

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

Reviewed Preprint

Revised by authors after peer review.

[About eLife's process](#)

Reviewed preprint version 2

February 23, 2024 (this version)

Reviewed preprint version 1

October 31, 2023

Posted to preprint server

June 28, 2023

Sent for peer review

June 28, 2023

Alexis J Breen , Dominik Deffner 

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

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

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.

eLife assessment

This **important** study uses a multi-pronged empirical and theoretical approach to advance our understanding of animal cognition. It presents **convincing** data on how differences in learning relate to differences in the ways that male versus female animals cope with urban environments, and more generally how reversal learning may benefit animals in urban habitats.

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 (Ronc, 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 Thornton2021; Szabo et al.2020; Barrett et al. 2019; but see: Vincze and Kovács2022). 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 commensally with humans in urban environments (Urbanec 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 Fromhage2021a, 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; [Lusciur, 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 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.

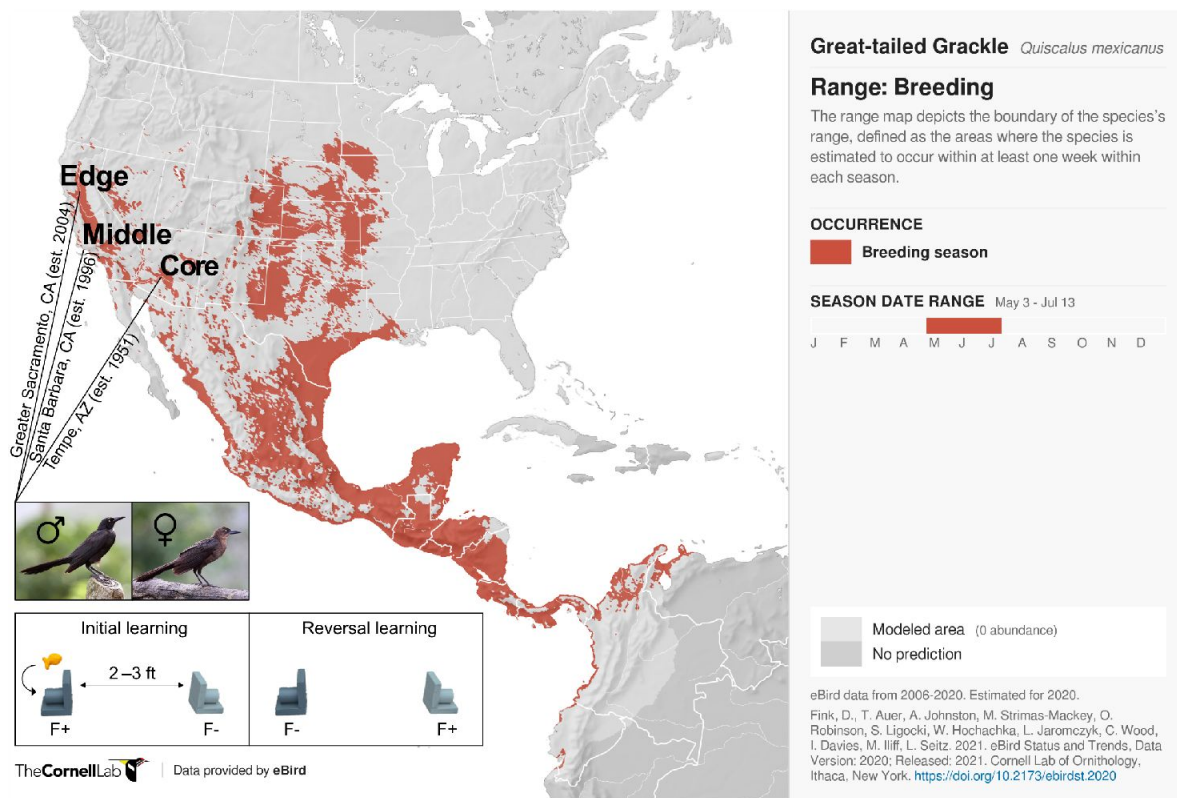


Figure 1.

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).

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** [2A](#) and Supplementary file 2a), and with comparable counts of choice-option switches (mean switches-at-finish: males, 14; females, 14; **Figure 2** [2B](#) and Supplementary file 2b). Indeed, the male-female (M-F) posterior contrasts for both behaviours centre around zero, evidencing no sex-effect (**Figure 2** [2C](#)). 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** [2B](#) and Supplementary file 2a), and by making fewer choice-option switches (mean switches-at-finish: males, 25; females, 36; **Figure 2** [2B](#) and Supplementary file 2b). The M-F posterior contrasts, which lie almost entirely below zero, clearly capture this sex-effect (**Figure 2** [2F](#) 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](#) [2](#)), 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](#) [2](#); but see also: [Farrell and Lewandowsky 2018](#) [2](#); [McElreath et al. 2005](#) [2](#); [Sutton and Barto 2018](#) [2](#)), which we validate *a priori* 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](#) [2](#)).

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** [2G](#) 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** [2G](#) 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** [2H](#) 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** [2G](#) and H and Supplementary file 2c).

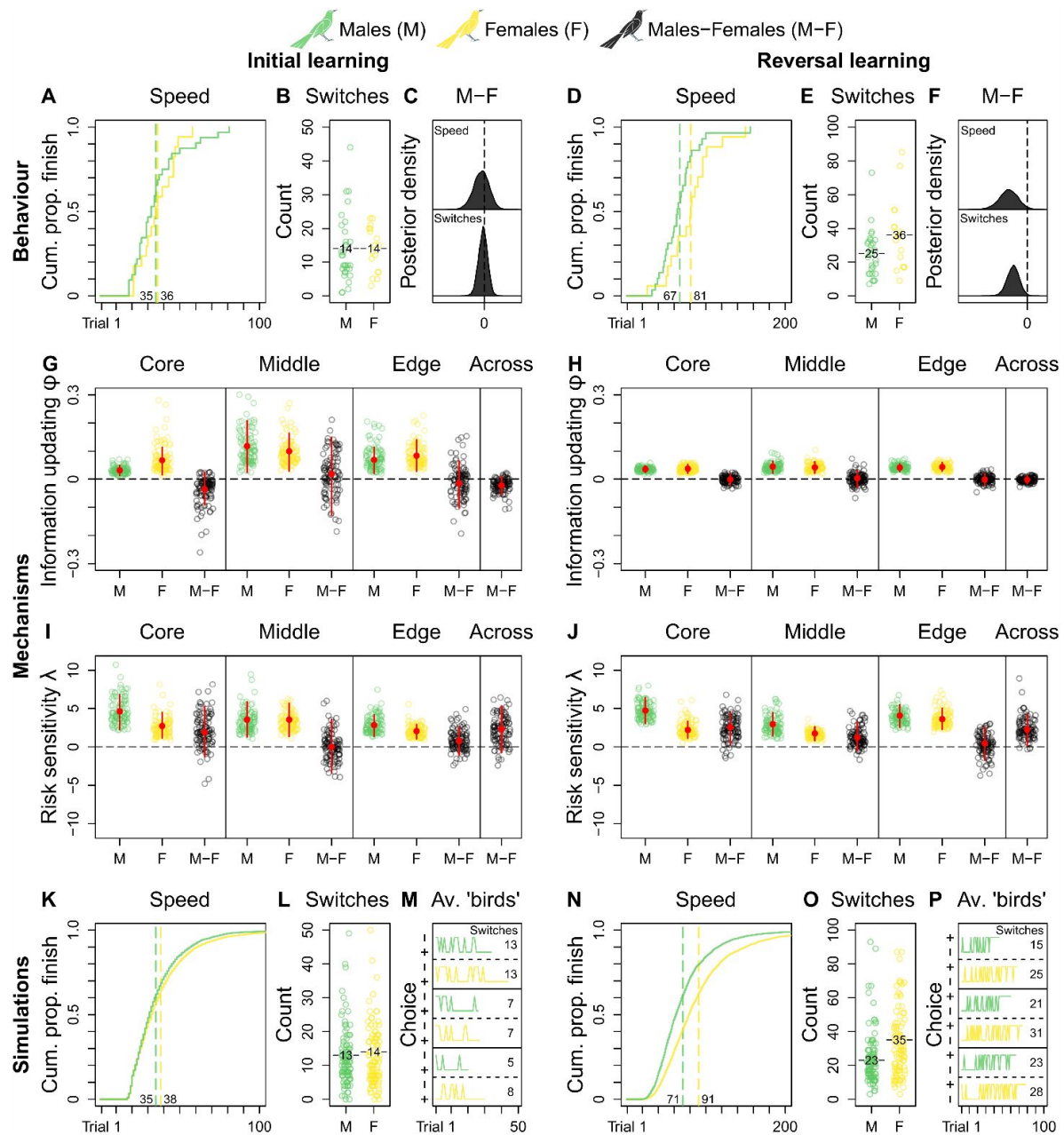


Figure 2.

Grackle reinforcement learning. *Behaviour*. Across-population reinforcement learning speed and choice-option switches in (A–B) initial learning. *Behaviour*. Across-population reinforcement learning speed and choice-option switches in (A–B) initial reversal learning (M, 29; F, 17), with (C, F) respective posterior estimates and M-F contrasts. *Mechanisms*. Within- and contrasts of information-updating rate ϕ and risk-sensitivity rate λ in (G, I) initial learning and (H, J) reversal w 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. tch (A, D) x-axes. Means are plotted/labelled in (A, B, D, E, K, L, N, O). Plots (G–P) are generated via model size (M, 32; F, 17).

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

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 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 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 [↗](#);

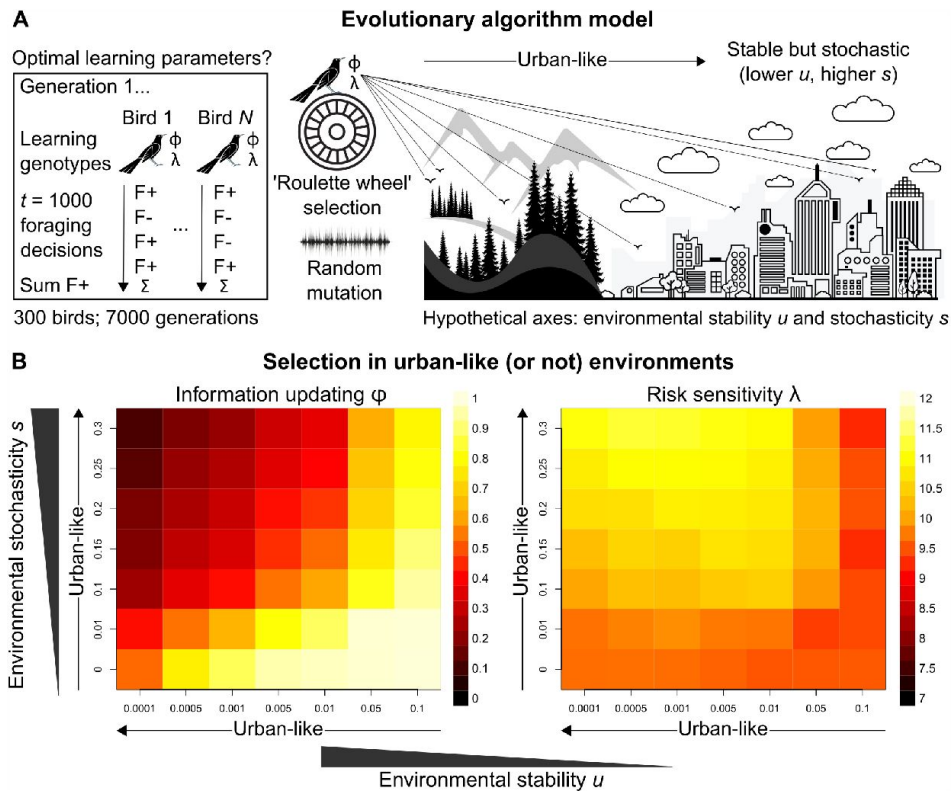


Figure 3.

