

1 Flexibility training does not increase behavioral diversity of Florida
2 scrub-jays

3 McCune, KB^{1,2} Barve, S³

4 2026-03-20



7 **Affiliations:**

- 8 1) Institute for Social, Behavioral and Economic Research, University of California Santa Barbara, Santa
9 Barbara, CA, USA
10 2) College of Forestry, Wildlife and Environment, Auburn University, Auburn, AL, USA
11 3) Archbold Biological Station, Venus, FL, USA.

12 *Corresponding author:* kelseybmccune@gmail.com

14 **This is a post-study manuscript of the programmatic [registered report](#) that has been pre-study
15 peer reviewed and received an In Principle Acceptance on 8 Sep 2022 by:**

16 Chris Chambers (2022) The role of behavioural flexibility in promoting resilience to human environmen-
17 tal impacts. *Peer Community in Registered Reports*. <https://rr.peercommunityin.org/articles/rec?id=200>.
18 Reviewers: Gloriana Chaverri, Vedrana Šlipogor, and Alizée Vernouillet



20 *See the reproducible manuscript ([Rmd](#)) version for the code*

22 ABSTRACT

23 Human modifications of environments are expanding, causing global changes that other species must adjust
24 to or suffer from. Behavioral flexibility (hereafter ‘flexibility’) could be key to coping with rapid change.
25 Behavioral research can contribute to conservation by determining which behaviors can predict the ability
26 to adjust to human modified environments and whether these can be manipulated. When research that
27 manipulates behavior in a conservation context occurs, it primarily trains a specific behavior to address one
28 potential threat improve individual success in the wild. However, training a domain general cognitive ability,
29 such as flexibility, has the potential to change a whole suite of behaviors, which could have a larger impact
30 on influencing success in adjusting to human modified environments. This project asks whether flexibility
31 can be increased by experimentally increasing environmental heterogeneity and whether such an increase can
32 help species succeed in human modified environments. We explore whether it is possible to take insights from
33 highly divergent species and apply them to address critical conservation challenges. This pushes the limits
34 in terms of understanding how conserved these abilities may be and to what extent they can be shaped
35 by the environment. We aim to 1) conduct flexibility interventions in flexible species that are successful
36 in human modified environments (California scrub-jays or blue jays) to understand how flexibility relates
37 to success; and 2) implement these interventions in a vulnerable species (Florida scrub-jays) to determine
38 whether flexibility as a generalizable cognitive ability can be trained and whether such training improves
39 success in human-modified environments. This research will significantly advance our understanding of the
40 causes and consequences of flexibility, linking behavior to environmental change, cognition, and success in
41 human modified environments through a comparative framework.

42 REGISTERED REPORT DETAILS

- 43 • **Level of bias = 6:** This registered report was written (Jul 2021-May 2022), and revised after round
44 one of peer review at Peer Community in Registered Reports (Jul 2022) prior to collecting any data.
- 45 • **Programmatic registered report:** This is one of three Stage 2 articles that will result from the
46 original, comprehensive Stage 1 registered report: one for toutouwai, one for grackles, and one for jays
47 (this current manuscript).
- 48 • **Deviations from the Stage 1 registered report:** We were unable to sample disturbance-resilient
49 blue jays or California scrub-jays and so we cannot address J.Q2 (one of four specific research questions)
50 at this time. However, to begin to test the question of whether behavior relates to success in human-
51 modified environments, we conducted an unregistered analysis comparing behavior of Florida scrub-jays
52 (FLSJ; a species that is overall disturbance-sensitive) that were persisting in suburban areas to FLSJ
53 in natural habitat managed without human disturbance.

54 INTRODUCTION

55 Human modified environments are expanding (Goldewijk, 2001; X. Liu et al., 2020; Wu et al., 2011), causing
56 global changes that other species must adjust to or suffer from (Alberti, 2015; Chejanovski et al., 2017; Ciani,
57 1986; Federspiel et al., 2017). Behavioral flexibility (hereafter ‘flexibility’) could be key for adjusting to such
58 change: individuals interact with their environment through behavior, making it crucial to an ecologically
59 valid understanding of how species adjust to environmental changes (Lee & Thornton, 2021). One of the top
60 priorities for behavioral research to maximize conservation progress is to determine which cognitive abilities
61 and behaviors can predict the ability to adjust to human modified environments and whether these can be
62 manipulated (Moseby et al., 2016). The rare research that manipulates behavior in a conservation context
63 usually focuses on training specific behaviors (for example, predator recognition through predator exposure)
64 to improve individual success in the wild (Jolly et al., 2018; Moseby et al., 2012; Ross et al., 2019; West et
65 al., 2018; see review in Tetzlaff et al., 2019). However, training a general cognitive ability, such as flexibility
66 – the ability to rapidly adapt behavior to changes through learning throughout the lifetime (see the theory
67 behind this definition in Mikhalevich et al., 2017) – has the potential to change a whole suite of behaviors

68 and more broadly influence success in adjusting to human modified environments. Recent evidence supports
69 this hypothesis: as far as we are aware, our previous research in the great-tailed grackle was the first to
70 show that flexibility can be manipulated using serial reversal learning of color preferences, and that the
71 manipulated individuals were more flexible during foraging (Logan et al., 2025), in a different experimental
72 context (locus switching on a puzzlebox) as well as being more innovative (solved more loci on a puzzlebox)
73 (Logan et al., 2023).

74 Environments where informational cues about resources vary in a heterogeneous (but non-random) way
75 across space and time are hypothesized to open a pathway for species to functionally detect and react to
76 such cues via flexibility (Mikhalevich et al., 2017). Human modified environments likely provide a different
77 set of informational cues that vary heterogeneously across space and time, and the species that are successful
78 in such environments are likely those who are able to detect and track such cues. Because heterogeneous
79 environments are hypothesized to select for flexibility (Wright et al., 2010), we expect that exposing indi-
80 viduals to experimentally increased environmental heterogeneity will train individuals to be more flexible,
81 which will then increase their success in such environments (Fig. 1). Success can relate to any number of
82 variables regarding the usage of and investment in resources and response to threats, from improved foraging
83 efficiency to increased dispersal and survival within human modified environments, to placing nests in more
84 protective locations. Whether a measure of success is predicted to relate to flexibility depends on what is
85 already known about the particular population and their particular environment.

86



87

88 **Figure 1.** The theory behind this research illustrated by a directed acyclic graph (DAG), which is a
89 theoretical model of the causal relationships among the key variables in our investigation. Based on the
90 theoretical background provided by Mikhalevich et al. (2017), we assume that more heterogeneity causes
91 more flexibility, which then causes more success in human modified environments.

92

93 This investigation asks whether a method of experimentally increasing environmental heterogeneity through
94 serial reversal learning leads to increased behavioral flexibility and greater fitness in human modified envi-
95 ronments. Serial reversal learning tasks have been performed with a wide diversity of species (birds: Bond
96 et al., 2007; bumblebees: Strang & Sherry, 2014; stingrays: Daniel & Schluessel, 2020). There is variation
97 across individuals and species in their performance, however almost all previous studies show that individuals
98 improve their flexibility if the reversal intervention is given multiple times in sequence (rats: Mackintosh et
99 al., 1968; guppies: Lucon-Xiccato & Bisazza, 2014; poison frogs: Y. Liu et al., 2016; grackles: Logan et al.,
100 2023). We aim to conduct a flexibility intervention in species that are successful in human modified envi-
101 ronments (California scrub-jays or blue jays) to understand how flexibility relates to success, and implement
102 these interventions in a vulnerable species (Florida scrub-jays).

103 While we do not examine the potential spread of the post-manipulation success behaviors from manipulated
104 individuals to individuals that are not involved in our studies, we acknowledge that this is a possibility worthy
105 of future investigation. Manipulating the flexibility of a few individuals could have population-level effects
106 because significant research on social information use in birds (e.g., Valente et al., 2021) demonstrates the
107 potential for the manipulated behavior to disseminate to conspecifics (for example, if manipulated individuals
108 are faster at locating new resources, which could attract the attention of conspecifics, or if unmanipulated
109 individuals copy the manipulated individuals' nesting or foraging locations). Indeed, there is evidence for
110 social learning in the focal species for this investigation (K. B. McCune et al., 2022; Midford et al., 2000).
111 In the event that social learning is not used by a given population to spread the behaviors of manipulated
112 individuals, investing in the training of specific individuals to increase their success in the wild could still
113 have conservation impacts. In some cases, it is possible to train many individuals in a population or a species
114 because there are not many individuals left (Greggor et al., 2021). It is also possible to train all individuals

115 involved in a conservation management event such as a translocation (Greggor et al., 2021). Therefore, there
116 can still be significant population consequences even if each individual needs to be trained to achieve the
117 goal.

118 This comparative approach will ultimately reveal the role of behavioral flexibility in adaptation to environ-
119 mental change and whether there is a causal relationship between this general cognitive trait and behavioral
120 interactions with human-modified environments, such as foraging and habitat use. Our previous research
121 on great-tailed grackles provided support that this causal relationship does exist in a generalist, invasive
122 species. However, the research presented here will evaluate the generalizability of this technique across dis-
123 tinct species differentially affected by anthropogenic change. Consequently, the results from this research
124 will substantially advance our understanding of the causes and consequences of flexibility, linking behavior
125 to environmental change, cognition, and success in human modified environments to facilitate conservation
126 of declining species.

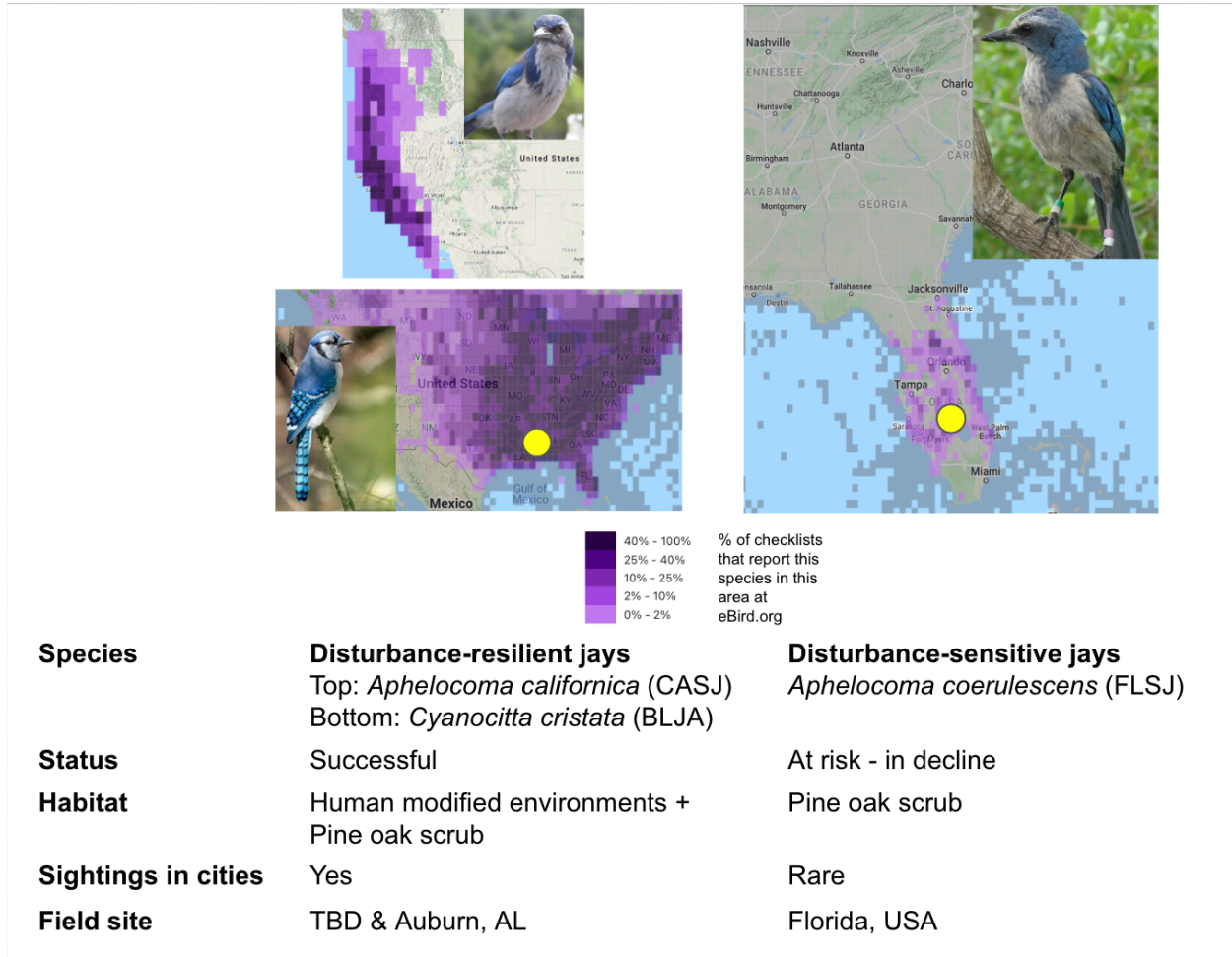
127 RESEARCH QUESTIONS

128 **Can behavioral flexibility in individuals be increased by increasing environmental**
129 **heterogeneity? If so, does increased flexibility help individuals succeed in human**
130 **modified environments?**

131 **Prediction 1:** Flexibility can be increased in individuals and such an increase **improves the likelihood**
132 **of success in human modified environments.** This would indicate that the abilities involved in tracking
133 changing resources in the environment are the same as or related to the abilities involved in succeeding in
134 human modified environments. It would also indicate that flexibility is trainable and that such training
135 could be a useful conservation tool for threatened and endangered species.

