Behavioral flexibility is manipulatable and it improves flexibility and problem solving in a new context: post-hoc analyses of the components of behavioral flexibility.

Lukas D[1]* McCune KB[3] Blaisdell AP[2] Johnson-Ulrich Z[3]
MacPherson M[3] Seitz B[2] Sevchik A[4] Logan CJ[1]*

2022-08-10

Open... access code peer review data

Affiliations: 1) Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, 2) University of California Los Angeles, USA, 3) University of California Santa Barbara, USA, 4) Arizona State University, Tempe, AZ USA. *Corresponding author: dieter_lukas@eva.mpg.de

This is one of three post-study manuscript of the preregistration that was pre-study peer reviewed and received an In Principle Recommendation on 26 Mar 2019 by:

Aurélie Coulon (2019) Can context changes improve behavioral flexibility? Towards a better understanding of species adaptability to environmental changes. *Peer Community in Ecology*, 100019. 10.24072/pci.ecology.100019. Reviewers: Maxime Dahirel and Andrea Griffin

Preregistration: html, pdf, rmd

Post-study manuscript (submitted to PCI Ecology for post-study peer review on 3 Jan 2022): html, preprint here

ABSTRACT

Behavioral flexibility, adapting behavior to changing situations, is hypothesized to be related to adapting to new environments and geographic range expansions. However, flexibility is rarely directly tested in a way that allows insight into how flexibility works. Research on great-tailed grackles, a bird species that has rapidly expanded their range into North America over the past 140 years, shows that grackle flexibility is manipulatable using colored tube reversal learning and that flexibility is generalizable across contexts multi-access box). Here, we use these grackle results to conduct a set of posthoc analyses using a model that breaks down performance on the reversal learning task into different components. We show that the rate of learning to be attracted to an option (phi) is a stronger predictor of reversal performance than the rate of deviating from learned attractions that were rewarded (lambda). This result was supported in simulations and in the data from the grackles: learning rates in the manipulated grackles doubled by the end of the manipulation compared to control grackles, while the rate of deviation slightly decreased. Grackles with intermediate rates of deviation in their last reversal, independently of whether they had gone through the serial reversal manipulation, solved fewer loci on the plastic and wooden multi-access boxes, and those with

intermediate learning rates in their last reversal were faster to attempt a new locus on both multi-access boxes. These findings provide additional insights into how grackles changed their behavior when conditions changed. Their ability to rapidly change their learned associations validates that the manipulation had an effect on the cognitive ability we think of as flexibility.

INTRODUCTION

The field of comparative cognition is strongly suspected to be in a replicability crisis, which calls into question the validity of the conclusions produced by this research (Brecht et al., 2021; Farrar, Boeckle, et al., 2020; Farrar, Altschul, et al., 2020; Farrar et al., 2021; Lambert et al., 2022; Tecwyn, 2021). The lack of replicability in experimental design, analyses, and results is, in part, because of the lack of clear theoretical frameworks (Frankenhuis et al., 2022), the resulting heavy reliance on measuring operationalized variables that are assumed to represent broad concepts, as well as small sample sizes (Farrar, Boeckle, et al., 2020). One solution is to start from *mechanistic models* informed by a theoretical framework that can represent and make predictions about how individuals behave in a given task, rather than just relying on *statistical models* that simply describe the observed data (McElreath, 2020). Statistical models cannot infer what leads to the differences in behavior, whereas mechanistic models offer the opportunity to infer the underlying processes (McElreath, 2020).

Here, we apply a mechanistic model to a commonly studied trait in animal cognition: behavioral flexibility. Recent work provides clearer conceptualizations of behavioral flexibility that allow us to apply such a mechanistic model. The theoretical framework argues that the critical element of behavioral flexibility is that individuals change their behavior when circumstances change (Mikhalevich et al., 2017), with freedom from instinctual constraints (Lea et al., 2020). These theoretical models point out that behavioral flexibility appears to contain two internal learning processes: the suppression of a previous behavioral choice and the simultaneous adoption of a new behavioral choice. Based on this framework, Blaisdell et al. (2021) showed how reversal learning experiments, where individuals have to choose between two options until they learn to prefer the rewarded option and then the reward is moved to the other option and they reverse their preference, reflect these learning processes. Blaisdell et al. (2021) built a mechanistic model by adapting Bayesian reinforcement learning models to infer the potential cognitive processes underlying behavioral flexibility.

As their name implies, Bayesian reinforcement learning models (Doya, 2007) assume that individuals will gain from learning which of the options leads to the reward. This learning is assumed to occur through reinforcement because individuals repeatedly experience that an option is either rewarded or not. The approach is represented as Bayesian because individuals continuously update their knowledge about the reward with each choice (Deffner et al., 2020). At their core, these models contain two individual-specific parameters that we aim to estimate from reversal performance: how quickly individuals update their attraction to an option based on the reward they received during their most recent choice relative to the rewards they received when choosing this option previously (their learning rate, termed “phi” ϕ), and whether individuals already act on small differences in their attraction or whether they continue to explore the less attractive option (the deviation rate, termed “lambda” λ). Applied to the serial reversal learning setup, where an individual’s preferences are reversed multiple times, the model assumes that, at the beginning of the experiment, individuals have equally low attractions to both options. Depending on which option they choose first, they either experience the reward or not. Experiencing the reward will potentially increase their attraction to this option: if ϕ is zero, their attraction remains unchanged; if ϕ is one, their attraction is completely dominated by the reward they just gained. In environments that are predictable for short periods of time, similar to the rewarded option during a single reversal in our experiment, individuals are likely to gain more rewards if they update their information based on their latest experience. In situations where rewards change frequently or novel options become available often, individuals are expected to deviate from their learned attractions to continue to explore, while in more stable environments individuals benefit from large λ values to exploit the associations they formed (Cohen et al., 2007). While performance in the reversal learning task has sometimes been decomposed between the initial association learning and the reversal learning phase (e.g. Federspiel et al., 2017), the reinforcement learning model does not make such a distinction. However, it does predict a difference between phases because individuals’ internal states, in particular their attraction toward the different options, are expected to continuously change throughout the experiment. We also expect individuals to "learn to learn" over subsequent reversals (Neftci & Averbeck, 2019), changing their learning and deviation rate over repeated reversals. The parameters of the serial reversal model can also capture broader concepts that have previously been used to describe variation in reversal learning performance, such as "proactive interference" (Morand-Ferron et al., 2022) as the tendency to continue to choose the previously rewarded option which would occur if individuals do not update their attractions quickly.

We applied this model to our great-tailed grackle (*Quiscalus mexicanus*, hereafter grackle) research on behavioral flexibility, which we measured as reversal learning of a color preference using two differently colored tubes (one light gray and one dark gray C. Logan et al., 2022). In one population, we conducted a flexibility manipulation using serial reversal learning - reversing individuals until their reversal speeds were consistently fast (at or less than 50 trials in two consecutive reversals). We randomly assigned individuals to a manipulated group who received serial reversals, or to a control group who received one reversal and then a similar amount of experience in making choices between two yellow tubes that both contained rewards (C. Logan et al., 2022). After the manipulation, grackles were given a flexibility and innovativeness test using one or two different multi-access boxes to determine whether improving flexibility in reversal learning also improved flexibility (the latency to attempt to solve a new locus) and innovativeness (the number of loci solved) in a different context (the multi-access boxes). We found that we were able to manipulate reversal learning performance (flexibility) and this improved flexibility and problem solving in a new context (multi-access boxes) (C. Logan et al., 2022). However, we were left with some lingering questions: what specifically did we manipulate about flexibility? And how might the cognitive changes induced by the manipulation transfer to influence performance in a new context? These questions are the focus of the current article.

RESEARCH QUESTIONS

1) How are the two parameters ϕ or λ linked to individual differences in reversal learning behavior in simulations? Can we reliably estimate ϕ or λ based on the performance of individuals in the reversal learning task?
   Prediction 1: We predicted that the Bayesian reinforcement learning model can reliably infer these two components based on the choices individuals make, which we tested by assigning individuals ϕ and λ values, simulating their choices based on these, and back-estimating ϕ and λ from the simulated choice data.
   Prediction 2: We predicted that both ϕ and λ influence the performance of individuals in a reversal learning task, with higher ϕ (faster learning rate) and lower λ (less exploration) values leading to individuals more quickly reaching the passing criterion after a reversal in the color of the rewarded option.

2) Which of the two parameters ϕ or λ explain more of the variation in the reversal performance of the tested grackles, and which changed more across the serial reversals?
   Prediction 3: We predicted that whichever of the two parameters, ϕ or λ, explains more of the variation in the first reversal performance is also the parameter that shows more change after the manipulation. However, in the serial reversals, birds need to be able to quickly learn the new reward location and also be ready to explore the other option. Accordingly, birds might end up with one of two solutions: they might adopt a strategy of weighting recent information more heavily while also showing low exploration, or they might show high exploration while being slow at updating their attractions.

