here, juveniles inherit both learning parameters, ϕ and λ, from their parent but with a small deviation (in random direction) due to mutation. Specifically, during each mutation event, a value drawn from zero-centered normal distributions $N(0, \mu_\phi)$ or $N(0, \mu_\lambda)$ is added to the parent value on the logit-/log-scale to ensure parameters remain within allowed limits (between 0 and 1 for ϕ; positive for λ). The mutation parameters μ_ϕ and μ_λ thus describe how much offspring values might deviate from parental values, which we set to 0.05. We restrict the risk-sensitivity rate λ to the interval 0 to 15, because greater values result in identical choice behaviour. All results reported in the main text are averaged over the last 5000 generations of 10 independent 7000-generation simulations per parameter combination. This duration is sufficient to reach steady state in all cases.

In summary, our evolutionary model is a necessary and useful first step towards addressing targeted research questions about the interplay between learning phenotype and environmental characteristics.

References

[1] Chuang, A. and Peterson, C. R. (2016) Expanding population edges: theories, traits, and trade-offs. *Global Change Biology* 22.2, 494–512. DOI: `10.1111/gcb.13107`.

[2] Ronce, O. (2007) How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics* 38, 231–253. DOI: `10.1146/annurev.ecolsys.38.091206.095611`.

[3] Trochet, A. et al. (2016) Evolution of sex-biased dispersal. *The Quarterly Review of Biology* 91.3, 297–320. DOI: `10.1086/688097`.

[4] Miller, T. E. et al. (2011) Sex-biased dispersal and the speed of two-sex invasions. *The American Naturalist* 177.5, 549–561. DOI: `10.1086/659628`.

[5] Miller, T. E. X. and Inouye, B. D. (2013) Sex and stochasticity affect range expansion of experimental invasions. *Ecology Letters* 16.3, 354–361. DOI: `10.1111/ele.12049`.

[6] Sol, D., Lapiedra, O., and González-Lagos, C. (2013) Behavioural adjustments for a life in the city. *Animal Behaviour* 85.5, 1101–1112. DOI: `10.1016/j.anbehav.2013.01.023`.

[7] Wright, T. F. et al. (2010) Behavioral flexibility and species invasions: the adaptive flexibility hypothesis. *Ethology Ecology & Evolution* 22.4, 393–404. DOI: `10.1080/03949370.2010.505580`.

[8] Barrett, L. P., Stanton, L. A., and Benson-Amram, S. (2019) The cognition of 'nuisance'species. *Animal Behaviour* 147, 167–177. DOI: `10.1016/j.anbehav.2018.05.005`.

[9] Lee, V. E. and Thornton, A. (2021) Animal cognition in an urbanised world. *Frontiers in Ecology and Evolution* 9, 120. DOI: `10.3389/fevo.2021.633947`.

[10] Batabyal, A. and Thaker, M. (2019) Lizards from suburban areas learn faster to stay safe. *Biology Letters* 15.2, 20190009. DOI: `10.1098/rsbl.2019.0009`.

[11] Liedtke, J. and Fromhage, L. (2021a) The joint evolution of learning and dispersal maintains intraspecific diversity in metapopulations. *Oikos* 130.5, 808–818. DOI: `10.1111/oik.08208`.

[12] — (2021b) Should dispersers be fast learners? Modeling the role of cognition in dispersal syndromes. *Ecology and Evolution* 11.20, 14293–14302. DOI: `10.1002/ece3.8145`.

[13] Li, G. et al. (2022) Global impacts of future urban expansion on terrestrial vertebrate diversity. *Nature communications* 13.1, 1–12. DOI: `10.1038/s41467-022-29324-2`.

[14] Eisenhauer, N. and Hines, J. (2021) Invertebrate biodiversity and conservation. *Current Biology* 31.19, R1214–R1218. DOI: `10.1016/j.cub.2021.06.058`.

[15] Summers, J. et al. (2023) The role of climate change and niche shifts in divergent range dynamics of a sister-species pair. *Peer Community Journal* 3. DOI: `10.24072/pcjournal.248`.

[16] Dinsmore, J. J. and Dinsmore, S. J. (1993) Range expansion of the great-tailed grackle in the 1900s. *Journal of the Iowa Academy of Science* 100.2, 54–59. DOI: `https://scholarworks.uni.edu/jias/vol100/iss2/4`.

[17] Wehtje, W. (2003) The range expansion of the great-tailed grackle (*Quiscalus mexicanus* Gmelin) in North America since 1880. *Journal of Biogeography* 30.10, 1593–1607. DOI: `10.1046/j.1365-2699.2003.00970.x`.

[18] Fink, D. et al. (n.d.) eBird Status and Trends. Version 2020 (). https://ebird.org/science/status-and-trends. DOI: `10.2173/ebirdst.2020`.

[19] Sevchik, A. et al. (2022) Investigating sex differences in genetic relatedness in great-tailed grackles in Tempe, Arizona to infer potential sex biases in dispersal. *Animal Behavior and Cognition* 9, 37–52. DOI: `10.32942/osf.io/t6beh`.

[20] Luscier, J. D. (2018) Effects of urbanization on great-tailed grackle habitat use and nest success in Sherman Texas. *Urban Naturalist* 15, 1–14.

[21] Phillips, B. L., Brown, G. P., and Shine, R. (2010) Life-history evolution in range-shifting populations. *Ecology* 91.6, 1617–1627. DOI: `10.1890/09-0910.1`.

[22] Duckworth, R. A. and Badyaev, A. V. (2007) Coupling of dispersal and aggression facilitates the rapid range expansion of a passerine bird. *Proceedings of the National Academy of Sciences* 104.38, 15017–15022. DOI: `10.1073/pnas.0706174104`.

[23] McElreath, R. (2018) *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

[24] Deffner, D., Kleinow, V., and McElreath, R. (2020) Dynamic social learning in temporally and spatially variable environments. *Royal Society Open Science* 7.12, 200734. DOI: `10.1098/rsos.200734`.

[25] Vinton, A. C. et al. (2022) Plasticity's role in adaptive evolution depends on environmental change components. *Trends in Ecology & Evolution*. DOI: `10.1016/j.tree.2022.08.008`.

[26] Caspi, T. et al. (2022) Behavioral plasticity can facilitate evolution in urban environments. *Trends in Ecology & Evolution* 37, 1092–1103. DOI: `10.1016/j.tree.2022.08.002`.

[27] Blackburn, T. M., Lockwood, J. L., and Cassey, P. (2009) *Avian invasions: the ecology and evolution of exotic birds*. Oxford University Press.

[28] Shochat, E. et al. (2006) From patterns to emerging processes in mechanistic urban ecology. *Trends in Ecology & Evolution* 21.4, 186–191. DOI: `10.1016/j.tree.2005.11.019`.

[29] Francis, C. D. and Barber, J. R. (2013) A framework for understanding noise impacts on wildlife: an urgent conservation priority. *Frontiers in Ecology and the Environment* 11.6, 305–313. DOI: `https://doi.org/10.1890/120183`.

[30] Gaston, K. J. et al. (2013) The ecological impacts of nighttime light pollution: a mechanistic appraisal. *Biological Reviews* 88.4, 912–927. DOI: `10.1111/brv.12036`.

[31] Fuller, R. A. et al. (2008) Garden bird feeding predicts the structure of urban avian assemblages. *Diversity and Distributions* 14.1, 131–137. DOI: `10.1111/j.1472-4642.2007.00439.x`.

[32] Davies, Z. G. et al. (2009) A national scale inventory of resource provision for biodiversity within domestic gardens. *Biological Conservation* 142.4, 761–771. DOI: `10.1016/j.biocon.2008.12.016`.

[33] Cowie, R. J. and Hinsley, S. A. (1988) The provision of food and the use of bird feeders in suburban gardens. *Bird Study* 35.3, 163–168. DOI: `10.1080/00063658809476985`.

[34] Frid, A. and Dill, L. (2002) Human-caused disturbance stimuli as a form of predation risk. *Conservation Ecology* 6.1.

[35] Lea, S. E. et al. (2020) Behavioral flexibility: a review, a model, and some exploratory tests. *Learning & Behavior* 48, 173–187. DOI: `10.3758/s13420-020-00421-w`.

[36] Coppens, C. M., Boer, S. F. de, and Koolhaas, J. M. (2010) Coping styles and behavioural flexibility: towards underlying mechanisms. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1560, 4021–4028. DOI: `10.1098/rstb.2010.0217`.

[37] Audet, J. N. and Lefebvre, L. (2017) What's flexible in behavioral flexibility? *Behavioral Ecology* 28.4, 943–947. DOI: `10.1093/beheco/arx007`.

[38] Strang, C. G. and Sherry, D. F. (2014) Serial reversal learning in bumblebees (*Bombus impatiens*). *Animal Cognition* 17, 723–734. DOI: `10.1007/s10071-013-0704-1`.