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'.

*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 permitted 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 tubechoice, 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 trial 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 over-sight. 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 trial 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 trials ([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 [McElreath2018](#)). 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:

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

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

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 *a priori* 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 *a priori* 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.

Acknowledgements

We thank Jean-François Gerard and Rachel Harrison for useful feedback on our study before data analyses; and James St Clair, Sue Healy, and Richard McElreath for similarly useful presubmission feedback. We are further and most grateful to Richard McElreath for study and overall full support. And we thank all members, past and present, of the Grackle Project for collecting and making available, either via public archive (core and middle population) or permitted 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.

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-sensitivity 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.

Supplementary materials

Investigating sex differences in learning in a range-expanding bird

Alexis J. Breen^{1,*} & Dominik Deffner^{1,2,3}

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

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

³Center 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.

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 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	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 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	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 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	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)

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.

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 [↗](#); Urdvary & 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.

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

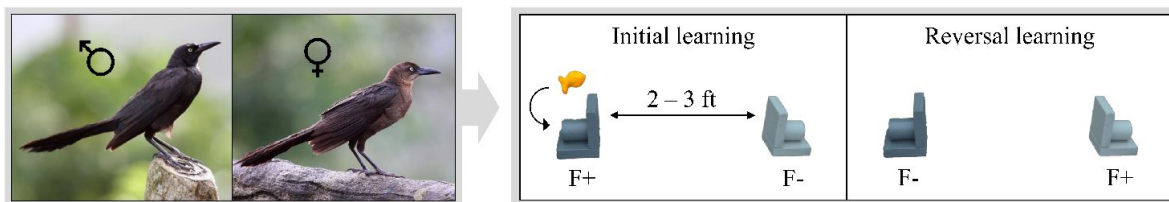


Figure 1.

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).

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:

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

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

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 similar 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).

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 grackles’ initial and reversal learning (it did; **Figure 3**).

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.

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.

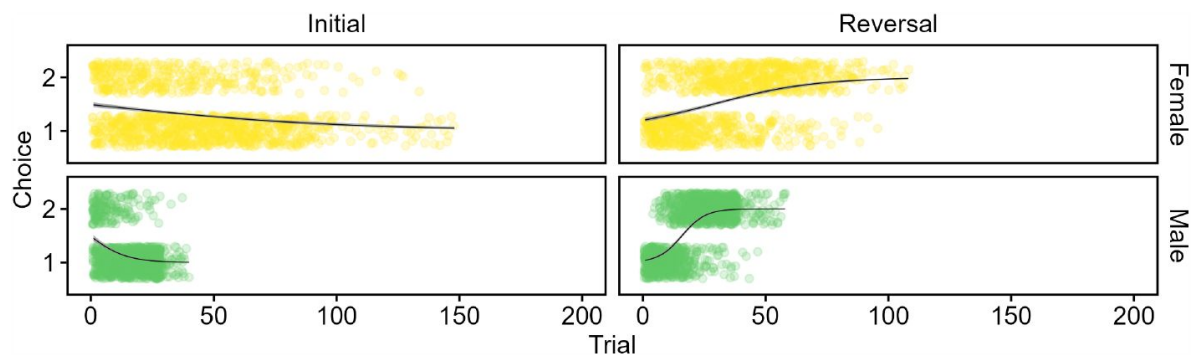


Figure 2.

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% compatibility intervals.

	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

Table 1

Comparison of assigned and recovered ϕ and λ values, following model validation step two. Eightynine percent highest posterior density intervals (HPDI) are shown for recovered values.

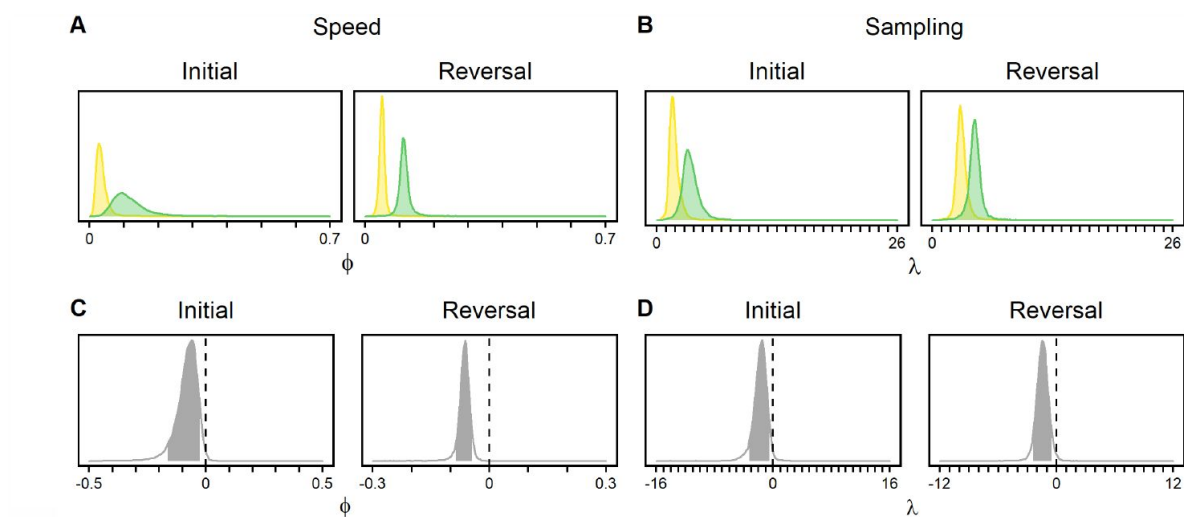
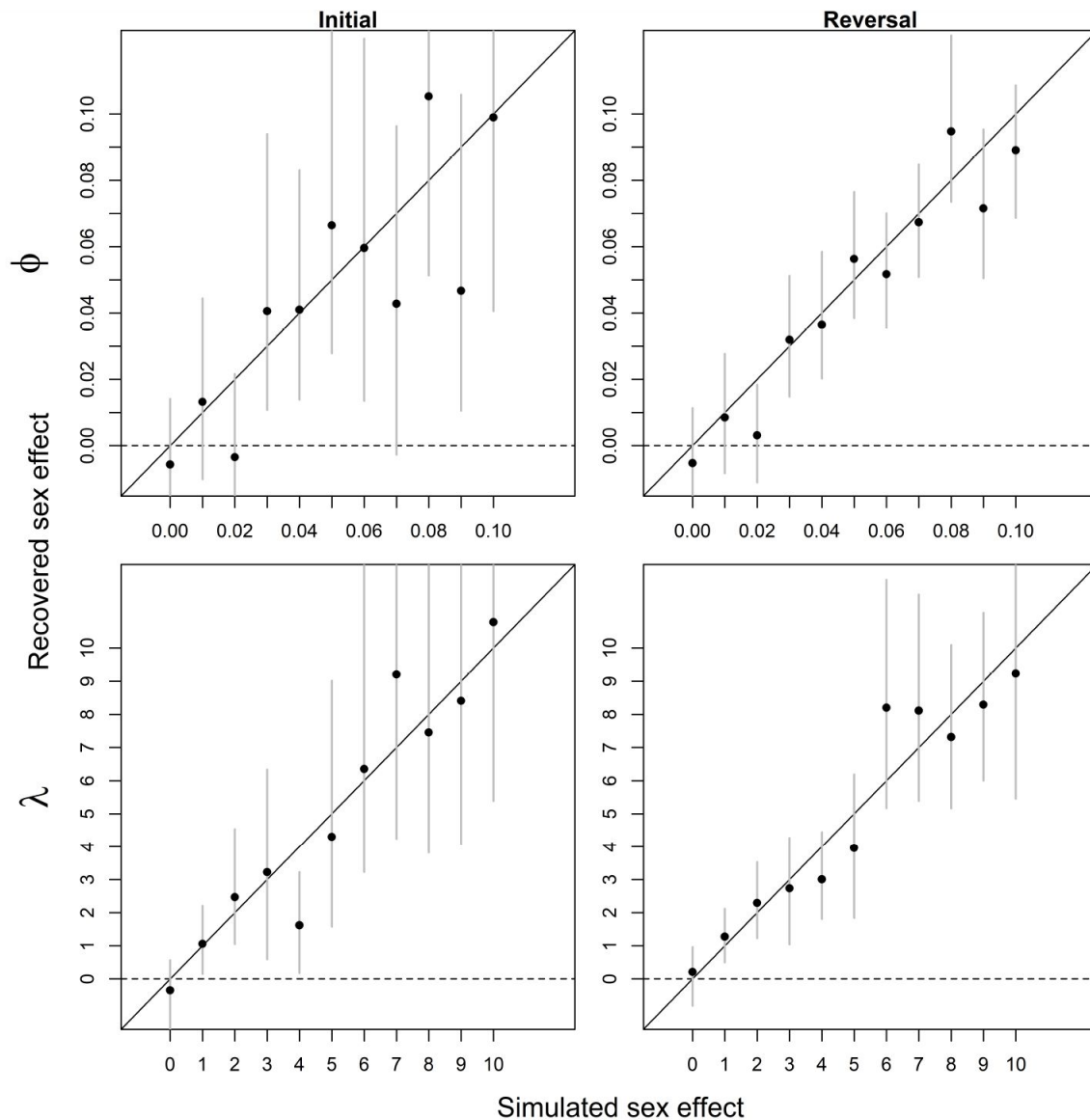


Figure 3.

Comparison of learning ability in simulated female (yellow; $n = 14$) and male (green; $n = 35$) greatailled 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 choiceoptions. **(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.



