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

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

***For correspondence:** alexis_breen@eva.mpg.de (AJB); deffner@mpib-berlin.mpg.de (DD)

Present address: [§]Department of Linguistic and Cultural Evolution, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany

[1]Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany; [2]Science of Intelligence Excellence Cluster, Technical University Berlin, 10623 Berlin, Germany; [3]Center for Adaptive Rationality, Max Planck Institute for Human Development, 14195 Berlin, Germany

Abstract In the unpredictable Anthropocene, a particularly pressing open question is how certain species invade urban environments. Sex-biased dispersal and learning arguably influence movement ecology, but their joint influence remains unexplored empirically, and might vary by space and time. We assayed reinforcement learning in wild-caught, temporarily-captive core-, middle- or edge-range great-tailed grackles—a bird species undergoing urban-tracking rapid range expansion, led by dispersing males. We show: across populations, both sexes initially perform similarly when learning stimulus-reward pairings, but, when reward contingencies reverse, male—versus female—grackles finish 'relearning' faster, making fewer choice-option switches. How do male grackles do this? Bayesian cognitive modelling revealed male grackles' choice behaviour is governed more strongly by the 'weight' of relative differences in recent foraging payoffs—i.e., they show more pronounced risk-sensitive learning. Confirming this mechanism, agent-based forward simulations of reinforcement learning—where we simulate 'birds' based on empirical estimates of our grackles' reinforcement learning—replicate our sex-difference behavioural data. Finally, evolutionary modelling revealed natural selection should favour risk-sensitive learning in hypothesised urban-like environments: stable but stochastic settings. Together, these results imply risk-sensitive learning is a winning strategy for urban-invasion leaders, underscoring the potential for life history *and* cognition to shape invasion success in human-modified environments.

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 (***Ronce, 2007***; ***Chuang and Peterson, 2016***). 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 (***Miller et al., 2011***). Female-biased dispersal, for instance, can 'speed up' staged invertebrate invasions by increasing offspring production (***Miller and Inouye, 2013***). Alongside sex-biased dispersal, cognition 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 (***Wright et al., 2010***; ***Sutter and Kawecki, 2009***). Indeed,

a growing number of studies show support for this supposition, at least for those animals thriving in urban environments (recent reviews: ***Lee and Thornton 2021***; ***Szabo et al. 2020***; ***Barrett et al. 2019***; but see: ***Vincze and Kovács 2022***). Carefully controlled choice tests show, for example, urban-dwelling individuals will learn novel stimulus-reward pairings more readily than do rural-dwelling counterparts (***Batabyal and Thaker, 2019***). Similarly, urban- compared to conspecific rural-dwellers will more frequently figure out how to overcome experimenter-placed obstacles blocking known food resources (***Mazza and Guenther, 2021***). And how often such hidden food rewards are successfully accessed positively relates to how long specific subspecies have lived commensually with humans in urban environments (***Vrbanec et al., 2021***). Taken together, these data in both resident urban and non-urban species as well as urban invasive species, generally support the view that urban environments favour particular learning phenotypes. It is perhaps surprising, then, that despite their apparent independent influence on movement ecology, the potential interactive influence of sex-biased dispersal and learning on successful urban invasion remains unexamined empirically (but for related theoretical study, see: ***Liedtke and Fromhage 2021a***,b), especially as the dynamics of urban environments appear distinctly demanding.

Urban environments are hypothesised to be both stable and stochastic: more specifically, urbanisation is thought to lead to stabilisation in some aspects of biotic structure, including predation pressure, thermal habitat, and resource availability, and to enhanced abiotic disruption, such as anthropogenic noise and light pollution (reviewed in: ***Shochat et al., 2006***; ***Francis and Barber, 2013***; ***Gaston et al., 2013***). 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; ***Cowie and Hinsley, 1988***; ***Davies et al., 2009***), the density of which can positively relate to avian abundance within an urban area (***Fuller et al., 2008***). But such supplemental feeding opportunities are theorised to be traded off against unpredictable and likely fitness-affecting anthropogenic disturbances (e.g., automobile and airplane traffic; as outlined in ***Frid and Dill, 2002***). Experimental data show, for example, the more variable are traffic noise and pedestrian presence, the more negative are such human-driven effects on birds' sleep (***Grunst et al., 2021***), mating (***Blickley et al., 2012***), and foraging behaviour (***Fernández-Juricic, 2000***). Understanding how and why particular species successfully inhabit seemingly stable but stochastic urban environments remains an open and timely question, as highlighted by recent concerns over (in)vertebrates' coping ability to current and projected levels of urbanisation (***Eisenhauer and Hines, 2021***; ***Li et al., 2022***)

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 ∼150 years, they have seemingly shifted their historically urban niche to include more variable urban environments (e.g., arid habitat; ***Summers et al., 2023***), moving from their native range in Central America into much of the United States, with several first-sightings spanning as far north as Canada (***Dinsmore and Dinsmore, 1993***; ***Wehtje, 2003***; ***Fink et al., 2020***). 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 of the twentieth century; ***Dinsmore and Dinsmore, 1993***). 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 (***Sevchik et al., 2022***). Considering these life history and genetic data in conjunction with data on grackle wildlife management efforts (e.g., pesticides, pyrotechnics, and sonic booms; ***Luscier, 2018***), it seems plausible that—regardless of their shared urban 'history'—urban invasion might 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; ***Phillips et al., 2010***). Indeed, irrespective of population membership, such sex differences could come about via differential reliance on learning strategies mediated by an interaction between grackles' polygynous mating system and

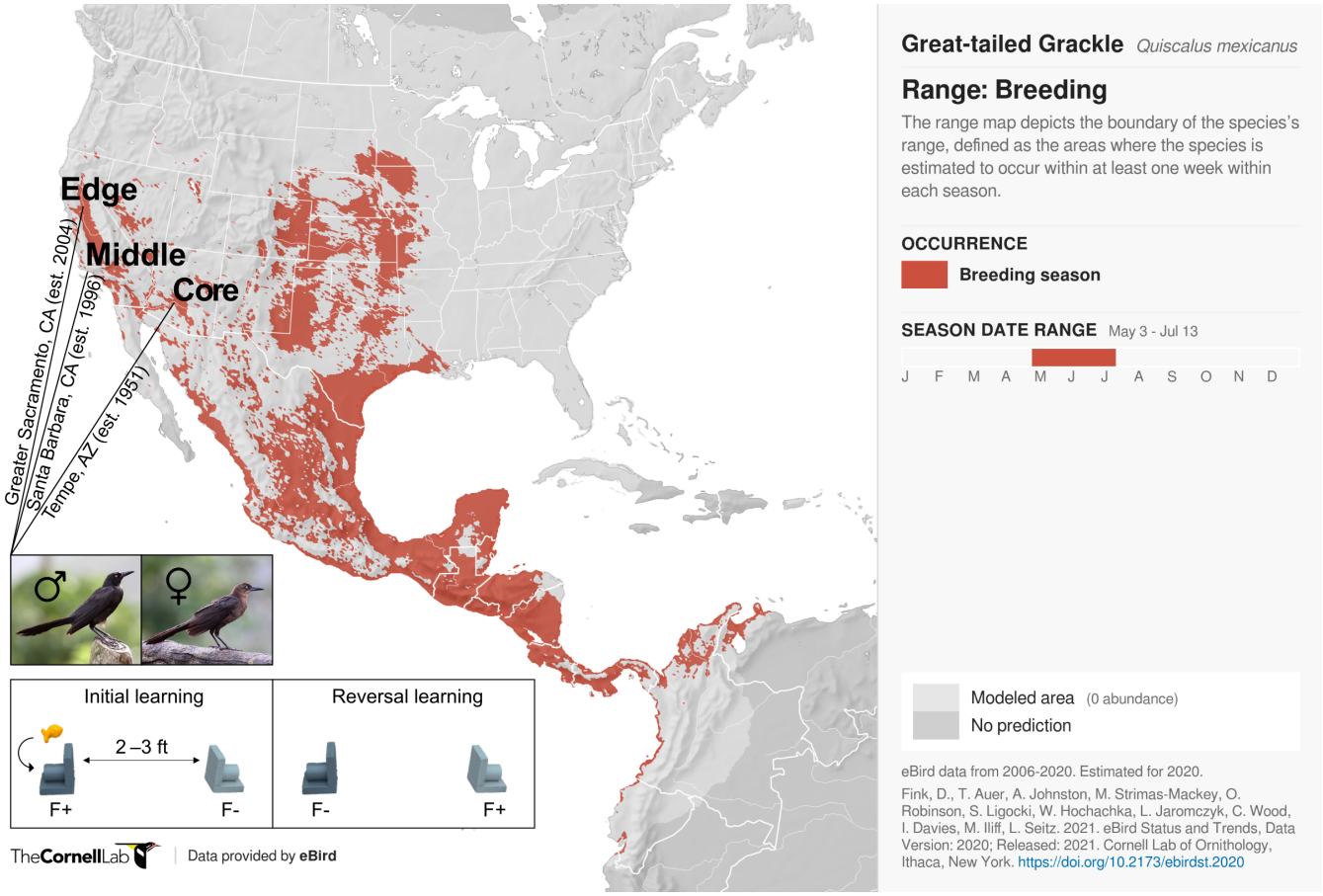

Figure 1. Participants and experimental protocol. Thirty-two male and 17 female wild-caught, temporarily-captive great-tailed grackles either inhabiting a core (17 males, 5 females), middle (4 males, 4 females) or edge (11 males, 8 females) population of their North American breeding range (establishment 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 Materials and methods).

male-biased dispersal system (see our in-depth discussion on this point). Population membership might, in turn, differentially moderate the magnitude of any such sex-effect since an edge population, even though urban, could still pose novel challenges—for example, by requiring grackles to learn novel daily temporal foraging patterns such as when and where garbage is collected (grackles appear to track this food resource: ***Rodrigo et al. 2021***).