[39] Raine, N. E. and Chittka, L. (2012) No trade-off between learning speed and associative flexibility in bumblebees: a reversal learning test with multiple colonies. *PloS one* 7, e45096. DOI: `10.1371/journal.pone.0045096`.

[40] Bond, A. B., Kamil, A. C., and Balda, R. P. (2007) Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus, Nucifraga columbiana, Aphelocoma californica*). *Journal of Comparative Psychology* 121.4, 372. DOI: `10.1037/0735-7036.121.4.372`.

[41] Morand-Ferron, J., Reichert, M. S., and Quinn, J. L. (2022) Cognitive flexibility in the wild: individual differences in reversal learning are explained primarily by proactive interference, not by sampling strategies, in two passerine bird species. *Learning & Behavior* 50.1, 153–166. DOI: `10.3758/s13420-021-00505-1`.

[42] Lucon-Xiccato, T. and Bisazza, A. (2014) Discrimination reversal learning reveals greater female behavioural flexibility in guppies. *Biology Letters* 10.6, 20140206. DOI: `10.1098/rsbl.2014.0206`.

[43] Bensky, M. K. and Bell, A. M. (2020) Predictors of individual variation in reversal learning performance in three-spined sticklebacks. *Animal cognition* 23.5, 925–938. DOI: `10.1007/s10071-020-01399-8`.

[44] Liu, Y. et al. (2016) Learning to learn: advanced behavioural flexibility in a poison frog. *Animal Behaviour* 111, 167–172. DOI: `10.1016/j.anbehav.2015.10.018`.

[45] Burmeister, S. S. (2022) Ecology, cognition, and the hippocampus: A tale of two frogs. *Brain, Behavior and Evolution* 97.3-4, 211–224. DOI: `10.1159/000522108`.

[46] Gaalema, D. E. (2011) Visual discrimination and reversal learning in rough-necked monitor lizards (*Varanus rudicollis*). *Journal of Comparative Psychology* 125.2, 246. DOI: `10.1037/a0023148`.

[47] Cantwell, A. et al. (2022) The origins of cognitive flexibility in chimpanzees. *Developmental science* 25.5, e13266. DOI: `10.1111/desc.13266`.

[48] Lacreuse, A. et al. (2018) Glutamine/glutamate (Glx) concentration in prefrontal cortex predicts reversal learning performance in the marmoset. *Behavioural Brain Research* 346, 11–15. DOI: `10.1016/j.bbr.2018.01.025`.

[49] Rochais, C., Hotte, H., and Pillay, N. (2021) Seasonal variation in reversal learning reveals greater female cognitive flexibility in African striped mice. *Scientific Reports* 11.1, 20061. DOI: `10.1038/s41598-021-99619-9`.

[50] Boulougouris, V., Dalley, J. W., and Robbins, T. W. (2007) Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behavioural brain research* 179.2, 219–228. DOI: `10.1016/j.bbr.2007.02.005`.

[51] Aljadeff, N. and Lotem, A. (2021) Task-dependent reversal learning dynamics challenge the reversal paradigm of measuring cognitive flexibility. *Animal Behaviour* 179, 183–197. DOI: `10.1016/j.anbehav.2021.07.002`.

[52] Blaisdell, A. et al. (2021) Do the more flexible individuals rely more on causal cognition? Observation versus intervention in causal inference in great-tailed grackles. *Peer Community Journal* 1.e50. DOI: `10.24072/pcjournal.44`.

[53] McNamara, J. M. and Houston, A. I. (2009) Integrating function and mechanism. *Trends in Ecology & Evolution* 24.12, 670–675. DOI: `10.1016/j.tree.2009.05.011`.

[54] Fawcett, T. W., Hamblin, S., and Giraldeau, L.-A. (2013) Exposing the behavioral gambit: the evolution of learning and decision rules. *Behavioral Ecology* 24.1, 2–11. DOI: `10.1093/beheco/ars085`.

[55] Dukas, R. (1998) *Cognitive ecology: the evolutionary ecology of information processing and decision making*. University of Chicago Press.

[56] Barrett, B. J. (2022) Inferential power in identifying frequency-dependent social learning strengthened by increasing behavioural options. *Journal of Animal Ecology*. DOI: `10.1111/1365-2656.13826`.

[57] McElreath, R. et al. (2005) Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior* 26.6, 483–508. DOI: `10.1016/j.evolhumbehav.2005.04.003`.

[58] Chimento, M. et al. (2022) Cultural diffusion dynamics depend on behavioural production rules. *Proceedings of the Royal Society B: Biological Sciences* 289, 20221001. DOI: `10.1098/rspb.2022.1001`.

[59] Aplin, L. M., Sheldon, B. C., and McElreath, R. (2017) Conformity does not perpetuate suboptimal traditions in a wild population of songbirds. *Proceedings of the National Academy of Sciences* 114.30, 7830–7837. DOI: `10.1073/pnas.1621067114`.

[60] Barrett, B. J., McElreath, R. L., and Perry, S. E. (2017) Pay-off-biased social learning underlies the diffusion of novel extractive foraging traditions in a wild primate. *Proceedings of the Royal Society B: Biological Sciences* 284.1856, 20170358. DOI: `10.1098/rspb.2017.0358`.

[61] Trump, A. N. et al. (2023) A cognitive computational approach to social and collective decision-making, https://doi.org/10.31219/osf.io/7aykm.

[62] Johnson, K. et al. (2000) Male mating strategies and the mating system of great-tailed grackles. *Behavioral Ecology* 11.2, 132–141. DOI: `10.1093/beheco/11.2.132`.

[63] Smolla, M. et al. (2019) Reproductive skew affects social information use. *Royal Society open science* 6.7, 182084. DOI: `10.1098/rsos.182084`.

[64] Selander, R. K. and Giller, D. R. (1961) Analysis of sympatry of great-tailed and boat-tailed grackles. *The Condor* 63.1, 29–86. DOI: `10.2307/1365420`.

[65] Davis, W. R. and Arnold, K. A. (1972) Food habits of the Great-tailed Grackle in Brazos county, Texas. *The Condor* 74.4, 439–446. DOI: `10.2307/1365896`.

[66] Bro-Jørgensen, J., Franks, D. W., and Meise, K. (2019) Linking behaviour to dynamics of populations and communities: application of novel approaches in behavioural ecology to conservation. *Philosophical Transactions of the Royal Society B* 374.1781, 20190008. DOI: `10.1098/rstb.2019.0008`.

[67] Breen, A. J., Sugasawa, S., and Healy, S. D. (2021) Manipulative and technological skills do not require a slow life history. *Frontiers in Ecology and Evolution* 9, 46. DOI: `10.3389/fevo.2021.635802`.

[68] Breen, A. J. (2021) Animal culture research should include avian nest construction. *Biology Letters* 17.7, 20210327. DOI: `10.1098/rsbl.2021.0327`.

[69] Deffner, D., Rohrer, J. M., and McElreath, R. (2022) A causal framework for cross-cultural generalizability. *Advances in Methods and Practices in Psychological Science* 5.3, 1–18.

[70] Somveille, M. et al. (2018) Movement and conformity interact to establish local behavioural traditions in animal populations. *PLoS Computational Biology* 14.12, e1006647. DOI: `10.1371/journal.pcbi.1006647`.

[71] Logan, C. J. et al. (2022) Is behavioral flexibility manipulatable and, if so, does it improve flexibility and problem solving in a new context? https://knb.ecoinformatics.org/view/doi:10.5063/F1862DWC.

[72] Logan, C. J. (2016) Great-tailed grackle behavioral flexibility and problem solving experiments, Santa Barbara, CA USA 2014-2015 https://knb.ecoinformatics.org/view/doi:10.5063/F1319SVV.

[73] Wehtje, W. (2004) The great-tailed grackle *(Quiscalus mexicanus* Gmelin) in the Western USA: Range Expansion and Secondary Contact Between Subspecies.

[74] Lehman, P. E. (2020) *The birds of Santa Barbara County, California*. DOI: `http://www.sbcobirding.com/lehmanbosbc.html`.

[75] Hampton, S. (2004) *Yolo County birding news*.

[76] Team, S. D. (2020) *RStan: the R interface to Stan*. Version 2.21.2.

[77] Wickham, H. et al. (2019) Welcome to the Tidyverse. *Journal of Open Source Software* 4.43, 1686. DOI: `10.21105/joss.01686`.

[78] Team, R. C. (2021) *R: A Language and environment for statistical computing*. Version 1.4.1106.

[79] Sutton, R. S. and Barto, A. G. (2018) *Reinforcement learning: An introduction*. MIT press.

[80] Stephens, D. W. and Krebs, J. R. (2019) Risk-Sensitive Foraging. In: *Foraging Theory*. Princeton University Press.

[81] Webster, M. M. and Rutz, C. (2020) How STRANGE are your study animals? *Nature* 582.7812, 337–340. DOI: `10.1038/d41586-020-01751-5`.

[82] Otto, S. P. and Day, T. (2011) *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press.