3) Are ϕ or λ, the two components of flexibility in reversal learning, associated with performance on the multi-access boxes across control and manipulated birds?
   Prediction 4: We predicted that birds that are more flexible, presumably those who have a high ϕ (faster learning rate), have shorter latencies to attempt a new locus and solve more loci on the two multi-access boxes. Given that birds might use different strategies to be flexible (see prediction 3), we also explore whether the relationship between ϕ or λ and the performance on the multi-access boxes is non-linear.

METHODS

The Bayesian reinforcement learning model We used the version of the Bayesian model that was developed by Blaisdell et al. (2021) and modified by Logan CJ et al. (2020) (see their Analysis Plan > "Flexibility analysis" for model specifications and validation). This model uses data from every trial of reversal learning (rather than only using the total number of trials to pass criterion) and represents behavioral flexibility using two parameters: the learning rate of attraction to either option (ϕ) and the rate of deviating from learned attractions (λ). The model repeatedly estimates the series of choices each bird made, based on two equations

Equation 1 (attraction and ϕ): $A_{j,i,t+1}=(1-\phi_j)A_{j,i,t}+\phi_j \; \pi_{j,i,t}$

Equation 1 tells us how attractions A of individual j to the two different options (i=1,2) change from one trial to the next (time t+1) as a function of previous attractions $A_{j,i,t}$ (how preferable option i is to the bird j at time t) and recently experienced payoffs π (i.e., 1 when they received a reward in a given trial, 0 when not). The (bird-specific) parameter ϕ_j describes the weight of recent experience. The higher the value of ϕ_j, the faster the bird updates their attraction. Attraction scores thus reflect the accumulated learning history up to this point. At the beginning of the experiment, we assume that individuals have the same low attraction to both options ($A_{j,1} = A_{j,2} = 0.1$).

Equation 2 (choice and λ): $P(j,i)_{t+1}=\dfrac{exp(\lambda_j A_{j,i,t})}{\sum_{i=1}^{2} exp(\lambda_j A_{j,i,t})}$

Equation 2 expresses the probability P that an individual j chooses option i in the next trial, t+1, based on the attractions. The parameter λ_j represents the rate of deviating from learned attractions of an individual. It controls how sensitive choices are to differences in attraction scores. As λ_j gets larger, choices become more deterministic and individuals consistently choose the option with the higher attraction even if attractions are very similar, as λ_j gets smaller, choices become more exploratory (random choice independent of the attractions if λ_j=0).

We implemented the Bayesian reinforcement learning model in the statistical language Stan (Team et al., 2019), calling the model and analyzing its output in R [current version 4.0.3; R Core Team (2017)]. The values for ϕ and λ for each individual are estimated as the mean from 2000 samples from the posterior.

1) Using simulations to check models estimating the role of the potential parameters underlying performance in the reversal experiment We ran the Bayesian model on simulated data to first understand whether we could recover the ϕ and λ values assigned to each individual from the choices individuals made based on their phis and lambdas in the initial and first reversal learning phases; and second to see whether inter-individual variation in ϕ or in λ contributed more to variation in their performance. The settings for the simulations were based on the previous analysis of data from grackles in a different population (Santa Barbara, Blaisdell et al. (2021)). We re-analyzed data we had simulated for power analyses to estimate sample sizes for population comparisons (Logan CJ et al., 2020). In brief, we simulated 20 individuals each from 32 different populations (640 individuals). The ϕ and λ values for each individual were drawn from a distribution representing that population, with different mean ϕ (8 different means) and mean λ (4 different values) for each population (32 populations as the combination of each ϕ and lambda). Based on their ϕ and λ value, each individual was simulated to pass first through the initial association learning phase and, after they reached criterion, a reversal learning phase. Each choice each individual made was simulated consecutively, updating their internal attraction to the two options based on their ϕ values and setting their next choice based on their λ weighing of their attractions. We first attempted to recover ϕ and λ for different subsets of the data (initial association learning and reversal learning separately or combined). Next, we determined how the ϕ and λ values that were assigned to the individuals influenced their performance in the reversal learning trial, building a regression model to determine which of the two parameters had a more direct influence on the number of trials individuals needed to reach criterion:
number of trials to reverse ~ normal(mu, sigma)
mu <- a + b * ϕ + c * λ

The model was also estimated in stan, using functions from the package 'rethinking' (McElreath, 2020) to build the model.

2) Estimating ϕ and λ from the observed serial reversal learning performances The collection of the great-tailed grackle data, as described in the main article (C. Logan et al., 2022), was based on our preregistration that received in principle acceptance at PCI Ecology (PDF version). All of the analyses of C. Logan et al. (2022) data reported here were not part of the original preregistration.

The data are available at the Knowledge Network for Biocomplexity's data repository: https://knb.ecoinformatics.org/view/doi:10.5063/F1H41PWS.

Great-tailed grackles were caught in the wild in Tempe, Arizona, USA for individual identification (colored leg bands in unique combinations). Some individuals were brought temporarily into aviaries for testing, and then released back to the wild. Individuals first participated in the reversal learning tasks. A subset of individuals was part of the control group, where they learned the association of reward with one color before experiencing one reversal to learn that the other color is rewarded. The other subset of individuals was part of the manipulated group. These individuals went through a series of reversals until they reached the criterion of having formed an association (17 out of 20 choices correct) in less than 50 trials in two consecutive reversals.

We fit the Bayesian reinforcement learning model to the data of both the control and the manipulated birds. For the manipulated birds, we calculated ϕ and λ separately for their performance in the beginning (initial association and first reversal) and at the end of the manipulation (final two reversals). Next, as with the simulated data, we fit a series of regression models to determine how ϕ and λ link to the number of trials birds needed during their reversals.

3) Linking ϕ and λ from the observed serial reversal learning performances to the performance on the multi-access boxes After the individuals had completed the reversal learning tasks, they were provided access to two multi-access boxes, one made of wood and one made of plastic. Both boxes had 4 possible ways (loci) to access food. Initially, individuals could explore all loci. After a preference for a locus was formed, this preferred choice became non-functional by closing access to the locus, and then the latency of the grackle to switch to a new locus was measured. If they again formed a preference, the second locus was also made non-functional, and so on. The outcome measures for each individual with each box were the average latency it took to switch to a new locus and the total number of loci they accessed. For details see (C. Logan et al., 2022).

We repeated the models in the original article (C. Logan et al., 2022) that linked performance on the serial reversal learning tasks to performance on the multi-access boxes, replacing the previously used independent variable of number of trials needed to reach criterion in the last reversal with the estimated ϕ and λ values from the last two reversals (manipulated birds) or the initial discrimination and the first reversal (control birds). The outcome variables were latency to attempt a locus on either the plastic or the wooden multi-access box, and the number of loci solved on the plastic and wooden multi-access boxes. With our observation that ϕ and λ could be negatively correlated (see results), we realized that birds might be using different strategies when facing a situation in which cues change: some birds might quickly discard previous information and rely on what they just experienced (high ϕ and low lambda), or they might rely on earlier information and continue to explore other options (low ϕ and high lambda). Accordingly, we assumed that there also might be non-linear, U-shaped relationships between ϕ and/or λ and the performance on the multi-access box. The regression models were again estimated in stan, using functions from the package 'rethinking' to build the model. We assumed that ϕ and/or λ were associated with the performance on the multi-access boxes if the 89% compatibility intervals of the posterior estimate did not cross zero.

Model: number of loci solved on the multi-access box ~ ϕ and λ
The model takes the form of:
locisolved ~ Binomial(4, p) [likelihood]
logit(p) ~ α[batch] + $\beta * \phi + \gamma * \lambda$ [model]

locisolved is the number of loci solved on the multi-access box, 4 is the total number of loci on the multi-access box, p is the probability of solving any one locus across the whole experiment, α is the intercept and each batch gets its own, β is the expected amount of change in locisolved for every one unit change in the learning rate ϕ in the reversal learning experiments, gamma is the expected amount of change in locisolved for every one unit change in the deviation rate λ in the reversal learning experiments.

Model: latency to attempt a new locus on the multi-access box ~ ϕ and λ
For the average latency to attempt a new locus on the multi-access box as it relates to trials to reverse (both are measures of flexibility), we simulated data and set the model as follows:
latency ~ gamma-Poisson(μ, σ) [likelihood]
log(μ) ~ α[batch] + β * ϕ + γ * λ [the model]
latency is the average latency to attempt a new locus on the multi-access box, μ is the rate (probability of attempting a locus in each second) per bird (and we take the log of it to make sure it is always positive; birds with a higher rate have a smaller latency), σ is the dispersion of the rates across birds, α is the intercept for the rate per batch, β is the expected amount of change in the rate of attempting to solve in any given second for every one unit change in the learning rate ϕ in the reversal learning experiments, γ is the expected amount of change in the rate of attempting to solve in any given second for every one unit change in the deviation rate λ in the reversal learning experiments.