Here, we test the hypothesis that sex differences in dispersal are related to sex differences in learning in an urban invader—grackles. Specifically, 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 Materials and methods; ***Figure 1***). To do this, we collated, cleaned, and curated existing reinforcement learning data (***Logan, 2016c***; ***Logan et al., 2022a***, ***2023a***), wherein novel stimulus-reward pairings are presented (i.e., *initial learning*), and, once successfully learned via reaching a criterion, these reward contingencies are reversed (i.e., *reversal learning*). 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 (trials to criterion) and choice-option switches (times alternating between available stimuli). Specifically, as documented

in our preregistration (see Supplementary file 1), we expect male—versus female—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.

To comprehensively examine links between sex-biased dispersal and learning in urban-invading grackles, we employ a combination of Bayesian computational and cognitive modelling methods, and both agent-based and evolutionary simulation techniques. Specifically, our paper proceeds as follows: (*i*) we begin by describing grackles' reinforcement learning and testing our predictions using multi-level Bayesian Poisson models; (*ii*) we next 'unblackbox' candidate learning mechanisms generating grackles' reinforcement learning using a multi-level Bayesian reinforcement learning model, as ultimately the mechanisms of behaviour should be the target of selection; (*iii*) we then try to replicate our behavioural data via agent-based forward simulations, to determine if our inferred learning mechanisms sufficiently explain our grackles' reinforcement learning; (*iv*) and we conclude by examining the evolutionary implications of variation in these learning mechanisms in hypothesised urban-like (or not) environments—that is, statistical settings that vary in both stability and stochasticity—via algorithmic optimisation.

Results

Reinforcement learning behaviour

We do not observe credible population-level differences in grackles' reinforcement learning (Supplementary file 2a and 2b). 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 (mean trials-to-finish: males, 35; females, 36; ***Figure 2***A and Supplementary file 2a), and with comparable counts of choice-option switches (mean switches-at-finish: males, 14; females, 14; ***Figure 2*** and Supplementary file 2b). Indeed, the male-female (M-F) posterior contrasts for both behaviours centre around zero, evidencing no sex-effect (***Figure 2***C). Once reward contingencies reverse, however, male—versus female—grackles finish this 'relearning' by taking fewer trials (mean trials-to-finish: males, 67; females, 81; ***Figure 2***B and Supplementary file 2a), and by making fewer choice-option switches (mean switches-at-finish: males, 25; females, 36; ***Figure 2***B and Supplementary file 2b). The M-F posterior contrasts, which lie almost entirely below zero, clearly capture this sex-effect (***Figure 2***F and Supplementary file 2a and b). Environmental unpredictability, then, dependably directs disparate reinforcement learning trajectories between male and female grackles, supporting our overall expectation of sex-mediated differential learning in urban-invading grackles.

Reinforcement learning mechanisms

Because (dis)similar behaviour can result from multiple latent processes (***McElreath, 2018***), we next employ computational methods to delimit reinforcement learning mechanisms. Specifically, we adapt a standard multi-level Bayesian reinforcement learning model (from ***Deffner et al. 2020***; but see also: ***Farrell and Lewandowsky 2018***; ***McElreath et al. 2005***; ***Sutton and Barto 2018***), which we validate *apriori* via agent-based simulation (see Materials and methods and Supplementary file 1), to estimate the contribution of two core latent learning parameters to grackles' reinforcement learning: 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, but they differentially direct any reward sensitivity back on choice-behaviour due to their distinct definitions (full mathematical details in Materials and methods). For ϕ, stronger reward sensitivity (bigger values) means faster internal

updating about stimulus-reward pairings, which translates behaviourally into faster learning about 'what to choose'. For λ, stronger reward sensitivity (bigger values) means stronger internal determinism about seeking the non-risk foraging option (i.e., the one with the higher expected payoffs based on prior experience), which translates behaviourally into less choice-option switching i.e., 'playing it safe'. In sum, conditional on our reinforcement learning model, we reverse engineer which values of our learning parameters most likely produce grackles' choice behaviour—an analytical advantage over less mechanistic methods (***McElreath, 2018***).

Looking at our reinforcement learning model's estimates between populations to determine replicability, 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 (M-F posterior contrasts lie more below zero; ***Figure 2***G and Supplementary file 2c), with smaller sample size likely explaining the middle population's more uncertain estimates (M-F posterior contrasts centre widely around zero; ***Figure 2***G and Supplementary file 2c); 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 (M-F posterior contrasts centre closely around zero; ***Figure 2***H and Supplementary file 2c). Therefore, the information-updating rate ϕ across male and female grackles is initially different (males $<$ females), but converges downwards over reinforcement learning phases (across-population M-F posterior contrasts first lie mostly below, and then, tightly bound zero; ***Figure 2***G and H and Supplementary file 2c).

These primary mechanistic findings are, at first glance, perplexing: if male grackles generally outperform female grackles in reversal learning (***Figure 2***D-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 (M-F posterior contrasts lie more, if not mostly, above zero; ***Figure 2***I and J and Supplementary file 2d), outwith the middle population in initial learning likely due to sample size (M-F posterior contrasts centre broadly around zero; ***Figure 2***I and Supplementary file 2d). In other words, choice behaviour in male grackles is more strongly governed by relative differences in predicted reward-payoffs, as spotlighted by across-population M-F posterior contrasts that lie mostly above zero in initial learning, and entirely above zero in reversal learning (***Figure 2***I and J and Supplementary file 2d). Thus, these combined mechanistic data reveal, when reward contingencies reverse, male—versus female—grackles 'relearn' faster via pronounced risk-sensitive learning.

Agent-based simulations and replication of reinforcement learning

To determine definitively whether our estimated learning parameters are sufficient to generate grackles' observed reinforcement learning, we conduct agent-based forward simulations. Agent-based forward simulations are posterior predictions, and they provide insight into the implied model dynamics and overall usefulness of our reinforcement learning model. Specifically, agent-based forward simulations allow us to ask—what would a 'new' grackle 'do', given our reinforcement learning model parameter estimates? It is important to ask this question because, in parameterising our model, we may have overlooked a critical contributing mechanism to grackles' reinforcement learning. Such an omission is invisible in the raw parameter estimates; it is only betrayed by the estimates *in actu*. The simulation thus assigns 'birds' random information-updating rate ϕ and risk-sensitivity rate λ draws (whilst maintaining their correlation structure), and tracks their reinforcement learning. 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. If no critical contributing mechanism(s) went overlooked, simulated 'birds' should behave similar to our real birds. A disparate mapping between simulated 'birds' and our real birds, however, would mean more work is needed with respect to model parameterisation that captures the causal, mechanistic dynamics behind real birds' reinforcement learning (for an example of this