[83] Yu, X. and Gen, M. (2010) *Introduction to evolutionary algorithms*. Springer Science & Business Media.

Acknowledgements: We thank Jean-François Gerard and Rachel Harrison for useful feedback on our study before data analyses; and James St Clair, Sue Healy, and Richard McElreath for similarly useful presubmission feedback. We are further and most grateful to Richard McElreath for study and overall full support. And we thank all members, past and present, of the Grackle Project for collecting and making available, either via public archive (core and middle population) or permissed (see Data provenance) advanced access (edge population), the data analysed herein. We note this material uses illustrations from Vecteezy.com. Finally, we further note this material uses data from the eBird Status and Trends Project at the Cornell Lab of Ornithology, eBird.org. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Cornell Lab of Ornithology. **Funding:** A.J.B. and D.D. received no independent funding for this research. **Author contributions:** A.J.B. conceived of the study; collated, cleaned, and curated all target data; D.D. led model and simulation building, with input from A.J.B.; A.J.B and D.D. contributed equally to mechanistic modelling and agent-based forward simulations; D.D. performed the Poisson and the evolutionary modelling; A.J.B. prepared all figures and tables, with help on Fig. 3 from D.D.; A.J.B. and D.D. annotated together all open source material; A.J.B. wrote the manuscript, with constructive contributions and revisions by D.D.; A.J.B. and D.D. agree over and approve the final draft of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Copies of the original data sheets, and all code to reproduce our manuscript, from data wrangling to analyses and figures, is available at our GitHub repository: https://github.com/alexisbreen/Sex-differences-in-grackles-learning.

Supplementary Materials for

Leading an urban invasion: risk-sensitive learning is a winning strategy

Alexis J. Breen[1] and Dominik Deffner[2,3]

[1]Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany

[2]Science of Intelligence Excellence Cluster, Technical University Berlin, Berlin 10623, Germany

[3]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin 14195, Germany

Correspondence: alexis_breen@eva.mpg.de; deffner@mpib-berlin.mpg.de

This PDF file includes:

Table S1 to S4
Figs S1
Study preregistration

Table S1. Reinforcement learning speed. Between- and across-population total-trials-in-test Poisson regression model estimates and male-female contrasts, with corresponding lower (L) and upper (U) 89% highest-posterior density intervals in parentheses.

Initial learning			
Population	Male	Female	Contrast
Core	36.43 (L: 31.60; U: 41.31)	37.38 (L: 28.74; U: 45.92)	-0.95 (L: -11.09; U: 8.99)
Middle	27.00 (L: 19.62; U: 33.61)	27.37 (L: 19.83; U: 34.22)	-0.37 (L: -10.78; U: 9.79)
Edge	33.85 (L: 28.21; U: 39.15)	38.65 (L: 31.21; U: 45.59)	-4.80 (L: -13.84; U: 4.28)
Across	33.77 (L: 30.41; U: 37.07)	35.08 (L: 30.14; U: 39.69)	-1.31 (L: -6.78; U: 4.79)
Reversal learning			
Population	Male	Female	Contrast
Core	61.60 (L: 53.60; U: 69.57)	80.83 (L: 62.32; U: 97.47)	-19.22 (L: -38.46; U: 0.65)
Middle	78.78 (L: 58.56; U: 98.84)	91.35 (L: 69.04; U: 111.88)	-12.57 (L: -41.64; U: 17.50)
Edge	66.30 (L: 56.63; U: 76.64)	71.80 (L: 58.87; U: 84.46)	-5.51 (L: -21.95; U: 10.95)
Across	60.00 (L: 54.25; U: 65.53)	78.18 (L: 68.26; U: 88.16)	-18.21 (L: -29.51; U: -6.99)

Table S2. Reinforcement learning switches Between- and across-population total-choice-option-switches-in-test Poisson regression model estimates and male-female contrasts, with corresponding lower (L) and upper (U) 89% highest-posterior density intervals in parentheses.

Initial learning

Population	Male	Female	Contrast
Core	13.64 (L: 10.69; U: 16.34)	16.04 (L: 10.17; U: 21.59)	-2.40 (L: -8.54; U: 4.38)
Middle	6.52 (L: 3.35; U: 9.15)	6.25 (L: 3.39; U: 9.15)	0.27 (L: -4.18; U: 4.33)
Edge	12.81 (L: 9.54; U: 16.08)	14.70 (L: 10.26; U: 18.72)	-1.89 (L: -7.20; U: 3.74)
Across	13.23 (L: 11.07; U: 15.18)	14.40 (L: 11.48; U: 17.31)	-1.17 (L: -4.72; U: 2.46)

Reversal learning

Population	Male	Female	Contrast
Core	22.38 (L: 17.88; U: 26.86)	37.02 (L: 23.96; U: 48.74)	-14.64 (L: -27.59; U: -1.17))
Middle	30.32 (L: 17.84; U: 42.10)	43.96 (L: 27.82; U: 59.76)	-13.65 (L: -34.30; U: 7.69)
Edge	21.71 (L: 16.58; U: 27.22)	27.75 (L: 20.10; U: 35.35)	-6.04 (L: -15.54; U: 3.39)
Across	22.18 (L: 18.72; U: 25.27)	36.74 (L: 29.78; U: 43.60)	-14.56 (L: -22.48; U: -7.15)

Table S3. Reinforcement learning information-updating rate ϕ. Between- and across-population computational model ϕ estimates and male-female contrasts, with posterior means and corresponding lower (L) and upper (U) 89% highest-posterior density intervals in parentheses.

Initial learning			
Population	Male	Female	Contrast
Core	0.03 (L: 0.01; U: 0.05)	0.07 (L: 0.02; U: 0.12)	-0.04 (L: -0.10; U: 0.02)
Middle	0.12 (L: 0.03; U: 0.20)	0.10 (L: 0.03; U: 0.17)	0.02 (L: -0.13; U: 0.15)
Edge	0.07 (L: 0.02; U: 0.11)	0.09 (L: 0.03; U: 0.14)	-0.02 (L: -0.10; U: 0.07)
Across	0.03 (L: 0.01; U: 0.04)	0.05 (L: 0.02; U: 0.08)	-0.02 (L: -0.06; U: 0.01)
Reversal learning			
Population	Male	Female	Contrast
Core	0.03 (L: 0.02; U: 0.05)	0.04 (L: 0.02; U: 0.05)	0.00 (L: -0.02; U: 0.02))
Middle	0.04 (L: 0.02; U: 0.06)	0.04 (L: 0.02; U: 0.06)	0.00 (L: -0.03; U: 0.03)
Edge	0.04 (L: 0.03; U: 0.05)	0.04 (L: 0.03; U: 0.06)	0.00 (L: -0.03; U: 0.02)
Across	0.03 (L: 0.01; U: 0.04)	0.03 (L: 0.02; U: 0.07)	0.00 (L: -0.01; U: -0.01)

Table S4. Reinforcement learning risk-sensitivity rate λ. Between- and across-population computational model λ estimates and male-female contrasts, with posterior means and corresponding lower (L) and upper (U) 89% highest-posterior density intervals in parentheses.

Initial learning			
Population	Male	Female	Contrast
Core	4.61 (L: 2.24; U: 6.84)	2.65 (L: 0.93; U: 4.20)	1.96 (L: -1.06; U: 5.12)
Middle	3.58 (L: 1.33; U: 5.86)	3.53 (L: 1.51; U: 5.66)	0.00 (L: -3.43; U: 3.52)
Edge	2.81 (L: 1.44; U: 4.14)	2.03 (L: 1.06; U: 2.89)	0.79 (L: -1.19; U: 2.34)
Across	5.65 (L: 3.15; U: 8.22)	3.30 (L: 1.68; U: 4.73)	2.35 (L: -0.77; U: 5.61)
Reversal learning			
Population	Male	Female	Contrast
Core	4.76 (L: 2.86; U: 6.33)	2.20 (L: 1.10; U: 3.28)	2.56 (L: 0.59; U: 4.82))
Middle	2.99 (L: 1.35; U: 4.40)	1.74 (L: 0.80; U: 2.57)	1.25 (L: -0.64; U: 3.12)
Edge	4.13 (L: 2.62; U: 5.75)	3.62 (L: 2.12; U: 4.94)	0.51 (L: -1.68; U: 2.67)
Across	5.86 (L: 3.30; U: 8.37)	3.50 (L: 1.98; U: 4.91)	2.36 (L: 0.14; U: 4.26)