To represent the potential U-shaped relationship, which assumes that birds with intermediate ϕ and λ values perform differently, we first transformed ϕ and λ to calculate for each individual how far their value is from the median. Second, we ran the models squaring ϕ and lambda. Both approaches gave the same results, and we only reported the estimates from the models with the transformed values.

RESULTS

1) Using simulations to check the validity of the Bayesian reinforcement learning models to estimate performance in the reversal learning task

We first ran the Bayesian reinforcement learning model on simulated data to better understand how the two parameters, */phi* and */lambda*, might lead to differences in performance, and whether we could detect meaningful differences between control and manipulated birds. When we used only the choices simulated individuals made during their first reversal, the estimated ϕ and λ values did not match those the individuals had been assigned. We realized that ϕ and λ values were consistently shifted in a correlated way. When estimating these values from only a single reversal, there was equifinality: multiple combinations of the two parameters ϕ and λ could potentially explain the performance of birds during this reversal, and the estimation adjusted both parameters towards the mean. However, when we combined data from across two reversal or from the initial discrimination learning and the first reversal, the model accurately recovered the ϕ and λ values that the simulated individuals had been assigned (Figure 1).

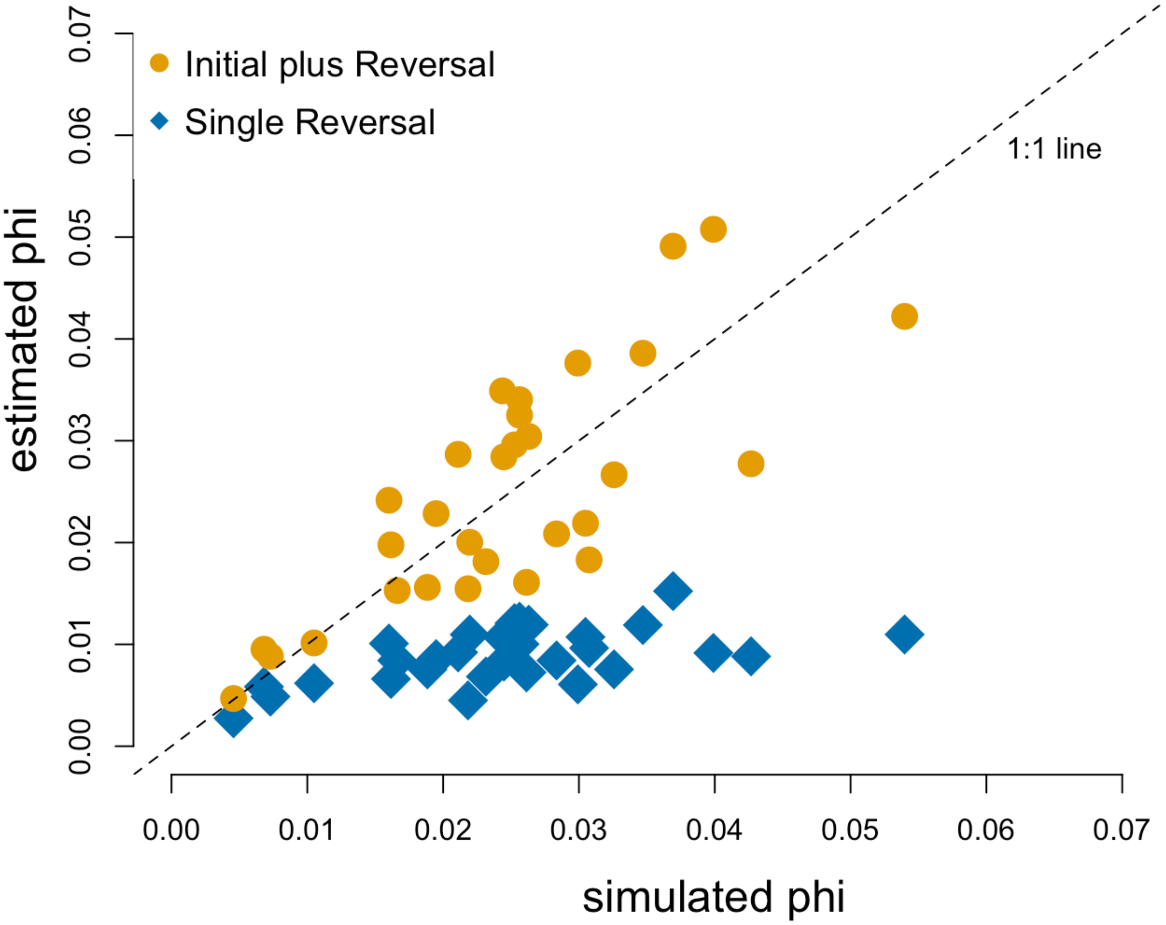


Figure 1: The ϕ values estimated by the model based on the choices made by 30 simulated individuals (x-axis) versus the ϕ values assigned to them (y-axis). Individuals were assigned the simulated phi, their choices were simulated and these values were used to back-estimate the ϕ. When ϕ was estimated based on the choices made only during one reversal, the estimates were consistently lower than the assigned values, particularly for large ϕ values (blue squares). However, when ϕ was estimated based on the choices made during the initial association and the first reversal, the estimates were close to the assigned values (yellow circles).

In terms of the influence of the two parameters ϕ and λ on the number of trials birds needed to reverse a color preference, the ϕ values assigned to simulated individuals had a stronger influence than the λ values (estimated association of number of trials with standardized values of ϕ: -21, 89% compatibility interval: -22 to -19; with standardized values of λ: -14, 89% CI: -16 to -13). In particular, low numbers of trials to reverse could be observed across the full range of λ values, though when λ was smaller than 8, simulated birds might need 150 or more trials to reverse a preference (Figure 2). In contrast, there was a more linear relationship between ϕ and the number of trials to reverse, with birds needing fewer trials the larger their ϕ.