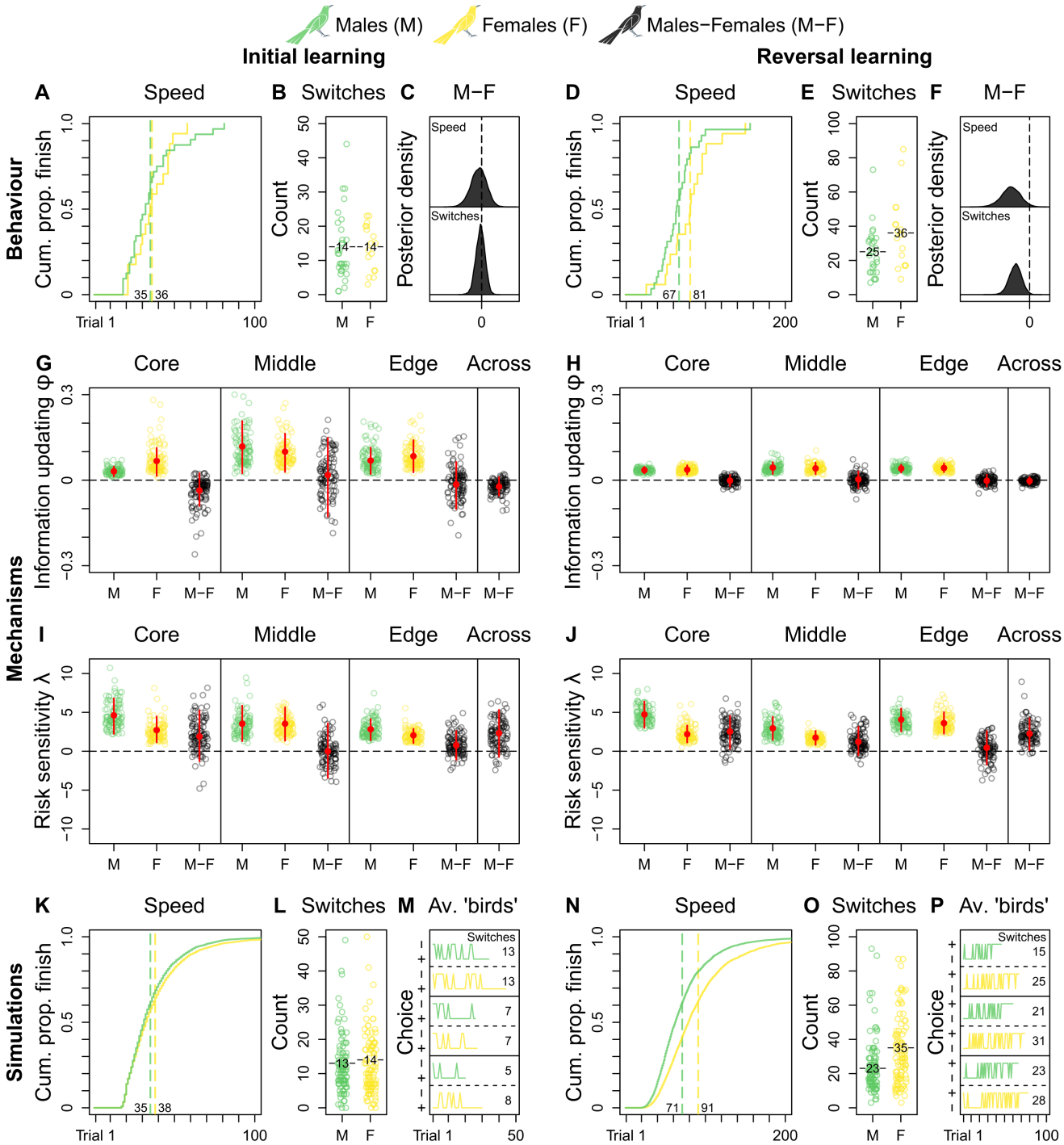

Figure 2. Grackle reinforcement learning. *Behaviour*. Across-population reinforcement learning speed and choice-option switches in (**A-B**) initial learning (M, 32; F, 17) and (**D-E**) reversal learning (M, 29; F, 17), with (**C, F**) respective posterior estimates and M-F contrasts. *Mechanisms*. Within- and across-population estimates and contrasts of *information-updating rate* ϕ and *risk-sensitivity rate* λ in (**G, I**) initial learning and (**H, J**) reversal learning. In (**G-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. In (**K, N**) the full simulation sample is plotted; in (**L, O**) open circles show 100 random simulant draws. Note (**K, N**) x-axes are cut to match (**A, D**) x-axes. Means are plotted/labelled in (**A, B, D, E, K, L, N, O**). Plots (**G-P**) are generated via model estimates using our full sample size (M, 32; F, 17).

Figure 2—figure supplement 1. Extra initial learning trials do not noticeably influence grackles' reversal learning.

happening in the human reinforcement learning literature, see ***Deffner et al. 2020***).

Ten thousand synthetic reinforcement learning trajectories (5000 'males' and 5000 'females'), together, compellingly show our 'birds' behave just like our grackles: 'males' outpace 'females' in reversal but not in initial learning (mean trials-to-finish initial and reversal learning: 'males', 35 and 71; 'females', 38 and 91; respectively; ***Figure 2***K and N); and 'males' make fewer choice-option switches in initial but not in reversal learning, compared to 'females' (mean switches-at-finish in initial and reversal learning: 'males', 13 and 23; 'females', 14 and 35; respectively; ***Figure 2***L and O). ***Figure 2***M and P show, respectively, synthetic initial 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. Such quantitative replication confirms our reinforcement learning model results sufficiently explain our behavioural sex-difference data.

Selection of reinforcement learning mechanisms under urban-like environments

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 (***Blackburn et al., 2009***; ***Sol et al., 2013***; ***Lee and Thornton, 2021***; ***Vinton et al., 2022***; ***Caspi et al., 2022***). Applying an evolutionary algorithm model (***Figure 3***A), 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 the two ecological axes thought to represent urban environmental characteristics (see Introduction): *environmental stability* u (How often does optimal behaviour change?) and *environmental stochasticity* s (How often does optimal behaviour fail to payoff?). In reality, much additional (a)biotic and life history information is relevant to thoroughly testing such urban eco-evolutionary theory. But our evolutionary model is a necessary and useful first step towards addressing targeted research questions about the interplay between learning phenotype and apparent environmental characteristics.

Strikingly, under hypothesised urban-like (i.e., stable but stochastic) environments, 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—should be favoured by natural selection (darker and lighter squares in, respectively the left and right plots in ***Figure 3***B). These results imply, in seemingly 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 such settings. We note that, for most rows of environmental stochasticity s, intermediate levels of environmental stability u generated the highest risk-sensitivity rate λ values, while risk-sensitivity rate λ values were slightly lower for both extremes of environmental stability u. The most plausible explanation for this pattern is that when the environment changes at intermediate levels, it is advantageous for learners to be fairly risk-sensitive (like our grackles), so that behaviour can be both effective and effectively adjusted. By contrast, in environments with either minimal or maximal environmental change, selection should be weaker for, or favouring lower degrees of, risk-sensitive learning because optimal behaviour is more-or-less constant or changing, respectively.

Discussion

Mapping a full pathway from behaviour to mechanisms through to selection and adaptation, we show risk-sensitive learning is a viable strategy to help explain how male grackles—the dispersing sex—currently lead their species' remarkable North American urban invasion. Specifically, in wild-caught, temporarily-captive core-, middle- or edge-range grackles, we show: (*i*) irrespective of population membership, male grackles outperform female counterparts on stimulus-reward reversal reinforcement learning, finishing faster and making fewer choice-option switches; (*ii*) they achieve

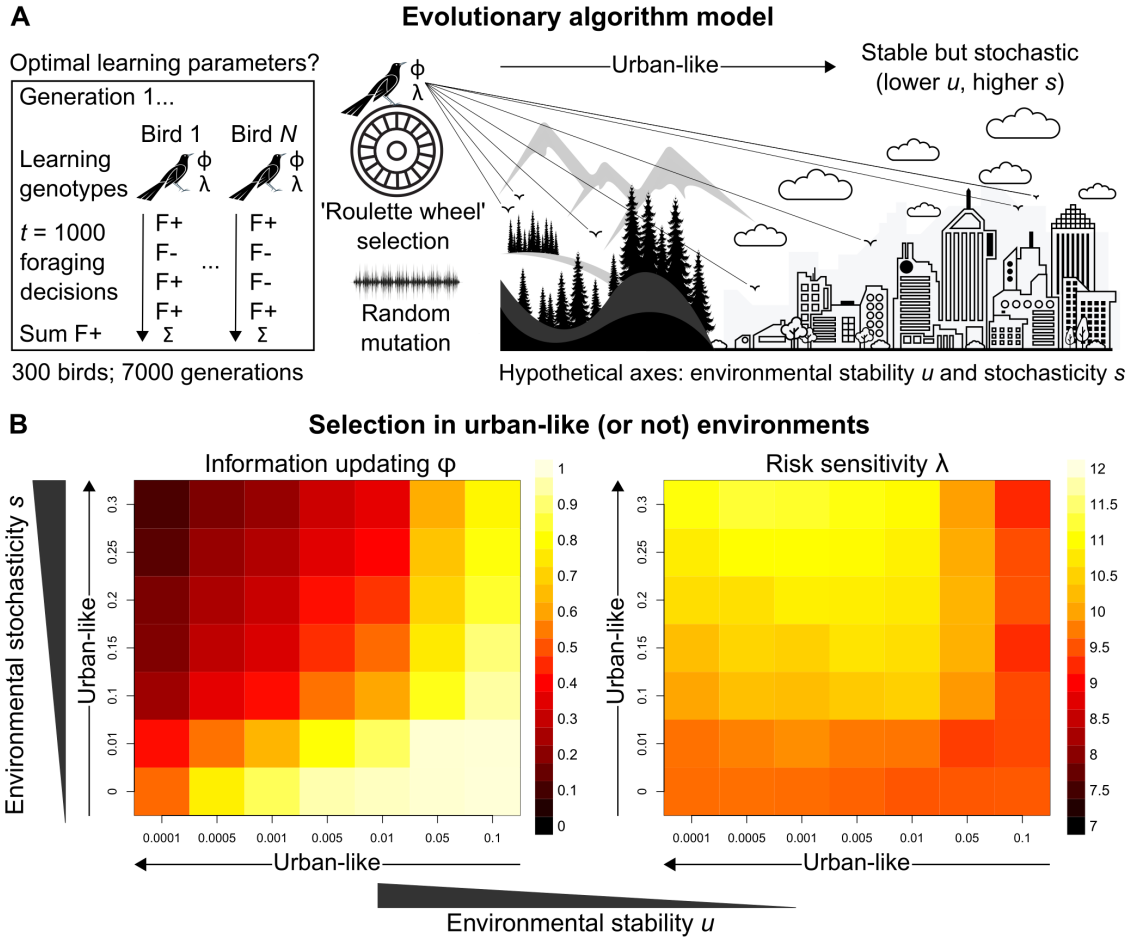

Figure 3. Evolutionary optimality of strategising risk-sensitive learning. (**A**) Illustration of our evolutionary algorithm model to estimate optimal learning parameters that evolve under systematically varied pairings of two hypothesised urban ecology axes: *environmental stability* u and *environmental 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 seemingly more urban-like (lower u and higher s values), selection should favour lower information-updating rate ϕ and higher risk-sensitivity rate λ (darker and lighter squares in left and right plot, respectively). We note arrows are intended as illustrative aids and do not correspond to a linear scale of 'urbanness'.

their speedier reversal learning performance via pronounced risk-sensitive learning (low ϕ and high λ), as 'unblackboxed' by our mechanistic model; (*iii*) these learning mechanisms sufficiently explain our sex-difference behavioural data: because we replicate our results using agent-based forward simulations; and (*iv*) risk-sensitive learning—i.e., low ϕ and high λ—appears advantageous in hypothesised urban-like environments (stable but stochastic settings), according to our evolutionary model. These results set the scene for future comparative research.