229

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.s

Supplementary material

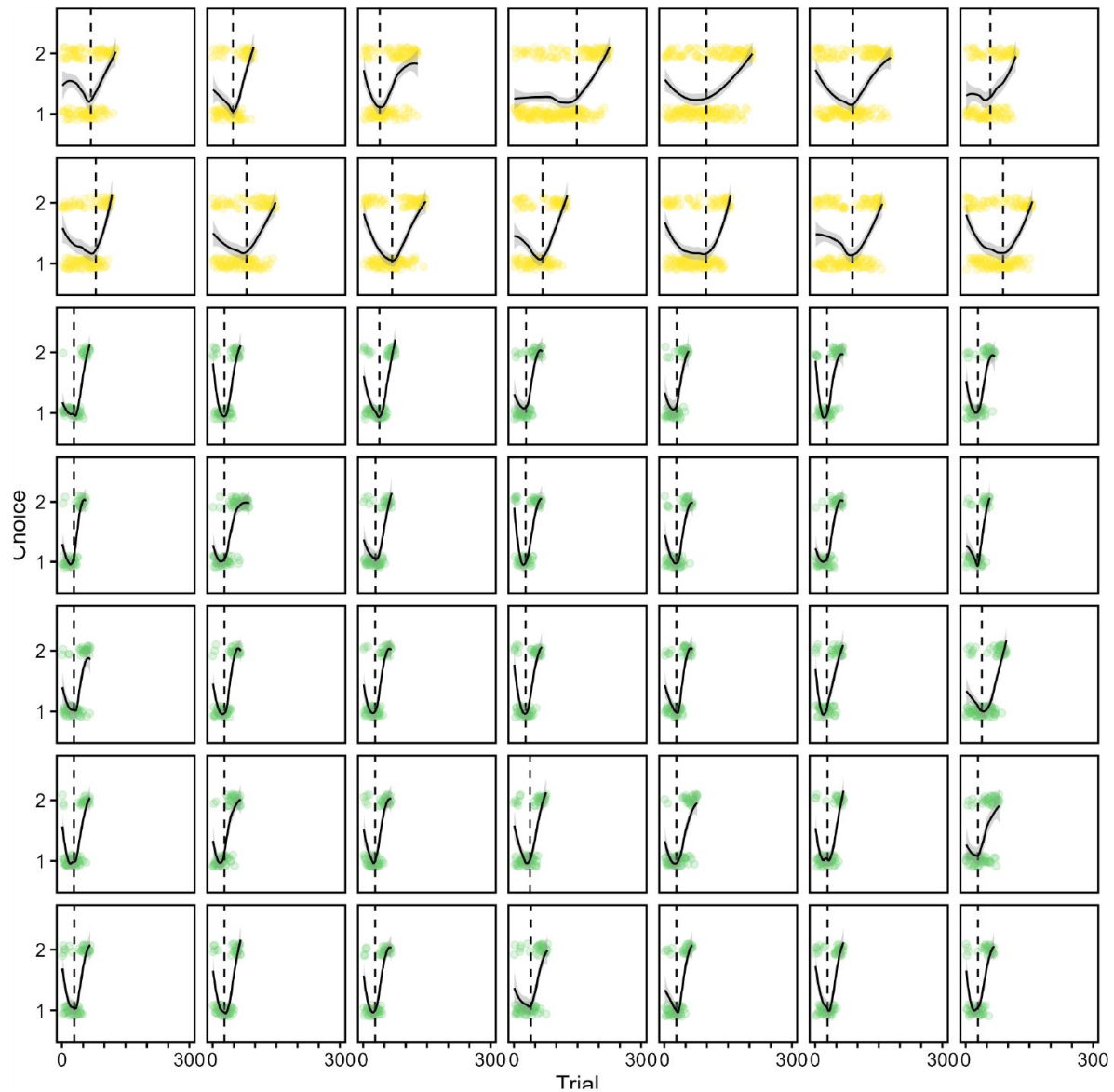


Figure S1.

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.

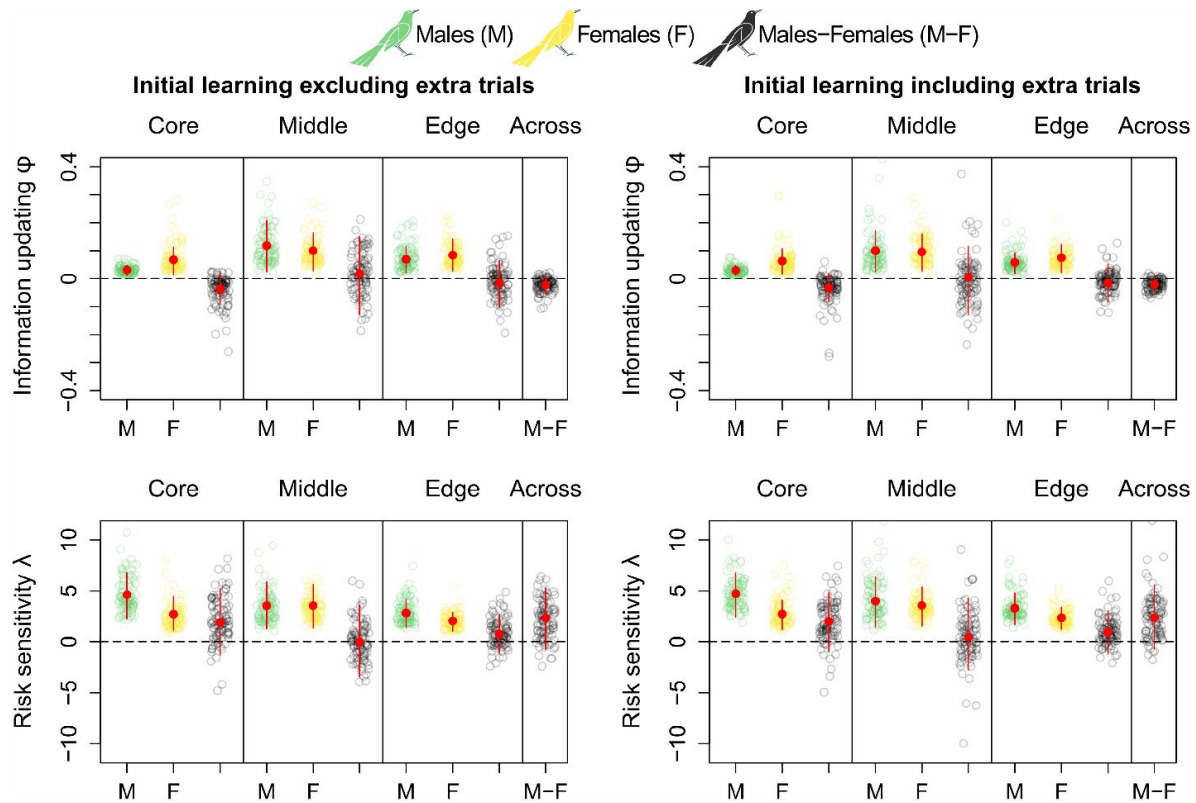


Figure S2—figure supplement 1

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.

References

- Aljadeff N, Lotem A. (2021) **Task-dependent reversal learning dynamics challenge the reversal paradigm of measuring cognitive flexibility** *Animal Behaviour* **179**:183–197 <https://doi.org/10.1016/j.anbehav.2021.07.002>
- Aplin LM, Sheldon BC, McElreath R. (2017) **Conformity does not perpetuate suboptimal traditions in a wild population of songbirds** *Proceedings of the National Academy of Sciences* **114**:7830–7837 <https://doi.org/10.1073/pnas.1621067114>
- Audet JN, Lefebvre L. (2017) **What’s flexible in behavioral flexibility?** *Behavioral Ecology* **28**:943–947 <https://doi.org/10.1093/beheco/arx007>
- Barrett BJ (2022) **Inferential power in identifying frequency-dependent social learning strengthened by increasing behavioural options** *Journal of Animal Ecology* <https://doi.org/10.1111/1365-2656.13826>
- Barrett BJ, McElreath RL, Perry SE (2017) **Pay-off-biased social learning underlies the diffusion of novel extractive foraging traditions in a wild primate** *Proceedings of the Royal Society B: Biological Sciences* **284** <https://doi.org/10.1098/rspb.2017.0358>
- Barrett LP, Stanton LA, Benson-Amram S. (2019) **The cognition of ‘nuisance’ species** *Animal Behaviour* **147**:167–177 <https://doi.org/10.1016/j.anbehav.2018.05.005>
- Batabyal A, Thaker M. (2019) **Lizards from suburban areas learn faster to stay safe** *Biology Letters* **15** <https://doi.org/10.1098/rsbl.2019.0009>
- Bateson M. (2002) **Recent advances in our understanding of risk-sensitive foraging preferences** *Proceedings of the Nutrition Society* **61**:509–516 <https://doi.org/10.1079/pns2002181>
- Bensky MK, Bell AM (2020) **Predictors of individual variation in reversal learning performance in three-spined sticklebacks** *Animal cognition* **23**:925–938 <https://doi.org/10.1007/s10071-020-01399-8>
- Blackburn TM, Lockwood JL, Cassey P. (2009) **Avian invasions: the ecology and evolution of exotic birds**
- Blaisdell A, Seitz B, Rowney C, Folsom M, MacPherson M, Deffner D, Logan CJ (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** <https://doi.org/10.24072/pcjournal.44>
- Blickley JL, Blackwood D, Patricelli GL (2012) **Experimental evidence for the effects of chronic anthropogenic noise on abundance of greater sage-grouse at leks** *Conservation Biology* **26**:461–471 <https://doi.org/10.1111/j.1523-1739.2012.01840.x>
- Bond AB, Kamil AC, Balda RP (2007) **Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus*, *Nucifraga columbiana*, *Aphelocoma californica*)** *Journal of Comparative Psychology* **121** <https://doi.org/10.1037/0735-7036.121.4.372>

- Boulougouris V, Dalley JW, Robbins TW (2007) **Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat** *Behavioural brain research* **179**:219–228 <https://doi.org/10.1016/j.bbr.2007.02.005>
- Breen AJ (2021) **Animal culture research should include avian nest construction** *Biology Letters* **17** <https://doi.org/10.1098/rsbl.2021.0327>
- Breen AJ, Sugasawa S, Healy SD (2021) **Manipulative and technological skills do not require a slow life history** *Frontiers in Ecology and Evolution* **9** <https://doi.org/10.3389/fevo.2021.635802>
- Bro-Jørgensen J, Franks DW, Meise K. (2019) **Linking behaviour to dynamics of populations and communities: application of novel approaches in behavioural ecology to conservation** *Philosophical Transactions of the Royal Society B* **374** <https://doi.org/10.1098/rstb.2019.0008>
- Burmeister SS (2022) **Ecology, cognition, and the hippocampus: A tale of two frogs** *Brain, Behavior and Evolution* **97**:211–224 <https://doi.org/10.1159/000522108>
- Cantwell A, Buckholtz JW, Atencia R, Rosati AG (2022) **The origins of cognitive flexibility in chimpanzees** *Developmental science* **25** <https://doi.org/10.1111/desc.13266>
- Caspi T, Johnson JR, Lambert MR, Schell CJ, Sih A. (2022) **Behavioral plasticity can facilitate evolution in urban environments** *Trends in Ecology & Evolution* **37**:1092–1103 <https://doi.org/10.1016/j.tree.2022.08.002>
- Chimento M, Barrett BJ, Kandler A, Aplin LM (2022) **Cultural diffusion dynamics depend on behavioural production rules** *Proceedings of the Royal Society B: Biological Sciences* **289** <https://doi.org/10.1098/rspb.2022.1001>
- Chuang A, Peterson CR (2016) **Expanding population edges: theories, traits, and trade-offs** *Global Change Biology* **22**:494–512 <https://doi.org/10.1111/gcb.13107>
- Coppens CM, de Boer SF, Koolhaas JM (2010) **Coping styles and behavioural flexibility: towards underlying mechanisms** *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**:4021–4028 <https://doi.org/10.1098/rstb.2010.0217>
- Cowie RJ, Hinsley SA (1988) **The provision of food and the use of bird feeders in suburban gardens** *Bird Study* **35**:163–168 <https://doi.org/10.1080/00063658809476985>
- Davies ZG, Fuller RA, Loram A, Irvine KN, Sims V, Gaston KJ (2009) **A national scale inventory of resource provision for biodiversity within domestic gardens** *Biological Conservation* **142**:761–771 <https://doi.org/10.1016/j.biocon.2008.12.016>
- Davis WR, Arnold KA (1972) **Food habits of the Great-tailed Grackle in Brazos county, Texas** *The Condor* **74**:439–446 <https://doi.org/10.2307/1365896>
- Deffner D, Kleinow V, McElreath R. (2020) **Dynamic social learning in temporally and spatially variable environments** *Royal Society Open Science* **7** <https://doi.org/10.1098/rsos.200734>
- Deffner D, Rohrer JM, McElreath R. (2022) **A causal framework for cross-cultural generalizability** *Advances in Methods and Practices in Psychological Science* **5**:1–18 <https://doi.org/10.1177/25152459221106366>