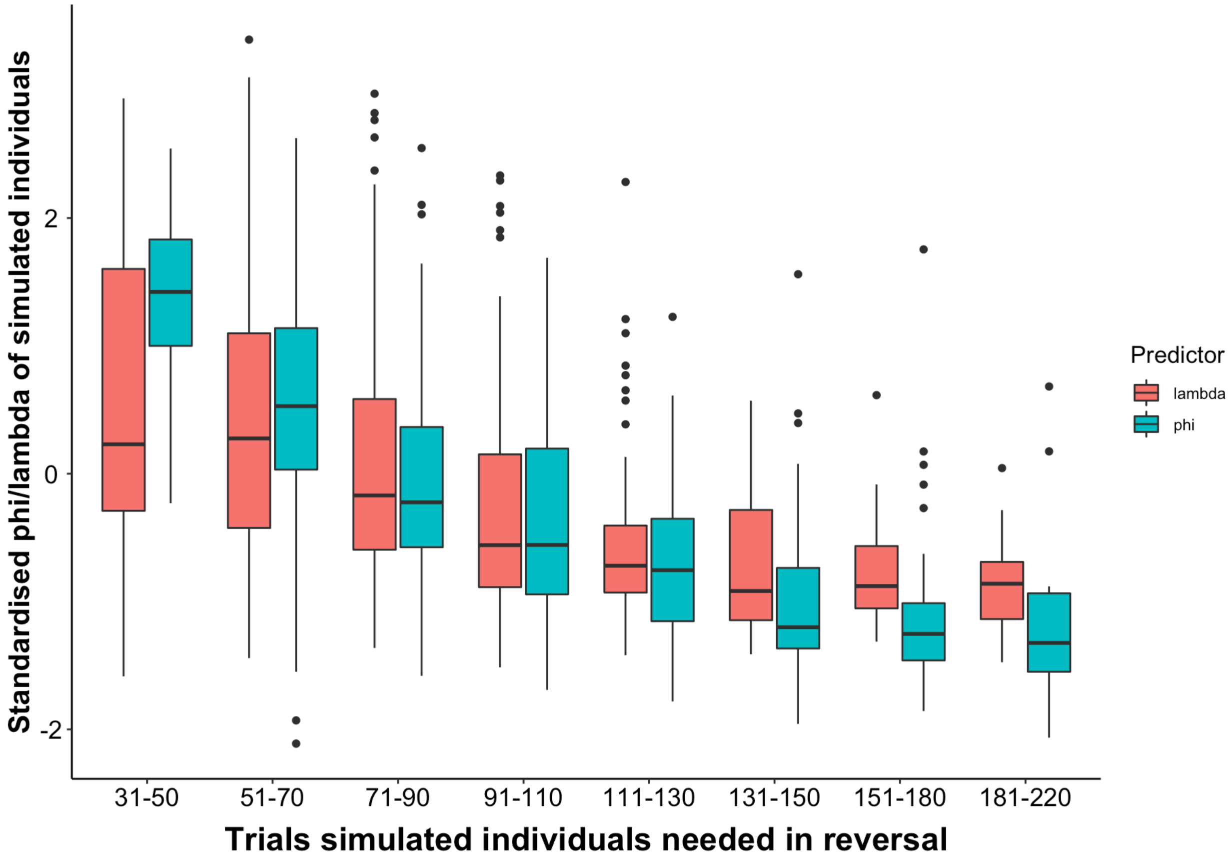


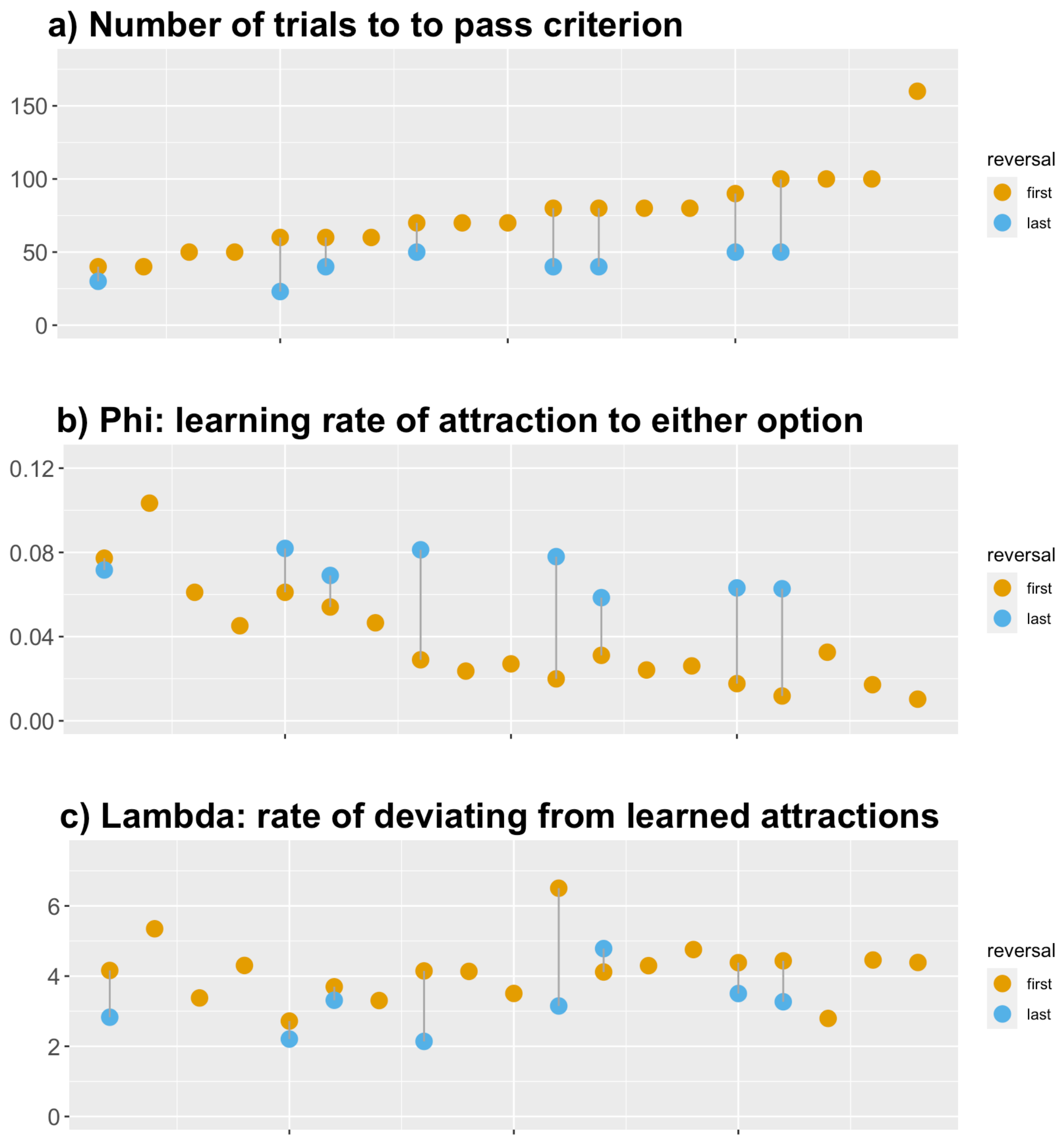
Figure 2. In the simulations, the ϕ values assigned to individuals (green) had a clearer influence on the number of trials these individuals needed to reverse than their λ values (red). ϕ and λ values were standardized for direct comparison. In general, individuals needed fewer trials to reverse if they had larger ϕ and λ values. However, relatively small λ values could be found across the range of reversal performances, whereas there was a more clear distinction with ϕ values.

2) Observed effects of the manipulation on reversal performance, ϕ, and λ

The findings from the simulated data indicated that λ and ϕ can only be estimated accurately when calculated across at least one switch, and we therefore estimated these values for the observed birds based on their performance in the initial discrimination plus first reversal, and for the manipulated birds additionally on their performance in the final two reversals. For the manipulated birds, the estimated ϕ more than doubled from 0.03 (for reference, control grackles=0.03) in their initial discrimination and first reversal to 0.07 in their last two reversals (model estimate of expected average change 89% compatibility interval: +0.02 to +0.05; Table 1: Model 17), while their λ went slightly down from 4.2 (for reference, control grackles=4.3) to 3.2 (model estimate of average change 89% compatibility interval: -1.63 to -0.56; Table 1: Model 18). The values we observed after the manipulation in the last reversal for the number of trials to reverse, as well as the ϕ and λ values estimated from the last reversal, all fall within the range of variation we observed among the control birds in their first and only reversal (Figure 3). This means that the manipulation did not push birds to new levels, but changed them within the boundaries of their natural environment. Some birds in the control group already had similar flexibility measures to the manipulated birds after going through serial reversal learning, presumably because some birds have had experiences in their natural environments that made them more flexible. Accordingly, birds in the manipulated group were not initially all better performers than all of the birds in the control group.

For ϕ, the increase during the manipulation fits with the observations in the simulations: larger ϕ values were associated with fewer trials to reverse. However, while in the simulations individuals needed fewer trials to reverse when we increased λ (less deviation from the learned association), the birds in the manipulation showed a decreased λ in their last reversal when they needed fewer trials to reverse. This suggests that λ is a constraint, rather than having a direct linear influence on the number of trials to reverse: birds with low λ still can reach the criterion in a small number of trials as long as they have a sufficiently high value of ϕ (see Figure 2). In line with this, across both manipulated and control birds, ϕ was more consistently associated with the number of trials individuals needed to reverse, and ϕ changed more than λ across reversals for the manipulated birds (Figure 3). The birds might have changed their learning rate ϕ because they repeatedly experienced an associative learning task, while the change in */lambda* might reflect that birds adapt to the serial reversal where the rewarded option changes every time they reach criterion so that their learned attractions are not completely reliable and it is beneficial to deviate from time to time.

For the ϕ values, we also observed a correlation between the ϕ estimated from an individual's performance in the first reversal and how much their ϕ changed toward the value for their performance in the last reversal (-0.4;Table 1: Model 17), while there is no such obvious relationship for λ (-0.15; Table 1: Model 18). For both ϕ and λ, unlike for the number of trials to reverse, we did not see that the individuals who had the largest values during the first reversal also always had the largest values during the last reversal. The manipulation changed both ϕ and λ, such that, across all birds, there was a negative correlation between ϕ and λ (mean estimate -0.39, 89% compatibility interval: -0.72 to -0.06).



a) Number of trials to to pass criterion

b) Phi: learning rate of attraction to either option

c) Lambda: rate of deviating from learned attractions

Figure 3. Comparisons of the different measures of performance in the reversal task for each of the 19 birds. The figure shows a) the number of trials to pass criterion for the first reversal (orange; all birds) and the last reversal (blue; only manipulated birds); b) the ϕ values reflecting the learning rate of attraction to the two options from the initial discrimination and first reversal (orange; all birds) and from the last two reversals (blue; manipulated birds); and c) the λ values reflecting the rate of deviating from the learned attractions to the two options from the initial discrimination and first reversal (orange; all birds) and from the last two reversals (blue; manipulated birds). Individual birds have the same position along the x-axis in all three panels. Birds that needed fewer trials to reverse their preference generally had higher ϕ values, whereas λ appeared to reflect whether any choices of the unrewarded color occurred throughout the trials or only at the beginning of a reversal. For the manipulated birds, their ϕ values changed more consistently than their λ values, and the ϕ values of the manipulated individuals were generally higher than those observed in the control individuals, while their λ values remained within the range observed in the control group.