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 (***Coppens et al., 2010***; ***Audet and Lefebvre, 2017***; ***Barrett et al., 2019***; ***Lea et al., 2020***)—has previously been hypothesised to explain invasion success (***Wright et al., 2010***), including that of grackles (***Summers et al., 2023***). But as eloquently argued elsewhere (***Audet and Lefebvre, 2017***), 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 assay in bees; ***Strang and Sherry 2014***; ***Raine and Chittka 2012***; birds; ***Bond et al. 2007***; ***Morand-Ferron et al. 2022***; fish; ***Lucon-Xiccato and Bisazza 2014***; ***Bensky and Bell 2020***; frogs; ***Liu et al. 2016***; ***Burmeister 2022***; reptiles; ***Batabyal and Thaker 2019***; ***Gaalema 2011***; primates; ***Cantwell et al. 2022***; ***Lacreuse et al. 2018***; and rodents; ***Rochais et al. 2021***; ***Boulougouris et al. 2007***). We have shown, however, at least for our grackles, faster reversal learning is governed primarily by risk-sensitive learning, so: firstly, these go-to reversal learning assays do not necessarily measure the unit they claim to measure (a point similarly highlighted in: ***Aljadeff and Lotem, 2021***; ***Federspiel et al., 2017***); and secondly, formal models based on the false premise that variation in learning speed relates to variation in behavioural flexibility require reassessment (***Lea et al., 2020***; ***Blaisdell et al., 2021***; ***Logan et al., 2022b***; ***Lukas et al., 2023***; ***Logan et al., 2023b***,c). Heeding previous calls (***Dukas, 1998***; ***McNamara and Houston, 2009***; ***Fawcett et al., 2013***), 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 examine individual decision making, two core underlying learning mechanisms, and their theoretical selection and benefit (see https://github.com/alexisbreen/Sex-differences-in-grackles-learning), which can be tailored to specific research questions. The reinforcement learning model, for example, generalises to, in theory, a variety of choice-option paradigms (***Barrett, 2022***), and these learning models can be extended to estimate asocial *and* social influence on individual decision making (e.g., ***McElreath et al., 2005***; ***Aplin et al., 2017***; ***Barrett et al., 2017***; ***Deffner et al., 2020***; ***Chimento et al., 2022***), facilitating insight into the multi-faceted feedback process between individual cognition and social systems (***Trump et al., 2023***). Our open-access analytical resource thus allows researchers to dispense with the umbrella term behavioural flexibility, and to biologically inform and interpret their science—only then can we begin to meaningfully examine the functional basis of behavioural variation across taxa and/or contexts.

Ideas and speculation

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 (***Johnson et al., 2000***). Recent learning 'style' simulations show the sex with high reproductive skew approaches pure individual learning, while the other sex approaches pure social learning (***Smolla et al., 2019***). During population establishment, then, later-arriving female grackles could rely heavily on vetted information provided by male grackles on 'what to do' (***Wright et al., 2010***), 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; ***Figure 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.

The lack of spatial replicates in the existing data set used herein inherently poses limitations on inference. Nevertheless, the currently available data do not show meaningful population-level behavioural or mechanistic differences in grackles' reinforcement learning, and we should thus be cautious about speculating on between-population variation. 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 (***Blackburn et al., 2009***; ***Sol et al., 2013***; ***Lee and Thornton, 2021***; ***Vinton et al., 2022***; ***Caspi et al., 2022***). For urban-invading grackles, both of these biological explanations seem strongly plausible, given: firstly, grackles' highly plastic foraging and nesting ecology (***Selander and Giller, 1961***; ***Davis and Arnold, 1972***;

Wehtje, 2003); 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 (***Summers et al., 2023***); and finally, our combined behavioural, mechanistic, and evolutionary modelling results showing environments apparently approaching grackles' general species niche—urban environments—select for particular traits that exist across grackle populations (here, sex-biased risk-sensitive learning). Admittedly, our evolutionary model is not a complete representation of urban ecology dynamics. Relevant factors—e.g., spatial dynamics and realistic life histories—are missing. These omissions are tactical ones. Our evolutionary model solely focuses on the response of reinforcement learning parameters to hypothesised 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., ***Duckworth and Badyaev, 2007***).

Conclusions

By revealing across-population interactive links between the dispersing sex and risk-sensitive learning in an urban invader (grackles), these analytically replicable insights, coupled with our finding that hypothesised urban-like environments favour pronounced risk-sensitivity, imply risk-sensitive learning is a winning strategy for urban-invasion leaders. 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 phenotypic distributions (***Somveille et al., 2018***; ***Bro-Jørgensen et al., 2019***; ***Lee and Thornton, 2021***; ***Breen et al., 2021***; ***Breen, 2021***; ***Deffner et al., 2022***).

Methods and Materials

Data provenance

The current study uses data from two types of sources: publicly archived data at the Knowledge Network for Biocomplexity (***Logan, 2016c***; ***Logan et al., 2022a***); or privately permissed access to A.J.B. of (at the time) unpublished data by Grackle Project principal investigator Corina Logan, who declined participation on this study. We note these shared data are now also available at the Knowledge Network for Biocomplexity (***Logan et al., 2023a***).

Data contents

The data used herein chart colour-reward reinforcement learning performance from 32 male and 17 female wild-caught, temporarily-captive grackles inhabiting one of 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 ***Chuang and Peterson, 2016***). Specifically: (*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; ***Wehtje, 2003***, ***2004***); (*ii*) Santa Barbara, California (four males and four females)—herein, the middle population (known to be breeding since 1996; ***Lehman, 2020***); and (*iii*) Greater Sacramento, California (eleven males and eight females)—herein, the edge population (known to be breeding since 2004; ***Hampton, 2004***).

Experimental protocol

A step-by-step description of the experimental protocol carried out by the original experiments is reported elsewhere (***Blaisdell et al., 2021***). 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 (***Figure 1***): (*i*) stimulus-reward initial learning and (*ii*) stimulus-reward 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 (***Figure 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 (***Figure 1***). The protocol for this second and final experimental phase is identical to that, described above, of initial learning.

Reinforcement 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 (***Logan, 2016c***) and Logan et al. (***Logan et al., 2022a***) 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 final trial number for that set (e.g., 20, 30, 40), which is problematic because this trial did not always coincide with the true passing trial (by a maximum of two additional trials; see below).

For edge-population grackles, Logan and colleagues (***Logan et al., 2023a***) used a sliding-window learning criterion, 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 20 onward) rather than at the end of discrete sets. This second method is also 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 (row 1816 in https://github.com/alexisbreen/Sex-differences-in-grackles-learning; as well as in ***Logan et al., 2023a***). 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 17/20 learning criterion ensures our assessment of grackles' reinforcement learning is informative, straightforward, and consistent.

As a consequence of applying our 17/20 learning criterion, grackles can remain in initial and/or reversal learning beyond reaching criterion. These extra learning trials, however, already exist for some core- and middle-population grackles originally assessed via the fixed-window learning criterion ($N = 18$ in initial [range: 1-2]; $N = 13$ in reversal [range: 1-2]), as explained above. And our cleaning of the original data (see our Data_Processing.R script at https://github.com/alexisbreen/Sex-differences-in-grackles-learning) detected additional cases where grackles remained in-test despite meeting the applied criterion (fixed-window: $N = 1$ in reversal for 11 trials; sliding-window: $N = 11$ in initial [range: 1-10]; $N = 7$ in reversal [range: 1-14]), presumably due to experimenter oversight. Similarly, our data cleaning detected two birds in the middle-population that were originally passed despite not meeting the assigned fixed-window learning criterion; instead, both chose 7/10 and 10/10 correct choices in the last two sets of 10 trials. Moreover, our data cleaning detected an additional four birds in the core-population that did not in fact meet the fixed-window learning criterion because of incorrect trail numbers entered by the original experimenters e.g., skipping trial 24. We note these data issues, as well as the problematic nature of both the fixed- and sliding-window learning criterion, continue to be unaddressed in work by Logan (et al.) (***Logan, 2016b***,a; ***Logan et al., 2022b***, ***2023b***,c). In any case, in our study we verified: (*i*) our 17/20 learning criterion results in a similar proportion of male and female grackles experiencing 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 (***Figure 2—figure Supplement 1***). Thus, we are confident that any carry-over effect of extra initial learning trials on grackles' reversal learning in our study is negligible if not nonexistent, and we therefore excluded extra learning trials.

Statistical analyses

We analysed, processed, and visually present our data using, respectively, the 'rstan' (***Team, 2020***), 'rethinking' (***McElreath, 2018***), and 'tidyverse' (***Wickham et al., 2019***) packages in R (***Team, 2021***). We note our reproducible code is available at https://github.com/alexisbreen/Sex-differences-in-grackles-learning. 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 the original 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 (***McElreath, 2018***). Our Poisson models, however, do exclude these three cases for our modelling of reversal learning, to conserve estimation. Finally, we note that in Bayesian statistics, there is no strict lower limit of required sample size as the inferences do not rely on asymptotic assumptions. With inferences remaining valid in principle, low sample size will of course be reflected in rather uncertain posterior estimates. We further note all of our multilevel models use partial pooling on individuals (the random-effects structure), which is a regularisation technique that especially improves inference in case of low sample sizes (see Ch. 13 in ***McElreath 2018***). The full output from each of our models, which use weakly informative priors, is available in Tables S1-S4, including posterior means and 89% highest posterior density intervals that convey the most likely values of our parameters over this range (***McElreath, 2018***).

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 ***Deffner et al., 2020***) 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 (***McElreath, 2018***).