- Dinsmore JJ, Dinsmore SJ (1993) **Range expansion of the great-tailed grackle in the 1900s** *Journal of the Iowa Academy of Science* **100**:54–59
- Duckworth RA, Badyaev AV (2007) **Coupling of dispersal and aggression facilitates the rapid range expansion of a passerine bird** *Proceedings of the National Academy of Sciences* **104**:15017–15022 <https://doi.org/10.1073/pnas.0706174104>
- Dukas R. (1998) **Cognitive ecology: the evolutionary ecology of information processing and decision making**
- Eisenhauer N, Hines J. (2021) **Invertebrate biodiversity and conservation** *Current Biology* **31**:R1214–R1218 <https://doi.org/10.1016/j.cub.2021.06.058>
- Farrell S, Lewandowsky S. (2018) **Computational modeling of cognition and behavior** <https://doi.org/10.1017/CBO9781316272503>
- Fawcett TW, Hamblin S, Giraldeau LA (2013) **Exposing the behavioral gambit: the evolution of learning and decision rules** *Behavioral Ecology* **24**:2–11 <https://doi.org/10.1093/beheco/ars085>
- Federspiel IG, Garland A, Guez D, Bugnyar T, Healy SD, Güntürkün O, Griffin AS (2017) **Adjusting foraging strategies: a comparison of rural and urban common mynas (*Acridotheres tristis*)** *Animal cognition* **20**:65–74 <https://doi.org/10.1007/s10071-016-1045-7>
- Fernández-Juricic E. (2000) **Local and regional effects of pedestrians on forest birds in a fragmented landscape** *The condor* **102**:247–255 <https://doi.org/10.1093/condor/102.2.247>
- Fink D *et al.* (2020) **Fink D, Auer T, Johnston A, Strimas-Mackey M, Robinson O, Ligoeki 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.** <https://doi.org/10.2173/ebirdst.2020>
- Francis CD, Barber JR (2013) **A framework for understanding noise impacts on wildlife: an urgent conservation priority** *Frontiers in Ecology and the Environment* **11**:305–313 <https://doi.org/10.1890/120183>
- Frid A, Dill L. (2002) **Human-caused disturbance stimuli as a form of predation risk** *Conservation Ecology* **6**
- Fuller RA, Warren PH, Armsworth PR, Barbosa O, Gaston KJ (2008) **Garden bird feeding predicts the structure of urban avian assemblages** *Diversity and Distributions* **14**:131–137 <https://doi.org/10.1111/j.1472-4642.2007.00439.x>
- Gaalema DE (2011) **Visual discrimination and reversal learning in rough-necked monitor lizards (*Varanus rudicollis*)** *Journal of Comparative Psychology* **125** <https://doi.org/10.1037/a0023148>
- Gaston KJ, Bennie J, Davies TW, Hopkins J. (2013) **The ecological impacts of nighttime light pollution: a mechanistic appraisal** *Biological Reviews* **88**:912–927 <https://doi.org/10.1111/brv.12036>
- Grunst ML, Grunst AS, Pinxten R, Eens M. (2021) **Variable and consistent traffic noise negatively affect the sleep behavior of a free-living songbird** *Science of the Total Environment* **778** <https://doi.org/10.1016/j.scitotenv.2021.146338>

Hampton S (2004) **Yolo County birding news**

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

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

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

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

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

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

Li G *et al.* (2022) **Global impacts of future urban expansion on terrestrial vertebrate diversity** *Nature communications* **13**:1–12 <https://doi.org/10.1038/s41467-022-29324-2>

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

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

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

Logan CJ (2016) **Behavioral flexibility and problem solving in an invasive bird** *PeerJ* **4** <https://doi.org/10.7717/peerj.1975>

Logan CJ (2016) **Behavioral flexibility in an invasive bird is independent of other behaviors** *PeerJ* **4** <https://doi.org/10.7717/peerj.2215>

Logan CJ (2016) **Great-tailed grackle behavioral flexibility and problem solving experiments, Santa Barbara, CA USA 2014-2015** *Knowledge Network for Biocomplexity* <https://doi.org/10.5063/F1319SVV>

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

- Logan CJ, McCune K, LeGrande-Rolls C, Marfori Z, Hubbard J, Lukas D. (2023) **Data: Implementing a rapid geographic range expansion - the role of behavior changes** *Knowledge Network for Biocomplexity* <https://doi.org/10.5063/F1QZ28FH>
- Logan C, Lukas D, Blaisdell A, Johnson-Ulrich Z, MacPherson M, Seitz B, Sevchik A, McCune K. (2023) **Behavioral flexibility is manipulable and it improves flexibility and innovativeness in a new context** *Peer Community Journal* **3** <https://doi.org/10.24072/pcjournal.284>
- Logan C, McCune K, LeGrande-Rolls C, Marfori Z, Hubbard J, Lukas D. (2023) **Implementing a rapid geographic range expansion-the role of behavior changes** *Peer Community Journal* **3** <https://doi.org/10.24072/pcjournal.320>
- Logan CJ, McCune K, MacPherson M, Johnson-Ulrich Z, Rowney C, Seitz B, Blaisdell A, Deffner D, Wascher C. (2022) **Are the more flexible great-tailed grackles also better at behavioral inhibition?** *Animal Behavior Cognition* **9**:14–36 <https://doi.org/10.26451/abc.09.01.03.2022>
- Lucon-Xiccato T, Bisazza A. (2014) **Discrimination reversal learning reveals greater female behavioural flexibility in guppies** *Biology Letters* **10** <https://doi.org/10.1098/rsbl.2014.0206>
- Lukas D, McCune K, Blaisdell AP, Johnson-Ulrich Z, MacPherson M, Seitz B, Sevchik A, Logan CJ (2023) **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* <https://doi.org/10.32942/osf.io/5z8xs>
- Luscier JD (2018) **Effects of urbanization on great-tailed grackle habitat use and nest success in Sherman Texas** *Urban Naturalist* **15**:1–14
- Mazza V, Guenther A. (2021) **City mice and country mice: innovative problem solving in rural and urban noncommensal rodents** *Animal Behaviour* **172**:197–210 <https://doi.org/10.1016/j.anbehav.2020.12.007>
- McElreath R. (2018) **Statistical rethinking: A Bayesian course with examples in R and Stan**
- McElreath R, Lubell M, Richerson PJ, Waring TM, Baum W, Edsten E, Efferson C, Paciotti B. (2005) **Applying evolutionary models to the laboratory study of social learning** *Evolution and Human Behavior* **26**:483–508 <https://doi.org/10.1016/j.evolhumbehav.2005.04.003>
- McNamara JM, Houston AI (2009) **Integrating function and mechanism** *Trends in Ecology & Evolution* **24**:670–675 <https://doi.org/10.1016/j.tree.2009.05.011>
- Miller TEX, Inouye BD (2013) **Sex and stochasticity affect range expansion of experimental invasions** *Ecology Letters* **16**:354–361 <https://doi.org/10.1111/ele.12049>
- Miller TE, Shaw AK, Inouye BD, Neubert MG (2011) **Sex-biased dispersal and the speed of two-sex invasions** *The American Naturalist* **177**:549–561 <https://doi.org/10.1086/659628>
- Morand-Ferron J, Reichert MS, Quinn JL (2022) **Cognitive flexibility in the wild: individual differences in reversal learning are explained primarily by proactive interference, not by sampling strategies, in two passerine bird species** *Learning & Behavior* **50**:153–166 <https://doi.org/10.3758/s13420-021-00505-1>
- Otto SP, Day T. (2011) **A biologist’s guide to mathematical modeling in ecology and evolution**

- Phillips BL, Brown GP, Shine R. (2010) **Life-history evolution in range-shifting populations** *Ecology* **91**:1617–1627 <https://doi.org/10.1890/09-0910.1>
- Raine NE, Chittka L. (2012) **No trade-off between learning speed and associative flexibility in bumblebees: a reversal learning test with multiple colonies** *PloS one* **7** <https://doi.org/10.1371/journal.pone.0045096>
- Rochais C, Hotte H, Pillay N. (2021) **Seasonal variation in reversal learning reveals greater female cognitive flexibility in African striped mice** *Scientific Reports* **11** <https://doi.org/10.1038/s41598-021-99619-9>
- Rodrigo A, Ávila-Chauvet L, Buriticá J. (2021) **Daily Patterns of Foraging and Aggressive Behaviors in Great-tailed Grackle (*Quiscalus mexicanus*) at an Urban Patch with Availability or Absence of Resources** *bioRxiv* :2021–6 <https://doi.org/10.1101/2021.06.14.448443>
- 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>
- Selander RK, Giller DR (1961) **Analysis of sympatry of great-tailed and boat-tailed grackles** *The Condor* **63**:29–86 <https://doi.org/10.2307/1365420>
- Sevchik A, Logan CJ, McCune KB, Blackwell A, Rowney C, Lukas D. (2022) **Investigating sex differences in genetic relatedness in great-tailed grackles in Tempe, Arizona to infer potential sex biases in dispersal** *Animal Behavior and Cognition* **9**:37–52 <https://doi.org/10.32942/osf.io/t6beh>
- Shochat E, Warren PS, Faeth SH, McIntyre NE, Hope D. (2006) **From patterns to emerging processes in mechanistic urban ecology** *Trends in Ecology & Evolution* **21**:186–191 <https://doi.org/10.1016/j.tree.2005.11.019>
- Smolla M, Rosher C, Gilman RT, Shultz S. (2019) **Reproductive skew affects social information use** *Royal Society open science* **6** <https://doi.org/10.1098/rsos.182084>
- Sol D, Lapiedra O, González-Lagos C. (2013) **Behavioural adjustments for a life in the city** *Animal Behaviour* **85**:1101–1112 <https://doi.org/10.1016/j.anbehav.2013.01.023>
- Somveille M, Firth JA, Aplin LM, Farine DR, Sheldon BC, Thompson RN (2018) **Movement and conformity interact to establish local behavioural traditions in animal populations** *PLoS Computational Biology* **14** <https://doi.org/10.1371/journal.pcbi.1006647>
- Strang CG, Sherry DF (2014) **Serial reversal learning in bumblebees (*Bombus impatiens*)** *Animal Cognition* **17**:723–734 <https://doi.org/10.1007/s10071-013-0704-1>
- Summers J, Lukas D, Logan CJ, Chen N. (2023) **The role of climate change and niche shifts in divergent range dynamics of a sister-species pair** *Peer Community Journal* **3** <https://doi.org/10.24072/pcjournal.248>
- Sutter M, Kawecki T. (2009) **Influence of learning on range expansion and adaptation to novel habitats** *Journal of Evolutionary Biology* **22**:2201–2214 <https://doi.org/10.1111/j.1420-9101.2009.01836.x>
- Sutton RS, Barto AG (2018) **Reinforcement learning: An introduction**