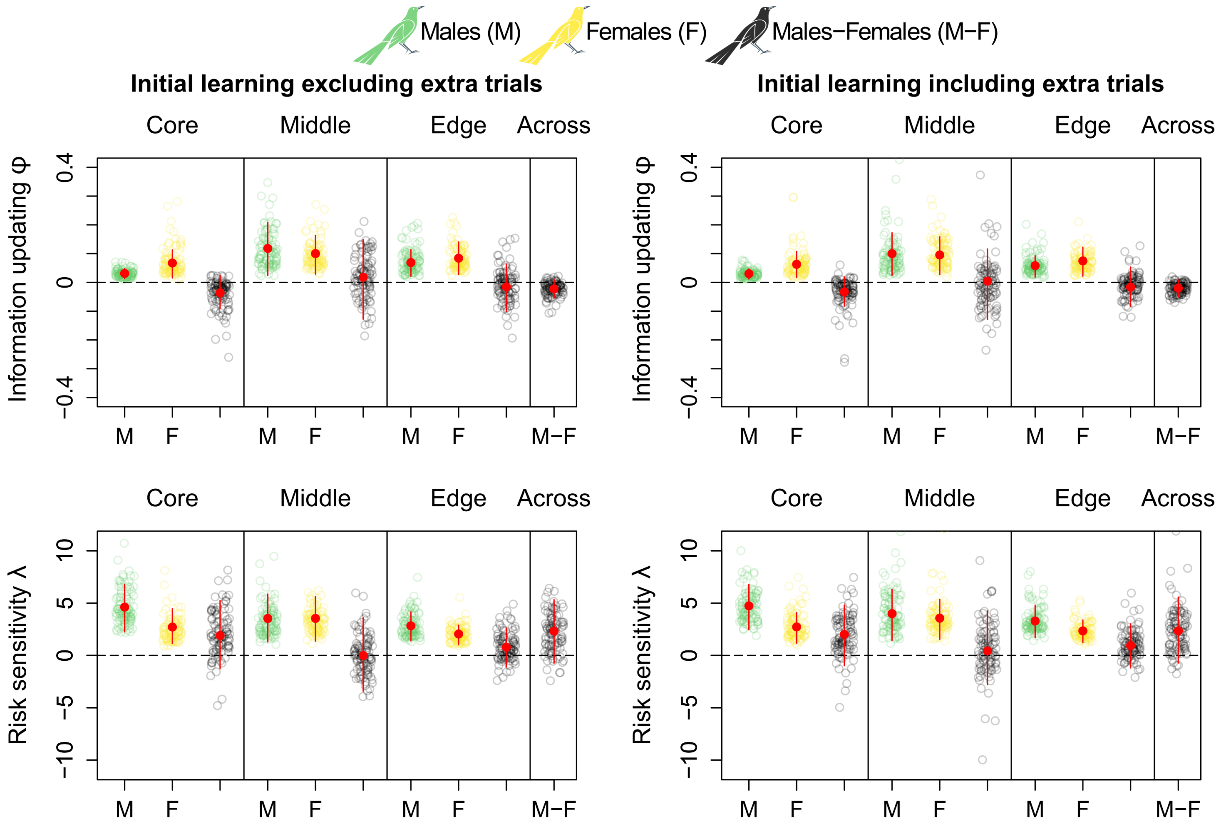


Fig. S1. Initial reinforcement learning and extra learning trials. Comparison of information-updating rate ϕ and risk-sensitivity rate λ estimates (top and bottom row, respectively) in initial learning excluding and including extra initial learning trials (left and right column, respectively), which are present in the original data set (see Methods). Because this comparison does not show any noticeable difference depending on their inclusion or exclusion, we excluded extra learning trials from our analyses. All plots are generated via model estimates using our full sample size: 32 males and 17 females.

This preregistration of our study—intended to serve as a record of our scientific method, which necessarily evolved beyond this original outline—was uploaded on 25 January 2022 to our GitHub repository:
https://github.com/alexisbreen/Sex-differences-in-grackles-learning

Investigating sex differences in learning in a range-expanding bird

Alexis J. Breen[1,*] & Dominik Deffner[1,2,3]

[1]Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany

[2]Science of Intelligence Excellence Cluster, Technical University Berlin, Berlin 10623, Germany

[3]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin 14195, Germany

*alexis_breen@eva.mpg.de

Abstract

How might differences in dispersal and learning interact in range expansion dynamics? To begin to answer this question, in this preregistration we detail the background, hypothesis plus associated predictions, and methods of our proposed study, including the development and validation of a mechanistic reinforcement learning model, which we aim to use to assay colour-reward reinforcement learning (and the influence of two candidate latent parameters—speed and sampling rate—on this learning) in great-tailed grackles—a species undergoing rapid range expansion, where males disperse.

Introduction

Dispersal and range expansion go 'hand in hand'; movement by individuals away from a population's core is a pivotal precondition of witnessed growth in species' geographic limits (Chuang & Peterson, 2016; Ronce, 2007). Because 'who' disperses—in terms of sex—varies both within and across taxa (for example, male-biased dispersal is dominant among fish and mammals, whereas female-biased dispersal is dominant among birds; see Table 1 in Trochet et al., 2016), skewed sex ratios are apt to arise at expanding range fronts, and, in turn, differentially drive invasion dynamics. Female-biased dispersal, for instance, can 'speed up' staged invertebrate invasions by increasing offspring production (Miller & Inouye, 2013). Alongside sex-biased dispersal, learning ability is also argued to contribute to species' colonisation capacity, as novel environments inevitably present novel (foraging, predation, shelter, and social) challenges that newcomers need to surmount in order to settle successfully (Sol et al., 2013; Wright et al., 2010). Indeed, a growing number of studies show support for this supposition (as recently reviewed in Lee & Thornton, 2021). Carefully controlled choice tests, for example, show that urban-dwelling individuals—that is, the 'invaders'—will both learn and unlearn novel reward-stimulus pairings more rapidly than their rural-dwelling counterparts (Batabyal & Thaker, 2019), suggesting that range expansion selects for enhanced learning ability at the dispersal and/or settlement stage(s). Given the independent influence of sex-biased dispersal and learning ability on range expansion, it is perhaps surprising, then, that their potential interactive influence on this aspect of movement ecology remains unexamined, particularly as interactive links between dispersal and other behavioural traits such as aggression are documented within the range expansion literature (Duckworth, 2006; Gutowsky & Fox, 2011).

That learning ability can covary with, for example, exploration (e.g., Auersperg et al., 2011; Guillette et al., 2011) and neophobia (e.g., Verbeek et al., 1994), two behaviours which may likewise play a role in range expansion (Griffin et al., 2017; Lee & Thornton, 2021), is one potential reason for the knowledge gap introduced above. Such correlations stand to mask what contribution, if any, learning ability lends to range expansion—an undoubtedly daunting research prospect. A second (and not mutually exclusive) reason is that, for many species, a detailed diary of their range expansion is lacking (Blackburn et al., 2009; Udvardy & Papp, 1969). And patchy population records inevitably introduce interpretive 'noise,' imaginably impeding population comparisons of learning ability (or the like).

In range-expanding great-tailed grackles (*Quiscalus mexicanus*), however, learning ability appears to represent a unique source of individual variation; more specifically, temporarily-captive great-tailed grackles' speed to solve colour-reward reinforcement learning tests does not correlate with measures of their exploration (time spent moving within a novel environment), inhibition (time to reverse a colour-reward preference), motor diversity (number of distinct bill and/or feet movements used in behavioural tests), neophobia (latency to approach a novel object), risk aversion (time spent stationary within a 'safe spot' in a novel environment), persistence (number of attempts to engage in behavioural tests), or problem solving (number of test-relevant functional and non-functional object-choices) (Logan, 2016a, 2016b). Moreover, careful combing by researchers of public records, such as regional bird reports and museum collections, means that great-tailed grackle range-expansion data is both comprehensive and readily available (Dinsmore & Dinsmore, 1993; Pandolfino et al., 2009; Wehtje, 2003). Thus, great-tailed grackles offer behavioural ecologists a useful study system to investigate the interplay between life-history strategies, learning ability, and range expansion.

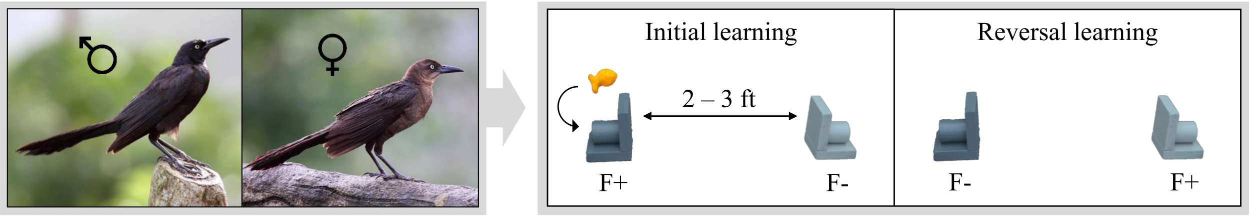


Figure 1 Left panel: images showing a male and female great-tailed grackle (credit: Wikimedia Commons). Right panel: schematic of the colour-reward reinforcement learning experimental protocol. In the *initial learning* phase, great-tailed grackles are presented with two colour-distinct tubes; however, only one coloured tube (e.g., dark grey) contains a food reward (F+ versus F-). In the *reversal learning* phase, the colour-reward tube-pairings are swapped. The passing criterion was identical in both phases (see main text for details).