Our reinforcement learning model consists of two equations:

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

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

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' (***Sutton and Barto, 2018***; ***Chimento et al., 2022***)—since the higher its value the more deterministic the choice behaviour of an individual becomes (note $\lambda = 0$ generates random choice). In line with risk-sensitive foraging theory—which focuses on how animals evaluate and choose between distinct food options, and how such foraging decisions are influenced by pay-off variance i.e., risk associated with alternative foraging options (seminal reviews: ***Bateson 2002***; ***Kacelnik and Bateson 1996***)—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: https://github.com/alexisbreen/Sex-differences-in-grackles-learning). 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 (***Webster and Rutz, 2020***). 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 https://github.com/alexisbreen/Sex-differences-in-grackles-learning).

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 file 1). 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 '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 ***Figure 2*** and associated main text; for an example of this approach in a different context, see ***Deffner et al., 2020***). We note agent-based forward simulation does not access goodness-of-fit—we assessed the fit of our model *apriori* in our preregistration (see Supplementary file 1)—but it does assess whether one did a comprehensive job of uncovering the mechanistic basis of target behaviour(s), as detailed in the main text.

Evolutionary model

To investigate the evolutionary significance of strategising risk-sensitive learning, we used algorithmic optimisation techniques (***Yu and Gen, 2010***; ***Otto and Day, 2011***). 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 (***Otto and Day, 2011***). 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 (***Otto and Day, 2011***).

To generate the next generation, we assume asexual, haploid reproduction, and use fitness-proportionate (or 'roulette wheel') selection to choose individuals for reproduction (***Yu and Gen, 2010***; ***Otto and Day, 2011***). 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.

Acknowledgments

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 advanced access (edge population; see Data provenance), 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.

Additional information

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 Figure 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.

Author ORCIDs

Alexis J Breen https://orcid.org/0000-0002-2331-0920
Dominik Deffner https://orcid.org/0000-0002-1649-3861

Additional files

Supplementary files

- Supplementary file 1. Study preregistration, including reinforcement learning model validation.
- Supplementary file 2. Supplementary tables. (a) Total-trials-in-test Poisson regression model output (b) Total-choice-option-switches-in-test Poisson regression model output. (c) Bayesian reinforcement learning model information-updating rate ϕ output (d) Bayesian reinforcement learning model risk-sensitvity rate λ output. For (a-d), both between- and across-population posterior means and corresponding 89% highest-posterior density intervals are reported for males, females, and male-female contrasts.

References

Aljadeff N, Lotem A. Task-dependent reversal learning dynamics challenge the reversal paradigm of measuring cognitive flexibility. Animal Behaviour. 2021; 179:183–197. doi: 10.1016/j.anbehav.2021.07.002.

Aplin LM, Sheldon BC, McElreath R. Conformity does not perpetuate suboptimal traditions in a wild population of songbirds. Proceedings of the National Academy of Sciences. 2017; 114(30):7830–7837. doi: 10.1073/pnas.1621067114.

Audet JN, Lefebvre L. What's flexible in behavioral flexibility? Behavioral Ecology. 2017; 28(4):943–947. doi: 10.1093/beheco/arx007.

Barrett BJ. Inferential power in identifying frequency-dependent social learning strengthened by increasing behavioural options. Journal of Animal Ecology. 2022; doi: 10.1111/1365-2656.13826.

Barrett BJ, McElreath RL, Perry SE. 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. 2017; 284(1856):20170358. doi: 10.1098/rspb.2017.0358.

Barrett LP, Stanton LA, Benson-Amram S. The cognition of 'nuisance'species. Animal Behaviour. 2019; 147:167–177. doi: 10.1016/j.anbehav.2018.05.005.

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

Bateson M. Recent advances in our understanding of risk-sensitive foraging preferences. Proceedings of the Nutrition Society. 2002; 61(4):509–516. doi: 10.1079/pns2002181.

Bensky MK, Bell AM. Predictors of individual variation in reversal learning performance in three-spined sticklebacks. Animal cognition. 2020; 23(5):925–938. doi: 10.1007/s10071-020-01399-8.

Blackburn TM, Lockwood JL, Cassey P. Avian invasions: the ecology and evolution of exotic birds, vol. 1. Oxford University Press; 2009.

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

Blickley JL, Blackwood D, Patricelli GL. Experimental evidence for the effects of chronic anthropogenic noise on abundance of greater sage-grouse at leks. Conservation Biology. 2012; 26(3):461–471. doi: 10.1111/j.1523-1739.2012.01840.x.

Bond AB, Kamil AC, Balda RP. 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. 2007; 121(4):372. doi: 10.1037/0735-7036.121.4.372.

Boulougouris V, Dalley JW, Robbins TW. Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. Behavioural brain research. 2007; 179(2):219–228. doi: 10.1016/j.bbr.2007.02.005.

Breen AJ. Animal culture research should include avian nest construction. Biology Letters. 2021; 17(7):20210327. doi: 10.1098/rsbl.2021.0327.

Breen AJ, Sugasawa S, Healy SD. Manipulative and technological skills do not require a slow life history. Frontiers in Ecology and Evolution. 2021; 9:46. doi: 10.3389/fevo.2021.635802.

Bro-Jørgensen J, Franks DW, Meise K. Linking behaviour to dynamics of populations and communities: application of novel approaches in behavioural ecology to conservation. Philosophical Transactions of the Royal Society B. 2019; 374(1781):20190008. doi: 10.1098/rstb.2019.0008.

Burmeister SS. Ecology, cognition, and the hippocampus: A tale of two frogs. Brain, Behavior and Evolution. 2022; 97(3-4):211–224. doi: 10.1159/000522108.

Cantwell A, Buckholtz JW, Atencia R, Rosati AG. The origins of cognitive flexibility in chimpanzees. Developmental science. 2022; 25(5):e13266. doi: 10.1111/desc.13266.

Caspi T, Johnson JR, Lambert MR, Schell CJ, Sih A. Behavioral plasticity can facilitate evolution in urban environments. Trends in Ecology & Evolution. 2022; 37:1092–1103. doi: 10.1016/j.tree.2022.08.002.

Chimento M, Barrett BJ, Kandler A, Aplin LM. Cultural diffusion dynamics depend on behavioural production rules. Proceedings of the Royal Society B: Biological Sciences. 2022; 289:20221001. doi: 10.1098/rspb.2022.1001.

Chuang A, Peterson CR. Expanding population edges: theories, traits, and trade-offs. Global Change Biology. 2016; 22(2):494–512. doi: 10.1111/gcb.13107.

Coppens CM, de Boer SF, Koolhaas JM. Coping styles and behavioural flexibility: towards underlying mechanisms. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010; 365(1560):4021–4028. doi: 10.1098/rstb.2010.0217.

Cowie RJ, Hinsley SA. The provision of food and the use of bird feeders in suburban gardens. Bird Study. 1988; 35(3):163–168. doi: 10.1080/00063658809476985.

Davies ZG, Fuller RA, Loram A, Irvine KN, Sims V, Gaston KJ. A national scale inventory of resource provision for biodiversity within domestic gardens. Biological Conservation. 2009; 142(4):761–771. doi: 10.1016/j.biocon.2008.12.016.

Davis WR, Arnold KA. Food habits of the Great-tailed Grackle in Brazos county, Texas. The Condor. 1972; 74(4):439–446. doi: 10.2307/1365896.

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

Deffner D, Rohrer JM, McElreath R. A causal framework for cross-cultural generalizability. Advances in Methods and Practices in Psychological Science. 2022; 5(3):1–18. doi: 10.1177/25152459221106366.

Dinsmore JJ, Dinsmore SJ. Range expansion of the great-tailed grackle in the 1900s. Journal of the Iowa Academy of Science. 1993; 100(2):54–59. doi: https://scholarworks.uni.edu/jias/vol100/iss2/4.

Duckworth RA, Badyaev AV. Coupling of dispersal and aggression facilitates the rapid range expansion of a passerine bird. Proceedings of the National Academy of Sciences. 2007; 104(38):15017–15022. doi: 10.1073/pnas.0706174104.

Dukas R. Cognitive ecology: the evolutionary ecology of information processing and decision making. University of Chicago Press; 1998.

Eisenhauer N, Hines J. Invertebrate biodiversity and conservation. Current Biology. 2021; 31(19):R1214–R1218. doi: 10.1016/j.cub.2021.06.058.

Farrell S, Lewandowsky S. Computational modeling of cognition and behavior. Cambridge University Press; 2018. doi: 10.1017/CBO9781316272503.

Fawcett TW, Hamblin S, Giraldeau LA. Exposing the behavioral gambit: the evolution of learning and decision rules. Behavioral Ecology. 2013; 24(1):2–11. doi: 10.1093/beheco/ars085.

Federspiel IG, Garland A, Guez D, Bugnyar T, Healy SD, Güntürkün O, Griffin AS. Adjusting foraging strategies: a comparison of rural and urban common mynas (Acridotheres tristis). Animal cognition. 2017; 20:65–74. doi: 10.1007/s10071-016-1045-7.

Fernández-Juricic E. Local and regional effects of pedestrians on forest birds in a fragmented landscape. The condor. 2000; 102(2):247–255. doi: 10.1093/condor/102.2.247.

Fink D, Auer T, Johnston A, Strimas-Mackey M, Robinson O, Ligocki S, Hochachka W, Jaromczyk L, Wood C, Davies I, Iliff M, Seitz L. eBird Status and Trends. . 2020; doi: 10.2173/ebirdst.2020, https://ebird.org/science/status-and-trends.

Francis CD, Barber JR. A framework for understanding noise impacts on wildlife: an urgent conservation priority. Frontiers in Ecology and the Environment. 2013; 11(6):305–313. doi: https://doi.org/10.1890/120183.

Frid A, Dill L. Human-caused disturbance stimuli as a form of predation risk. Conservation Ecology. 2002; 6(1).