The pairwise analyses above indicated that the number of trials in the last reversal was correlated with the number of trials in the first reversal, with ϕ, and with λ. The number of trials in the first reversal, ϕ, and λ were also correlated with each other (Figure 4). With the Bayesian approach, we used one model to estimate all potential links simultaneously to identify the pathways through which the variables interacted with each other (e.g., some variables might be correlated because both are influenced by a third variable). We therefore simultaneously estimated support for the following pathways:

- trials last reversal ~ trials first reversal + ϕ last reversal + λ last reversal
- trials first reversal ~ ϕ first reversal + λ first reversal
- ϕ last reversal ~ ϕ first reversal
- λ last reversal ~ λ first reversal

Results from this simultaneous estimation of the potential pathways show that our data best support that the ϕ from the initial learning and first reversal link to the number of trials to pass the first reversal, which, in turn, appear associated with how many trials they needed to pass their last reversal. The ϕ for the last reversal did not appear to provide any additional information about the number of trials in the last reversal, and λ was not directly associated with the number of trials birds needed to reverse (Table 1: Model 20) (Figure 4).

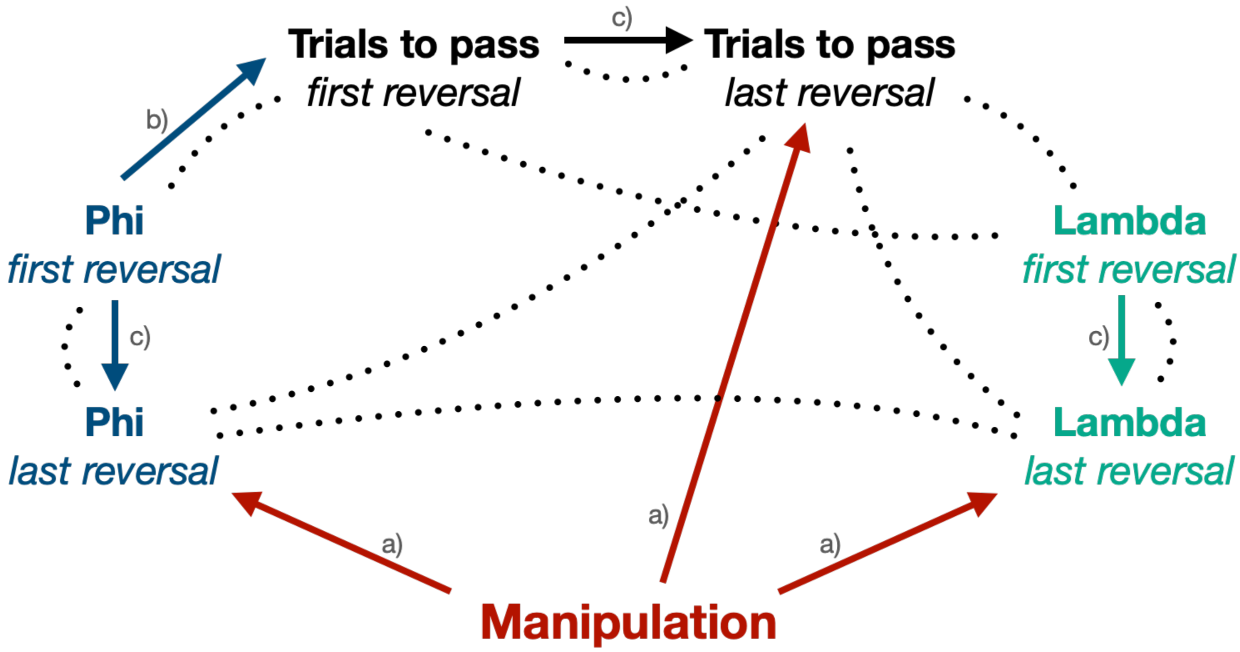


Figure 4. Graph showing the pathways between the number of trials to pass a reversal, ϕ, λ, and the flexibility manipulation (serial reversals). In the pairwise assessments (dotted lines), most of the variables were indicated as being associated with each other. The combined model identified which of these associations were likely to be direct (solid lines with arrows). The results from the combined model indicate that a) the manipulation worked, b) ϕ had a more direct influence on performance in the reversals than λ did, and c) individuals had some consistency both in their abilities and in their performance.

Table 1. Model outputs for the pairwise comparisons (models 1-5) and for the combined model (model 6) explaining the changes during the manipulation. SD=standard deviation, the 89% compatibility intervals are shown, n_eff=effective sample size, Rhat4=an indicator of model convergence (1.00 is ideal).

	Mean	SD	5.5%	94.5%	n_eff	Rhat4
model 1 trials improvement						
intercept per bird	-30.30	3.51	-35.65	-24.65	109	1.00
b * reversal	2.13	2.93	0.17	9.77	9	1.00
variance	6.54	2.42	0.23	9.41	10	1.00
model 2 trials improvement						
correlation intercept slope	0.34	0.39	-0.40	0.85	2452	1.00
model 3 phi improvement						
intercept per bird	0.00	0.02	-0.02	0.03	620	1.00
b * reversal	0.03	0.01	0.02	0.05	207	1.01
correlation intercept slope	-0.29	0.46	-0.93	0.52	1492	1.00
variance	0.02	0.01	0.01	0.03	184	1.01
model 4 lambda improvement						
intercept per bird	5.36	0.35	4.57	6.18	255	1.01
b * reversal	-1.10	0.30	-1.57	-0.64	260	1.01
correlation intercept slope	-0.08	0.44	-0.77	0.64	566	1.01
variance	0.85	0.20	0.58	1.19	648	1.00
model 5 performanceimprovement						
intercept	32.74	2.52	28.76	36.79	1362	1.00
b * phi improvement	10.63	3.09	5.68	15.31	1155	1.00
c * lambda improvement	5.58	3.03	0.73	10.20	1223	1.00
sigma	7.22	1.36	5.31	9.56	1322	1.00
MODEL 6 (combined)						
trials last ~ trials first	0.62	0.36	0.04	1.17	1166	1.00
trials last ~ phi last	-0.28	0.51	-1.07	0.54	1095	1.00
trials last ~ lambda last	-0.22	0.48	-0.98	0.55	1278	1.00
trials first ~ phi first	-1.04	0.15	-1.26	-0.80	1059	1.00
trials first ~ lambda first	0.18	0.16	-0.41	0.06	890	1.00
phi last ~ phi first	0.29	0.37	-0.31	0.86	1696	1.00
lambda last ~ lambda first	0.19	0.38	-0.41	0.79	1806	1.00

3) Association between ϕ and λ with performance on the multi-access boxes

We first modified the analyses from the preregistered analyses in the original article that assessed potential linear links between reversal learning and performance on the multi-access boxes by replacing the number of trials it took individuals to reverse with ϕ (learning rate of attraction to either option) and λ (rate of deviating from learned attractions) estimated from the reversal performances. These modified analyses did not find matches with any of the three previously detected correlations between reversal learning and performance on the two multi-access boxes (latency to attempt a locus on the plastic multi-access box, number of loci solved on the plastic and wooden multi-access boxes) (Table 2,3). We detected a different correlation: the latency to attempt a new locus on the wooden multi-access box was positively correlated with ϕ in the last reversal (Table 2: Model 28). This correlation appears to arise not because of a linear increase of the latency with increasing ϕ values, but because there were several individuals who had both a long latency and a large ϕ. However, there were also some individuals who had a long latency with a low ϕ (see below for additional analyses). This indicates that individuals who were faster to update their associations in reversal learning (higher ϕ, therefore needed fewer trials in their last reversal) took more time to attempt a new locus. Even though ϕ was closely associated with the number of trials a bird needed to reach the reversal criterion, we presumably could not recover the previous correlations because of our small sample sizes. In addition, we estimated ϕ and λ across at least one reversal (initial discrimination plus first reversal, or last two reversals for manipulated birds), whereas the previous analyses using the number of trials to reverse were based on a single reversal (first or last reversal).