136 **Prediction 1 alternative 1:** Flexibility can be increased in individuals, but such an increase **does not**
137 **improve the likelihood of success** in human modified environments. This would indicate that species
138 associated with human modified environments form this association for reasons other than their flexibility,
139 and that threatened species are likely not very successful in human modified environments for reasons
140 unrelated to their ability to change their behavior with changing circumstances. An alternative could be
141 that the changes induced by the increase in flexibility do not persist for sufficiently long times to make a
142 difference on the subsequent likelihood of success (changes in great-tailed grackles were still present for four
143 weeks after the manipulation and longer time periods were not attempted so the threshold is unknown Logan
144 et al., 2022).

145 **Prediction 1 alternative 2:** Flexibility can be increased in some populations, but not others. This would
146 indicate that **flexibility manipulations may not work for all populations**, and that the effectiveness
147 of such experiments should first be tested in the population of interest before including such an intervention
148 in a conservation plan. If flexibility is not manipulatable in threatened populations, this would indicate that
149 they are likely not very successful in human modified environments because of their inability to change their
150 behavior with changing circumstances, and that flexibility is not trainable. If flexibility is not manipulatable
151 in populations that are successful in human modified environments, this could indicate that they might have
152 used flexibility in the past when originally forming the association, but the need to maintain flexibility in
153 their repertoire is no longer necessary (Wright et al., 2010). In populations where flexibility is not manipu-
154 latable, this would indicate that the abilities involved in tracking changing resources in the environment are
155 independent of the abilities involved in succeeding in human modified environments.



159 **Figure 2.** Comparing the species involved in this investigation relative to their geographic range and
 160 association with human modified habitats. The yellow dots represent field site locations. Photo credit:
 161 CASJ, Corina Logan; BLJA, Rhododendrites; FLSJ, VvAndromedavV.

163 **Background**

164 Jay species exhibit a diversity of social systems and success in colonizing suburban and urban areas (Fig.
 165 2). California scrub-jays (*Aphelocoma californica*, hereafter “CASJ”) and blue jays (*Cyanocitta cristata*,
 166 hereafter “BLJA”) are singular, monogamous breeders that are increasing in abundance, expanding their
 167 range sizes, and highly successful in natural, suburban, and urban areas (Blair, 1996; Curry et al., 2017).
 168 We therefore consider these “disturbance-resilient” (DR) jay species. In contrast, the Florida scrub-jay (*A.*
 169 *coerulescens*; hereafter “FLSJ”) is a “disturbance-sensitive” (DS) jay species that is threatened, endemic,
 170 and range-restricted to xeric oak scrub habitat in Florida (Woolfenden & Fitzpatrick, 1996).

171 These species forage primarily on mast (acorns, hazelnuts, etc.) that they cache throughout their territory,
 172 which makes it available to eat year-round. They are also opportunistic omnivores and specifically need
 173 high-fat and high-protein arthropods to feed to nestlings and fledglings (Curry et al., 2017). Nesting and
 174 foraging substrates can be drastically different in human modified environments compared to natural areas

175 [e.g., predominance of non-native vegetation; Tuomainen & Candolin (2011)], and it is unknown whether
176 suburban and urban jays are able to persist in these environments through behavioral adjustments. The
177 DS jay species, the FLSJ, can persist in suburban habitats after conversion from xeric oak scrub, however
178 suburban populations of FLSJ steadily decline (Bowman pers. comm.). This is potentially due to the
179 presence of suboptimal habitat resulting from fire suppression (Woolfenden & Fitzpatrick, 1996), higher
180 rates of brood reduction through nestling starvation (Shawkey et al., 2004), and the lack of nutritionally
181 complete prey items (Shawkey et al., 2004) in suburban habitats. It is possible that behavioral flexibility in
182 habitat use and foraging breadth underlies the ability of some FLSJ to persist in human-dominated areas.

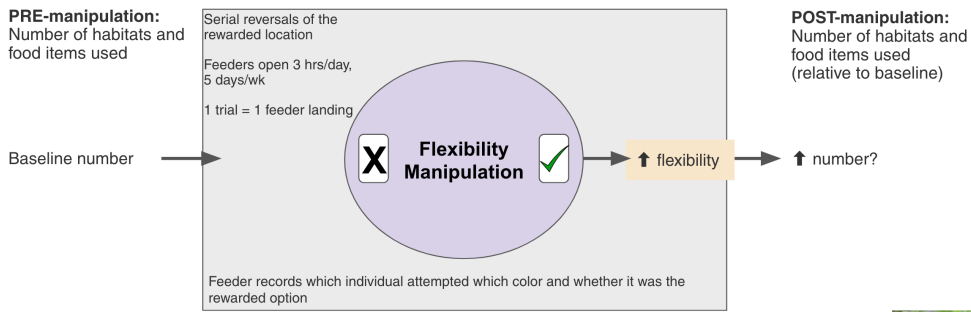
183 We aimed to compare behavioral flexibility within species, between suburban and natural populations to
184 determine whether variation in flexibility relates to variation in presence in these habitats. Subsequently, we
185 planned to compare flexibility between DS and DR jay species to determine whether this trait is related to the
186 greater success of DR jay species, like the CASJ and BLJA, in human-dominated areas (but see Deviations
187 section). Lastly, we tested whether manipulating flexibility increases the foraging and microhabitat breadth
188 of jays in suburban and natural environments. Manipulating the flexibility of a subset of individuals has
189 the potential to affect the population because previous research demonstrates that these species have the
190 capacity to use foraging information discovered by others (social learning) to flexibly change their behavior
191 (K. B. McCune et al., 2022; Midford et al., 2000).

192 Tailored Research Questions

193 For all research questions, Table 1 summarizes our predictions, analysis plans, interpretations for the various
194 directions the results could go, and the hypotheses that could be contradicted given the various outcomes.

- 195 • **J.Q1: Do jay populations in human modified areas differ in baseline behavioral flexibility**
196 **compared to populations in natural areas?** We will investigate this question by comparing
197 performance on serial reversal learning in the wild between jays in natural areas and jays in human
198 modified areas.
- 199 • **J.Q2: Are disturbance-resilient (DR) jays more behaviorally flexible than disturbance-**
200 **sensitive (DS) jays?** We will investigate this question by comparing performance on serial reversal
201 learning in the wild between DR and DS jay species.
- 202 • **J.Q3: Does manipulating behavioral flexibility alter the number of microhabitats used?**
203 We will investigate this question by tracking their presence in a variety of microhabitats before and
204 after manipulating their flexibility using serial reversal learning in the wild. We only count that a
205 microhabitat was used if the individual had at least 5% of their data points there. This prevents a
206 microhabitat from being counted even if an individual was simply moving through it, and therefore
207 not necessarily using it.
- 208 • **J.Q4: Does manipulating flexibility alter the number of different food items taken by**
209 **jays?** We will investigate this question by tracking the various food items they take before and after
210 manipulating their flexibility using serial reversal learning in the wild.

Do flexibility manipulated individuals differ in how they interact with their environment?



211 *Within individual comparison before and after flexibility manipulation* All Individuals color banded and radio tagged

212 **Figure 3.** The reversal learning experiment in a group context (Design 2 from the ManyIndividuals reg-
 213 istered report) tailored to the jay research questions. The white rectangles represent feeder locations, the
 214 feeder with the X is in the unrewarded location while the feeder with the green check is the rewarded location.

215 **Table 1.** Study design for the jay research. Note that we are unable to test Q2 because it was
 216 not possible to sample CASJ or BLJA within the timeframe of this study. References that were
 217 not already cited in the introduction: Galbraith et al. (2015), Lapiedra et al. (2017), Rice et
 218 al. (2003), Emery & Clayton (2004), Sol et al. (2002), Peterson et al. (2011), Grinnell (1917).

Question	Hypothesis	Sampling plan	Analysis plan	Rationale for deciding test sensitivity to confirm/disconfirm hypothesis	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes	
1. Do jay populations in human modified areas differ in baseline behavioral flexibility compared to populations in natural areas?	Prediction 1.1: Suburban jays are more flexible than jays in natural areas	Simulations using bespoke Bayesian models in Logan et al. (2021) showed a high likelihood of detecting	Bayesian model: Response: phi and lambda Explanatory: Habitat (suburban/ natural) (see Analysis Plan)	Contrasts will determine whether the before and after conditions differed from each other. We will conclude there is a difference if the confidence interval does not cross zero. False positives: the power analyses suggest that false positives are unlikely even with small sample sizes. Accordingly, we will interpret any contrast that does not cross zero as indicating an effect. False negatives: the power analyses suggest that, especially with small sample sizes, we will not have sufficient power to exclude the possibility that an effect exists even though our model does not indicate an effect (contrast crosses zero).	This implies that flexibility is related to the ability to occupy human modified environments where spatial and temporal heterogeneity of resources is high. This implies that human modification of the environment has led to less spatial and temporal heterogeneity of resources. For example, the prevalence of bird feeders in suburban areas leads to consistently available food. This implies that additional behavioral (e.g. boldness, Lapiedra et al., 2017) or genetic traits may facilitate success in human modified environments.	Selection for exploitation of supplementary food (Galbraith et al. 2015), where individuals learn to depend on anthropogenic food sources and are less likely to flexibly sample alternative resources. (1) Flexibility facilitates adapting to environmental change (see Introduction) (1) The urban filter (Lapiedra et al. 2017) where novel anthropogenic pressures select for flexible individuals.	
	Prediction 1.2: Suburban jays are less flexible than jays in natural areas	differences with a sample size of 15 when mean differences in phi were at least 0.01 and lambda at least 3	(see Analysis Plan)				
	Prediction 1.3: There is no difference in flexibility between suburban jays and jays in natural areas	(see Analysis Plan)					
2. Are disturbance resilient (DR) jays more behaviorally flexible than disturbance sensitive (DS) jays?	Prediction 2.1: DR jays are more flexible than DS jays	Simulations using bespoke Bayesian models showed a high likelihood of detecting differences with a sample size of 20 when mean difference in the proportion of microhabitats used was at least 0.1 (standard deviation=0.1) or 0.15 (SD=0.2) (see Analysis Plan)	Bayesian model: Response: phi and lambda Explanatory: Species (see Analysis Plan)	Accordingly, we will interpret any contrast that does not cross zero as indicating an effect. False negatives: the power analyses suggest that, especially with small sample sizes, we will not have sufficient power to exclude the possibility that an effect exists even though our model does not indicate an effect (contrast crosses zero).	This difference may explain the range expansion and greater success of DR jays in human modified environments. This implies that flexibility is not related to success in human modified environments and that flexibility may instead be related to a different, unknown social or environmental characteristic. For example, the cooperative breeding system of the DS jay species, the Florida scrub-jay, may favor increased flexibility for responding to group mates' behavior. This implies flexibility is not related to success in human modified environments and the level of flexibility is potentially an evolutionary conserved trait from a corvid common ancestor (Emery & Clayton 2004).	Range expansion instead relates to ecological niche differentiation (Rice et al. 2003), where DR jays evolved to occupy a niche that more closely resembles human modified environments than DS jays. (1) (1)	
	Prediction 2.2: DR jays are less flexible than DS jays						(see Analysis Plan)
	Prediction 2.3: DR jays and DS jays are equally flexible						(see Analysis Plan)
3. Habitat: Does manipulating behavioral flexibility alter the number of microhabitats used?	Prediction 3.1: Flexibility can be increased and such an increase alters daily habitat use to include more variety of habitats	Simulations using bespoke Bayesian models showed a high likelihood of detecting differences with a sample size of 20 when mean difference in the proportion of microhabitats used was at least 0.1 (standard deviation=0.1) or 0.15 (SD=0.2) (see Analysis Plan)	Bayesian model: Response: Number of microhabitats used per individual Explanatory: Condition (before/after) Random: Condition ID (see Analysis Plan)	Increasing behavioral flexibility, with a serial reversal learning manipulation, increases the likelihood the individual will sample new areas while foraging. Increasing behavioral flexibility potentially leads to increased foraging breadth or use of resources within one habitat, rather than leading to sampling across habitat types. This suggests that the cognitive ability behavioral flexibility may not generalize to all domains (e.g. may relate to foraging but not habitat use). Alternatively, it could indicate that they might have used flexibility in the past when originally forming the association, but the need to maintain flexibility in their repertoire is no longer necessary, or that changes induced by the increase in flexibility do not persist for sufficiently long times to make a difference on the subsequent likelihood of success.	Increasing behavioral flexibility, with a serial reversal learning manipulation, increases the likelihood the individual will sample new areas while foraging. Increasing behavioral flexibility potentially leads to increased foraging breadth or use of resources within one habitat, rather than leading to sampling across habitat types. This suggests that the cognitive ability behavioral flexibility may not generalize to all domains (e.g. may relate to foraging but not habitat use). Alternatively, it could indicate that they might have used flexibility in the past when originally forming the association, but the need to maintain flexibility in their repertoire is no longer necessary, or that changes induced by the increase in flexibility do not persist for sufficiently long times to make a difference on the subsequent likelihood of success.	(2) Habitat preferences and the foraging niche are fixed within species because each species evolves within a specific ecological niche (Grinnell 1917; Peterson et al. 2011). Behavioral flexibility facilitates the use of novel habitats and invasion success through dietary generalism (Sol et al. 2002) (1) (3) Behavioral flexibility is a general cognitive ability (see Introduction)	
	Prediction 3.2: Flexibility can be increased and such an increase alters daily habitat use to decrease the variety of habitats used						(see Analysis Plan)
	Prediction 3.3: Flexibility can be increased but has no effect on the variety of habitats used						(see Analysis Plan)
4. Foraging: Does manipulating flexibility alter the number of different food items taken by jays?	Prediction 4.1: Flexibility can be increased and such an increase alters daily foraging breadth to include more variety of food items	Simulations using bespoke Bayesian models showed a high likelihood of detecting differences with a sample size of 20 when mean differences in the number of foods taken were at least 1 (and a standard deviation of 2) (see Analysis Plan)	Bayesian model: Response: Number of foods taken per individual Explanatory: Condition (before/after) Random: Condition ID (see Analysis Plan)	Increasing behavioral flexibility, with a serial reversal learning manipulation, increases the likelihood the individual will sample new food sources while foraging. Increasing behavioral flexibility potentially leads to increased use of the same type of food items across habitat types, rather than leading to sampling more food items within one habitat. This suggests that the cognitive ability behavioral flexibility may not generalize to all domains (e.g. may relate to habitat use but not foraging breadth). Alternatively, it could indicate that they might have used flexibility in the past when originally forming the association, but the need to maintain flexibility in their repertoire is no longer necessary, or that changes induced by the increase in flexibility do not persist for sufficiently long times to make a difference on the subsequent likelihood of success.	Increasing behavioral flexibility, with a serial reversal learning manipulation, increases the likelihood the individual will sample new food sources while foraging. Increasing behavioral flexibility potentially leads to increased use of the same type of food items across habitat types, rather than leading to sampling more food items within one habitat. This suggests that the cognitive ability behavioral flexibility may not generalize to all domains (e.g. may relate to habitat use but not foraging breadth). Alternatively, it could indicate that they might have used flexibility in the past when originally forming the association, but the need to maintain flexibility in their repertoire is no longer necessary, or that changes induced by the increase in flexibility do not persist for sufficiently long times to make a difference on the subsequent likelihood of success.	(2) (1) (3)	
	Prediction 4.2: Flexibility can be increased and such an increase alters daily foraging breadth to decrease the variety of food items taken						(see Analysis Plan)
	Prediction 4.3: Flexibility can be increased but has no effect on foraging breadth						(see Analysis Plan)