- Szabo B, Damas-Moreira I, Whiting MJ (2020) **Can cognitive ability give invasive species the means to succeed? A review of the evidence** *Frontiers in ecology and evolution* <https://doi.org/10.3389/fevo.2020.00187>
- Team RC (2021) **R: A Language and environment for statistical computing**
- Team SD (2020) **RStan: the R interface to Stan**
- Trochet A, Courtois EA, Stevens VM, Baguette M, Chaine A, Schmeller DS, Clobert J, Wiens JJ (2016) **Evolution of sexbiased dispersal** *The Quarterly Review of Biology* **91**:297–320 <https://doi.org/10.1086/688097>
- Trump AN, Deffner D, Pleskac TJ, Romanczuk P, Kurvers R. (2023) **A cognitive computational approach to social and collective decision-making** *Perspectives on Psychological Science* **0** <https://doi.org/10.1177/17456916231186964>
- Vincze E, Kovács B. (2022) **Urbanization's Effects on Problem Solving Abilities: A Meta-Analysis** *Frontiers in Ecology and Evolution* **10** <https://doi.org/10.3389/fevo.2022.834436>
- Vinton AC, Gascoigne SJ, Sepil I, Salguero-Gómez R. (2022) **Plasticity's role in adaptive evolution depends on environmental change components** *Trends in Ecology & Evolution* <https://doi.org/10.1016/j.tree.2022.08.008>
- Urbanec L, Matijević V, Guenther A. (2021) **Enhanced problem-solving ability as an adaptation to urban environments in house mice** *Proceedings of the Royal Society B* **288** <https://doi.org/10.1098/rspb.2020.2504>
- Webster MM, Rutz C. (2020) **How STRANGE are your study animals?** *Nature* **582**:337–340 <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**:1593–1607 <https://doi.org/10.1046/j.1365-2699.2003.00970.x>
- Wehtje W. (2004) **The great-tailed grackle (Quiscalus mexicanus Gmelin) in the Western USA: Range Expansion and Secondary Contact Between Subspecies**
- Wickham H *et al.* (2019) **Welcome to the Tidyverse** *Journal of Open Source Software* **4** <https://doi.org/10.21105/joss.01686>
- Wright TF, Eberhard JR, Hobson EA, Avery ML, Russello MA (2010) **Behavioral flexibility and species invasions: the adaptive flexibility hypothesis** *Ethology Ecology & Evolution* **22**:393–404 <https://doi.org/10.1080/03949370.2010.505580>
- Yu X, Gen M. (2010) **Introduction to evolutionary algorithms**
- 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** <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** <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**
- 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** <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**: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** <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**:54–59
- Duckworth R. A. (2006) **Behavioral correlations across breeding contexts provide a mechanism for a cost of aggression** *Behavioral Ecology* **17**: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: Slowexploring birds are more sensitive to changes in a vocal discrimination task** *Proceedings of the Royal Society B: Biological Sciences* **278**: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**: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** <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** <https://doi.org/10.3389/fevo.2021.633947>
- Lehman P. E. (2020) **Lehman, P. E. (2020). The birds of Santa Barbara County, California (2nd ed.).** <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. (2016) **Behavioral flexibility and problem solving in an invasive bird** *PeerJ* **4** <https://doi.org/10.7717/peerj.1975>
- Logan C. J. (2016) **Behavioral flexibility in an invasive bird is independent of other behaviors** *PeerJ* **4** <https://doi.org/10.7717/peerj.2215>
- McElreath R. (2018) **Statistical rethinking: A Bayesian course with examples in R and Stan**
- McElreath R., Smaldino P. E. (2015) **Replication, communication, and the population dynamics of scientific discovery** *PloS One* **10** <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**: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 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. **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**:1101–1112 <https://doi.org/10.1016/j.anbehav.2013.01.023>
- Stan Development Team (2020) **Stan Development Team. (2020). RStan: The R interface to Stan (Version 2.21.2) [Computer software].** <http://mc-stan.org>
- Stepney P. H. R. (1975) **First recorded breeding of the great-tailed grackle in Colorado** *The Condor* **77**: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**: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**: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]**
- Webster M. M., Rutz C. (2020) **How STRANGE are your study animals?** *In Nature* :337–340 <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**: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**:162–169 <https://doi.org/10.1016/j.applanim.2010.11.010>
- Wickham H. *et al.* (2019) **Welcome to the Tidyverse** *Journal of Open Source Software* **4** <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**:393–404 <https://doi.org/10.1080/03949370.2010.505580>

Article and author information

Alexis J Breen

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

For correspondence: alexis_breen@eva.mpg.de

ORCID iD: [0000-0002-2331-0920](https://orcid.org/0000-0002-2331-0920)

Dominik Deffner

Science of Intelligence Excellence Cluster, Technical University Berlin, 10623 Berlin, Germany, Center for Adaptive, Rationality, Max Planck Institute for Human Development, 14195 Berlin, Germany

For correspondence: deffner@mpib-berlin.mpg.de

ORCID iD: [0000-0002-1649-3861](https://orcid.org/0000-0002-1649-3861)

Copyright

© 2023, Alexis J Breen & Dominik Deffner

This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Editors

Reviewing Editor

Rosalyn Gloag

University of Sydney, Sydney, Australia

Senior Editor

Michael Frank

Brown University, Providence, United States of America

Reviewer #2 (Public Review):

Summary: The study is titled "Leading an urban invasion: risk-sensitive learning is a winning strategy", and consists of three different parts. First, the authors analyse data on initial and reversal learning in Grackles confronted with a foraging task, derived from three populations labeled as "core", "middle" and "edge" in relation to the invasion front. The suggested difference between study populations does not surface, but the authors do find support for a difference between male and female individuals. Secondly, the authors confirm that the proposed mechanism can actually generate patterns such as observed in the Grackle data through agent-based forward simulations. In the third part, the authors present an evolutionary model, in which they show that learning strategies, as observed in male Grackles, do evolve in simplified urban conditions including different levels of environmental stability and environmental stochasticity.

Strengths: The manuscript's strength is that it combines real learning data collected across different populations of the Great-tailed grackle (*Quiscalus mexicanus*) with theoretical approaches to better understand the processes with which grackles learn and how such learning processes might be advantageous during range expansion and invasion. Furthermore, the authors also take sex into account revealing that males, the dispersing sex, show better reversal learning through higher reward-payoff sensitivity. I also find it refreshing to see that the authors took the time to preregister their study to improve transparency especially regarding data analysis.

Weakness: The small sample size of grackles across populations increases uncertainty as to parameter estimates and the conclusions drawn from these estimates.

After revision, the introduction is appropriate, and in the methods, the authors take great care in explaining the rational behind decisions as to the selection of analysis methods and parameters. I very much appreciate that the authors took such care in revising their paper, the quality of which has now greatly improved.

<https://doi.org/10.7554/eLife.89315.2.sa0>

Author Response

The following is the authors' response to the original reviews.

eLife assessment

This study uses a multi-pronged empirical and theoretical approach to advance our understanding of how differences in learning relate to differences in the ways that male versus female animals cope with urban environments, and more generally how reversal learning may benefit animals in urban habitats. The work makes an important contribution and parts of the data and analyses are solid, although several of the main claims are only partially supported or overstated and require additional support.

We thank the Editor and both Reviewers for their time and for their constructive evaluation of our manuscript. We worked to address each comment and suggestion offered by the Reviewers in our revision—please see our point-by-point responses below.

Public Reviews:

We thank the Editor and both Reviewers for their time and for their constructive evaluation of our manuscript. We worked to address each comment and suggestion offered by the Reviewers in our revision—please see our point-by-point responses below.

Reviewer #1 (Public Review):

Summary:

In this highly ambitious paper, Breen and Deffner used a multi-pronged approach to generate novel insights on how differences between male and female birds in their learning strategies might relate to patterns of invasion and spread into new geographic and urban areas.

The empirical results, drawn from data available in online archives, showed that while males and females are similar in their initial efficiency of learning a standard color-food association (e.g., color X = food; color Y = no food) scenario when the associations are switched (now, color Y = food, X = no food), males are more efficient than females at adjusting to the new situation (i.e., faster at 'reversal learning'). Clearly, if animals live in an unstable world, where associations between cues (e.g., color) and what is good versus bad might change unpredictably, it is important to be good at reversal learning. In these grackles, males tend to disperse into new areas before females. It is thus fascinating that males appear to be better than females at reversal learning. Importantly, to gain a better understanding of underlying learning mechanisms, the authors use a Bayesian learning model to assess the relative role of two mechanisms (each governed by a single parameter) that might contribute to differences in learning. They find that what they term 'risk sensitive' learning is the key to explaining the differences in reversal learning. Males tend to exhibit higher risk sensitivity which explains their faster reversal learning. The authors then tested the validity of their empirical results by running agent-based simulations where 10,000 computer-simulated 'birds' were asked to make feeding choices using the learning parameters estimated from real birds. Perhaps not surprisingly, the computer birds exhibited learning patterns that were strikingly similar to the real birds. Finally, the authors ran evolutionary algorithms that simulate evolution by natural selection where the key traits that can evolve are the two learning parameters. They find that under conditions that might be common in urban environments, high-risk sensitivity is indeed favored.

Strengths:

The paper addresses a critically important issue in the modern world. Clearly, some organisms (some species, some individuals) are adjusting well and thriving in the modern, human-altered world, while others are doing poorly. Understanding how organisms cope with human-induced environmental change, and why some are particularly good at adjusting to change is thus an important question.

The comparison of male versus female reversal learning across three populations that differ in years since they were first invaded by grackles is one of few, perhaps the first in any species, to address this important issue experimentally.

Using a combination of experimental results, statistical simulations, and evolutionary modeling is a powerful method for elucidating novel insights.

Thank you—we are delighted to receive this positive feedback, especially regarding the inferential power of our analytical approach.

Weaknesses:

The match between the broader conceptual background involving range expansion, urbanization, and sex-biased dispersal and learning, and the actual comparison of three urban populations along a range expansion gradient was somewhat confusing. The fact that three populations were compared along a range expansion gradient implies an expectation that they might differ because they are at very different points in a range expansion. Indeed, the predicted differences between males and females are largely couched in terms of population differences based on their 'location' along the range expansion gradient. However, the fact that they are all urban areas suggests that one might not expect the populations to differ. In addition, the evolutionary model suggests that all animals, male or female, living in urban environments (that the authors suggest are stable but unpredictable) should exhibit high-risk sensitivity. Given that all grackles, male and female, in all populations, are both living in urban environments and likely come from an urban background, should males and females differ in their learning behavior? Clarification would be useful.

Thank you for highlighting a gap in clarity in our conceptual framework. To answer the Reviewer's question—yes, even with this shared urban 'history', it seems plausible that males and females could differ in their learning. For example, 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 male-biased dispersal system, as we discuss in L254–265 (now L295–306). 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 [DOI: 10.1101/2021.06.14.448443]). We now introduce this important conceptual information—please see L89–96.

Reinforcement learning mechanisms:

Although the authors' title, abstract, and conclusions emphasize the importance of variation in 'risk sensitivity', most readers in this field will very possibly misunderstand what this means biologically. Both the authors' use of the term 'risk sensitivity' and their statistical methods for measuring this concept have potential problems.

Please see our below responses concerning our risk-sensitivity term.

First, most behavioral ecologists think of risk as predation risk which is not considered in this paper. Secondly, some might think of risk as uncertainty. Here, as discussed in more detail below, the 'risk sensitivity' parameter basically influences how strongly an option's attractiveness affects the animal's choice of that option. They say that this is in line with foraging theory (Stephens and Krebs 2019) where sensitivity means seeking higher expected payoffs based on prior experience. To me, this sounds like 'reward sensitivity', but not what most think of as 'risk sensitivity'. This problem can be easily fixed by changing the name of the term.

We apologise for not clearly introducing the field of risk-sensitive foraging, 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 [DOI: 10.1079/PNS2002181]; Kacelnik & Bateson 1996 [DOI: 10.1093/ICB/36.4.402]). We have added this information to our manuscript in L494–497. We further apologise for not clearly explaining how our lambda parameter estimates such

risk-sensitive foraging. To do so here, we need to consider our Bayesian reinforcement learning model in full. This model uses observed choice-behaviour during reinforcement learning to infer our ϕ (information-updating) and λ (risk sensitivity) learning parameters. Thus, payoffs incurred through choice simultaneously influence estimation of each learning parameter—that is, in a sense, they are *both* sensitive to rewards. But ϕ and λ differentially direct any reward sensitivity back on choice behaviour due to their distinct definitions. Glossing over the mathematics, for ϕ , stronger reward sensitivity (bigger ϕ 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 λ 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’. We hope this information, which we have incorporated into our revised manuscript (please see L153–161), clarifies the rationale and mechanics of our reinforcement learning model, and why λ measures risk-sensitivity.