Next, we additionally assessed whether ϕ and λ were associated with performance on the multi-access boxes in a non-linear way. For the manipulated birds, we found that during their last reversal there was a negative correlation between ϕ and λ, with individuals with higher ϕ values showing lower λ values. This negative correlation could lead to worse performance on the multi-access boxes for birds with intermediate values. Exploration of our data shows that, for the number of loci solved on both the plastic and the wooden multi-access boxes, there was a U-shaped association, particularly with λ values in the last reversal (Table 3: models 39 & 46) (Figure 5), with birds with intermediate values of λ solving fewer loci on both multi-access boxes. For the latency to attempt a new locus, there was also a U-shaped association, particularly with ϕ, with birds with intermediate values of ϕ showing shorter latencies to attempt a new locus (Table 2: models 25 & 32). Given that there was also a positive correlation between number of loci solved and the latency to attempt a new locus, there might be a trade off, where birds with extreme ϕ and λ values solve more loci, but need more time, whereas birds with intermediate values have shorter latencies, but solve fewer loci.

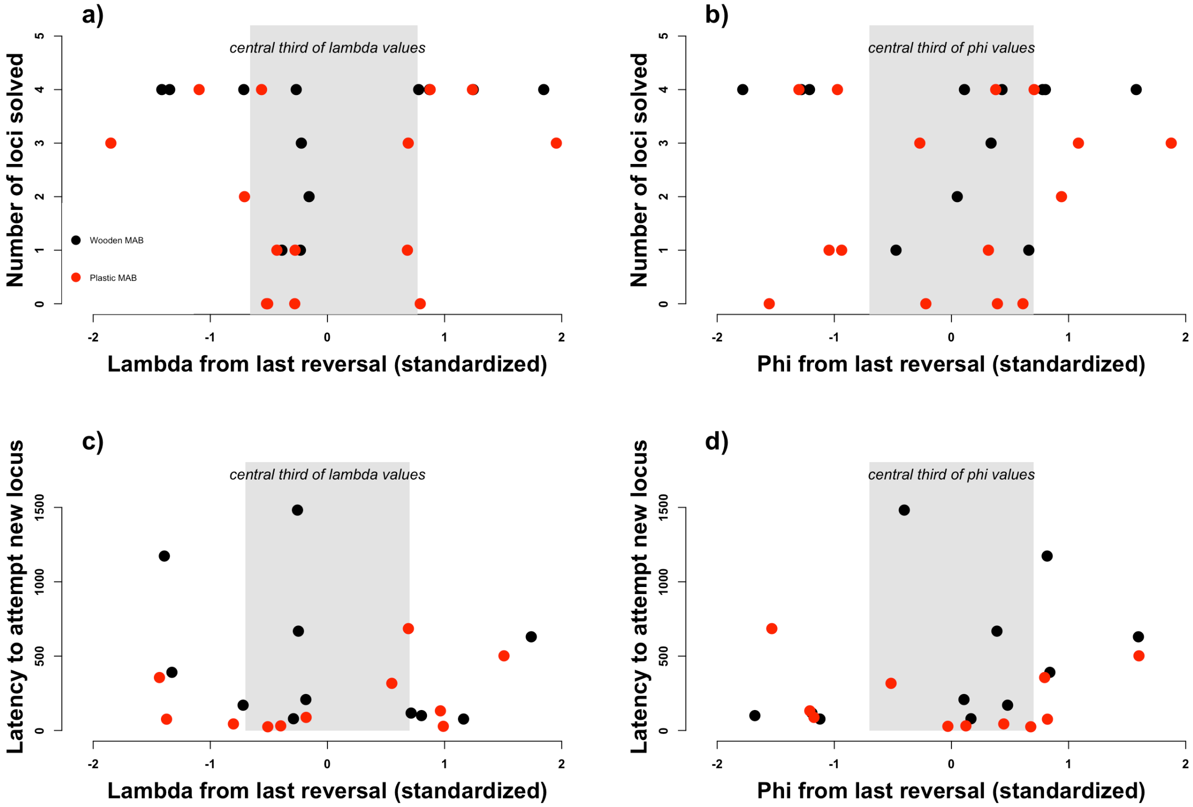


Figure 5. Relationships between ϕ and λ from the last reversal and performance on the wooden (black dots) and plastic (red dots) multi-access boxes. Birds with intermediate λ values in their last reversal (a) were less likely to solve all four loci on the multi-access boxes than birds with either high or low λ values. Birds who solved two or fewer loci on either box all fall within the central third of the λ values observed for the last reversal, while 12 of the 14 birds who solved all four loci fall outside this central range. An individual's ϕ and λ values change slightly between the top and bottom rows because values were standardized for each plot and not all individuals were tested on both boxes, therefore values changed relative to the mean of the points included in each plot. There are no clear relationships between (b) ϕ and the number of loci solved, (c) λ and the latency to attempt a locus, or (d) ϕ and the latency to attempt a new locus.

Table 2. Model outputs for the **latency** to switch loci after passing criterion on a different locus on the plastic (models 7-13) and wooden (models 14-20) multi-access boxes in relation to ϕ and λ. SD=standard deviation, the 89% compatibility intervals are shown, n_eff=effective sample size, Rhat4=an indicator of model convergence (1.00 is ideal), b=the slope of the relationship between loci solved or average switch latency and ϕ or λ.

	Mean	SD	5.5%	94.5%	n_eff	Rhat4
model 7 plastic						
intercept	4.99	0.31	4.51	5.48	1354	1
b * phi	-0.07	0.24	-0.45	0.31	1769	1
variance	0.80	0.31	0.39	1.34	1527	1
model 8 plastic						
intercept	4.97	0.30	4.50	5.46	1547	1
b * lambda	0.32	0.27	-0.10	0.74	1260	1
variance	0.87	0.34	0.40	1.46	1425	1
model 9 plastic						
intercept	4.99	0.31	4.52	5.46	1183	1
b * phi	0.33	0.27	-0.09	0.76	1736	1
c * lambda	-0.01	0.25	-0.41	0.42	1556	1
variance	0.83	0.32	0.39	1.42	1321	1
model 10 plastic						
intercept	5.02	0.31	4.51	5.49	886	1
b * phi * lambda	0.07	0.21	-0.25	0.42	1256	1
variance	0.80	0.30	0.39	1.33	1493	
model 11 plastic						
intercept	3.07	0.52	2.29	3.91	1210	1
b * abs(phi)	0.82	0.53	-0.02	1.68	1353	1
c * abs(lambda)	1.49	0.47	0.76	2.27	1226	1
variance	1.27	0.48	0.61	2.12	1456	1
model 12 plastic						
intercept	4.97	0.30	4.49	5.44	1105	1
b * phi first	0.16	0.26	-0.24	0.60	1376	1
variance	0.80	0.30	0.39	1.32	1218	1
model 13 plastic						
intercept	4.95	0.34	4.40	5.47	1284	1
b * lambda first	0.20	0.27	-0.53	0.88	1334	1
variance	0.80	0.34	0.36	1.41	1614	1
model 14 wooden						
intercept	5.73	0.28	5.27	6.15	1064	1
b * phi	0.47	0.30	0.00	0.94	1144	1
variance	1.06	0.44	0.48	1.86	1364	1
model 15 wooden						
intercept	5.76	0.30	5.28	6.21	1373	1
b * lambda	-0.25	0.25	-0.63	0.15	1415	1
variance	0.96	0.37	0.35	1.62	1532	1
model 16 wooden						
intercept	5.72	0.31	4.52	5.46	1183	1
b * phi	-0.29	0.27	-0.09	0.76	1736	1
c * lambda	0.47	0.25	-0.41	0.42	1556	1
variance	1.07	0.32	0.39	1.42	1321	1
model 17 wooden						
intercept	5.80	0.30	5.31	6.23	1259	1
b * phi * lambda	0.15	0.24	-0.22	0.56	1448	1
variance	0.92	0.35	0.44	1.54	1342	1
model 18 wooden						
intercept	5.07	0.53	4.20	5.90	739	1
b * abs(phi)	0.68	0.59	-0.23	1.68	867	1
c * abs(lambda)	0.39	0.77	-0.81	1.62	931	1
variance	0.78	0.34	0.34	1.42	932	1
model 19 wooden						
intercept	5.75	0.30	5.27	6.22	1172	1
b * phi first	0.30	0.33	-0.22	0.82	1467	1
variance	0.95	0.40	0.43	1.65	1216	1
model 20 wooden						
intercept	5.76	0.30	5.28	6.21	1250	1
b * lambda first	-0.21	0.25	-0.60	0.21	1233	1
variance	0.94	0.37	0.45	1.59	1537	1

Table 3. Model outputs for the **number of loci solved** on the plastic (models 21-27) and wooden (models 28-34) multi-access boxes in relation to ϕ and λ. SD=standard deviation, the 89% compatibility intervals are shown, n_eff=effective sample size, Rhat4=an indicator of model convergence (1.00 is ideal), b=the slope of the relationship between loci solved or average switch latency and ϕ or λ.