219

220 METHODS

221 Planned sample

222 We studied Florida scrub-jays in unfragmented, fire-maintained habitat at Archbold Biological Station (nat-
 223 ural areas), as well as in the adjacent suburbs of Lake Placid, FL (human modified areas). Much of the
 224 study population at Archbold Biological Station was already individually identifiable with color leg bands.
 225 We caught an additional 24 jays using walk-in traps and mist nets; collected their biometric data, and a

226 blood sample; applied colored leg bands; and released them at their point of capture. We collected data
227 on pre- and post-manipulation success measures (foraging and habitat use, see below) and conducted the
228 experiment within the non-breeding season to control for potential temporal differences in the environment
229 and behavior. We planned to use radio tags to find jays for observing before and after manipulation success
230 measures. However, Florida scrub-jays are highly territorial, and we were always able to find the jays to
231 collect observational data without the radio tags.

232 To determine whether the flexibility manipulation has influenced the ability of jays to persist in human
233 modified environments, we planned for half of the total sample of jays from areas with access to human-
234 supplemented food (i.e. private property, university campus, parks adjacent to neighborhoods with feeders)
235 and the other half from natural areas (wildlife management areas, reserves). To address Q1, testing for
236 differences in baseline behavioral flexibility between jays in natural or human-modified environments, the
237 power analysis (see Supplementary materials S1) showed a high likelihood of detecting differences with a
238 minimum sample of 15. We were able to quantify baseline behavioral flexibility in 14 Florida scrub-jays (8
239 in natural areas, 6 in human-modified areas).

240 To address Q3 and Q4, we tracked baseline behavior and changes after the flexibility manipulation via all-
241 occurrence sampling during focal observations (Altmann, 1974) that last for 60 min. During observations, we
242 documented all occurrences of microhabitats used within the territory and foraging behavior (see Observa-
243 tional methods, below). Our planned minimum sample size was 20 individuals per species with a minimum
244 of 4 observations per individual (at least 2 per month) pre-manipulation and a minimum of 4 per individual
245 (at least 2 per month) post-manipulation (at least 320 observations in total). We were only able to conduct
246 the serial reversal manipulation on 9 color-banded individuals, but we met or exceeded the minimum sample
247 size for observations before and after the manipulation. There was one exception where an individual disap-
248 peared after the manipulation and was not found during extensive searching of the territory and surrounding
249 areas and so was presumed dead. We only collected one-third of the necessary post-manipulation data on
250 this individual.

251 **Observational methods**

252 During observations, we recorded each microhabitat the individual is present in and all food items consumed.
253 Before data analysis, to ensure that we were only including the microhabitats individuals use (rather than
254 just pass through), we filtered the data to include only microhabitats that account for at least 5% of their
255 data points. Although we may not observe every possible microhabitat or food item the individual may
256 use, by equally sampling before and after the manipulation we can detect changes in habitat use. We
257 counterbalanced observations across the morning and afternoon for all individuals in the study. In this way,
258 we collected a random sample of data from active and inactive time periods for all individuals.

259 Microhabitat types in the suburban habitat (<100m from human structure) included: vertical human struc-
260 ture (e.g. building, bench), native vegetation, non-native vegetation, grass, impervious surface, and dirt. In
261 the natural habitat (>100m from human structure), microhabitat types included all previous categories, but
262 not human structure or impervious surface. All categories were further defined by whether the subject was
263 high (>3m) or low (<3m). For example, grass and impervious surfaces can occur above 3m if grass is on
264 the roof of a building, and if an individual is walking on the impervious surface of an upper floor parking
265 garage.

266 Food types were broken down into plant (seed, fruit, human-provided, or unknown plant) and animal (insect
267 larva, adult insect, amphibian, crustacean, reptile, mammal, bird, egg, human-provided, or unknown animal).
268 “Human-provided” indicates any food item that was acquired from a store at some point and is left out by
269 humans. For example, sunflower seeds would be considered human-provided if they are in the form of bird
270 seed or a human snack. Sunflower seeds would only be counted in the “seed” category if the bird is seen
271 eating it from a plant.

272 Additionally, for all observations we used binoculars so that we could remain far enough from the focal
273 individual that our presence was not affecting their behavior. Because that distance can be different for each
274 individual and species, we hesitate to give a specific number. However, if at any point the focal individual

275 showed that it was affected by our presence, by looking directly at the researcher, alarm calling or startling,
276 then we ended the observation immediately, dropped all data from that observation session and attempted
277 the observation again the next day.

278 **Flexibility manipulation (Design 2 from the registered report)**

279 Our approach involved individual jays participating in a serial reversal learning task in the wild. We set up
280 remote-controlled feeders at multiple study sites, each containing several jay territories, spaced at least 2km
281 apart: 2 sites in natural habitat, 2 in human-modified habitat.

282 **Feeder habituation**

283 One feeder at each location contained food and the experimenter used the remote to feed the focal individual
284 every time it approached. Habituation consisted of 2-hour sessions between 8a-3p on a minimum of 4
285 consecutive days or until at least 1 banded jay per territory had visited the feeder at least 3 times without
286 showing fear behaviors.

287 **The experiment**

288 Prior to the flexibility manipulation, we collected the 4 minimum observation sessions to determine the
289 baseline values for microhabitat use and foraging breadth. Afterwards, we set up the remote-controlled
290 feeders to initiate the serial reversal learning flexibility manipulation phase. All feeders contained one type
291 of high value food (i.e., peanuts). In addition, feeders were opaque and always had food in them to eliminate
292 a confound with performance due to olfactory differences between the feeders that could be introduced if only
293 the rewarded feeders had food in them. If a feeder needed to be refilled, we refilled all feeders consecutively
294 in the same time period and for the same amount of time even if that feeder did not need much or any food
295 (in these cases, we pretended to fill the feeder). This eliminated potential confounds in jay performance from
296 cues provided by a differential amount of experimenter attention given to the feeders during refilling.

297 At the beginning of serial reversal learning, one of the two feeders, at a consistent location (north or south,
298 east or west) within the territory, was randomly selected to dispense food when the focal jay approached. If
299 the subject visited the non-rewarded feeder, this incorrect choice was recorded, but the feeder did not open.
300 Reversal sessions for the manipulation treatment consisted of 30-120 min sessions per day per territory, up
301 to 4 days per week. Each visit by the focal jay to the correct or incorrect feeder was considered a trial. Jays
302 passed a given reversal when they correctly chose the rewarded feeder in at least 10 trials out of the most
303 recent 12 trials (see S1 of the supplementary materials for details on the simulations that determined that
304 threshold). After passing one reversal, we switched to rewarding jays at the other feeder location. Serial
305 reversals continued until jays passed two consecutive reversals in 12 trials or less (see S2 of the supplementary
306 materials for details on the simulations that determined that threshold). At this criterion, the jays were
307 considered to have increased their behavioral flexibility.

308 After the manipulation was complete for the focal individual, we again conducted the all-occurrences focal ob-
309 servation sessions to measure microhabitat use and foraging breadth. We began collecting post-manipulation
310 data on an individual as soon as it passed the manipulation because it is unknown for how long any potential
311 effects of the manipulation will last.

312 For more social species, like the Florida scrub-jay, who may experience the experiment in groups, there
313 is the potential that individuals who were not in the manipulation condition or who have not yet passed
314 the manipulation condition will learn about post-manipulation success behaviors from the manipulated
315 individuals who passed the experiment. We are ultimately interested in determining whether we can change
316 success behaviors of threatened populations or species as a result of the flexibility manipulation. If part of
317 this change is the result of social learning from some of the manipulated individuals, it will still result in a
318 change even if we do not quantify what percentage of the mechanism comes from individual learning as a

319 result of the manipulation or social learning after the manipulation. If there is a change in success measures
320 between before and after the manipulation, the manipulation will have been the cause in either case.

321 Open data

322 The data will be published in the Knowledge Network for Biocomplexity’s data repository.

323 Protocols

- 324 • [Jay protocol and data sheet templates](#)
- 325 • [Protocol for applying radio tags and conducting GPS tracks](#) from McCune KB et al. (2020)
- 326 • [Jay processing protocol](#)

327 Analysis plan

328 We run analyses in R [current version 4.5.0; R Core Team (2017)] using the following R packages: rethinking
329 (McElreath, 2020a), rstan (Stan Development Team, 2020), cmdstanr (Gabry & Češnovar, 2022), knitr (Xie,
330 2018), and irr (Gamer et al., 2012). We describe below the Bayesian models we developed to address each
331 of the specific research questions. All Bayesian models were developed using McElreath (2020b) as a guide.
332 Code for running these models is available in supplementary materials S4.