Here, for the first time (to our knowledge), we propose to investigate potential differences in colour-reward reinforcement learning performance between male and female great-tailed grackles (Figure 1), to test the hypothesis that sex differences in learning ability are related to sex differences in dispersal. Since the late nineteenth century, great-tailed grackles have been expanding their range at an unprecedented rate, moving northward from their native range in Central America into the United States (breeding in at least 20 states), with several first-sightings spanning as far north as Canada (Dinsmore & Dinsmore, 1993; Wehtje, 2003). Notably, the record of this range expansion in great-tailed grackles is heavily peppered with first-sightings involving a single or multiple male(s) (Dinsmore & Dinsmore, 1993; Kingery, 1972; Littlefield, 1983; Stepney, 1975; Wehtje, 2003). Moreover, recent genetic data show that, when comparing great-tailed grackles within a population, average relatedness: (*i*) is higher among females than among males; and (*ii*) decreases with increasing geographic distance among females; but (*iii*) is unrelated to geographic distance among males; hence, confirming a role for male-biased dispersal in great-tailed grackles (Sevchik et al., in press). Considering these natural history and genetic data, then, we expect male and female great-tailed grackles to differ across at least two colour-reward reinforcement learning parameters: speed and sampling rate (here, sampling is defined as switching between choice-options). Specifically, we expect male—versus female—great-tailed grackles: (prediction 1 & 2) to be faster to, firstly, learn a novel colour-reward pairing, and secondly, reverse their colour preference when the colour-reward pairing is swapped; and (prediction 3) to be more deterministic—that is, sample less often—in their colour-reward learning; if learning ability and dispersal relate. Indeed, since invading great-tailed grackles face agribusiness-led wildlife management strategies, including the use of chemical crop repellents (Werner et al., 2011, 2015), range expansion should disfavour slow, error-prone learning strategies, resulting in a spatial sorting of learning ability in great-tailed grackles (Wright et al., 2010). Related to this final point, we further expect (prediction 4) such sex differences in learning ability to be more pronounced in great-tailed grackles living at the edge, rather than the intermediate and/or core, region of their range (e.g., Duckworth, 2006).

Methods

Data

This preregistration aims to use colour-reward reinforcement learning data collected (or being collected) in great-tailed grackles across three study sites that differ in their range-expansion demographics; that is, belonging to a core, intermediate, or edge population (based on time-since-settlement population growth dynamics, as outlined in Chuang & Peterson, 2016). Specifically, data will be utilised from: (*i*) Tempe, Arizona—hereafter, the core population (estimated—by adding the average time between first sighting and first breeding to the year first sighted—to be breeding since 1951) (Walter, 2004; Wehtje, 2003); (*ii*) Santa Barbara, California—hereafter, the intermediate population (known to be breeding since 1996) (Lehman, 2020); and (*iii*) Woodland, California—hereafter, the edge population (known to be breeding since 2004) (Hampton, 2001). Data collection at both the Tempe, Arizona and Santa Barbara, California study sites has been completed prior to the submission of this preregistration (total sample size across sites: nine females and 25 males); however, data collection at the Woodland, California study site is ongoing (current sample size: three females and nine males; anticipated minimum total sample size: five females and ten males). Thus, the final data set should contain colour-reward reinforcement learning data from at least 14 female and 35 male great-tailed grackles.

Experimental protocol

General

A step-by-step description of the experimental protocol is reported elsewhere (e.g., Blaisdell et al., 2021). As such, below we detail only the protocol for the colour-reward reinforcement learning tests that we propose to analyse herein.

Colour-reward reinforcement learning tests

The reinforcement learning tests consist of two phases (Figure 1, right panel): (*i*) colour-reward learning (hereafter, initial learning) and (*ii*) colour-reward reversal learning (hereafter, reversal learning). In both phases, two different coloured tubes are used: for Santa Barbara great-tailed grackles, gold and grey (Logan, 2016b, 2016a); for all other great-tailed grackles: light and dark grey (Blaisdell et al., 2021). Each tube consists of an outer and inner diameter of 26 mm and 19 mm, respectively; and each is mounted to two pieces of plywood attached at a right angle (entire apparatus: 50 mm wide × 50 mm tall × 67 mm deep); thus resulting in only one end of each coloured tube being accessible (Figure 1, right panel).

In the *initial learning phase*, great-tailed grackles are required to learn that only one of the two coloured tubes contains a food reward (e.g., dark grey; this colour-reward pairing is counterbalanced across great-tailed grackles within each study site). Specifically, the rewarded and unrewarded coloured tubes are placed—either on a table or on the floor—in the centre of the aviary run (distance apart: table, 2 ft; floor, 3 ft), with the open tube-ends facing, and perpendicular to, their respective aviary side-wall. Which coloured tube is placed on which side of the aviary run (left or right) is pseudorandomised across trials. A trial begins at tube-placement, and ends when a great-tailed grackle has either made a tube-choice or the maximum trial time has elapsed (eight minutes). A tube-choice is defined as a great-tailed grackle bending down to examine the contents (or lack thereof) of a tube. If the chosen tube contains food, the great-tailed grackle is allowed to retrieve and eat the food, before both tubes are removed and the rewarded coloured tube is rebaited out of sight (for the great-tailed grackle). If a chosen tube does not contain food, both tubes are immediately removed. Each great-tailed grackle is given, first, up to three minutes to make a tube-choice (after which a piece of food is placed equidistant between the tubes to entice participation); and then, if no choice has been made, an additional five minutes maximum, before both tubes are removed. All trials are recorded as either correct (choosing the rewarded colour tube), incorrect (choosing the unrewarded colour tube), or incomplete (no choice made); and are presented in 10-trial blocks. To pass initial learning, a great-tailed grackle must make a correct choice in at least 17 out of the most recent 20 trials, with a minimum of eight and nine correct choices across the last two blocks.

In the *reversal learning phase*, great-tailed grackles are required to learn that the colour-reward pairing has been switched; that is, the previously unrewarded coloured tube (e.g., light grey) now contains a food reward. The protocol for this second and final learning phase is identical to that, described above, of the initial learning phase.

Analysis plan

General

Here, we will analyse, process, and visually present our data using, respectively, the 'rstan' (Stan Development Team, 2020), 'rethinking' (McElreath, 2018), and 'tidyverse' (Wickham et al., 2019) packages in R (R Core Team, 2021). Our reproducible code is available on GitHub (https://github.com/alexisbreen/Sex-differences-in-grackles-learning).

Reinforcement learning model

In this preregistration, we propose to employ an adapted (from Deffner et al., 2020) Bayesian reinforcement learning model, to examine the influence of sex on great-tailed grackles' initial and reversal learning performance. The reinforcement learning model, defined below, allows us to link observed coloured tube-choices to latent individual-level knowledge-updating (of attractions towards, learning about, and sampling of, either coloured tube) based on recent tube-choice reward-payoffs, and to translate such latent knowledge-updating into individual tube-choice probabilities; in other words, we can reverse engineer the probability that our parameters of interest (speed and sampling rate) produce great-tailed grackles' observed tube-choice behaviour by formulating our scientific model as a statistical model (McElreath, 2018, p. 537). This method can therefore capture whether, and, if so, how multiple latent learning strategies simultaneously guide great-tailed grackles' decision making—an analytical advantage over more traditional methods (e.g., comparing trials to passing criterion) that ignore the potential for equifinality (Barrett, 2019; Kandler & Powell, 2018).

Our reinforcement learning model consists of two equations:

$$A_{i,j,t+1} = (1 - \phi_{k,l})A_{i,j,t} + \phi_{k,l}\pi_{i,j,t}, \tag{1}$$

$$P(i)_{t+1} = \frac{\exp(\lambda_{k,l}A_{i,j,t})}{\sum\limits_{m=1}^{2}\exp(\lambda_{k,l}A_{m,j,t})}. \tag{2}$$

Equation 1 expresses how attraction (A) to a choice-option (i) changes for an individual (j) across time ($t+1$) based on their prior attraction to that choice-option ($A_{i,j,t}$) plus their recently experienced choice-payoff ($\pi_{i,j,t}$), whilst accounting for the weight given to recent payoffs ($\phi_{k,l}$). As $\phi_{k,l}$ increases in value, so, too, does the rate of individual attraction-updating; thus, $\phi_{k,l}$ represents the individual learning rate. We highlight that the k,l indexing denotes that we estimate separate ϕ parameters for each phase of the experiment ($k = 1$ for initial, $k = 2$ for reversal) and each sex ($l = 1$ for females, $l = 2$ for males).

Equation 2 is a softmax function that expresses the probability (P) that option (i) is selected in the next choice-round ($t+1$) as a function of the attractions and a parameter ($\lambda_{k,l}$) that governs how much relative differences in attraction scores guide individual choice-behaviour. The higher the value of $\lambda_{k,l}$, the more deterministic (less option-switching) the choice-behaviour of an individual becomes (note $\lambda_{k,l} = 0$ generates random choice); thus, $\lambda_{k,l}$ represents the individual sampling rate for phase k and sex l.