	Mean	SD	5.5%	94.5%	n_eff	Rhat4
model 21 plastic						
intercept	0.02	0.30	-0.45	0.50	1153	1
b * phi	0.24	0.26	-0.16	0.65	1463	1
model 22 plastic						
intercept	0.00	0.25	-0.40	0.41	1369	1
b * lambda	0.14	0.22	-0.21	0.49	1200	1
model 23 plastic						
intercept	4.99	0.31	4.52	5.46	1183	1
b * phi	0.33	0.27	-0.09	0.76	1736	1
c * lambda	-0.01	0.25	-0.41	0.42	1556	1
model 24 plastic						
intercept	5.02	0.31	4.51	5.49	886	1
b * phi * lambda	0.07	0.21	-0.25	0.42	1256	1
model 25 plastic						
intercept	-0.66	0.50	-1.45	0.15	947	1
b * abs(phi)	1.51	0.60	0.61	2.48	845	1
c * abs(lambda)	-0.55	0.58	-1.45	0.37	861	1
model 26 plastic						
intercept	0.02	0.26	-0.41	0.42	1313	1
b * phi first	0.20	0.22	-0.17	0.54	1624	1
model 27 plastic						
intercept	0.01	0.26	-0.41	0.42	1346	1
b * lambda first	0.29	0.23	-0.08	0.66	1536	1
model 28 wooden						
intercept	1.35	0.34	0.83	1.90	1329	1
b * phi	-0.08	0.27	-0.52	0.37	1268	1
model 29 wooden						
intercept	1.34	0.33	0.83	1.87	1566	1
b * lambda	0.20	0.27	-0.24	0.63	1444	1
model 30 wooden						
intercept	0.75	0.42	0.07	1.43	1186	1
b * phi	0.37	0.34	-0.18	0.92	1354	1
c * lambda	0.56	0.36	-0.01	1.14	1131	1
model 31 wooden						
intercept	0.92	0.38	0.34	1.53	966	1
b * phi * lambda	0.67	0.32	0.17	1.19	952	1
model 32 wooden						
intercept	0.40	0.50	-0.43	1.20	902	1
b * abs(phi)	1.52	0.75	0.33	2.70	827	1
c * abs(lambda)	0.43	0.67	-0.60	1.52	1002	1
model 33 wooden						
intercept	1.34	0.34	0.82	1.19	1259	1
b * phi first	0.05	0.28	-0.37	0.48	1434	1
model 34 wooden						
intercept	1.34	0.33	0.82	1.88	1283	1
b * lambda first	-0.11	0.27	-0.52	0.32	1111	1

DISCUSSION

Our post-hoc analyses indicate that applying a more mechanistic model to understand the behavior of great-tailed grackles in a serial reversal learning experiment can provide additional insights into the potential components of behavioral flexibility. The simulations showed that the Bayesian reinforcement learning model captures variation in the behavior and that the two key components phi, the learning rate, and lambda, the deviation rate, can be reliably estimated from the choices individuals make. The post-hoc Bayesian analyses of the grackle data revealed that the primary component of flexibility that was manipulated was the learning rate (ϕ), which more than doubled between the first and last reversals. The learning rate also explained more of the interindividual variation in how many trials individuals needed to reach criterion during a reversal. Finally, linking these two components of behavioral flexibility to the performance on the multi-access boxes suggests that birds with intermediate values of λ solve fewer loci on both multi-access boxes, and birds with intermediate values of ϕ have shorter latencies to attempt a new locus. The two key components of the Bayesian reinforcement learning model, the learning rate ϕ and the deviation rate lambda, appear to reflect differences and changes in behavioral flexibility: individuals with a higher learning rate are more likely to update their previously learned associations and individuals with a higher deviation rate are more likely to explore new options.

The Bayesian reinforcement learning model we applied in these post-hoc analyses appears to be an accurate representation of the behavior of grackles in the serial reversal experiment. In the previous application of this model to reversal learning data from a different population, Blaisdell et al. (2021) found that the choices of grackles were consistent with what this model predicts. Here, we add to this by showing that the model can identify variation in performance, and in particular reveal how individuals change their behavior through the manipulation series of multiple reversals. Previous analyses of reversal learning performance of wild-caught animals have often focused on summaries of the choices individuals make (Bond et al., 2007), setting criteria to define success and how much individuals sample/explore versus acquire/exploit (Federspiel et al., 2017). These approaches are more descriptive, making it difficult to predict how variation in behavior might transfer to other tasks. While there have been attempts to identify potential rules that individuals might learn during serial reversal learning (Spence, 1936; Warren, 1965), it is unclear how to use these rule-based approaches for cases like the grackles, who, while apparently shifting toward a win-stay/lose-shift rule, did not fully land on this rule (C. Logan et al., 2022). More recent analyses of serial reversal learning experiments of laboratory animals have specifically focused on determining when individuals might switch to more specialized rules (Jang et al., 2015). In such analyses, some individuals were found to learn more specific rules about the serial reversal because such specialized reversal rule models seemed to fit the behavior better than the reinforcement learning models because individuals appeared to switch toward the win-stay/lose-shift strategy rather than continuously updating their attractions Metha et al. (2020). However, these specialized strategies only seem to emerge in over-trained animals who have experienced a very large number of trials (Bartolo & Averbeck, 2020), whereas individuals such as the grackles in our experiment are more likely to use the more general learning strategies that are reflected in the reinforcement learning models. Accordingly, the changes in behavior that can be observed in the serial reversal experiments we analyzed are likely better captured by the changes in the learning rate and the deviation rate than by switches in rules.

The increase in the learning rate during the manipulation might reflect that birds recognize that this is an environment where new information should be prioritized over previously learned associations. This change in the learning rate over the serial reversal experiment in the grackles matches what has been reported for squirrel monkeys (Bari et al., 2022). In contrast, the rate of deviating from learned preferences (λ) did not correlate with the number of trials to reverse. The change in the rate of deviation during the manipulation might indicate that individuals learned about the serial nature of the reversal experiment, that they should deviate from their previous attractions as soon as the reward changes. While there were individual differences in learning and deviation rates (McCune et al., 2022), all individuals appeared to change depending on their experiences. The manipulation pushed individuals to levels that were already observed in some individuals at the beginning of the experiment, suggesting that individuals might also change their behavioral flexibility in response to their experiences in their natural habitats.

The analyses linking ϕ and λ to the performance on the multi-access boxes suggest that birds might use different strategies to solve a larger number of loci on the multi-access box. The negative correlation between

ϕ and λ prompted us to explore whether the relationship between these two variables and the performance on the multi-access boxes could be non-linear. We did detect U-shaped relationships between ϕ and λ and how individuals performed on the multi-access boxes. First, birds with intermediate ϕ values showed shorter latencies to attempt a new locus. This could reflect that birds with high ϕ values take longer because they formed very strong attractions to the previously rewarded locus, while birds with small ϕ values take longer because they do not update their attraction even though the first locus is no longer rewarded. Second, we found that birds with intermediate values of λ solved fewer loci. This could indicate that birds with a small λ are more likely to explore new loci while birds with a large λ are more likely to stop returning to an option that is no longer rewarded. We also found further interactions in that birds who solved more loci had longer latencies to attempt a locus. An alternative interpretation could therefore be that birds with high ϕ and small λ solve fewer loci because they do not switch attractions quickly but need less time because they explore more, while birds with small ϕ and high λ solve fewer loci because they do not learn but need less time because they already act on small differences in attractions. In addition, it is also possible that performance on the multi-access boxes relies on other cognitive abilities in which individuals may differ. For example, we previously found that grackles who are faster to complete go no-go, an inhibition task, were slower to switch loci on the multi-access boxes (Logan et al., 2021). As such, variation in self control may affect performance on flexibility and innovation tasks by decreasing exploratory behaviors. However, all these analyses are exploratory and based on a small sample, so these interpretations are speculative and further investigation is needed to understand how potential cognitive abilities shape performance on these different tasks.

Overall, these post-hoc analyses indicate the potential benefits of applying a more mechanistic model to the serial reversal learning paradigm. Inferring the potential underlying cognitive processes can allow us to make clearer predictions about how the experiments link to behavioral flexibility. In particular, we could expect that the previously observed differences in whether reversal learning performance links with performance in other traits like innovation or inhibition Logan (2016) could be linked to differences in whether the learning rate or the deviation plays a larger role in the reversal performance in a given species and in particular for the other trait. The mechanistic model can also help with setting criteria to better design the serial reversal experiments because the changes in attraction can be used to reflect whether individuals have formed a sufficient association to reverse the rewarded option (Logan et al., 2022). We believe that considering such mechanistic models more generally can offer an opportunity for the field of comparative cognition to better fulfill its potential.