333 **J.Q1 Do jay populations in human modified areas differ in baseline behavioral flexibility com- 334 pared to populations in natural areas?**

335 The model

336 We used the reversal learning Bayesian model in Logan CJ et al. (2020) to simulate and analyze population
337 differences in reversal learning, and calculate our ability to detect differences between populations. The
338 model “accounts for every choice made in the reversal learning experiment and updates the probability of
339 choosing either option after the choice was made depending on whether that choice contained a food reward
340 or not. It does this by updating three main components for each choice: an attraction score, a learning rate
341 (ϕ), and a rate of deviating from learned attractions (λ)” (Logan CJ et al., 2020).

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

343 Equation 1 “tells us how attractions to different behavioral options $A_{i,j,t+1}$ (i.e., how preferable option i is to
344 the bird j at time $t + 1$) change over time as a function of previous attractions $A_{i,j,t}$ and recently experienced
345 payoffs $\pi_{i,j,t}$ (i.e., whether they received a reward in a given trial or not). Attraction scores thus reflect
346 the accumulated learning history up to this point. The (bird-specific) parameter ϕ_j describes the weight
347 of recent experience. The higher the value of ϕ_j , the faster the bird updates their attraction. It thus can
348 be interpreted as the *learning or updating rate of an individual*. A value of $\phi_j = 0.04$, for example, means
349 that receiving a single reward for one of the two options will shift preferences by 0.02 from initial 0.5-0.5
350 attractions, a value of $\phi_j = 0.06$ will shift preferences by 0.03 and so on” (Blaisdell et al., 2021).

351 **Equation 2 (λ):** $P(i)_{t+1} = \frac{\exp(\lambda_j A_{i,j,t})}{\sum_{m=1}^2 \exp(\lambda_j A_{m,j,t})}$.

352 Equation 2 “expresses the probability an individual j chooses option i in the next round, $t + 1$, based on the
353 latent attractions. The parameter λ_j represents the *rate of deviating from learned attractions* of an individual
354 (also called inverse temperature). It controls how sensitive choices are to differences in attraction scores.
355 As λ_j gets larger, choices become more deterministic, as it gets smaller, choices become more exploratory
356 (random choice if $\lambda_j = 0$). For instance, if an individual has a 0.6-0.4 preference for option A, a value of

357 $\lambda_j = 3$ means they choose A 65% of the time, a value of $\lambda_j = 10$ means they choose A 88% of the time and
358 a value of $\lambda_j = 0.5$ means they choose A only 53% of the time” (Blaisdell et al., 2021).

359 We used the c values as the response variable in the Bayesian model to examine whether there were differences
360 in flexibility between the habitats:

$$361 y \sim \alpha[\text{habitat}]$$

362 y is the response variable (ϕ_j and λ_j , which are extracted from the correct and incorrect choices in the serial
363 reversals). There is one intercept, α , per habitat (suburban or natural) and we will estimate the habitat’s
364 average and standard deviation of the response variable.

365 **J.Q2 Are disturbance-resilient jays more flexible than disturbance-resistant jays?**

366 We were not able to collect data on disturbance-resilient species within the timeframe of this study and the
367 funding of our grant. However, we were able to conduct an *unregistered analysis* testing for differences in
368 behavior between disturbance-resilient Florida scrub-jay individuals persisting in human-modified areas and
369 conspecifics in natural habitat lacking human disturbance.

370 **The model**

371 We used poisson regression to assess whether jays in human-modified areas foraged on more kinds of foods and
372 used more substrate types relative to jays in natural areas. Models were not overdispersed or heteroscedastic.

373 **J.Q3 More flexible = use more microhabitats?**

374 **The model**

375 *Bayesian model with a normal distribution:*

$$376 \text{habitatuse} \sim \alpha[\text{ind}] + \beta[\text{ind}]*\text{before}$$

377 habitatuse is the response variable: the total number of different microhabitats used per individual. There
378 was one intercept, α , and one slope β per individual, which was estimated for the two conditions, before
379 (and after) the manipulation. ID is nested within condition as a random effect because there was more than
380 one data point per individual: each individual has a data point in the before condition and in the after
381 condition. A normal distribution was used because the response variable is a sum without an expected skew
382 to the curve (see Figure 10.6 in McElreath, 2020b).

383 **J.Q4 More flexible = more food types?**

384 **The model**

385 *Bayesian model with a normal distribution:*

$$386 y \sim \alpha[\text{ind}] + \beta[\text{ind}]*\text{before}$$

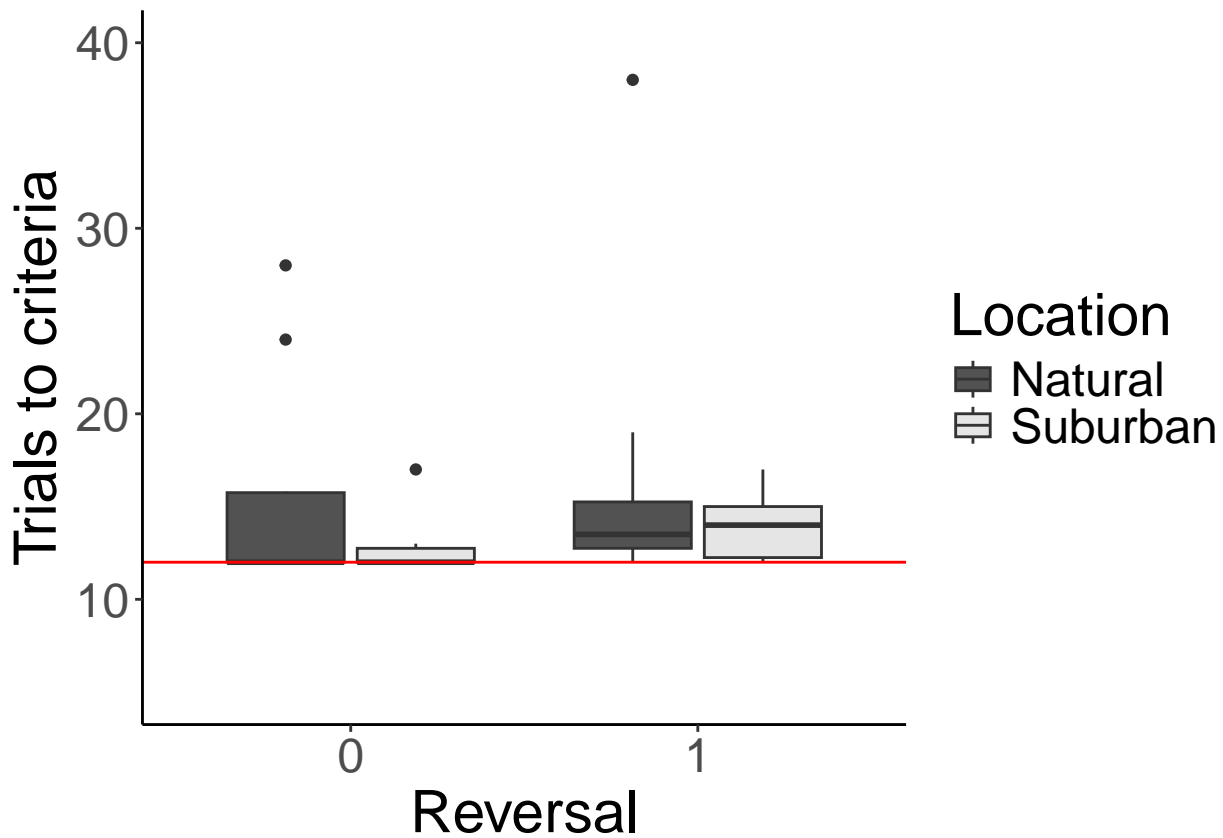
387 y is the response variable: the total number of different food types taken per individual. There will be one
388 intercept, α , and one slope β per individual, which will be estimated for the two conditions, before (and
389 after) the manipulation. ID is nested within condition as a random effect because there is more than one
390 data point per individual: each individual has a data point in the before condition and in the after condition.
391 A normal distribution was used because the response variable is a sum without an expected skew to the
392 curve (see Figure 10.6 in McElreath, 2020b).

393 **RESULTS**

394 Very few Florida scrub-jay were found in human-modified areas at our study site in Lake Placid, FL.
395 Across two years we collected focal observations and baseline behavioral flexibility on 6 individuals in these
396 human-modified areas and 8 jays in natural areas at Archbold Biological Station. A smaller subset of 9
397 jays (all 8 jays in the areas and 1 from the human-modified areas) received the serial research learning
398 flexibility manipulation because it was not feasible to conduct this longer experiment on more jays within
399 the constraints of our field season.

400 **J.Q1 Do jay populations in human modified areas differ in baseline behavioral**
401 **flexibility compared to populations in natural areas?**

402 There was no difference in baseline behavioral flexibility performance between jays in suburban or natural
403 areas. This was true when we considered either of the two behavioral flexibility components, ϕ_j , the rate
404 of learning the newly rewarded option ($\beta = -0.23$, $sd = 1.27$, $CI = -2.22 - 1.4$, $p > 0.05$), or λ_j , the rate of
405 deviation from previously rewarded options ($\beta = -0.01$, $sd = 0.18$, $CI = -0.28 - 0.28$, $p > 0.05$).



406 Figure 3: Before the flexibility training, jays in the suburban and natural habitats performed exceptionally
407 well on the baseline behavioral flexibility task and did not differ in ability. These boxplots show the
408 performance as the number of trials to reach the passing criterion on baseline associative learning (Reversal
409 0) and behavioral flexibility (Reversal 1) for jays in the natural (dark grey) and human modified (light grey)
410 areas. We use trials to pass, rather than or in this plot for more interpretable illustration of the speed of
411 performance on this task. The horizontal red line illustrates the minimum number of trials to pass criteria
412 (10 out of the most recent 12 trials correct).
413

414 **J.Q2 Are disturbance-resilient jays more flexible than disturbance-resistant jays?**

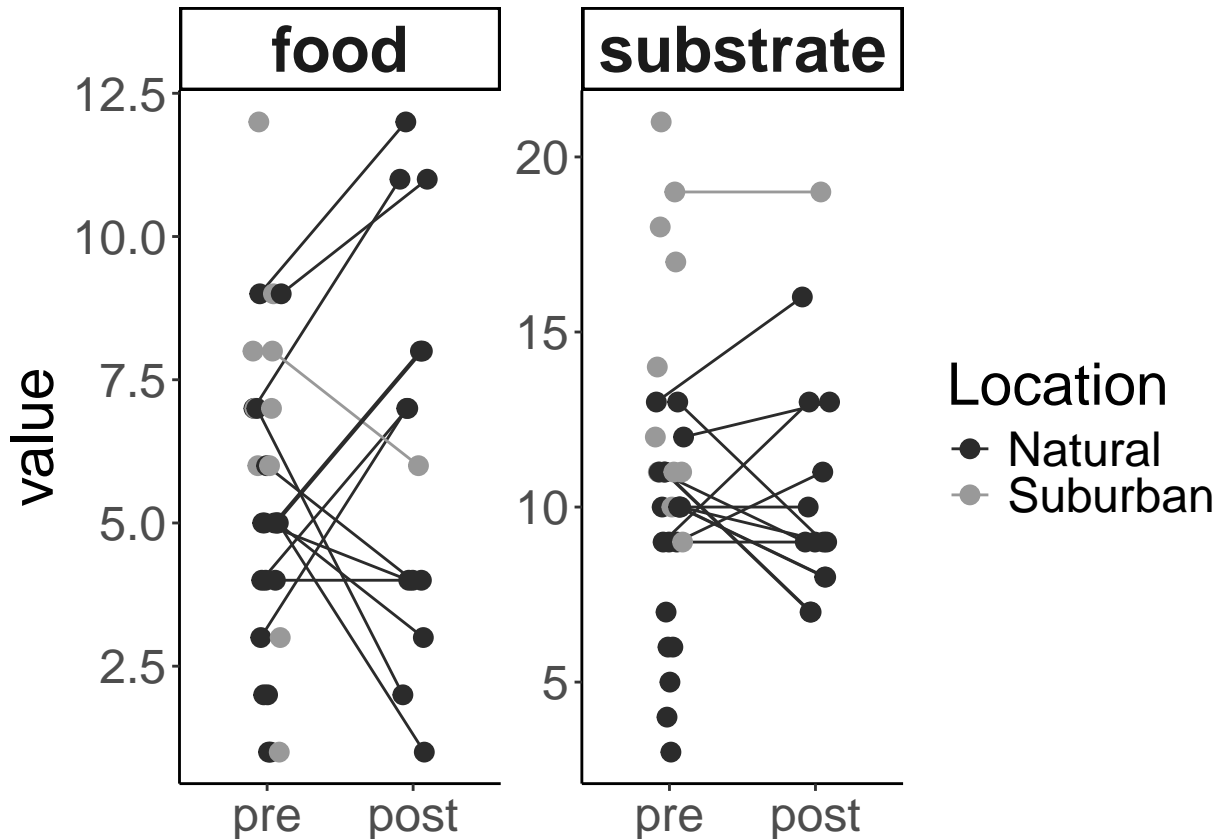
415 We were not able to test this question within the timeframe and funding of the grant. However, we conducted
416 an unregistered analysis using Poisson models to compare foraging and habitat use behavior of Florida scrub-
417 jays in human modified areas to those in natural areas. In other words, within this disturbance sensitive
418 species, a subset of individuals were resilient to disturbance, at least during the timeline of this study.
419 The disturbance-resilient individuals in the suburbs were not more diverse foragers ($\beta = 0.26 \pm 0.19$, $p =$
420 0.15), but did use a greater variety of microhabitats ($\beta = 0.52 \pm 0.16$, $p = 0.001$) compared to the jays in
421 undisturbed, natural habitat at Archbold Biological Station.