From the above reinforcement learning model, then, we will generate inferences about the effect of sex on $\phi_{k,l}$ and $\lambda_{k,l}$ from at least 1000 effective samples of the posterior distribution (see our model validation below). We note that our reinforcement learning model also includes both individual bird and study site as random effects (to account for repeated measures within both individuals and populations); however, for clarity, these parameters are omitted from our equations (but not our code: https://github.com/alexisbreen/Sex-differences-in-grackles-learning). Regarding our study site random effect, we further note that, as introduced above, we will also explore population-mediated sex-effects on ϕ and λ, by comparing these learning parameters both within and between sexes at each study site. Finally, our reinforcement learning model excludes trials where a great-tailed grackle did not make a tube-choice, as this measure cannot clearly speak to individual learning ability—for example, satiation rather than any learning of 'appropriate' colour tube-choice could be invoked as an explanation in such cases. Indeed, there are, admittedly, a number of intrinsic and extrinsic factors (e.g., temperament and temperature, respectively) that might bias great-tailed grackles' tube-choice behaviour, and, in turn, the output from our reinforcement learning model (Webster & Rutz, 2020). Nonetheless, our reinforcement learning model serves as a useful first step towards addressing if learning ability and dispersal relate in great-tailed grackles (for a similiar rationale, see McElreath & Smaldino, 2015).

Model validation

We validated our reinforcement learning model in three steps. First, we performed agent-based simulations. Specifically, we followed the tube-choice behaviour of simulated great-tailed grackles—that is, 14 females and 35 males from one of three populations (where population membership matched known study site sex distributions)—across the described initial learning and reversal learning phases. The tube-choice behaviour of the simulated great-tailed grackles was governed by a set of rules identical to those defined by our mathematical equations—for example, coloured tube attractions were independently updated based on the reward outcome of tube choices. Because we assigned higher average ϕ and λ values to simulated male (versus female) great-tailed grackles, the resulting data set should show males outperform females on initial and reversal learning, at both the group and individual-level; it did (Figure 2 & S1, respectively).

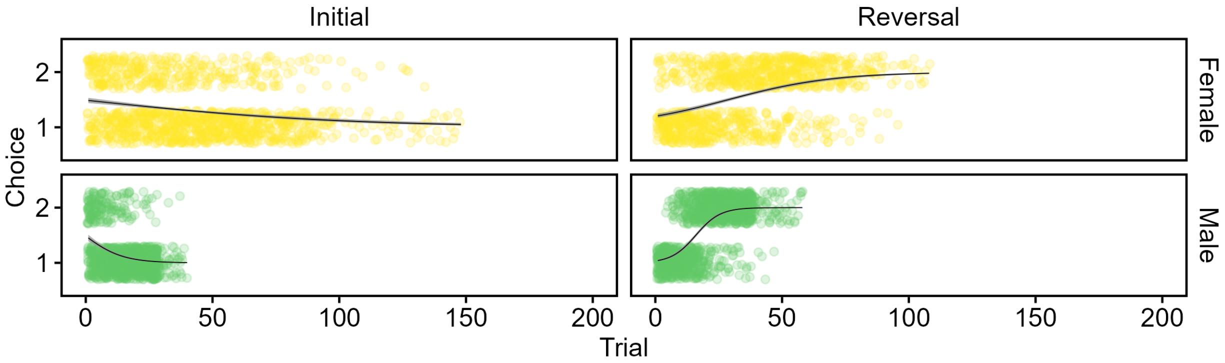


Figure 2 Group-level tube-choice behaviour of simulated great-tailed grackles across colour-reward reinforcement learning trials (females: yellow, $n = 14$; males: green, $n = 35$), following model validation step one. Tube option 1 (e.g., dark grey) was the rewarded option in the initial learning phase; conversely, tube option 2 (e.g., light grey) contained the food reward in the reversal learning phase. Each open circle represents an individual tube-choice; black lines indicate binomial smoothed conditional means fitted with grey 89% compatability intervals.

Next, we ran our simulated data set on our reinforcement learning model. Here, we endeavored to determine whether our reinforcement learning model: (*i*) recovered our assigned $\phi_{k,l}$ and $\lambda_{k,l}$ values (it did; Table 1); and (*ii*) produced 'correct' qualitative inferences—that is, detected the simulated sex differences in great-tailed grackles' initial and reversal learning (it did; Figure 3).

Table 1: Comparison of assigned and recovered ϕ and λ values, following model validation step two. Eighty-nine percent highest posterior density intervals (HPDI) are shown for recovered values.

	ϕ				λ			
	Female		Male		Female		Male	
	Initial	Reversal	Initial	Reversal	Initial	Reversal	Initial	Reversal
Assigned	0.03	0.05	0.09	0.11	2.00	3.00	4.00	5.00
Recovered	0.03	0.05	0.07	0.10	2.16	2.82	4.31	5.68
89% HPDI	0.01 - 0.04	0.04 - 0.06	0.03 - 0.11	0.08 - 0.12	1.29 - 2.99	2.05 - 3.58	2.65 - 6.00	4.41 - 6.97

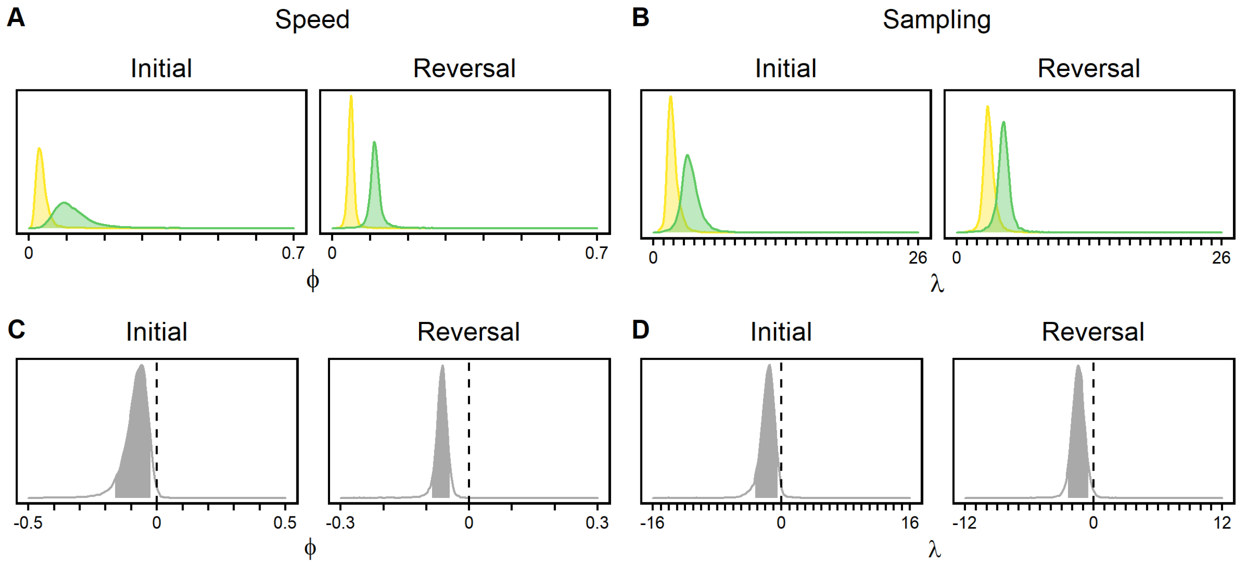


Figure 3 Comparison of learning ability in simulated female (yellow; $n = 14$) and male (green; $n = 35$) great-tailed grackles across initial and reversal colour-reward reinforcement learning, following model validation step two. (**A**) ϕ, the rate of learning i.e., speed. (**B**) λ, the rate of sampling i.e., switching between choice-options. (**C**) and (**D**) show posterior distributions for respective contrasts between female and male learning. Eighty-nine percent highest posterior density intervals are shaded in grey; that this interval does not cross zero evidences a simulated effect of sex on learning ability.

Finally, we repeated step one and step two, using a range of realistically plausible ϕ and λ sex differences (note that values for female great-tailed grackles were left unchanged from Table 1), to determine whether our reinforcement learning model could detect different effect sizes of sex on our target learning parameters. This final step confirmed that, for our anticipated minimum sample size, our reinforcement learning model: (*i*) detects sex differences in ϕ values $>=$ 0.03 and λ values $>=$ 1; and (*ii*) infers a null effect for ϕ values $<$ 0.03 and λ values $<$ 1 i.e., very weak simulated sex differences (Figure 4). Both of these points together highlight how our reinforcement learning model allows us to say that null results are not just due to small sample size. Additionally, estimates obtained from step three were more precise in the reversal learning phase compared to the initial learning phase (Figure 4), and we can expect to detect even smaller sex differences if we analyse learning across both phases—an approach we will apply if we detect no effect of phase. In sum, model validation steps one through three confirm that our reinforcement learning model is reasonably fit.