G. AUTHOR CONTRIBUTIONS

Lukas: Hypothesis development, simulation development, data interpretation, write up, revising/editing.

McCune: Added MAB log experiment, protocol development, data collection, data interpretation, revising/editing, materials.

Blaisdell: Prediction revision, assisted with programming the reversal learning touchscreen experiment, protocol development, data interpretation, revising/editing.

Johnson-Ulrich: Prediction revision, programming, data collection, data interpretation, revising/editing.

MacPherson: Data collection, data interpretation, revising/editing.

Seitz: Prediction revision, programmed the reversal learning touchscreen experiment, protocol development, data interpretation, revising/editing.

Sevchik: Data collection, revising/editing.

Logan: Hypothesis development, protocol development, data collection, data analysis and interpretation, write up, revising/editing, materials/funding.

H. FUNDING

This research is funded by the Department of Human Behavior, Ecology and Culture at the Max Planck Institute for Evolutionary Anthropology.

I. CONFLICT OF INTEREST DISCLOSURE

We, the authors, declare that we have no financial conflicts of interest with the content of this article. CJ Logan is a Recommender and, until 2022, was on the Managing Board at PCI Ecology. D Lukas is a Recommender at PCI Ecology.

J. ACKNOWLEDGEMENTS

We thank our PCI Ecology recommender, Aurelie Coulon, and reviewers, Maxime Dahirel and Andrea Griffin, for their feedback on this preregistration; Julia Cissewski for tirelessly solving problems involving financial transactions and contracts; Sophie Kaube for logistical support; and Richard McElreath for project support.

K. REFERENCES

Bari, B. A., Moerke, M. J., Jedema, H. P., Effinger, D. P., Cohen, J. Y., & Bradberry, C. W. (2022). Reinforcement learning modeling reveals a reward-history-dependent strategy underlying reversal learning in squirrel monkeys. *Behavioral Neuroscience*, *136*(1), 46.

Bartolo, R., & Averbeck, B. B. (2020). Prefrontal cortex predicts state switches during reversal learning. *Neuron*, *106*(6), 1044–1054.

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*.

Bond, A. B., Kamil, A. C., & Balda, R. P. (2007). Serial reversal learning and the evolution of behavioral flexibility in three species of north american corvids (gymnorhinus cyanocephalus, nucifraga columbiana, aphelocoma californica). *Journal of Comparative Psychology*, *121*(4), 372.

Brecht, K. F., Legg, E. W., Nawroth, C., Fraser, H., & Ostojić, L. (2021). The status and value of replications in animal behavior science. *Animal Behavior and Cognition*, *8*(2), 97–106.

Chow, P. K. Y., Lea, S. E., & Leaver, L. A. (2016). How practice makes perfect: The role of persistence, flexibility and learning in problem-solving efficiency. *Animal Behaviour*, *112*, 273–283. https://doi.org/10.1016/j.anbehav.2015.11.014

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should i stay or should i go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933–942.

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

Dhawan, S. S., Tait, D. S., & Brown, V. J. (2019). More rapid reversal learning following overtraining in the rat is evidence that behavioural and cognitive flexibility are dissociable. *Behavioural Brain Research*, *363*, 45–52.

Doya, K. (2007). Reinforcement learning: Computational theory and biological mechanisms. *HFSP Journal*, *1*(1), 30. https://doi.org/10.2976/1.2732246/10.2976/1

Farrar, B. G., Altschul, D. M., Fischer, J., Van Der Mescht, J., Placì, S., Troisi, C. A., Vernouillet, A., Clayton, N. S., & Ostojić, L. (2020). Trialling meta-research in comparative cognition: Claims and statistical inference in animal physical cognition. *Animal Behavior and Cognition*, *7*(3), 419.

Farrar, B. G., Boeckle, M., & Clayton, N. S. (2020). Replications in comparative cognition: What should we expect and how can we improve? *Animal Behavior and Cognition*, *7*(1), 1.

Farrar, B. G., Ostojić, L., & Clayton, N. S. (2021). The hidden side of animal cognition research: Scientists' attitudes toward bias, replicability and scientific practice. *PloS One*, *16*(8), e0256607.

Federspiel, I. G., Garland, A., Guez, D., Bugnyar, T., Healy, S. D., Güntürkün, O., & Griffin, A. S. (2017). Adjusting foraging strategies: A comparison of rural and urban common mynas (acridotheres tristis). *Animal Cognition*, *20*(1), 65–74.

Frankenhuis, W., Panchanathan, K., & Smaldino, P. E. (2022). *Strategic ambiguity in the social sciences.*

Griffin, A. S., Guez, D., Lermite, F., & Patience, M. (2013). Tracking changing environments: Innovators are fast, but not flexible learners. *PloS One*, *8*(12), e84907.

Jang, A. I., Costa, V. D., Rudebeck, P. H., Chudasama, Y., Murray, E. A., & Averbeck, B. B. (2015). The role of frontal cortical and medial-temporal lobe brain areas in learning a bayesian prior belief on reversals. *Journal of Neuroscience*, *35*(33), 11751–11760.

Lambert, M., Farrar, B., Garcia-Pelegrin, E., Reber, S., & Miller, R. (2022). ManyBirds: A multi-site collaborative open science approach to avian cognition and behavior research. *Animal Behavior and Cognition*, *9*(1), 133–152.

Lea, S. E., Chow, P. K., Leaver, L. A., & McLaren, I. P. (2020). Behavioral flexibility: A review, a model, and some exploratory tests. *Learning & Behavior*, *48*(1), 173–187.

Logan, C. J. (2016). Behavioral flexibility in an invasive bird is independent of other behaviors. *PeerJ*, *4*, e2215.

Logan, C. J., McCune, K., MacPherson, M., Johnson-Ulrich, Z., Rowney, C., Seitz, B., Blaisdell, A., Deffner, D., & Wascher, C. (2021). *Are the more flexible great-tailed grackles also better at behavioral inhibition?* https://doi.org/10.31234/osf.io/vpc39

Logan, C. J., Shaw, R., Lukas, D., & McCune, K. B. (2022). *How to succeed in human modified environments.* http://corinalogan.com/ManyIndividuals/mi1.html

Logan, C., Blaisdell, A., Johnson-Ulrich, Z., Lukas, D., MacPherson, M., Seitz, B., Sevchik, A., & McCune, K. (2022). Behavioral flexibility is manipulatable and it improves flexibility and problem solving in a new context. *EcoEvoRxiv*. https://doi.org/https://doi.org/10.32942/osf.io/5z8xs

Logan, CJ, McCune, KB, Chen, N, & Lukas, D. (2020). Implementing a rapid geographic range expansion - the role of behavior and habitat changes. *In Principle Acceptance by PCI Ecology of the Version on 6 Oct 2020*. http://corinalogan.com/Preregistrations/gxpopbehaviorhabitat.html

McCune, K., Blaisdell, A., Johnson-Ulrich, Z., Lukas, D., MacPherson, M., Seitz, B., Sevchik, A., & Logan, C. (2022). Repeatability of performance within and across contexts measuring behavioral flexibility. *EcoEvoRxiv*. https://doi.org/10.32942/osf.io/kevqp

McElreath, R. (2020). *Rethinking: Statistical rethinking book package.*

Metha, J. A., Brian, M. L., Oberrauch, S., Barnes, S. A., Featherby, T. J., Bossaerts, P., Murawski, C., Hoyer, D., & Jacobson, L. H. (2020). Separating probability and reversal learning in a novel probabilistic reversal learning task for mice. *Frontiers in Behavioral Neuroscience*, *13*, 270.

Mikhalevich, I., Powell, R., & Logan, C. (2017). Is behavioural flexibility evidence of cognitive complexity? How evolution can inform comparative cognition. *Interface Focus*, *7*(3), 20160121. https://doi.org/10.1098/rsfs.2016.0121

Morand-Ferron, J., Reichert, M. S., & Quinn, J. L. (2022). Cognitive flexibility in the wild: Individual differences in reversal learning are explained primarily by proactive interference, not by sampling strategies, in two passerine bird species. *Learning & Behavior*, *50*(1), 153–166.

Neftci, E. O., & Averbeck, B. B. (2019). Reinforcement learning in artificial and biological systems. *Nature Machine Intelligence*, *1*(3), 133–143.

R Core Team. (2017). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org

Spence, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, *43*(5), 427.

Team, S. D.others. (2019). Stan modeling language users guide and reference manual. *Technical Report*.

Tecwyn, E. C. (2021). Doing reliable research in comparative psychology: Challenges and proposals for improvement. *Journal of Comparative Psychology*, *135*(3), 291.

Warren, J. M. (1965). The comparative psychology of learning. *Annual Review of Psychology*, *16*(1), 95–118.