422 **J.Q3 More flexible = use more microhabitats?**

423 Even though all jays increased the speed with which they switched their preference in the serial reversal
424 learning training, there was no effect on subsequent habitat use diversity (before relative to after log-odds
425 difference = -0.06 , CI = $-0.41 - 0.30$, $p > 0.05$; Fig. 4).

426 **J.Q4 More flexible = more food types?**

427 Similarly, after the serial reversal flexibility training, jays were not more likely to eat a greater diversity of
428 food types ($\beta = 0.26$, CI = $-0.45 - 0.96$, $p > 0.05$; Fig. 4).



429 Figure 4: Jays did not predictably change their foraging or microhabitat use behaviors after they underwent
430 the flexibility training manipulation. Values indicate the number of different foods or substrates jays
431 were observed to use during the all-occurrences focal observation session before (pre) and after (post) we
432 conducted serial reversal learning to increase behavioral flexibility. Each dot represents a jay and lines
433

434 connect the dots of jays with data both pre and post. We collected pre observations on some jays for which
435 we were not able to conduct the flexibility training and so we did not need to collect post observations.

436 DISCUSSION

437 Anthropogenic change is creating unpredictable heterogeneity in the environment that requires some level
438 of fast, adaptable learning in wildlife populations. In some species, the general cognitive trait, behavioral
439 flexibility, is thought to underlie the ability to cope with rapid changes in human-modified environments
440 (Sol et al., 2002; Vardi & Berger-Tal, 2022). Our previous work in the great-tailed grackle supported this
441 hypothesis because grackles that were trained to be more flexible foraged on more diverse food and used more
442 distinct behaviors to forage in the wild (Logan et al., 2025). The grackle is an urban-adapted generalist species
443 that has demonstrated adaptability to environmental change through a rapid range expansion (Wehtje, 2003).
444 We tested whether the same training to increase behavioral flexibility in a threatened species, the Florida
445 scrub-jay (FLSJ), would lead to more adaptable foraging and habitat use behaviors in the wild. However,
446 we found no effect of the behavioral flexibility training in this species.

447 Unlike great-tailed grackles, FLSJ populations have rapidly declined coincident with anthropogenic change
448 and there is little evidence for adaptation to human-modified environments. Yet, for many species, suburban
449 environments can also provide significantly more foraging opportunities through the provision of novel foods,
450 ornamental plants or agricultural activities that increase survival and reproductive success. For example, use
451 of food at bird feeders has been linked to population growth and range expansions (Greig et al., 2017; Veech
452 et al., 2011). One such species is the California scrub-jay, a congener to the FLSJ, that similarly prefers
453 xeric scrub-oak habitat along the west coast of the United States. In contrast to the FLSJ, this species has
454 been able to adapt to persist in highly developed areas (e.g., Los Angeles, CA) and expand its range north,
455 while taking advantage of bird feeders (Curry et al., 2017).

456 In our population of FLSJ, a small population does persist in suburban areas. Anecdotally, homeowners
457 report frequently seeing these individuals at their backyard bird feeders or other human-provided food
458 sources. We compared foraging and habitat use of these individuals to FLSJ in pristine, unfragmented and
459 fire-maintained scrub habitat at Archbold Biological Station. Jays in both habitat types foraged on equally
460 diverse food items, but suburban jays used more distinct kinds of microhabitats (substrate type and height)
461 relative to the jays at Archbold. In other words, suburban jays did not completely restrict their habitat
462 use to the small remnant fragments of undeveloped scrub in the suburbs. This demonstrates that FLSJ are
463 capable of adapting their habitat use behavior to include novel, human-modified environments. However,
464 when we compared baseline behavioral flexibility between the two jay populations, suburban jays were not
465 more behaviorally flexible than Archbold jays. Consequently, it is likely that behavioral flexibility, as we
466 measured it through spatial reversal learning, does not underlie adaptation to human-disturbance in this
467 species.

468 This was supported by our finding of no change in foraging and habitat use behavior after undergoing
469 behavioral flexibility training. Our results support our Prediction 1 alternative 1; all jays improved their
470 reversal learning speed to meet the criterion, indicating they had increased their behavioral flexibility. But
471 there was no consistent trend in the change in food item or microhabitat diversity measured before and after
472 the training. There are a few potential explanations for why the training worked in great-tailed grackles
473 but led to no behavior change in FLSJ. First, it is possible that the serial reversal learning task and passing
474 criterion were too easy for FLSJ and so the training did not lead to a large enough change in flexibility to be
475 seen in resulting behavior. We determined the criteria for when an individual was considered to have passed
476 a given reversal from simulations using the performance of captive great-tailed grackles in color cue reversal
477 in the previous experiment (supplementary material S1). While the criterion to pass in the previous study
478 was 17 correct out of the 20 most recent trials, simulations indicated we could detect a significant change in
479 preference for the newly rewarded option after 10 correct choices out of the most recent 12 trials. Half of
480 the jays reached this criterion in 12 or 13 trials (average = 15.6, range = 12 - 38) which is almost as fast
481 as possible (Fig. 3). FLSJ have excellent spatial memory because they are dependent on cached food for
482 survival and reproductive success (Beauchamp et al., 2025). Consequently, the spatial cue was likely easier

483 to learn than a color cue. Testing individuals in the wild can also impact performance (K. B. McCune et al.,
484 2019) to yield the faster reversal speeds relative to the captive grackles. In support of this, a new study on
485 grackles in the wild is more comparable to the FLSJ results because they used the same passing criterion.
486 Preliminary evidence shows wild grackles are slower than FLSJ to pass the first reversal ($n = 8$, average =
487 33 trials, range = 13 - 50; Logan, pers comm), though the range of performance overlaps significantly with
488 FLSJ. Moreover, in a previous study of color reversal learning, captive FLSJ required a similar amount of
489 time (3 days) to pass the reversal (Bebus et al., 2016). Together, this suggests that the spatial cue, rather
490 than testing context, led to the fast performance of jays on reversal learning.

491 Another potential reason we failed to see a change in natural behavior after the flexibility training is that
492 jays were already using the majority of foods and substrates available. Due to time constraints, size and
493 accessibility of the suburban population, the majority of jays (8 out of 9 total) that underwent the flexibility
494 training were from Archbold Biological Station where the scrub-oak habitat is maintained at high quality
495 and mid-successional state through prescribed fire and other management techniques. Thus, this habitat is
496 far more homogenous than jays experience in the suburban areas. Consequently, it is possible that there
497 was no opportunity for flexibility training to alter the foods or microhabitats used by these jays. However,
498 this is unlikely to be a large confound because the suburban and Archbold jays did not differ in the number
499 of different food items they ate before training and the one jay from the suburban area that we were able
500 to train, also did not show a change in the variety of substrates used and showed a slight decrease in the
501 diversity of food items used after the training (Fig. 4, single grey line).

502 To better understand the role of behavior in adaptation or resilience to human disturbance, ongoing and fu-
503 ture work in FLSJ will evaluate variation in other relevant behavioral traits like boldness, foraging role
504 (producer-scrounger dynamics), exploration and movement behavior (Snell-Rood & Steck, 2019). The
505 ManyIndividuals registered report also creates a versatile structure for testing behavioral flexibility in other
506 species or systems. Although we were unable to test disturbance resilient jay species within the timeline of
507 our current funding, we hope the research we present here will provide guidance and motivation for other
508 investigators to address this question. As humans accelerate the frequency and degree of environmental
509 change, experimental interventions such as this, that train species in behaviors that facilitate resilience, may
510 become increasingly important.

511 ETHICS

512 This research is carried out in accordance with permits from the: **1.** US Fish and Wildlife Service (scientific
513 collecting permit number ES824723). **2.** US Geological Survey Bird Banding Laboratory (federal bird
514 banding permit number 07732). **3.** Institutional Animal Care and Use Committee (University of California
515 Santa Barbara protocol number 958 and Auburn University protocol number 2022-5132).

516 AUTHOR CONTRIBUTIONS

517 **McCune:** Hypothesis development, data collection, data analysis and interpretation, write up, revis-
518 ing/editing.

519 **Barve:** Interpretation, revising/editing, materials/funding

520 FUNDING

521 This research is funded by the Department of Human Behavior, Ecology and Culture at the Max Planck
522 Institute for Evolutionary Anthropology (to Logan and McCune).

CONFLICT OF INTEREST DISCLOSURE

We, the authors, declare that we have no financial conflicts of interest with the content of this article.

ACKNOWLEDGEMENTS

We are grateful to our funders: Richard McElreath at the Max Planck Institute for Evolutionary Anthropology (whole project), as well as Corina Logan and Dieter Lukas for assistance with analyses and coding. In addition, we are grateful for the ManyIndividuals co-founders, Drs. Corina Logan and Rachel Shaw. We are thankful for the input and assistance of Dr. Reed Bowman when designing the Florida scrub-jay data collection.

SUPPLEMENTARY MATERIALS

S1 - Determining when to switch each individual to the next reversal: reversal passing criterion

Different criteria exist to decide whether an individual has learned an association between the presence of a reward and some other feature (e.g., color or shape). The two main two criteria used are to switch an individual after it either has chosen 10 out of 12 choices correct (e.g., Shaw et al., 2015) or 17 out of 20 choices correct (e.g., C. J. Logan, 2016). The criteria are further modified depending on whether choices are assessed continuously or grouped in predetermined blocks.

Here, we assess whether achieving 10 correct choices out of the last 12 continuously counted choices can be used as a reliable reversal passing criterion. To determine reliability and suitability, we investigated five questions (see below) by generalizing previously simulated reversal learning data from Logan CJ et al. (2020), based on data from great-tailed grackles. We simulated the choices individuals with different learning rates (ϕ) and rates of deviating from learned associations (λ) would make in the initial discrimination and in the first reversal. Grackles are fast to reverse preferences compared with many other species (C. J. Logan, 2016), therefore we generalized the simulations to other species by setting the parameters that guide performance (ϕ and λ) to lead to slower performances.

The findings from these simulated data indicate that deciding that an individual has passed the reversal when they choose 10 out of the last 12 consecutive trials correctly is functional and reliable because of the following:

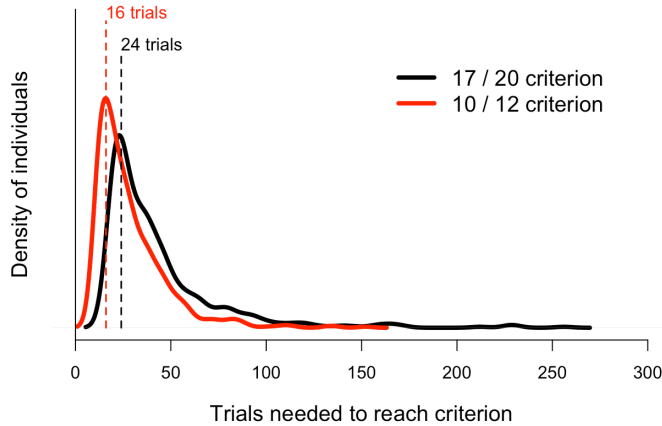
1) individuals will be finished after fewer trials than with other criteria

With the 10 out of 12 criterion, individuals pass the reversal 8 trials faster (median) than with the 17 out of 20 criterion. This means that, for most individuals, the two rules are equally effective because they will pass both in the same amount of trials (i.e., the individual who met the 17/20 criterion in 50 trials would have met the 10/12 criterion in 42 trials), but because the 10 out of 12 criterion is restricted to 12 trials instead of 20, individuals need 8 fewer trials to meet the passing criterion. No individual needs more trials with the 10 out of 12 criterion. When trials are grouped into blocks of 10 such that they could only pass on trial 20, 30, etc., individuals need a median of 5 more trials compared to when choices are assessed continuously.