Fuller RA, Warren PH, Armsworth PR, Barbosa O, Gaston KJ. Garden bird feeding predicts the structure of urban avian assemblages. Diversity and Distributions. 2008; 14(1):131–137. doi: 10.1111/j.1472-4642.2007.00439.x.

Gaalema DE. Visual discrimination and reversal learning in rough-necked monitor lizards (*Varanus rudicollis*). Journal of Comparative Psychology. 2011; 125(2):246. doi: 10.1037/a0023148.

Gaston KJ, Bennie J, Davies TW, Hopkins J. The ecological impacts of nighttime light pollution: a mechanistic appraisal. Biological Reviews. 2013; 88(4):912–927. doi: 10.1111/brv.12036.

Grunst ML, Grunst AS, Pinxten R, Eens M. Variable and consistent traffic noise negatively affect the sleep behavior of a free-living songbird. Science of the Total Environment. 2021; 778:146338. doi: 10.1016/j.scitotenv.2021.146338.

Hampton S, Yolo County birding news; 2004.

Johnson K, DuVal E, Kielt M, Hughes C. Male mating strategies and the mating system of great-tailed grackles. Behavioral Ecology. 2000; 11(2):132–141. doi: 10.1093/beheco/11.2.132.

Kacelnik A, Bateson M. Risky theories—the effects of variance on foraging decisions. American zoologist. 1996; 36(4):402–434. doi: 10.1093/icb/36.4.402.

Lacreuse A, Moore CM, LaClair M, Payne L, King JA. Glutamine/glutamate (Glx) concentration in prefrontal cortex predicts reversal learning performance in the marmoset. Behavioural Brain Research. 2018; 346:11–15. doi: 10.1016/j.bbr.2018.01.025.

Lea SE, Chow PK, Leaver LA, McLaren IP. Behavioral flexibility: a review, a model, and some exploratory tests. Learning & Behavior. 2020; 48:173–187. doi: 10.3758/s13420-020-00421-w.

Lee VE, Thornton A. Animal cognition in an urbanised world. Frontiers in Ecology and Evolution. 2021; 9:120. doi: 10.3389/fevo.2021.633947.

Lehman PE, The birds of Santa Barbara County, California; 2020. doi: http://www.sbcobirding.com/lehmanbosbc.html.

Li G, Fang C, Li Y, Wang Z, Sun S, He S, Qi W, Bao C, Ma H, Fan Y, et al. Global impacts of future urban expansion on terrestrial vertebrate diversity. Nature communications. 2022; 13(1):1–12. doi: 10.1038/s41467-022-29324-2.

Liedtke J, Fromhage L. The joint evolution of learning and dispersal maintains intraspecific diversity in metapopulations. Oikos. 2021; 130(5):808–818. doi: 10.1111/oik.08208.

Liedtke J, Fromhage L. Should dispersers be fast learners? Modeling the role of cognition in dispersal syndromes. Ecology and Evolution. 2021; 11(20):14293–14302. doi: 10.1002/ece3.8145.

Liu Y, Day LB, Summers K, Burmeister SS. Learning to learn: advanced behavioural flexibility in a poison frog. Animal Behaviour. 2016; 111:167–172. doi: 10.1016/j.anbehav.2015.10.018.

Logan CJ. Behavioral flexibility and problem solving in an invasive bird. PeerJ. 2016; 4:e1975. doi: 10.7717/peerj.1975.

Logan CJ. Behavioral flexibility in an invasive bird is independent of other behaviors. PeerJ. 2016; 4:e2215. doi: 10.7717/peerj.2215.

Logan CJ. Great-tailed grackle behavioral flexibility and problem solving experiments, Santa Barbara, CA USA 2014-2015. Knowledge Network for Biocomplexity. 2016; doi: 10.5063/F1319SVV.

Logan CJ, Blaisdell AP, Johnson-Ulrich Z, Lukas D, MacPherson M, Seitz B, Sevchik A, McCune KB. Is behavioral flexibility manipulatable and, if so, does it improve flexibility and problem solving in a new context? Knowledge Network for Biocomplexity. 2022; doi: 10.5063/F1862DWC.

Logan CJ, McCune K, LeGrande-Rolls C, Marfori Z, Hubbard J, Lukas D. Data: Implementing a rapid geographic range expansion - the role of behavior changes. Knowledge Network for Biocomplexity. 2023; doi: 10.5063/F1QZ28FH.

Logan C, Lukas D, Blaisdell A, Johnson-Ulrich Z, MacPherson M, Seitz B, Sevchik A, McCune K. Behavioral flexibility is manipulable and it improves flexibility and innovativeness in a new context. Peer Community Journal. 2023; 3. doi: 10.24072/pcjournal.284.

Logan C, McCune K, LeGrande-Rolls C, Marfori Z, Hubbard J, Lukas D. Implementing a rapid geographic range expansion-the role of behavior changes. Peer Community Journal. 2023; 3. doi: 10.24072/pcjournal.320.

Logan CJ, McCune K, MacPherson M, Johnson-Ulrich Z, Rowney C, Seitz B, Blaisdell A, Deffner D, Wascher C. Are the more flexible great-tailed grackles also better at behavioral inhibition? Animal Behavior Cognition. 2022; 9(1):14–36. doi: 10.26451/abc.09.01.03.2022.

Lucon-Xiccato T, Bisazza A. Discrimination reversal learning reveals greater female behavioural flexibility in guppies. Biology Letters. 2014; 10(6):20140206. doi: 10.1098/rsbl.2014.0206.

Lukas D, McCune K, Blaisdell AP, Johnson-Ulrich Z, MacPherson M, Seitz B, Sevchik A, Logan CJ. Behavioral flexibility is manipulatable and it improves flexibility and problem solving in a new context: post-hoc analyses of the components of behavioral flexibility. EcoEvoRxiv. 2023; doi: 10.32942/osf.io/5z8xs.

Luscier JD. Effects of urbanization on great-tailed grackle habitat use and nest success in Sherman Texas. Urban Naturalist. 2018; 15:1–14.

Mazza V, Guenther A. City mice and country mice: innovative problem solving in rural and urban noncommensal rodents. Animal Behaviour. 2021; 172:197–210. doi: 10.1016/j.anbehav.2020.12.007.

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

McElreath R, Lubell M, Richerson PJ, Waring TM, Baum W, Edsten E, Efferson C, Paciotti B. Applying evolutionary models to the laboratory study of social learning. Evolution and Human Behavior. 2005; 26(6):483–508. doi: 10.1016/j.evolhumbehav.2005.04.003.

McNamara JM, Houston AI. Integrating function and mechanism. Trends in Ecology & Evolution. 2009; 24(12):670–675. doi: 10.1016/j.tree.2009.05.011.

Miller TEX, Inouye BD. Sex and stochasticity affect range expansion of experimental invasions. Ecology Letters. 2013; 16(3):354–361. doi: 10.1111/ele.12049.

Miller TE, Shaw AK, Inouye BD, Neubert MG. Sex-biased dispersal and the speed of two-sex invasions. The American Naturalist. 2011; 177(5):549–561. doi: 10.1086/659628.

Morand-Ferron J, Reichert MS, Quinn JL. 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. 2022; 50(1):153–166. doi: 10.3758/s13420-021-00505-1.

Otto SP, Day T. A biologist's guide to mathematical modeling in ecology and evolution. Princeton University Press; 2011.

Phillips BL, Brown GP, Shine R. Life-history evolution in range-shifting populations. Ecology. 2010; 91(6):1617–1627. doi: 10.1890/09-0910.1.

Raine NE, Chittka L. No trade-off between learning speed and associative flexibility in bumblebees: a reversal learning test with multiple colonies. PloS one. 2012; 7:e45096. doi: 10.1371/journal.pone.0045096.

Rochais C, Hotte H, Pillay N. Seasonal variation in reversal learning reveals greater female cognitive flexibility in African striped mice. Scientific Reports. 2021; 11(1):20061. doi: 10.1038/s41598-021-99619-9.

Rodrigo A, Ávila-Chauvet L, Buriticá J. Daily Patterns of Foraging and Aggressive Behaviors in Great-tailed Grackle (Quiscallus mexicanus) at an Urban Patch with Availability or Absence of Resources. bioRxiv. 2021; p. 2021–06. doi: 10.1101/2021.06.14.448443.

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

Selander RK, Giller DR. Analysis of sympatry of great-tailed and boat-tailed grackles. The Condor. 1961; 63(1):29–86. doi: 10.2307/1365420.

Sevchik A, Logan CJ, McCune KB, Blackwell A, Rowney C, Lukas D. Investigating sex differences in genetic relatedness in great-tailed grackles in Tempe, Arizona to infer potential sex biases in dispersal. Animal Behavior and Cognition. 2022; 9:37–52. doi: 10.32942/osf.io/t6beh.

Shochat E, Warren PS, Faeth SH, McIntyre NE, Hope D. From patterns to emerging processes in mechanistic urban ecology. Trends in Ecology & Evolution. 2006; 21(4):186–191. doi: 10.1016/j.tree.2005.11.019.

Smolla M, Rosher C, Gilman RT, Shultz S. Reproductive skew affects social information use. Royal Society open science. 2019; 6(7):182084. doi: 10.1098/rsos.182084.

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

Somveille M, Firth JA, Aplin LM, Farine DR, Sheldon BC, Thompson RN. Movement and conformity interact to establish local behavioural traditions in animal populations. PLoS Computational Biology. 2018; 14(12):e1006647. doi: 10.1371/journal.pcbi.1006647.

Strang CG, Sherry DF. Serial reversal learning in bumblebees (*Bombus impatiens)*. Animal Cognition. 2014; 17:723–734. doi: 10.1007/s10071-013-0704-1.