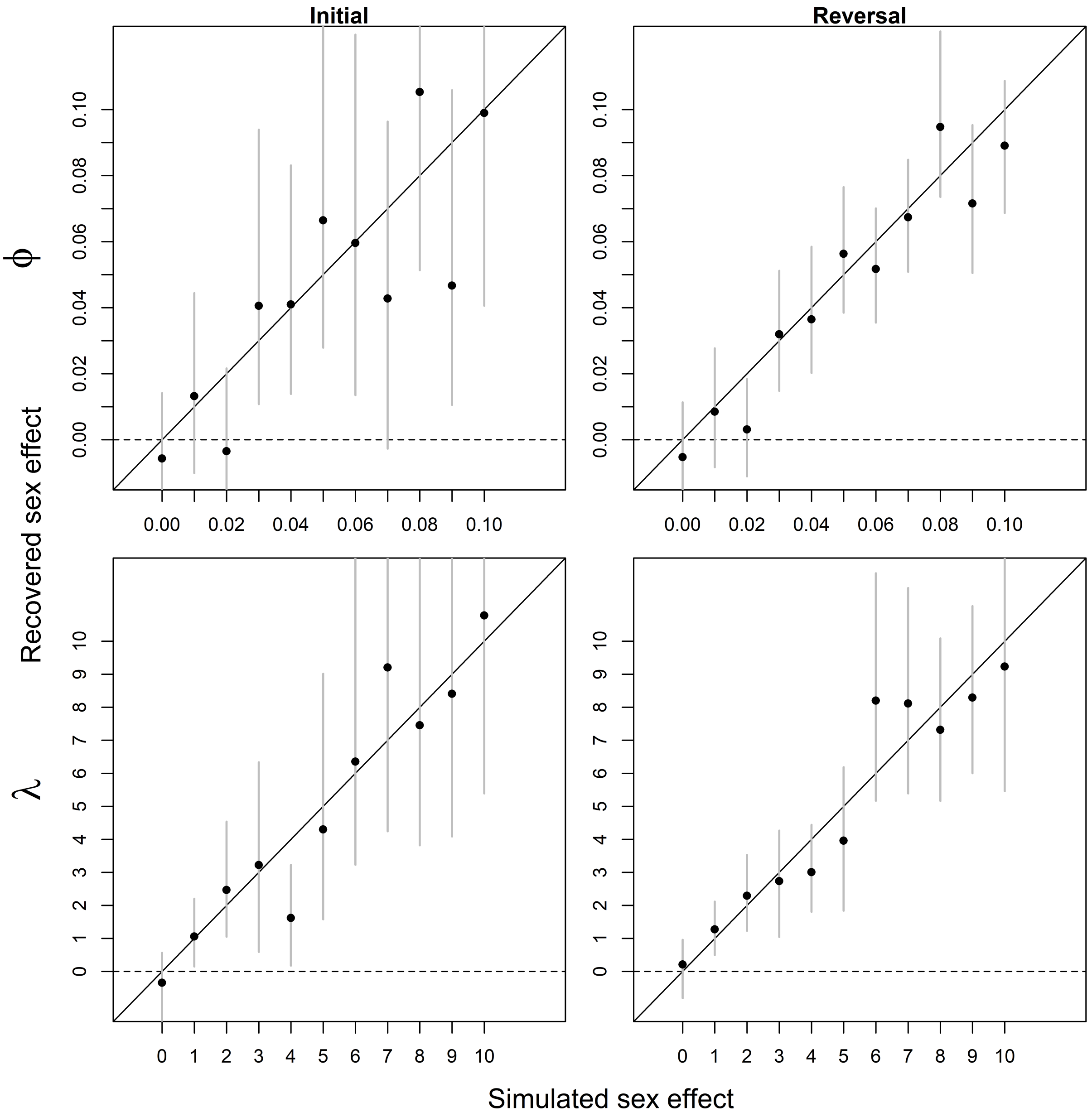


Figure 4 Parameter recovery test for different sizes of simulated sex differences. Plots show posterior estimates of the effect of sex (contrasts between simulated male and female great-tailed grackles; n = 14 and 35, respectively) on speed (ϕ) and sampling (λ) learning parameters, following model validation step three. Black circles represent the mean recovered sex effect estimates with grey eighty-nine percent highest posterior density intervals (HPDIs); black solid diagonal lines represent a 'perfect' match between assigned and recovered parameter estimates (note that we would not expect a perfect correspondence due to stochasticity of agent-based simulations); and black dashed horizontal lines represent a recovered null sex effect.

Bias

AJB and DD are (at the time of submitting this preregistration) blind with respect to all but two aspects of the target data: the sex and population membership of each grackle that has, thus far, completed, or is expected to complete, the colour-reward reinforcement learning tests (because these parameters were used in model validation simulations—see above).

Open materials

https://github.com/alexisbreen/Sex-differences-in-grackles-learning

Acknowledgements

We thank all members, past and present, of the Grackle Project for collecting and sharing the data that we propose to analyse herein. We further thank Richard McElreath for study support.

Ethics

All data utilised herein were collected with ethical approval.

References

Auersperg, A. M. I., Von Bayern, A. M. P., Gajdon, G. K., Huber, L., & Kacelnik, A. (2011). Flexibility in problem solving and tool use of kea and New Caledonian crows in a multi access box paradigm. *PloS ONE*, *6*(6), e20231. https://doi.org/10.1371/journal.pone.0020231

Barrett, B. J. (2019). Equifinality in empirical studies of cultural transmission. *Behavioural Processes*, *161*, 129–138. https://doi.org/10.1016/j.beproc.2018.01.011

Batabyal, A., & Thaker, M. (2019). Lizards from suburban areas learn faster to stay safe. *Biology Letters*, *15*(2), 20190009. https://doi.org/10.1098/rsbl.2019.0009

Blackburn, T. M., Lockwood, J. L., & Cassey, P. (2009). *Avian invasions: The ecology and evolution of exotic birds* (Vol. 1). Oxford University Press.

Blaisdell, A., Seitz, B., Rowney, C., Folsom, M., MacPherson, M., Deffner, D., & Logan, C. J. (2021). Do the more flexible individuals rely more on causal cognition? Observation versus intervention in causal inference in great-tailed grackles. *Peer Community Journal*, *1*(e50). https://doi.org/10.24072/pcjournal.44

Chuang, A., & Peterson, C. R. (2016). Expanding population edges: Theories, traits, and trade-offs. *Global Change Biology*, *22*(2), 494–512. https://doi.org/10.1111/gcb.13107

Deffner, D., Kleinow, V., & McElreath, R. (2020). Dynamic social learning in temporally and spatially variable environments. *Royal Society Open Science*, *7*(12), 200734. https://doi.org/10.1098/rsos.200734

Dinsmore, J. J., & Dinsmore, S. J. (1993). Range expansion of the great-tailed grackle in the 1900s. *Journal of the Iowa Academy of Science*, *100*(2), 54–59.

Duckworth, R. A. (2006). Behavioral correlations across breeding contexts provide a mechanism for a cost of aggression. *Behavioral Ecology*, *17*(6), 1011–1019. https://doi.org/10.1093/beheco/arl035

Griffin, A. S., Netto, K., & Peneaux, C. (2017). Neophilia, innovation and learning in an urbanized world: A critical evaluation of mixed findings. *Current Opinion in Behavioral Sciences*, *16*, 15–22. https://doi.org/10.1016/j.cobeha.2017.01.004

Guillette, L. M., Reddon, A. R., Hoeschele, M., & Sturdy, C. B. (2011). Sometimes slower is better: Slow-exploring birds are more sensitive to changes in a vocal discrimination task. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1706), 767–773. https://doi.org/10.1098/rspb.2010.1669

Gutowsky, L. F. G., & Fox, M. G. (2011). Occupation, body size and sex ratio of round goby *Neogobius melanostomus* in established and newly invaded areas of an ontario river. *Hydrobiologia*, *671*(1), 27–37. https://doi.org/10.1007/s10750-011-0701-9

Hampton, S. (2001). *Yolo County birding news*.

Kandler, A., & Powell, A. (2018). Generative inference for cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1743), 20170056. https://doi.org/10.1098/rstb.2017.0056

Kingery, H. E. (1972). The nesting season: June 1-August 15, 1972. *American Birds*, *26*, 882–887.

Lee, V. E., & Thornton, A. (2021). Animal cognition in an urbanised world. *Frontiers in Ecology and Evolution*, *9*, 120. https://doi.org/10.3389/fevo.2021.633947

Lehman, P. E. (2020). *The birds of Santa Barbara County, California* (2nd ed.). https://doi.org/http://www.sbcobirding.com/lehmanbosbc.html

Littlefield, C. D. (1983). Oregon's first records of the great-tailed grackle. *Western Birds*, *14*, 201–202.