2) classification of individuals using the 10/12 criterion is less noisy because there is less of a chance for individuals to approach the criterion and not pass or never pass

The average improvement in the number of trials individuals need to reach the respective criterion is larger than the median of 8 trials. This occurs because there are no individuals who are faster with the 17 out of 20 criterion, and because there is a subset of individuals who need considerably fewer trials with the 10/12 criterion (Figure 7). Individuals who require a larger number of trials (>100) to pass almost never occur

564 with the 10/12 criterion, whereas they are more common with the 17/20. With more trials, there is a higher
 565 chance that an individual will deviate from their preference by chance. This is also reflected in that 65 of the
 566 626 simulated individuals never reached the 17/20 criterion within the maximum 300 trials, whereas there
 567 were only 4 individuals with the 10/12 criterion. Accordingly, an additional benefit of choosing the 10/12
 568 criterion is that it is more likely that data for all individuals, even those who are slow to learn an association,
 569 can be collected.



570

571 **Figure S1.** There is less variation with the 10/12 reversal learning passing criterion and it requires fewer
 572 trials to reach than the 17/20 passing criterion. The lines represent the densities of individuals (estimated
 573 with smoothing which means values can go down to zero) across the 626 simulated individuals that needed
 574 a certain number of trials to either reach the 10/12 (red) or the 17/20 (black) criterion. With the 10/12
 575 criterion, most individuals need 8 fewer trials (indicated by the lines showing the mode of the two density
 576 distributions). In addition, there are only very few individuals who need 100 or more trials with the 10/12
 577 criterion, while there are several individuals that needed such large numbers with the 17/20 criterion.

578 **3) variation among individuals with the 10/12 criterion is still present and similar to the**
 579 **variation detected with other criteria**

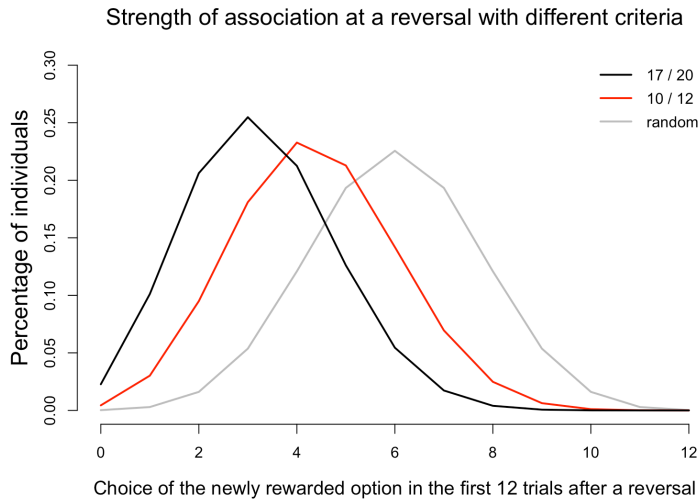
580 As described in point 1, when changing the criterion from 17/20 to 10/12, most individuals need 8 fewer
 581 trials. This also means that the differences among individuals, which might contain relevant information
 582 about variation among them, is preserved. When transforming performance with the two criteria to ranks,
 583 individuals are sorted essentially in the same order independent of which criterion is used. This is shown in
 584 Figure 7: most points are shifted up by exactly 8 trials.

585 **4) individuals can be assumed to have reliably learned the association using the 10/12 criterion**

586 Based on the two reversal passing criteria (10/12 and 17/20), we can extract the attractions that simulated
 587 individuals have formed toward both the rewarded and the unrewarded option at the point at which they
 588 meet each of these criteria. Comparing the two attractions (to the rewarded and unrewarded options), we can
 589 determine whether individuals are likely to have learned an association or not. Independent of the criterion,
 590 individuals generally formed a preference for the rewarded option: 89% of individuals favor the rewarded
 591 option between 2.5 and 14 times more than the unrewarded option. With both criteria, individuals always
 592 have a stronger attraction to the rewarded than the unrewarded option. The smallest difference between the
 593 attraction scores to the rewarded and unrewarded options we observe at the point of passing is the same with
 594 both criteria. With the 10/12 criterion, individuals would in the next trial, on average, choose the rewarded
 595 option with a probability of 76% (3 times more likely to choose rewarded over unrewarded option), whereas
 596 this is 84% with the 17/20 criterion (5 times more likely).

597 **5) the learned association means that individuals who move to the next reversal are unlikely**
 598 **to solve the reversed association by chance**

599 As expected, based on the relative attraction scores at the end of the previous reversal, most individuals
600 are unlikely to choose the now rewarded option. We expect that, on average, individuals will choose the
601 newly rewarded option in 4 or fewer trials out of the first 12 trials (red line in Figure 8). This is a lower
602 number of trials compared to individuals who have no association with either option (gray line in Figure
603 8), and a slightly higher number compared to individuals who use the 17/20 criterion (black line in Figure
604 8). The probability that an individual would, after a reversal, immediately choose the rewarded option 10
605 times during the first 12 trials (and pass) by chance is 0.001. However, even such rare individuals will have
606 actually reversed their preference during their first 12 trials because they update their attractions on every
607 trial.



608

609 **Figure S2.** Individuals form strong enough preferences using the 10/12 passing criterion as indicated by the fact that they are unlikely to pass in the first 12 trials of their next reversal (red line). These individuals
610 would take longer to switch their preference than individuals who have no preference (gray line), and they
611 would be slightly faster at switching their preference than individuals who formed their previous association
612 using the 17/20 criterion (black line).
613

614 **S2 - Determining after which reversal an individual has completed the experi-**
615 **ment: serial reversal passing criterion**

616 Data from previous serial reversal experiments suggests that individuals who go through multiple reversals
617 will end up with a performance that is similar to the individuals who needed the fewest trials on the first
618 reversal (Logan et al., 2022; Lucon-Xiccato & Bisazza, 2014). This suggests that the manipulation changes
619 individuals within their natural range of variation rather than pushing them to new limits. This means that
620 we can use the performance of the fastest individuals in the first reversal to set the criterion for passing
621 the serial reversal experiment. Accordingly, we can only set the serial reversal passing criterion after the
622 data from the first reversal begins to become available. Some species might already have data from previous
623 studies on reversal learning, however it is important to set the passing criterion for this experiment using
624 this particular setup. Therefore, the criterion must be established from scratch for each species using this
625 setup.

626 **The serial reversal passing criterion:** reach the reversal passing criterion (10 out of 12 trials correct) in
627 X trials or fewer in two consecutive reversals.

628 X = the number of trials required that marks the fastest 20% of individuals in the first reversal. For example,
629 if you test 20 individuals, the number of trials for the 4th fastest individual will be the criterion. For 10
630 individuals, use the number of trials for the 2nd fastest individual. The fastest 20% was validated using the
631 grackle data (Logan et al., 2022): it aligns with the one sigma rule from a normal distribution, indicating

632 the percentage of individuals who are faster than the mean number of trials minus one standard deviation.
633 If more than 20% of individuals reach this number of trials in their first reversal (because there might be a
634 tie), choose the next fastest number of trials to pass. Particularly near the beginning of the experiment, it
635 will be important to set the passing criterion to a lower number to ensure that individuals will be overtrained
636 rather than undertrained.

637 As the data for additional individuals becomes available, this number can change accordingly. If the number
638 changes across the experiment, we will check whether any currently participating individuals would have
639 already passed according to this criterion and end their experiment.

640 Individuals need to meet this criterion in two consecutive reversals to pass the serial reversal experiment to
641 ensure that their behavior is consistent and that their speedy performance did not occur by chance. Previous
642 serial reversal experiments show that reversal performance plateaus after a certain number of reversals (e.g.,
643 6-8 reversals in great-tailed grackles Logan et al., 2022). If individuals show no consistent improvement
644 after 12 reversals and have not yet met the serial reversal passing criterion, they will be excluded from the
645 experiment. We will plan to start with many more individuals than the minimum sample size to allow for
646 potential drop outs.

647 We do not expect that the serial reversal manipulation will introduce new negative effects because the passing
648 criterion is set such that the manipulated individuals are only as fast as the fastest 20% of tested individuals.
649 This means that we are not introducing an unnatural amount of flexibility because we are not making any
650 individuals more flexible than what already exists in their population.

651 There will be individual variation in terms of baseline flexibility before the manipulation such that the
652 flexibility training might influence individuals differently. For example, individuals who are already flexible
653 before the manipulation will not benefit much from the manipulation, while the less flexible individuals will
654 benefit more. Individuals who are already flexible and pass the serial reversals in fewer reversals will still
655 meet the experiment’s passing criterion and be considered to have completed the manipulation, even if they
656 did not improve. Baseline flexibility differences could also be reflected in their pre-manipulation success
657 measures (i.e., individuals with high baseline flexibility might already be successful before the manipulation)
658 if these success measures relate to flexibility. Our statistical models account for these baseline individual
659 differences in success as they might relate to performance on the flexibility manipulation because they include
660 an interaction between the intercept (the value at which individuals start) and slope (by how much they
661 change).

662 **S3 - Planned Sample**

663 For each population, depending on the response variable, we ran separate power analyses to determine
664 the planned sample size (see Analysis Plan). For each population, we will aim to reach the minimum
665 sample size required to detect the expected effects of the intervention on the response variables. However,
666 given the difficulties of working with wild individuals, there might be instances where we might not reach
667 a particular target. In such a case, we will interpret the result in light of the power that the particular
668 sample size provides, as indicated by the power analyses. The minimum sample size depends on whether the
669 intervention is performed as a within-subject design (higher power), measuring the response for the same
670 individuals before and after the intervention, or whether it is performed as a between-subjects design, where
671 half the individuals are randomly assigned to the intervention group. In addition, it will depend on whether
672 the response variable has a binary or a continuous outcome (higher power), and in the latter case whether
673 the measure is open-ended (lower power) and therefore individuals will show a large range of values (e.g.,
674 dispersal distance).

675 We will stop collecting data for the flexibility manipulation experiment when a buffer above the minimum
676 sample size is reached or when the season in which the minimum sample size is reached comes to an end,
677 or when the minimum sample sizes for the success measures have been reached. When conducting the
678 manipulation experiment, it is important to aim to test more than the minimum number of individuals
679 because some might not have data in the post-intervention stage.

680 S4 - Power analyses simulations and model code

681 We run analyses in R [current version 4.5.0; R Core Team (2017)] using the following R packages: rethinking
682 (McElreath, 2020a), rstan (Stan Development Team, 2020), cmdstanr (Gabry & Češnovar, 2022), knitr (Xie,
683 2018), and irr (Gamer et al., 2012). We describe below the Bayesian models we developed to address each
684 of the specific research questions. All Bayesian models were developed using McElreath (2020b) as a guide.
685 Code for running these models is available in supplementary materials S4.

686 J.Q1 Do jay populations in human modified areas differ in baseline behavioral 687 flexibility compared to populations in natural areas?

688 The model

689 We used the reversal learning Bayesian model in Logan CJ et al. (2020) to simulate and analyze population
690 differences in reversal learning, and calculate our ability to detect differences between populations. The
691 model “accounts for every choice made in the reversal learning experiment and updates the probability of
692 choosing either option after the choice was made depending on whether that choice contained a food reward
693 or not. It does this by updating three main components for each choice: an attraction score, a learning rate
694 (ϕ), and a rate of deviating from learned attractions (λ)” (Logan CJ et al., 2020).

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

696 Equation 1 “tells us how attractions to different behavioral options $A_{i,j,t+1}$ (i.e., how preferable option i is to
697 the bird j at time $t + 1$) change over time as a function of previous attractions $A_{i,j,t}$ and recently experienced
698 payoffs $\pi_{i,j,t}$ (i.e., whether they received a reward in a given trial or not). Attraction scores thus reflect
699 the accumulated learning history up to this point. The (bird-specific) parameter ϕ_j describes the weight
700 of recent experience. The higher the value of ϕ_j , the faster the bird updates their attraction. It thus can
701 be interpreted as the *learning or updating rate of an individual*. A value of $\phi_j = 0.04$, for example, means
702 that receiving a single reward for one of the two options will shift preferences by 0.02 from initial 0.5-0.5
703 attractions, a value of $\phi_j = 0.06$ will shift preferences by 0.03 and so on” (Blaisdell et al., 2021).

704 **Equation 2 (λ):** $P(i)_{t+1} = \frac{\exp(\lambda_j A_{i,j,t})}{\sum_{m=1}^2 \exp(\lambda_j A_{m,j,t})}$.

705 Equation 2 “expresses the probability an individual j chooses option i in the next round, $t + 1$, based on the
706 latent attractions. The parameter λ_j represents the *rate of deviating from learned attractions* of an individual
707 (also called inverse temperature). It controls how sensitive choices are to differences in attraction scores.
708 As λ_j gets larger, choices become more deterministic, as it gets smaller, choices become more exploratory
709 (random choice if $\lambda_j = 0$). For instance, if an individual has a 0.6-0.4 preference for option A, a value of
710 $\lambda_j = 3$ means they choose A 65% of the time, a value of $\lambda_j = 10$ means they choose A 88% of the time and
711 a value of $\lambda_j = 0.5$ means they choose A only 53% of the time” (Blaisdell et al., 2021).