Summers J, Lukas D, Logan CJ, Chen N. The role of climate change and niche shifts in divergent range dynamics of a sister-species pair. Peer Community Journal. 2023; 3. doi: 10.24072/pcjournal.248.

Sutter M, Kawecki T. Influence of learning on range expansion and adaptation to novel habitats. Journal of Evolutionary Biology. 2009; 22(11):2201–2214. doi: 10.1111/j.1420-9101.2009.01836.x.

Sutton RS, Barto AG. Reinforcement learning: An introduction. MIT press; 2018.

Szabo B, Damas-Moreira I, Whiting MJ. Can cognitive ability give invasive species the means to succeed? A review of the evidence. Frontiers in ecology and evolution. 2020; p. 187. doi: 10.3389/fevo.2020.00187.

Team RC, R: A Language and environment for statistical computing; 2021.

Team SD, RStan: the R interface to Stan; 2020.

Trochet A, Courtois EA, Stevens VM, Baguette M, Chaine A, Schmeller DS, Clobert J, Wiens JJ. Evolution of sex-biased dispersal. The Quarterly Review of Biology. 2016; 91(3):297–320. doi: 10.1086/688097.

Trump AN, Deffner D, Pleskac TJ, Romanczuk P, Kurvers R. A cognitive computational approach to social and collective decision-making. Perspectives on Psychological Science. 2023; 0(0):17456916231186964. doi: 10.1177/17456916231186964.

Vincze E, Kovács B. Urbanization's Effects on Problem Solving Abilities: A Meta-Analysis. Frontiers in Ecology and Evolution. 2022; 10:834436. doi: 10.3389/fevo.2022.834436.

Vinton AC, Gascoigne SJ, Sepil I, Salguero-Gómez R. Plasticity's role in adaptive evolution depends on environmental change components. Trends in Ecology & Evolution. 2022; doi: 10.1016/j.tree.2022.08.008.

Vrbanec L, Matijević V, Guenther A. Enhanced problem-solving ability as an adaptation to urban environments in house mice. Proceedings of the Royal Society B. 2021; 288(1945):20202504. doi: 10.1098/rspb.2020.2504.

Webster MM, Rutz C. How STRANGE are your study animals? Nature. 2020; 582(7812):337–340. doi: 10.1038/d41586-020-01751-5.

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

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

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, l Müller, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, et al. Welcome to the Tidyverse. Journal of Open Source Software. 2019; 4(43):1686. doi: 10.21105/joss.01686.

Wright TF, Eberhard JR, Hobson EA, Avery ML, Russello MA. Behavioral flexibility and species invasions: the adaptive flexibility hypothesis. Ethology Ecology & Evolution. 2010; 22(4):393–404. doi: 10.1080/03949370.2010.505580.

Yu X, Gen M. Introduction to evolutionary algorithms. Springer Science & Business Media; 2010.

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

Table a. 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 b. 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 c. 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 d. 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)

This preregistration of our study, which is intended to serve as a record of our scientific method, and 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; and to OSF on 21 February 2022: https://doi.org/10.17605/OSF.IO/V3WXB

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.

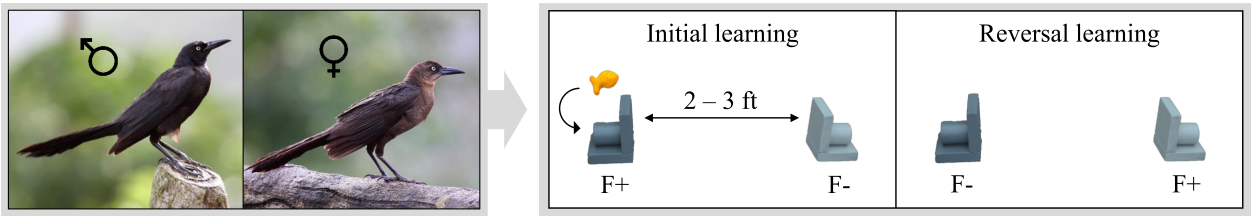

Figure 1 Left panel: images showing a male and female great-tailed grackle (credit: Wikimedia Commons). Right panel: schematic of the colour-reward reinforcement learning experimental protocol. In the *initial learning* phase, great-tailed grackles are presented with two colour-distinct tubes; however, 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:

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

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

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).

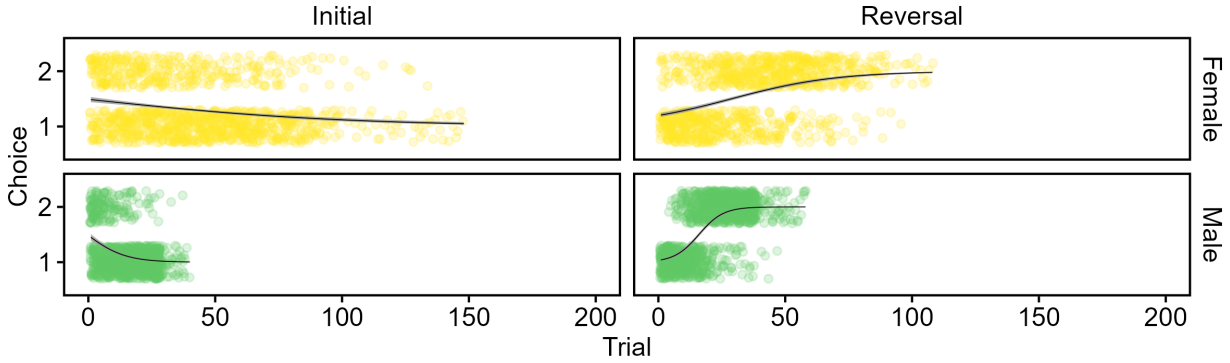

Figure 2 Group-level tube-choice behaviour of simulated great-tailed grackles across colour-reward reinforcement learning trials (females: yellow, $n = 14$; males: green, $n = 35$), following model validation step one. Tube option 1 (e.g., dark grey) was the rewarded option in the initial learning phase; 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

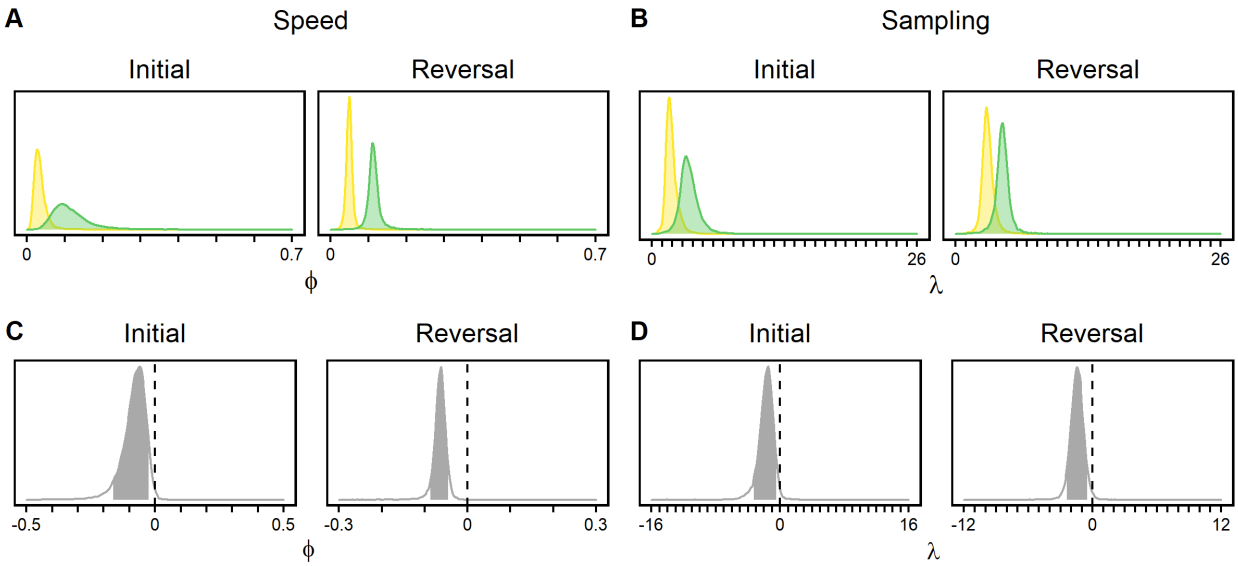

Figure 3 Comparison of learning ability in simulated female (yellow; $n = 14$) and male (green; $n = 35$) great-tailed grackles across initial and reversal colour-reward reinforcement learning, following model validation step two. (**A**) ϕ, the rate of learning i.e., speed. (**B**) λ, the rate of sampling 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.