In addition, however, the parameter does not measure sensitivity to rewards per se - rewards are not in equation 2. As noted above, instead, equation 2 addresses the sensitivity of choice to the attraction score which can be sensitive to rewards, though in complex ways depending on the updating parameter. Second, equations 1 and 2 involve one specific assumption about how sensitivity to rewards vs. to attraction influences the probability of choosing an option. In essence, the authors split the translation from rewards to behavioral choices into 2 steps. Step 1 is how strongly rewards influence an option's attractiveness and step 2 is how strongly attractiveness influences the actual choice to use that option. The equation for step 1 is linear whereas the equation for step 2 has an exponential component. Whether a relationship is linear or exponential can clearly have a major effect on how parameter values influence outcomes. Is there a justification for the form of these equations? The analyses suggest that the exponential component provides a better explanation than the linear component for the difference between males and females in the sequence of choices made by birds, but translating that to the concepts of information updating versus reward sensitivity is unclear. As noted above, the authors' equation for reward sensitivity does not actually include rewards explicitly, but instead only responds to rewards if the rewards influence attraction scores. The more strongly recent rewards drive an update of attraction scores, the more strongly they also influence food choices. While this is intuitively reasonable, I am skeptical about the authors' biological/cognitive conclusions that are couched in terms of words (updating rate and risk sensitivity) that readers will likely interpret as concepts that, in my view, do not actually concur with what the models and analyses address.

To answer the Reviewer's question—yes, these equations are very much standard and the canonical way of analysing individual reinforcement learning (see: Ch. 15.2 in Computational Modeling of Cognition and Behavior by Farrell & Lewandowsky 2018 [DOI: 10.1017/CBO9781316272503]; McElreath et al. 2008 [DOI: 10.1098/rstb/2008/0131]; Reinforcement Learning by Sutton & Barto 2018). To provide a “justification for the form of these equations”, equation 1 describes a convex combination of previous values and recent payoffs. Latent values are updated as a linear combination of both factors, there is no simple linear mapping between payoffs and behaviour as suggested by the reviewer. Equation 2 describes the standard softmax link function. It converts a vector of real numbers (here latent values) into a simplex vector (i.e., a vector summing to 1) which represents the probabilities of different outcomes. Similar to the logit link in logistic regression, the softmax simply maps the model space of latent values onto the outcome space of choice probabilities which enter the categorical likelihood distribution. We can appreciate how we did not make this clear in our manuscript by not highlighting the standard nature of our analytical

approach—we now do so in our revised manuscript (please see L148–149). As far as what our reinforcement learning model measures, and how it relates cognition and behaviour, please see our previous response.

To emphasize, while the authors imply that their analyses separate the updating rate from 'risk sensitivity', both the 'updating parameter' and the 'risk sensitivity' parameter influence both the strength of updating and the sensitivity to reward payoffs in the sense of altering the tendency to prefer an option based on recent experience with payoffs. As noted in the previous paragraph, the main difference between the two parameters is whether they relate to behaviour linearly versus with an exponential component.

Please see our two earlier responses on the mechanics of our reinforcement learning model.

Overall, while the statistical analyses based on equations (1) and (2) seem to have identified something interesting about two steps underlying learning patterns, to maximize the valuable conceptual impact that these analyses have for the field, more thinking is required to better understand the biological meaning of how these two parameters relate to observed behaviours, and the 'risk sensitivity' parameter needs to be re-named.

Please see our earlier response to these suggestions.

Agent-based simulations:

The authors estimated two learning parameters based on the behaviour of real birds, and then ran simulations to see whether computer 'birds' that base their choices on those learning parameters return behaviours that, on average, mirror the behaviour of the real birds. This exercise is clearly circular. In old-style, statistical terms, I suppose this means that the R-square of the statistical model is good. A more insightful use of the simulations would be to identify situations where the simulation does not do as well in mirroring behaviour that it is designed to mirror.

Based on the Reviewer's summary of agent-based forward simulation, we can see we did a poor job explaining the inferential value of this method—we apologise. Agent-based forward simulations are posterior predictions, and they provide insight into the implied model dynamics and overall usefulness of our reinforcement learning model. R-squared calculations are retrodictive, and they say nothing about the causal dynamics of a model. Specifically, agent-based forward simulation allows 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 parameters *in actu*. Agent-based forward simulation is 'designed' to facilitate this call to action—not to mirror behavioural results. The simulation has no *a priori* 'opinion' about computer 'birds' behavioural outcomes; rather, it simply assigns these agents random ϕ and λ draws (whilst maintaining their correlation structure), and tracks their reinforcement learning. The exercise only appears circular *if* no critical contributing mechanism(s) went overlooked—in this case computer 'birds' should behave similar to real birds. A disparate mapping between computer 'birds' and 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 happening in the human reinforcement learning literature, see Deffner et al. 2020 [DOI: 10.1098/rsos.200734]). In sum, agent-based forward simulation does not assess goodness-of-fit—we assessed the fit of our model *a priori* in our preregistration (<https://osf.io/v3wxb>)—but it does assess whether one did a

comprehensive job of uncovering the mechanistic basis of target behaviour(s). We have worked to make the above points on the method and the insight afforded by agent-based forward simulation explicitly clear in our revision—please see L192–207 and L534–537.

Reviewer #2 (Public Review):

Summary:

The study is titled "Leading an urban invasion: risk-sensitive learning is a winning strategy", and consists of three different parts. First, the authors analyse data on initial and reversal learning in Grackles confronted with a foraging task, derived from three populations labeled as "core", "middle" and "edge" in relation to the invasion front. The suggested difference between study populations does not surface, but the authors do find moderate support for a difference between male and female individuals. Secondly, the authors confirm that the proposed mechanism can actually generate patterns such as those observed in the Grackle data. In the third part, the authors present an evolutionary model, in which they show that learning strategies as observed in male Grackles do evolve in what they regard as conditions present in urban environments.

Strengths:

*The manuscript's strength is that it combines real learning data collected across different populations of the Great-tailed grackle (*Quiscalus mexicanus*) with theoretical approaches to better understand the processes with which grackles learn and how such learning processes might be advantageous during range expansion. Furthermore, the authors also take sex into account revealing that males, the dispersing sex, show moderately better reversal learning through higher reward-payoff sensitivity. I also find it refreshing to see that the authors took the time to preregister their study to improve transparency, especially regarding data analysis.*

Thank you—we are pleased to receive this positive evaluation, particularly concerning our efforts to improve scientific transparency via our study's preregistration (<https://osf.io/v3wxb>).

Weaknesses:

One major weakness of this manuscript is the fact that the authors are working with quite low sample sizes when we look at the different populations of edge (11 males & 8 females), middle (4 males & 4 females), and core (17 males & 5 females) expansion range. Although I think that when all populations are pooled together, the sample size is sufficient to answer the questions regarding sex differences in learning performance and which learning processes might be used by grackles but insufficient when taking the different populations into account.

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 note all of our multilevel models use partial pooling on individuals (the random-effects structure), which is a regularisation technique that generally reduces the inference constraint imposed by a low sample size (see Ch. 13 in Statistical Rethinking by Richard McElreath [PDF: <https://bit.ly/3RXCy8c>]). We further note that, in our study preregistration (<https://osf.io/v3wxb>), we formally tested our reinforcement learning model for different effect sizes of sex on learning for both target parameters (ϕ and λ) across populations, using a similarly modest N (edge: 10 M, 5 F; middle: 22 M, 5 F; core: 3 M, 4 F) to our actual final N , that we anticipated to be our final N at that time. This apriori analysis shows 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 (see Figure 4 in <https://osf.io/v3wxb>). Thus, both of these points together highlight how our reinforcement learning model allows us to say that across-population null results are not just due to small sample size. Nevertheless the Reviewer is not wrong to wonder whether a bigger N might change our population-level results (it might; so might muchneeded population replicates—see L310), but our Bayesian models still allow us to learn a lot from our current data. We now explain this in our revised manuscript—please see L452–457.

Another weakness of this manuscript is that it does not set up the background well in the introduction. Firstly, are grackles urban dwellers in their natural range and expand by colonising urban habitats because they are adapted to it? The introduction also fails to mention why urban habitats are special and why we expect them to be more challenging for animals to inhabit. If we consider that one of their main questions is related to how learning processes might help individuals deal with a challenging urban habitat, then this should be properly introduced.

In L74–75 (previously L53–56) we introduce that the estimated historical niche of grackles is urban environments, and that shifts in habitat breadth—e.g., moving into more arid, agricultural environments—is the estimated driver of their rapid North American colonisation. We hope this included information sufficiently answers the Reviewer's question. We have worked towards flushing out how urban-imposed challenges faced by grackles, such as the wildlife management efforts introduced in L64–65 (now L85–86), may apply to animals inhabiting urban environments more broadly; for example, we now include an entire paragraph in our Introduction detailing how urban environments may be characterised differently to nonurban environments, and thus why they are perhaps more challenging for animals to inhabit— please see L56–71.

Also, the authors provide a single example of how learning can differ between populations from more urban and more natural habitats. The authors also label the urban dwellers as the invaders, which might be the case for grackles but is not necessarily true for other species, such as the Indian rock agama in the example which are native to the area of study. Also, the authors need to be aware that only male lizards were tested in this study. I suggest being a bit more clear about what has been found across different studies looking at: (1) differences across individuals from invasive and native populations of invasive species and (2) differences across individuals from natural and urban populations.

We apologise for not including more examples of such learning differences. We now include three examples (please see L43–49), and we are careful to call attention to the fact that these data cover both resident urban and non-urban species as well as urban invasive species (please see L49–50). We also revised our labelling of the lizard species (please see L44). We are aware only male lizards were tested but this information is not relevant to substantiating our use of this study; that is, to highlight that learning can differ between urbandwelling and non-urban counterparts. We hope the changes we did make to our manuscript satisfy the Reviewer's general suggestion to add biological clarity.

Finally, the introduction is very much written with regard to the interaction between learning and dispersal, i.e. the 'invasion front' theme. The authors lay out four predictions, the most important of which is No. 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." The authors, however, never return to this prediction, at least not in a transparent way that clearly pronounces this pattern not being found. The model looking at the evolution of risk-sensitive learning in urban environments is based on the assumption that urban and natural environments "differ

along two key ecological axes: environmental stability μ (How often does optimal behaviour change?) and environmental stochasticity s (How often does optimal behaviour fail to pay off?). Urban environments are generally characterised as both stable (lower μ) and stochastic (higher s)". Even though it is generally assumed that urban environments differ from natural environments the authors' assumption is just one way of looking at the differences which have generally not been confirmed and are highly debated. Additionally, it is not clear how this result relates to the rest of the paper: The three populations are distinguished according to their relation to the invasion front, not with respect to a gradient of urbanization, and further do not show a meaningful difference in learning behaviour possibly due to low sample sizes as mentioned above.