712 We used the ϕ_j and λ_j values as the response variable in the Bayesian model to examine whether there were
713 differences in flexibility between the habitats:

714 $y \sim \alpha[\text{habitat}]$

715 y is the response variable (ϕ_j and λ_j , which are extracted from the correct and incorrect choices in the serial
716 reversals). There is one intercept, α , per habitat (suburban or natural) and we will estimate the habitat’s
717 average and standard deviation of the response variable.

718 Power analysis

719 Simulations using bespoke Bayesian models in Logan CJ et al. (2020) (the same model structure we use here)
720 showed a high likelihood of detecting differences with a minimum sample size of 15 when mean differences
721 in ϕ were at least 0.01 and mean differences in λ at least 3.

722 Run this model on the actual data

723 Run the code below to determine whether there were differences between the two habitats in their phi and
724 lambda flexibility measures.

725 **J.Q2 Are disturbance-resilient jays more flexible than disturbance-resistant jays? The model**

726 Same as in J.Q1 above.

727 **Power analysis**

728 Same as in J.Q1 above.

729 **J.Q3 More flexible = use more microhabitats? The model**

730 *Bayesian model with a normal distribution:*

731 $\text{habitatuse} \sim \alpha[\text{ind}] + \beta[\text{ind}] * \text{before}$

732 habitatuse is the response variable: the total number of different microhabitats used per individual. There
733 will be one intercept, α , and one slope β per individual, which will be estimated for the two conditions,
734 before (and after) the manipulation. ID is nested within condition as a random effect because there is more
735 than one data point per individual: each individual has a data point in the before condition and in the after
736 condition. A normal distribution was used because the response variable is a sum without an expected skew
737 to the curve (see Figure 10.6 in McElreath, 2020b). The Bayesian model was developed using McElreath
738 (2020b) as a guide.

739 **Power analysis**

740 We estimated our power to detect differences between conditions at different sample sizes and with different
741 mean changes in the proportion of different microhabitats used per individual in the before vs. after conditions
742 (Figure 11). We simulated the proportion of habitats used for different sample sizes of individuals before
743 and after the flexibility manipulation. We analyzed these simulated data with the model we will use to
744 analyze the actual data, estimating the change in the proportion of habitats used between the before and
745 after conditions. From the posterior estimates of the model, we extracted both the mean change as well as
746 the ratio of the posterior estimates that were below zero.

747 If the mean ratio of estimates below zero is close to 0.5, the model assumes that the change in the proportion
748 of habitats used before the flexibility manipulation is similar to after. If the ratio is close to zero, the model
749 assumes individuals have changed their behavior. For changes in the proportion of habitats used smaller than
750 0.15 (standard deviation=0.2) or 0.1 (SD=0.1), models are likely to assume that no changes occurred even
751 with large sample sizes. If the change in the proportion of habitats used before the flexibility manipulation
752 vs. after is 0.15 with a standard deviation of 0.2, on average 94% of the posterior of the model based on a
753 sample size of 20 individuals will be larger than zero (93% with a standard deviation of 0.1). This means
754 that the model is quite certain there is a difference that is larger than zero. In addition, only four of the 30
755 models for a sample size of 20 at the mean change of 0.15 have a ratio larger than 0.3, meaning that the risk
756 of having a false negative is not very likely.

757 In general, with sample sizes at or above 20 and mean changes in the proportion of habitats used at 0.1 or
758 larger, it is highly likely that the model will indicate that individuals have changed their behavior. Mean
759 changes below 0.1 can still be detected, however there is a higher risk that there will be a false negative and
760 this risk is independent of sample size.

761 With small mean changes in the response variable, some individuals might not increase or even decrease their
762 response after the manipulation because there is variation around the mean change in individual responses.
763 With small sample sizes, there is a risk that only individuals who did not clearly increase their response will
764 be studied, whereas larger sample sizes are more likely to include a wider spectrum of individuals.

765 To estimate the risk of detecting false positives, we set the mean change in the proportion of habitats used
766 to zero so there was no change between the before and after conditions. As expected, the average ratio of
767 estimates below zero is close to but below 0.5 and independent of sample size. The estimates went generally

below 0.5 because the maximum number of habitats used was set to 10 and we had a condition where individuals before the manipulation used a mean of 7 habitats. Accordingly, if individuals randomly either increase or decrease their number of habitats used, decreases will be more severe because individuals can only increase by 3 habitats, but potentially decrease by 6 habitats. With a sample size of 20, 27% have a ratio smaller than 0.3, meaning that the risk of having a false positive is high. The risk would be lower if the variation among individuals was lower than what we assumed (the standard deviation of the mean change in number of habitats was 0.2, which is a conservative estimate).

Figure 11. Risk of false positives and false negatives depending on sample and effect sizes. Curves of the mean ratio of estimates below zero (vertical lines show the minimum and maximum ratios), which illustrates the power we have to detect differences between conditions at different sample sizes. Across all models, the standard deviation of the mean change in proportion of habitats used was 0.2 (A) or 0.1 (B). A mean change in proportion of habitats of 0.3 is associated with a difference of 3 habitats (when the maximum number of habitats is 10). The curves show the model estimates as the effect increases (larger changes in the mean proportion time spent after versus before on the x-axis) for different potential sample sizes (10-60, illustrated by different colors). When there is no change (x value of zero), estimates suggest that, as expected, half of the estimates are below zero because the before and after conditions are not different from each other. As the change increases, the estimates decrease because models are able to reliably tell that the before and after conditions differ from each other.

Run this model on the actual data

Run the code below to determine whether there were differences between the before and after conditions in the proportion of microhabitats used. Only count that a microhabitat was used if the individual had at least 5% of their data points there. This prevents a microhabitat from being counted even if an individual was simply moving through it, and therefore not necessarily using it.

J.Q4 More flexible = more food types? The model

Bayesian model with a normal distribution:

$$y \sim \alpha[\text{ind}] + \beta[\text{ind}] * \text{before}$$

y is the response variable: the total number of different food types taken per individual. There will be one intercept, α , and one slope β per individual, which will be estimated for the two conditions, before (and after) the manipulation. ID is nested within condition as a random effect because there is more than one data point per individual: each individual has a data point in the before condition and in the after condition. A normal distribution was used because the response variable is a sum without an expected skew to the curve (see Figure 10.6 in McElreath, 2020b). The Bayesian model was developed using McElreath (2020b) as a guide.

Power analysis

We estimated our power to detect differences between conditions at different sample sizes and with different mean changes in the total number of different food types taken per individual in the before vs. after conditions (Figure 12). We simulated the number of food types taken for different sample sizes of individuals before and after the flexibility manipulation. We analyzed these simulated data with the model we will use to analyze the actual data, estimating the change in the number of food types taken between the before and after conditions. From the posterior estimates of the model, we extracted both the mean change as well as the ratio of the posterior estimates that were below zero.

If the mean ratio of estimates below zero is close to 0.5, the model assumes that the change in the number of food types taken before the flexibility manipulation is similar to after. If the ratio is close to zero, the model assumes individuals have changed their behavior. For changes in the number of food types taken smaller than 1, models are likely to assume that no changes occurred even with large sample sizes. If the

815 change in the number of food types taken before the flexibility manipulation vs. after is 1 with a standard
816 deviation of 2, on average 99.96% of the posterior of the model based on a sample size of 20 individuals will
817 be larger than zero. This means that the model is quite certain there is a difference that is larger than zero.
818 In addition, none of the 30 models for a sample size of 20 at the mean change of 1 have a ratio larger than
819 0.3, meaning that the risk of having a false negative is not very likely.

820 In general, with sample sizes at or above 20 and mean changes in the number of food types taken at 1 or
821 larger, it is likely that the model will indicate that individuals have changed their behavior. Mean changes
822 below 1 can still be detected, however there is a higher risk that there will be a false negative and this risk
823 is independent of sample size. For example, 17% of the models for a sample size of 20 at the mean change
824 of 0.5 have a ratio larger than 0.3, meaning there is a risk of having a false negative.

825 With small mean changes in the response variable, some individuals might not increase or even decrease their
826 response after the manipulation because there is variation around the mean change in individual responses.
827 With small sample sizes, there is a risk that only individuals who did not clearly increase their response will
828 be studied, whereas larger sample sizes are more likely to include a wider spectrum of individuals.

829 To estimate the risk of detecting false positives, we set the mean change in the number of food types taken
830 to zero so there was no change between the before and after conditions. As expected, the average ratio of
831 estimates below zero is close to 0.5 and independent of sample size. With a sample size of 20, 30% have a
832 ratio smaller than 0.3, meaning that the risk of having a false positive is high. The risk would be lower if the
833 variation among individuals was lower than what we assumed (the standard deviation of the mean change
834 in number of foods was 2, which is a conservative estimate).

835

836 **Figure 12.** Risk of false positives and false negatives depending on sample and effect sizes. Curves of the
837 mean ratio of estimates below zero (vertical lines show the minimum and maximum ratios), which illustrates
838 the power we have to detect differences between conditions at different sample sizes. Across all models, the
839 standard deviation of the mean change in number of food types taken was 2 (A) or 1 (B), and the number
840 of food types taken before the flexibility manipulation was 6.5. A mean change in proportion of habitats of
841 0.3 is associated with a difference of 3 habitats (when the maximum number of habitats is 10). The curves
842 show the model estimates as the effect increases (larger changes in the mean proportion time spent after
843 versus before on the x-axis) for different potential sample sizes (10-60, illustrated by different colors). When
844 there is no change (x value of zero), estimates suggest that, as expected, half of the estimates are below
845 zero because the before and after conditions are not different from each other. As the change increases, the
846 estimates decrease because models are able to reliably tell that the before and after conditions differ from
847 each other.

848

849 **Run this model on the actual data**

850 Run the code to determine whether there were differences between the before and after conditions in the
851 number of food types taken.

852 **REFERENCES**

- 853 Alberti, M. (2015). Eco-evolutionary dynamics in an urbanizing planet. *Trends in Ecology & Evolution*,
854 30(2), 114–126. <https://doi.org/10.1016/j.tree.2014.11.007>
- 855 Beauchamp, G., Bakley, T. D., Fitzpatrick, J. W., & Barve, S. (2025). Food, weather, and population
856 density, not number of helpers, drive overwinter survival in florida scrub-jays. *Oecologia*, 207(5), 68.
- 857 Bebus, S. E., Small, T. W., Jones, B. C., Elderbrock, E. K., & Schoech, S. J. (2016). Associative learning is
858 inversely related to reversal learning and varies with nestling corticosterone exposure. *Animal Behaviour*,
859 111, 251–260.
- 860 Blair, R. B. (1996). Land use and avian species diversity along an urban gradient. *Ecological Applications*,
861 6(2), 506–519.