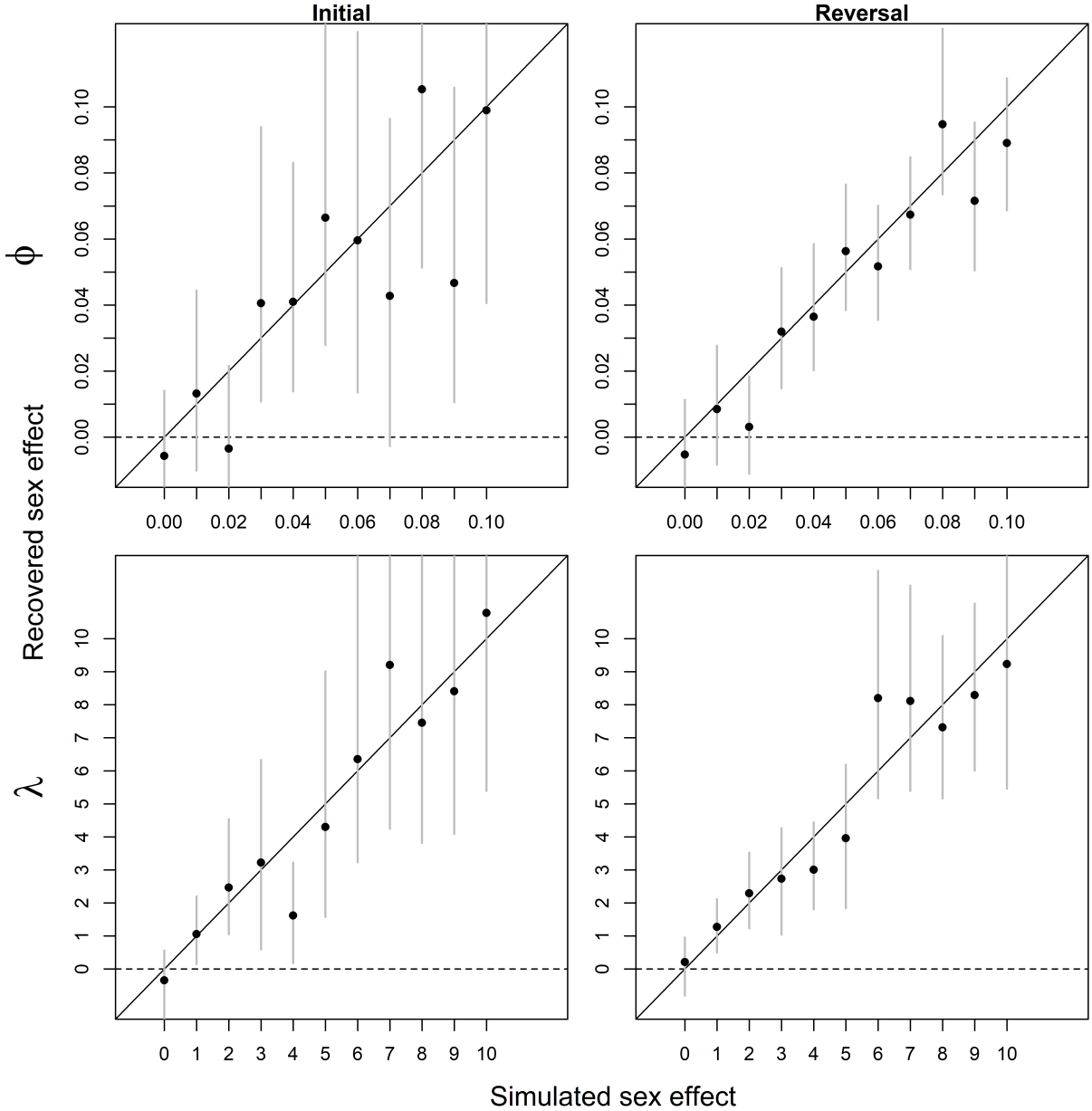

Figure 4 Parameter recovery test for different sizes of simulated sex differences. Plots show posterior estimates of the effect of sex (contrasts between simulated male and female great-tailed grackles; n = 14 and 35, respectively) on speed (ϕ) and sampling (λ) learning parameters, following model validation 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

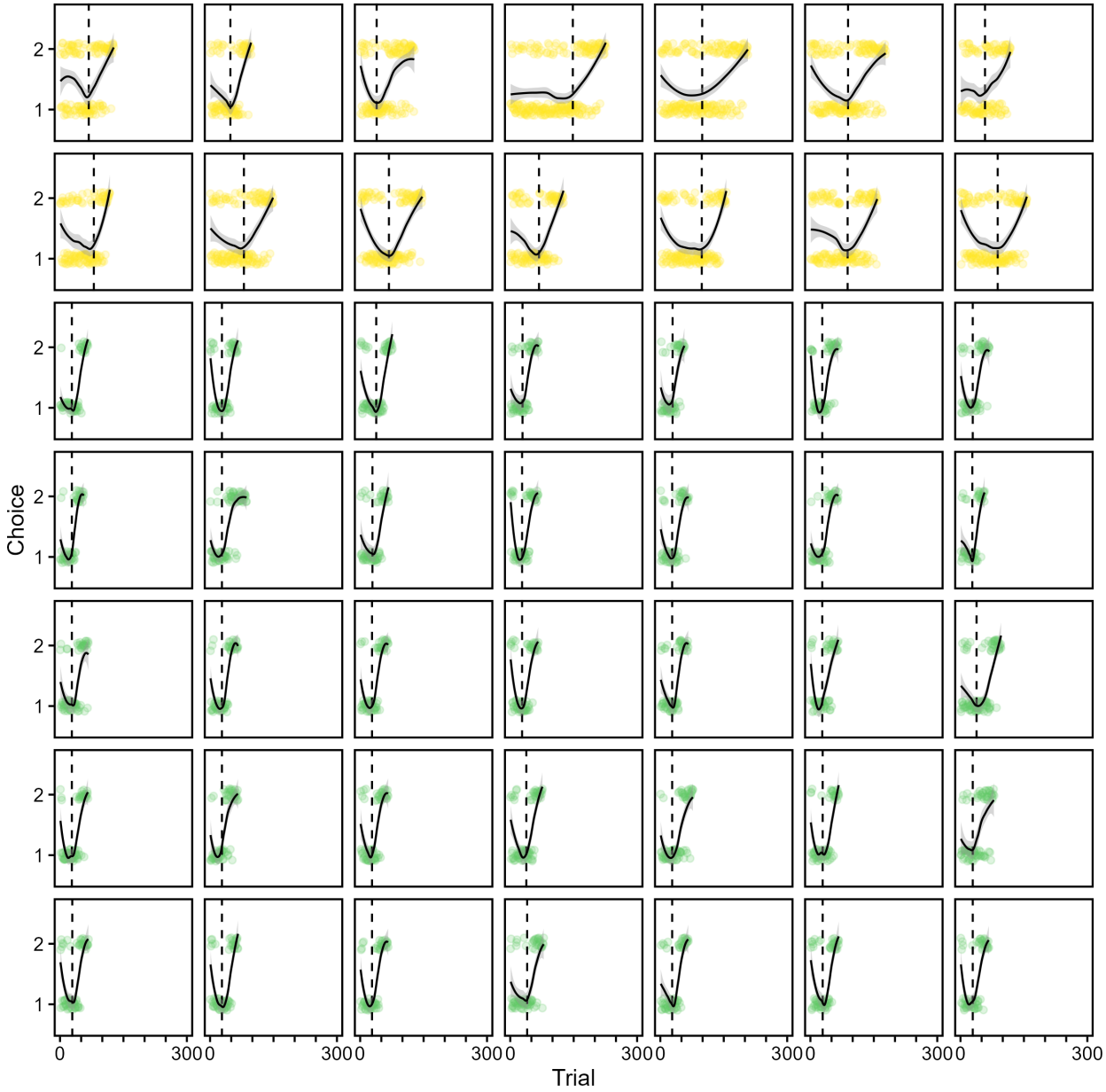

Figure S1 Individual-level tube-choice behaviour of simulated great-tailed grackles across colour-reward reinforcement learning trials (females: yellow, $n = 14$; males: green, $n = 35$). Tube option 1 (e.g., dark grey) was the rewarded option in the initial learning phase; conversely, tube option 2 (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.

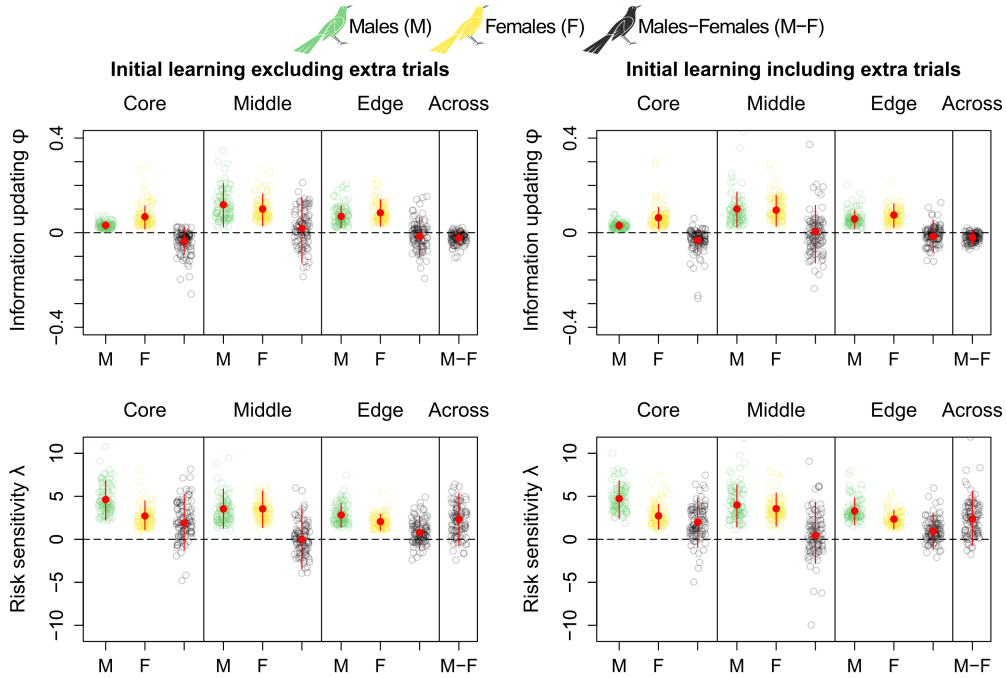

Figure 2—figure supplement 1. Comparison of information-updating rate ϕ and risk-sensitivity rate λ estimates (top and bottom row, respectively) in initial learning excluding and including extra initial learning trials (left and right column, respectively), which are present in the original data set (see Methods and materials). 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.