Thank you for highlighting a gap in our reporting clarity. We now take care to transparently report our null result regarding our fourth prediction; more specifically, that we did not detect credible population-level differences in grackles' learning (please see L130). Regarding our evolutionary model, we agree with the Reviewer that this analysis is only one way of looking at the interaction between learning phenotype and apparent urban environmental characteristics. Indeed, in L282–288 (now L325–329) we state: “Admittedly, our evolutionary model is not a complete representation of urban ecology dynamics. Relevant factors—e.g., spatial dynamics and realistic life histories—are missed out. These omissions are tactical ones. Our evolutionary model solely focuses on the response of reinforcement learning parameters to two core urban-like (or not) environmental statistics, providing a baseline for future study to build on”. But we can see now that ‘core’ is too strong a word, and instead ‘supposed’, ‘purported’ or ‘theorised’ would be more accurate—we have revised our wording throughout our manuscript to say as much (please see, for example, L24; L56; L328). We also further highlight the preliminary nature of our evolutionary model, in terms of allowing a narrow but useful first-look at urban eco-evolutionary dynamics—please see L228–232. Finally, we now detail the theorised characteristics of urban environments in our Introduction (rather than in our Results; please see L56–71), and we hope that by doing so, how our evolutionary results relate to the rest of our paper is now better set up and clear.

In conclusion, the manuscript was well written and for the most part easy to follow. The format of eLife having the results before the methods makes it a bit harder to follow because the reader is not fully aware of the methods at the time the results are presented. It would, therefore, be important to more clearly delineate the different parts and purposes. Is this article about the interaction between urban invasion, dispersal, and learning? Or about the correct identification of learning mechanisms? Or about how learning mechanisms evolve in urban and natural environments? Maybe this article can harbor all three, but the borders need to be clear. The authors need to be transparent about what has and especially what has not been found, and be careful to not overstate their case.

Thank you, we are pleased to read that the Reviewer found our manuscript to be generally digestible. We have worked to add further clarity, and to temper our tone (please see our above and below responses).

Reviewer #1 (Recommendations For The Authors):

Several of the results are based on CIs that overlap zero. Tone these down somewhat.

We apologise for overstating our results, which we have worked to tone down in our revision. For instance, in L185–186 we now differentiate between estimates that did or did not overlap zero (please also see our response to Reviewer 2 on this tonal change). We note we do not report confidence intervals (i.e., the range of values expected to contain the true estimate if one redoes the study/analysis many times). Rather, we report 89% highest posterior density

intervals (i.e., the most likely values of our parameters over this range). We have added this definition in L459, to improve clarity.

The literature review suggesting that urban environments are more unpredictable is not convincing. Yes, they have more noise and light pollution and more cars and planes, but does this actually relate to the unpredictability of getting a food reward when you choose an option that usually yields rewards?

To answer the Reviewer's question—yes. But we can see that by not including empirical examples from the literature, we did a poor job of arguing such links. In L43–49 we now give three empirical examples; more specifically, we state: “[...] experimental data show 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).” We note we now detail such apparently stable but stochastic urban environmental characteristics in our Introduction rather than our Results section, to hopefully improve the clarity of our manuscript (please see L56–71). We further note that we cite three literature reviews—not one—suggesting urban environments are stable in certain characteristics and more unpredictable in others (please see L59–60). Finally, we appreciate such characterisation is not certain, and so in our revision we have qualified all writing about this potential dynamic with words such as “apparent”, “supposed”, “theorised”, “hypothesised” etc.

It would be interesting to see if other individual traits besides sex affect their learning/reversal learning ability and/or their learning parameters. Do you have data on age, size, condition, or personality? Or, the habitat where they were captured?

We do not have these data. But we agree with the Reviewer that examining the potential influence of such covariates on grackles' reinforcement learning would be interesting in future study, especially habitat characteristics (please see L306–309).

For most levels of environmental noise, there appears to be an intermediate maximum for the relationship between environmental stability and the risk sensitivity parameter. What does this mean?

There is indeed an intermediate maximum for certain values of environmental stochasticity (although the differences are rather small). The most plausible reason for this is that for very stable environments, simulated birds essentially always “know” the rewarded solution and never need to “relearn” behaviour. In this case, differences in latent values will tend to be large (because they consistently get rewarded for the same option), and different lambda values (in the upper range) will produce the same choice behaviour, which results in very weak selection. While in very unstable environments, optimal choice behaviour should be more exploratory, allowing learners to track frequently-changing environments. We now note this pattern in L240–248.

Reviewer #2 (Recommendations For The Authors):

L2: I'd encourage the authors to reconsider the term "risk-sensitive learning", at least in the title. It's not apparent to me how 'risk' relates to the investigated foraging behaviour. Elsewhere, risk-reward sensitivity is used which may be a better term.

We apologise for not clearly introducing the field of risk-sensitive foraging, 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 [DOI: 10.1079/PNS2002181]; Kacelnik & Bateson 1996

[DOI: 10.1093/ICB/36.4.402]). We have added this information to our manuscript in L494–497. In explaining our reinforcement model, we also now detail how risk relates to foraging behaviour. Specifically, in L153–161 we now state: “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 nonrisk 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’.” We hope this information clarifies why lambda measures risk-sensitivity, and why we continue to use this term.

L1-3: The title is a bit misleading with regard to the empirical data. From the data, all that can be said is that male grackles relearn faster than females. Any difference between populations actually runs the other way, with the core population exhibiting a larger difference between males and females than the mid and edge populations.

It is customary for a manuscript title to describe the full scope of the study. In our study, we have empirical data, cognitive modelling, and evolutionary simulations of the background theory all together. And together these analytical approaches show: (1) across three populations, male grackles—the dispersing sex in this historically urban-dwelling and currently urban-invading species—outperform female counterparts in reversal learning; (2) they do this via risk-sensitive learning, so they’re more sensitive to relative differences in reward payoffs and choose to stick with the ‘safe’ i.e., rewarding option, rather than continuing to ‘gamble’ on an alternative option; and (3) risk-sensitive learning should be favoured in statistical environments characterised by purported urban dynamics. So, we do not feel our title “Leading an urban invasion: risk-sensitive learning is a winning strategy” is misleading with regard to our empirical data; it just doesn’t summarise only our empirical data. Finally, as we now state in L312–313, we caution against speculating about any between-population variation, as we did not infer any meaningful behavioural or mechanistic population-level differences.

L13: “Assayed”, is that correctly put, given that the authors did not collect the data?

Merriam-Webster defines assay as “to analyse” or “examination or determination as to characteristics”, and so to answer the Reviewer’s question—yes, we feel this is correctly put. We note we explicitly introduce in L102–103 that we did not collect the data, and we have an explicit “Data provenance” section in our methods (please see L342–347).

L42-46: The authors provide a single example of how learning can differ between populations from more urban and more natural habitats. I would like to point out that many of these studies do not directly confirm that the ability in question has indeed led to the success of the species tested (e.g. show fitness consequences). Then the authors could combine these insights to form a solid prediction for the grackles. As of now, this looks like cherry-picking supportive literature without considering negative results.

Here are some references that might be helpful in identifying relevant literature to cite:

Szabo, B., Damas-Moreira, I., & Whiting, M. J. (2020). Can cognitive ability give invasive species the means to succeed? A review of the evidence. Frontiers in Ecology and Evolution, 8, 187.

Griffin AS, Tebbich S, Bugnyar T, 2017. *Animal cognition in a human-dominated world. Anim Cogn* 20(1):1-6.

Kark, S., Iwaniuk, A., Schalimtzek, A., & Banker, E. (2007). *Living in the city: Can anyone become an "urban exploiter"?* *Journal of Biogeography*, 34(4), 638-651.

We apologise for not including more examples of such learning differences. We now include three examples (please see L43–49). We are aware that direct evidence of fitness consequences is entirely lacking in the scientific literature on cognition and successful urban invasion; hence why such data is not present in our paper. But we now explicitly point out a role for likely fitness-affecting anthropogenic disturbances on sleep, mate, and foraging behaviour on animals inhabiting urban environments (please see L63–68). We hope these new data bolster our predictions for our grackles. Finally, the Reviewer paints a (in our view) inaccurate picture of our use of available literature. Nevertheless, to address their comment, we now highlight a recent meta-analysis advocating for further research to confirm apparent ‘positive’ trends between animal ‘smarts’ and successful ‘city living’ (please see L43).

L64: *Is their niche historically urban, or have they recently moved into urban areas?*

In L74–75 (previously L53–56) we introduce that the estimated historical niche of grackles is urban environments, and that shifts in habitat breadth—e.g., moving into more arid, agricultural environments—is the estimated driver of their rapid North American colonisation. We hope this included information sufficiently answers the Reviewer’s question.

L66-67: *This is an important point that is however altogether missing from the discussion.*

We thank the Reviewer for highlighting a gap in our discussion regarding populationlevel differences in grackles’ reinforcement learning. In L310–312 we now state: “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”.

L68-71: *The paper focuses on cognitive ability. The whole paragraph sets up the prediction of why male grackles should be better learners due to their dispersal behaviour. This example, however, focuses on aggression, not cognition. Here is a study showing differences in learning in male and female mynas that might be better suited:*

Federspiel IG, Garland A, Guez D, Bugnyar T, Healy SD, Güntürkün O, Griffin AS, 2017. *Adjusting foraging strategies: a comparison of rural and urban common mynas (Acridotheres tristis).* *Anim Cogn* 20(1):65-74.

We thank the Reviewer for suggesting this paper. We feel it is better suited to substantiating our point in the Discussion about reversal learning not being indicative of cognitive ability—please see L276–277.

L73: *Generally, I suggest not writing "for the first time" as this is not a valid argument for why a study should be conducted. Furthermore, except for replication studies, most studies investigate questions that are novel and have not been investigated before.*

The Reviewer makes a fair point—we have removed this statement.

L80-81: Here again, this is left undiscussed later on.

By ‘this’ we assume the Reviewer is referring to our hypothesis, which is that sex differences in dispersal are related to sex differences in learning in an urban invader— grackles. At the beginning of our Discussion, we state how we found support for this hypothesis (please see L250–261); and in our ‘Ideas and speculation’ section, we discuss how these hypothesis-supporting data fit into the literature more broadly (please see L294–331). We feel this is therefore sufficiently discussed.

L77-81: This sentence is very long and therefore hard to read. I suggest trying to split it into at least 2 separate sentences which would improve readability.

Per the Reviewer’s useful suggestion, we have split this sentence into two separate sentences—please see L97–115.

L83: Please explain choice-option switches. I am not aware of what that is and it should be explained at first mention.

We apologise for this operational oversight. We now include a working definition of speed and choice-option switches at first mention. Specifically, in L107–108 we state: “[...] 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)”.

L83-87: Again, a very long sentence. Please split.

We thank the Reviewer for their suggestion. In this case we feel it is important to not change our sentence structure because we want our prediction statements to match between our manuscript and our preregistration.

L96-97: Important to not overstate this. It merely demonstrates the potential of the proposed (not detected) mechanism to generate the observed data.

As in any empirical analysis, our drawn conclusions depend on causal assumptions about the mechanisms generating behaviour (Pearl, J. (2009). Causality). Therefore, we “detected” specific learning mechanisms assuming a certain generative model, namely reinforcement learning. As there is overwhelming evidence for the widespread importance of value-based decision making and Rescorla-Wagner updating rules across numerous different animals (Sutton & Barto (2018) Reinforcement Learning), we would argue that this assumed model is highly plausible in our case. Still, we changed the text to “inferred” instead of “detected” learning mechanisms to account for this concern—please see L123–124.

L99: "urban-like settings" again a bit confusing. The authors talk about invasion fronts, but now also about an urbanisation gradient. Is the main difference between the size and the date of establishment, or is there additionally a gradient in urbanisation to be considered?

We now include a paragraph in our Introduction detailing apparent urban environmental characteristics (please see 56–71), and we now refer to this dynamic specifically when we define urban-like settings (please see L126–127). To answer the Reviewer’s question—we consider both differences. Specifically, we consider the time since population establishment in our paper (with respect to our behavioural and mechanistic modelling), as well as how statistical environments that vary in how similar they are to apparently characteristically

urban-like environments, might favour particular learning phenotypes (with respect to our evolutionary modelling). We hope the edits to our Introduction as a whole now make both of the aims clear.