- 862 Blaisdell, A., Seitz, B., Roney, C., Folsom, M., MacPherson, M., Deffner, D., & Logan, C. J. (2021). *Do*
863 *the more flexible individuals rely more on causal cognition? Observation versus intervention in causal*
864 *inference in great-tailed grackles (version 5 of this preprint has been peer reviewed and recommended by*
865 *peer community in ecology [https://doi.org/10.24072/pci.ecology.100076]).* <https://doi.org/10.31234/osf.io/z4p6s>
866
- 867 Bond, A. B., Kamil, A. C., & Balda, R. P. (2007). Serial reversal learning and the evolution of behavioral
868 flexibility in three species of North American corvids (*Gymnorhynchus cyanocephalus*, *Nucifraga columbiana*,
869 *Aphelocoma californica*). *Journal of Comparative Psychology*, 121(4), 372–379. <https://doi.org/10.1037/0735-7036.121.4.372>
870
- 871 Chejanovski, Z. A., Avilés-Rodríguez, K. J., Lapedra, O., Preisser, E. L., & Kolbe, J. J. (2017). An
872 experimental evaluation of foraging decisions in urban and natural forest populations of anolis lizards.
873 *Urban Ecosystems*, 20(5), 1011–1018. <https://doi.org/10.1007/s11252-017-0654-5>
874
- 874 Ciani, A. C. (1986). Intertroop agonistic behavior of a feral rhesus macaque troop ranging in town and
875 forest areas in india. *Aggressive Behavior*, 12(6), 433–439. [https://doi.org/10.1002/1098-2337\(1986\)12:6%3C433::AID-AB2480120606%3E3.0.CO;2-C](https://doi.org/10.1002/1098-2337(1986)12:6%3C433::AID-AB2480120606%3E3.0.CO;2-C)
876
- 877 Curry, R., Townsend Peterson, A., & Langen, T. (2017). *California scrub-jay (aphelocoma californica), birds*
878 *of north america*.
- 879 Daniel, M. M., & Schluessel, V. (2020). Serial reversal learning in freshwater stingrays (potamotrygon
880 motoro). *Animal Cognition*, 23(1), 109–119. <https://doi.org/10.1007/s10071-019-01321-x>
881
- 881 Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in
882 corvids and apes. *Science*, 306(5703), 1903–1907.
- 883 Federspiel, I. G., Garland, A., Guez, D., Bugnyar, T., Healy, S. D., Güntürkün, O., & Griffin, A. S. (2017).
884 Adjusting foraging strategies: A comparison of rural and urban common mynas (*acridotheres tristis*).
885 *Animal Cognition*, 20(1), 65–74. <https://doi.org/10.1007/s10071-016-1045-7>
886
- 886 Gabry, J., & Češnovar, R. and. (2022). *Cmdstanr: R interface to 'CmdStan'*. <https://mc-stan.org/cmdstanr/>
887
- 888 Galbraith, J. A., Beggs, J. R., Jones, D. N., & Stanley, M. C. (2015). Supplementary feeding restructures
889 urban bird communities. *Proceedings of the National Academy of Sciences*, 112(20), E2648–E2657.
- 890 Gamer, M., Lemon, J., Gamer, M. M., Robinson, A., & Kendall's, W. (2012). Package “irr.” *Various*
891 *Coefficients of Interrater Reliability and Agreement*, 22.
- 892 Goldewijk, K. K. (2001). Estimating global land use change over the past 300 years: The HYDE database.
893 *Global Biogeochemical Cycles*, 15(2), 417–433. <https://doi.org/10.1029/1999GB001232>
894
- 894 Greggor, A. L., Masuda, B., Gaudioso-Levita, J. M., Nelson, J. T., White, T. H., Shier, D. M., Farabaugh, S.
895 M., & Swaisgood, R. R. (2021). Pre-release training, predator interactions and evidence for persistence
896 of anti-predator behavior in reintroduced alalā, hawaiian crow. *Global Ecology and Conservation*, 28,
897 e01658. <https://doi.org/10.1016/j.gecco.2021.e01658>
898
- 898 Greig, E. I., Wood, E. M., & Bonter, D. N. (2017). Winter range expansion of a hummingbird is associated
899 with urbanization and supplementary feeding. *Proceedings of the Royal Society B: Biological Sciences*,
900 284(1852).
- 901 Grinnell, J. (1917). The niche-relationships of the california thrasher. *The Auk*, 34(4), 427–433. <https://doi.org/10.2307/4072271>
902
- 903 Jolly, C. J., Kelly, E., Gillespie, G. R., Phillips, B., & Webb, J. K. (2018). Out of the frying pan: Reintro-
904 duction of toad-smart northern quolls to southern kakadu national park. *Austral Ecology*, 43(2), 139–149.
905 <https://doi.org/10.1111/aec.12551>
906
- 906 Lapedra, O., Chejanovski, Z., & Kolbe, J. J. (2017). Urbanization and biological invasion shape animal
907 personalities. *Global Change Biology*, 23(2), 592–603.
- 908 Lee, V. E., & Thornton, A. (2021). Animal cognition in an urbanised world. *Frontiers in Ecology and*
909 *Evolution*, 9, 120. <https://doi.org/10.3389/fevo.2021.633947>
910
- 910 Liu, X., Huang, Y., Xu, X., Li, X., Li, X., Ciais, P., Lin, P., Gong, K., Ziegler, A. D., Chen, A., et al. (2020).
911 High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. *Nature Sustainability*,
912 1–7. <https://doi.org/10.1038/s41893-020-0521-x>
913
- 913 Liu, Y., Day, L. B., Summers, K., & Burmeister, S. S. (2016). Learning to learn: Advanced behavioural
914 flexibility in a poison frog. *Animal Behaviour*, 111, 167–172. <https://doi.org/10.1016/j.anbehav.2015.10.018>
915

- 916 Logan, C. J. (2016). Behavioral flexibility and problem solving in an invasive bird. *PeerJ*, 4, e1975. <https://doi.org/10.7717/peerj.1975>
- 917
- 918 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>
- 919
- 920
- 921 Logan, C.J., McCune, K.B., Breen, A., 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>
- 922
- 923
- 924 Logan, C., Lukas, D., Blaisdell, A., Johnson-Ulrich, Z., MacPherson, M., Seitz, B., Sevchik, A., & McCune, K. (2023). Behavioral flexibility is manipulatable and it improves flexibility and innovativeness in a new context. *Peer Community Journal*, 3. <https://doi.org/10.24072/pcjournal.284>
- 925
- 926
- 927 Logan, C., Lukas, D., Geng, X., Hardy, K., LeGrande, C., Marfori, Z., MacPherson, M., Rowney, C., Smith, C., & McCune, K. (2025). Behavioral flexibility is related to foraging, but not social or habitat use behaviors, in a species that is rapidly expanding its range. *Peer Community Journal*, 5.
- 928
- 929
- 930 Lucon-Xiccato, T., & Bisazza, A. (2014). Discrimination reversal learning reveals greater female behavioural flexibility in guppies. *Biology Letters*, 10(6), 20140206. <https://doi.org/10.1098/rsbl.2014.0206>
- 931
- 932 Mackintosh, N., Mcgonigle, B., & Holgate, V. (1968). Factors underlying improvement in serial reversal learning. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 22(2), 85. <https://doi.org/10.1037/h0082753>
- 933
- 934
- 935 McCune, K. B., Jablonski, P., Lee, S., & Ha, R. R. (2019). Captive jays exhibit reduced problem-solving performance compared to wild conspecifics. *Royal Society Open Science*, 6(1), 181311.
- 936
- 937 McCune, K. B., Valente, J. J., Jablonski, P. G., Lee, S., & Ha, R. R. (2022). Social behavior mediates the use of social and personal information in wild jays. *Scientific Reports*, 12(1), 2494.
- 938
- 939 McCune, K.B., Folsom, M., Ross, C., Bergeron, L., & Logan, C. (2020). Does great-tailed grackle space use behavior reflect individual differences in exploration? *In Principle Acceptance by PCI Ecology of the Version on 23 Sep 2020*. <http://corinalogan.com/Preregistrations/gspaceuse.html>
- 940
- 941
- 942 McElreath, R. (2020a). *Rethinking: Statistical rethinking book package*.
- 943
- 944 McElreath, R. (2020b). *Statistical rethinking: A bayesian course with examples in r and stan*. Chapman; Hall/CRC.
- 945
- 946 Midford, P. E., Hailman, J. P., & Woolfenden, G. E. (2000). Social learning of a novel foraging patch in families of free-living florida scrub-jays. *Animal Behaviour*, 59(6), 1199–1207.
- 947
- 948 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>
- 949
- 950 Moseby, K. E., Blumstein, D. T., & Letnic, M. (2016). Harnessing natural selection to tackle the problem of prey naïveté. *Evolutionary Applications*, 9(2), 334–343. <https://doi.org/10.1111/eva.12332>
- 951
- 952 Moseby, K. E., Cameron, A., & Crisp, H. A. (2012). Can predator avoidance training improve reintroduction outcomes for the greater bilby in arid australia? *Animal Behaviour*, 83(4), 1011–1021. <https://doi.org/10.1016/j.anbehav.2012.01.023>
- 953
- 954
- 955 Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). Ecological niches and geographic distributions (MPB-49). In *Ecological niches and geographic distributions (MPB-49)*. Princeton University Press. <https://doi.org/10.1515/9781400840670>
- 956
- 957
- 958 R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org>
- 959
- 960 Rice, N. H., Martinez-Meyer, E., & Peterson, A. T. (2003). Ecological niche differentiation in the aphelocoma jays: A phylogenetic perspective. *Biological Journal of the Linnean Society*, 80(3), 369–383.
- 961
- 962 Ross, A. K., Letnic, M., Blumstein, D. T., & Moseby, K. E. (2019). Reversing the effects of evolutionary prey naïveté through controlled predator exposure. *Journal of Applied Ecology*, 56(7), 1761–1769. <https://doi.org/10.1111/1365-2664.13406>
- 963
- 964
- 965 Shaw, R. C., Boogert, N. J., Clayton, N. S., & Burns, K. C. (2015). Wild psychometrics: Evidence for ‘general’ cognitive performance in wild new zealand robins, *petroica longipes*. *Animal Behaviour*, 109, 101–111.
- 966
- 967
- 968 Shawkey, M. D., Bowman, R., & Woolfenden, G. E. (2004). Why is brood reduction in florida scrub-jays higher in suburban than in wildland habitats? *Canadian Journal of Zoology*, 82(9), 1427–1435.
- 969

- 970 Snell-Rood, E. C., & Steck, M. K. (2019). Behaviour shapes environmental variation and selection on
971 learning and plasticity: Review of mechanisms and implications. *Animal Behaviour*, *147*, 147–156.
- 972 Sol, D., Timmermans, S., & Lefebvre, L. (2002). Behavioural flexibility and invasion success in birds. *Animal*
973 *Behaviour*, *63*(3), 495–502. <https://doi.org/10.1006/anbe.2001.1953>
- 974 Stan Development Team. (2020). *RStan: The R interface to Stan*. <http://mc-stan.org/>
- 975 Strang, C. G., & Sherry, D. F. (2014). Serial reversal learning in bumblebees (*bombus impatiens*). *Animal*
976 *Cognition*, *17*(3), 723–734. <https://doi.org/10.1007/s10071-013-0704-1>
- 977 Tetzlaff, S. J., Sperry, J. H., & DeGregorio, B. A. (2019). Effects of antipredator training, environmental
978 enrichment, and soft release on wildlife translocations: A review and meta-analysis. *Biological Conser-*
979 *vation*, *236*, 324–331. <https://doi.org/10.1016/j.biocon.2019.05.054>
- 980 Tuomainen, U., & Candolin, U. (2011). Behavioural responses to human-induced environmental change.
981 *Biological Reviews*, *86*(3), 640–657.
- 982 Valente, J. J., LeGrande-Rolls, C. L., Rivers, J. W., Tucker, A. M., Fischer, R. A., & Betts, M. G. (2021).
983 Conspecific attraction for conservation and management of terrestrial breeding birds: Current knowl-
984 edge and future research directions. *The Condor*, *123*(2), duab007. [https://doi.org/10.1093/ornithapp/](https://doi.org/10.1093/ornithapp/duab007)
985 [duab007](https://doi.org/10.1093/ornithapp/duab007)
- 986 Vardi, R., & Berger-Tal, O. (2022). Environmental variability as a predictor of behavioral flexibility in urban
987 environments. *Behavioral Ecology*, *33*(3), 573–581.
- 988 Veech, J. A., Small, M. F., & Baccus, J. T. (2011). The effect of habitat on the range expansion of a native
989 and an introduced bird species. *Journal of Biogeography*, *38*(1), 69–77.
- 990 Wehtje, W. (2003). The range expansion of the great-tailed grackle (*quiscalus mexicanus gmelin*) in north
991 america since 1880. *Journal of Biogeography*, *30*(10), 1593–1607.
- 992 West, R., Letnic, M., Blumstein, D. T., & Moseby, K. E. (2018). Predator exposure improves anti-predator
993 responses in a threatened mammal. *Journal of Applied Ecology*, *55*(1), 147–156. [https://doi.org/10.](https://doi.org/10.1111/1365-2664.12947)
994 [1111/1365-2664.12947](https://doi.org/10.1111/1365-2664.12947)
- 995 Woolfenden, G. E., & Fitzpatrick, J. W. (1996). *Florida scrub-jay (aphelocoma coerulescens), birds of north*
996 *america*.
- 997 Wright, T. F., Eberhard, J. R., Hobson, E. A., Avery, M. L., & Russello, M. A. (2010). Behavioral flexibility
998 and species invasions: The adaptive flexibility hypothesis. *Ethology Ecology & Evolution*, *22*(4), 393–404.
999 <https://doi.org/10.1080/03949370.2010.505580>
- 1000 Wu, J., Jenerette, G. D., Buyantuyev, A., & Redman, C. L. (2011). Quantifying spatiotemporal patterns of
1001 urbanization: The case of the two fastest growing metropolitan regions in the united states. *Ecological*
1002 *Complexity*, *8*(1), 1–8. <https://doi.org/10.1016/j.ecocom.2010.03.002>
- 1003 Xie, Y. (2018). Knitr: A comprehensive tool for reproducible research in r. In *Implementing reproducible*
1004 *research* (pp. 3–31). Chapman; Hall/CRC.