Logan, C. J. (2016a). Behavioral flexibility and problem solving in an invasive bird. *PeerJ*, *4*, e1975. https://doi.org/10.7717/peerj.1975

Logan, C. J. (2016b). Behavioral flexibility in an invasive bird is independent of other behaviors. *PeerJ*, *4*, e2215. https://doi.org/10.7717/peerj.2215

McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman; Hall/CRC.

McElreath, R., & Smaldino, P. E. (2015). Replication, communication, and the population dynamics of scientific discovery. *PloS One*, *10*(8), e0136088. https://doi.org/10.1371/journal.pone.0136088

Miller, T. E. X., & Inouye, B. D. (2013). Sex and stochasticity affect range expansion of experimental invasions. *Ecology Letters*, *16*(3), 354–361. https://doi.org/10.1111/ele.12049

Pandolfino, E. R., Deuel, B. E., & Young, L. (2009). Colonization of the california's central valley by the great-tailed grackle. *Central Valley Bird Club Bull*, *12*, 77–95.

R Core Team. (2021). *R: A language and environment for statistical computing* (Version 1.4.1106) [Computer software]. https://www.R-project.org/

Ronce, O. (2007). How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics*, *38*, 231–253. https://doi.org/10.1146/annurev.ecolsys.38.091206.095611

Sevchik, A., Logan, C. J., McCune, K. B., Blackwell, A., Rowney, C., & Lukas, D. (in press). Investigating sex differences in genetic relatedness in great-tailed grackles in Tempe, Arizona to infer potential sex biases in dispersal. In *Peer community in ecology*. https://doi.org/10.32942/osf.io/t6beh

Sol, D., Lapiedra, O., & González-Lagos, C. (2013). Behavioural adjustments for a life in the city. *Animal Behaviour*, *85*(5), 1101–1112. https://doi.org/10.1016/j.anbehav.2013.01.023

Stan Development Team. (2020). *RStan: The R interface to Stan* (Version 2.21.2) [Computer software]. http://mc-stan.%20org

Stepney, P. H. R. (1975). First recorded breeding of the great-tailed grackle in Colorado. *The Condor*, *77*(2), 208–210. https://doi.org/10.2307/1365794

Trochet, A., Courtois, E. A., Stevens, V. M., Baguette, M., Chaine, A., Schmeller, D. S., Clobert, J., & Wiens, J. J. (2016). Evolution of sex-biased dispersal. *The Quarterly Review of Biology*, *91*(3), 297–320. https://doi.org/10.1086/688097

Udvardy, M. D. F., & Papp, C. S. (1969). *Dynamic zoogeography*. Van Nostrand Reinhold.

Verbeek, M. E. M., Drent, P. J., & Wiepkema, P. R. (1994). Consistent individual differences in early exploratory behaviour of male great tits. *Animal Behaviour*, *48*(5), 1113–1121. https://doi.org/10.1006/anbe.1994.1344

Walter, W. (2004). *The great-tailed grackle* (Quiscalus mexicanus *Gmelin) in the Western USA: Range expansion and secondary contact between subspecies* [PhD thesis]. University of California Riverside.

Webster, M. M., & Rutz, C. (2020). How STRANGE are your study animals? In *Nature* (No. 7812; Vol. 582, pp. 337–340). Nature Publishing Group. https://doi.org/10.1038/d41586-020-01751-5

Wehtje, W. (2003). The range expansion of the great-tailed grackle *(Quiscalus mexicanus* Gmelin) in North America since 1880. *Journal of Biogeography*, *30*(10), 1593–1607. https://doi.org/10.1046/j.1365-2699.2003.00970.x

Werner, S. J., DeLiberto, S. T., Mangan, A. M., Pettit, S. E., Ellis, J. W., & Carlson, J. C. (2015). Anthraquinone-based repellent for horned larks, great-tailed grackles, American crows and the protection of California's specialty crops. *Crop Protection*, *72*, 158–162. https://doi.org/10.1016/j.cropro.2015.03.020

Werner, S. J., Linz, G. M., Carlson, J. C., Pettit, S. E., Tupper, S. K., & Santer, M. M. (2011). Anthraquinone-based bird repellent for sunflower crops. *Applied Animal Behaviour Science*, *129*(2-4), 162–169. https://doi.org/10.1016/j.applanim.2010.11.010

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, l, Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wright, T. F., Eberhard, J. R., Hobson, E. A., Avery, M. L., & Russello, M. A. (2010). Behavioral flexibility and species invasions: The adaptive flexibility hypothesis. *Ethology Ecology & Evolution*, *22*(4), 393–404. https://doi.org/10.1080/03949370.2010.505580

Supplementary material

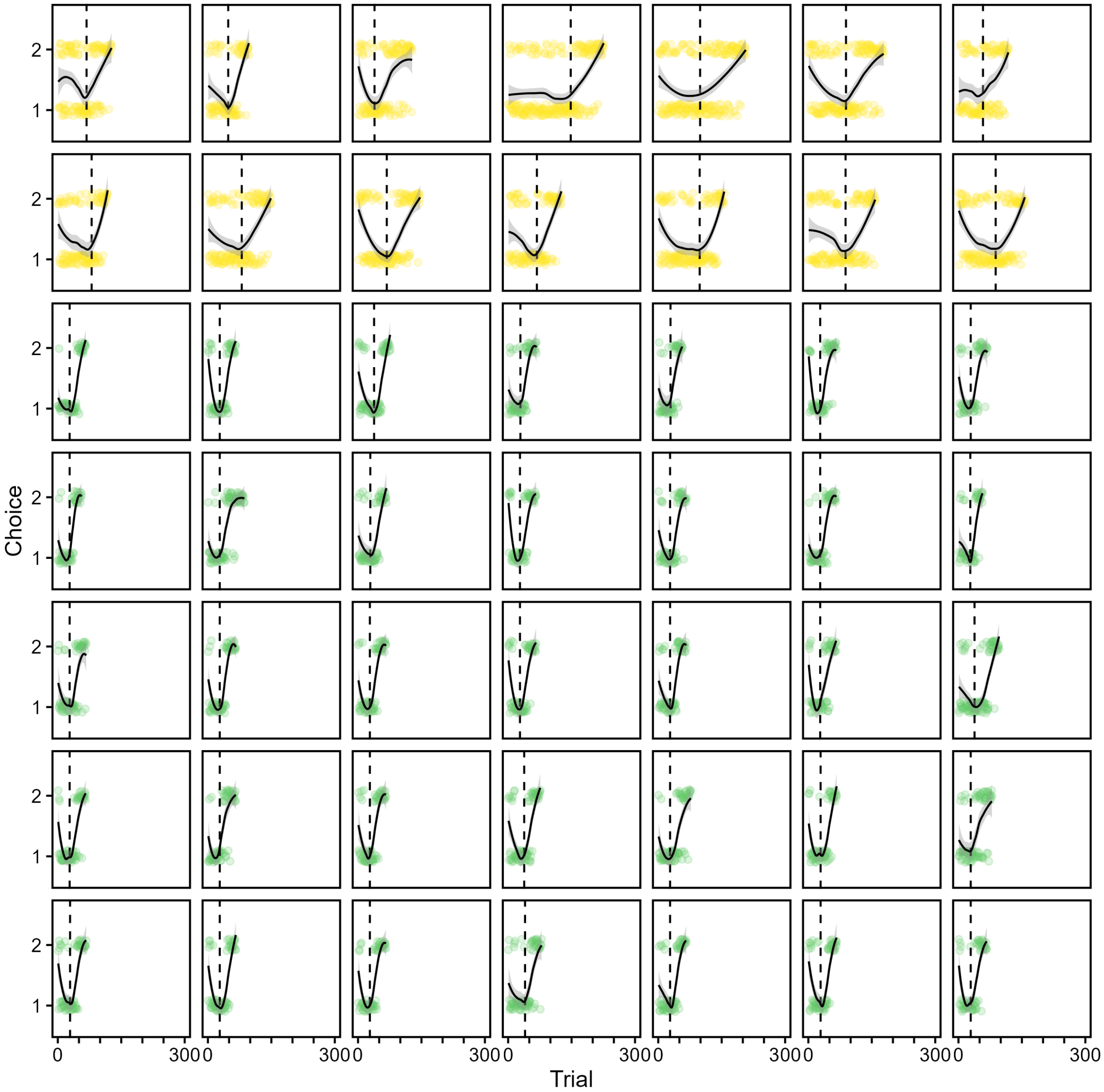


Figure S1 Individual-level tube-choice behaviour of simulated great-tailed grackles across colour-reward reinforcement learning trials (females: yellow, $n = 14$; males: green, $n = 35$). Tube option 1 (e.g., dark grey) was the rewarded option in the initial learning phase; conversely, tube option 2 (e.g., light grey) contained the food reward in the reversal learning phase. Each open circle shows an individual tube-choice; black solid lines show loess smoothed conditional means fitted with grey 89% compatibility intervals; and dashed black lines show individual-unique transitions between learning phases.