L11-112: Above the authors talk about a comparable number of switches ($10.5/15=0.7$), and here of fewer number of switches ($25/35=0.71$), even though the magnitude of the difference is almost identical and actually runs the other way. The authors are probably misled by their conservative priors, which makes the difference appear greater in the second case than in the first. Using flat priors would avoid this particular issue.

Mathematically, the number of trials-to-finish and the number of choice-option-switches are both a Poisson distributed outcome with rate λ (we note lambda here is not our risk-sensitivity parameter; just standard notation). As such, our Poisson models infer the rate of these outcomes by sex and phase—not the ratio of these outcomes by sex and phase. So comparing the magnitude of divided medians of choice-option-switches between the sexes by phase is not a meaningful metric with respect to the distribution of our data, as the Reviewer does above. For perspective, 1 vs. 2 switches provides much less information about the difference in rates of a Poisson distribution than 50 vs 100 (for the former, no difference would be inferred; for the latter, it would), but both exhibit a 1:2 ratio. To hopefully prevent any such further confusion, and to focus on the fact that our Poisson models estimate the expected value i.e., the mean, we now report and graph (please see Fig. 2) mean and not median trialsto-finish and total-switch-counts. Finally, we can see that our use of the word “conservative” to describe our weakly informative priors is confusing, because conservative could mean either strong priors with respect to expected effect size (not our parameterisation) or weak priors with respect to such assumptions (our parameterisation). To address this lack of clarity, we now state that we use “weakly informative priors” in L457–458.

L126: It is not clear what risk sensitivity means in the context of these experiments.

Thank you for pointing out our lack of clarity. In L153–161 we now state: “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 nonrisk 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’.” We hope this information clarifies what risk sensitivity means and measures, with respect to our behavioural experiments.

L128-129: I find this statement too strong. A plethora of other mechanisms could produce similar patterns, and you cannot exclude these by way of your method. All you can show is whether the mechanism is capable of producing broadly similar outcomes as observed

In describing the inferential value of our reinforcement learning model, we now qualify that the insight provided is of course conditional on the model, which is tonally accurate. Please see L161.

L144: As I have already mentioned above, here is the first time we hear about unpredictability related to urban environments. I suggest clearly explaining in the introduction how urban and natural environments are assumed to be different which

leads to animals needing different cognitive abilities to survive in them which should explain why some species thrive and some species die out in urbanised habitats.

Thank you for this suggestion. We now include a paragraph in our Introduction detailing as much—please see L56–71.

L162: "almost entirely above zero" again, this is worded too strongly.

In reporting our lambda across-population 89% HPDI contrasts in L185–186, we now state: “[...] across-population contrasts that lie mostly above zero in initial learning, and entirely above zero in reversal learning”. Our previous wording stated: “[...] across-population contrasts that lie almost entirely above zero”. The Reviewer was correct to point out that this previous wording was too strong if we considered the contrasts together, as, indeed, we find the range of the contrast in initial learning does minimally overlap zero (L: -0.77; U: 5.61), while the range of the contrast in reversal learning does not (L: 0.14; U: 4.26). This rephrasing is thus tonally accurate.

L178-179: I think it should be said instead that the model accounts well for the observed data.

We have rephrased in line with the Reviewer’s suggestion, now stating in L217–218 that “Such quantitative replication confirms our reinforcement learning model results sufficiently explain our behavioural sex-difference data.”

L188-190: I am not convinced this is a general pattern. It is quite a bold claim that I don't find to be supported by the citations. Why should biotic and abiotic factors differ in how they affect behavioural outcomes? Also, events in urban environments such as weekend/weekday could lead to highly regular optimal behaviour changes.

Please see our response to Reviewer 1 on this point. We note we now touch on such regular events in L94–96.

L209-211: The first sentence is misleading. The authors have found that males and females differ in 'risk sensitivity', that their learning model can fit the data rather well, and that under certain, not necessarily realistic assumptions, the male learning type is favoured by natural selection in urban environments. A difference between core, middle, and edge habitats however is barely found, and in fact seems to run the other way than expected.

In our study, we found: (1) across three populations, male grackles—the dispersing sex in this historically urban-dwelling and currently urban-invading species—outperform female counterparts in reversal learning; (2) they do this via risk-sensitive learning, so they’re more sensitive to relative differences in reward payoffs and choose to stick with the ‘safe’ i.e., rewarding option, rather than continuing to ‘gamble’ on an alternative option; (3) we are sufficiently certain risk-sensitive learning generates our sex-difference data, as our agentbased forward simulations replicate our behavioural results (not because our model ‘fits’ the data, but because we inferred meaningful mechanistic differences—see our response to Reviewer 1 on this point); and (4) under theorised dynamics of urban environments, natural selection should favour risk-sensitive learning. We therefore do not feel it is misleading to say that we mapped a full pathway from behaviour to mechanisms through to selection and adaptation. Again, as we now state in L311–313, we caution against speculating about any between-population variation, as we did not infer any meaningful behavioural or mechanistic population-level differences. And we note the Reviewer is wrong to assume an interaction between learning, dispersal, and sex requires population-level differences on the

outcome scale—please see our discussion on phenotypic plasticity and inherent species trait(s) in L313–324.

| L216: *"indeed explain" again worded too strongly.*

We have tempered our wording. Specifically, we now state in L218: “sufficiently explain”. This wording is tonally accurate with respect to the inferential value of agent-based forward simulations—please see L192–207 on this point.

| L234: *"reward-payoff sensitivity" might be a better term than risk-sensitivity?*

Please see our earlier response to this suggestion. We note we have changed this text to state “risk-sensitive learning” rather than “reward-payoff sensitivity”, to hopefully prevent the reader from concluding only our lambda term is sensitive to rewards—a point we now include in L153–154.

| L234-237: *I think these points may be valuable, but come too much out of the blue. Many readers will not have a detailed knowledge of the experimental assays. It therefore also does not become clear how they measure the wrong thing, what this study does to demonstrate this, or whether a better alternative is presented herein. It almost seems like this should be a separate paper by itself.*

We apologise for this lack of context. We now explicitly state in L275 that we are discussing reversal learning assays, to give all readers this knowledge. In doing so, we hope the logic of our argument is now clear: reversal learning assays do not measure behavioural flexibility, whatever that even is. The Reviewer’s suggestion of a separate paper focused on what reversal learning assays actually measure, in terms of mechanism(s), is an interesting one, and we would welcome this discussion. But any such paper should build on the points we make here.

| L270-288: *Somewhere here the authors have to explain how they have not found differences between populations, or that in so far as they found them, they run against the originally stated hypothesis.*

We thank the Reviewer for these suggestions. In L310–313 we now state: “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”.

| L284: *should be "missing" not "missed out"*

We have made this change.

| L290-291: *It is unclear what "robust interactive links" were found. A pattern of sexbiased learning was found, which can potentially be attributed to evolutionary pressures in urban environments. An interaction e.g. between learning, dispersal, and sex can only be tentatively suggested (no differences between populations). Also "fully replicable" is a bit misleading. The analysis may be replicable, but the more relevant question of whether the findings are replicable we cannot presently answer.*

We apologise for our lack of clarity. By “robust” we mean “across population”, which we now state in L333. We again note the Reviewer is wrong to assume an interaction between learning, dispersal, and sex requires population-level differences on the outcome scale—

please see our discussion on phenotypic plasticity and inherent species trait(s) in L313–324. Finally, the Reviewer makes a good point about our analyses but not our findings being replicable. In L334 we now make this distinction by stating “analytically replicable”.

L306-315: I think you have a bit of a sample size issue not so much when populations are pooled but when separated. This might also factor in the fact that you do not really find differences across the populations in your analysis. When we look at the results presented in Figure 2 (and table d), we can see a trend towards males having better risk sensitivity in core (HPDI above 0) and middle populations (HPDI barely crossing 0) but the difference is very small. Especially the results on females are based on the performance of only 8 and 4 females respectively. I suggest making this clear in the manuscript.

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 note all of our multilevel models use partial pooling on individuals (the random-effects structure), which is a regularisation technique that generally reduces the inference constraint imposed by a low sample size (see Ch. 13 in Statistical Rethinking by Richard McElreath [PDF: <https://bit.ly/3RXCy8c>]). We further note that, in our study preregistration (<https://osf.io/v3wxb>), we formally tested our reinforcement learning model for different effect sizes of sex on learning for both target parameters (phi and lambda) across populations, using a similarly modest N (edge: 10 M, 5 F; middle: 22 M, 5 F; core: 3 M, 4 F) to our actual final N, that we anticipated to be our final N at that time. This apriori analysis shows our reinforcement learning model: (i) detects sex differences in phi values ≥ 0.03 and lambda values ≥ 1 ; and (ii) infers a null effect for phi values < 0.03 and lambda values < 1 i.e., very weak simulated sex differences (see Figure 4 in <https://osf.io/v3wxb>). Thus, both of these points together highlight how our reinforcement learning model allows us to say that across-population null results are not just due to small sample size. Nevertheless the Reviewer is not wrong to wonder whether a bigger N might change our population-level results; it might; so might much-needed population replicates—see L310. But our Bayesian models still allow us to learn a lot from our current data, and, at present, we infer no meaningful population-level behavioural or mechanistic differences in grackles’ behaviour. To make clear the inferential sufficiency of our analytical approach, we now include some of the above points in our Statistical analyses section in L452–457. Finally, we caution against speculating on any between-population variation, as we now highlight in L311–313 of our Discussion.

Figure 2: I think the authors should rethink their usage of colour in this graph. It is not colour-blind friendly or well-readable when printed in black and white.

We used the yellow (hex code: #fde725) and green (hex code: #5ec962) colours from the viridis package. As outlined in the viridis package vignette (<https://cran.rproject.org/web/packages/viridis/index.html>), this colour package is “designed to improve graph readability for readers with common forms of color blindness and/or color vision deficiency. The color maps are also perceptually-uniform, both in regular form and also when converted to black-and-white for printing”.

Figure 3B: Could the authors turn around the x-axis and the colour code? It would be easier to read this way.

We appreciate that aesthetic preferences may vary. In this case, we prefer to have the numbers on the x-axis run the standard way i.e., from small to large. We note we did remove the word ‘Key’ from this Figure, in line with the Reviewer’s point about these characteristics not being totally certain.

I also had a look at the preregistration. I do think that there are parts in the preregistration that would be worth adding to the manuscript:

L36-40: This is much easier to read here than in the manuscript.

We changed this text generally in the Introduction in our revision, so we hope the Reviewer will again find this easier to read.

L49-56: This is important information that I would also like to see in the manuscript.

We no longer have confidence in these findings, as our cleaning of only one part of these data revealed considerable experimenter oversight (see ‘Learning criterion’).

L176: Why did you remove the random effect study site from the model? It is not part of the model in the manuscript anymore.

The population variable is part of the RL_Comp_Full.stan model that we used in our manuscript to assess population differences in grackles’ reinforcement learning, the estimates from which we report in Table C and D (please note we never coded this variable as “study cite”). But rather than being specified as a random effect, in our RL_Comp_Full.stan model we index phi and lambda by population as a predictor variable, to explicitly model population-level effects. Please see our code:

https://github.com/alexisbreen/Sex-differences-in-grackles-learning/blob/main/Models/Reinforcement%20learning/RL_Comp_Full.stan

L190-228: I am wondering if the model validation should also be part of the manuscript as well, rather than just being in the preregistration?

We are not sure how the files were presented to the Reviewer for review, but our study preregistration, which includes our model validation, should be part of our manuscript as a supplementary file.