Behavioral flexibility and problem solving in an invasive bird

Corina J Logan

Behavioral flexibility is considered an important trait for adapting to environmental change, but it is unclear what it is, how it works, and whether it is a problem solving ability. I investigated behavioral flexibility and problem solving experimentally in great-tailed grackles, an invasive bird species and thus a likely candidate for possessing behavioral flexibility. Grackles demonstrated behavioral flexibility in two contexts, the Aesop's Fable paradigm and a color association test. Contrary to predictions, behavioral flexibility did not correlate across contexts. Four out of 6 grackles exhibited efficient problem solving abilities, but problem solving efficiency did not appear to be directly linked with behavioral flexibility. Problem solving speed also did not significantly correlate with reversal learning scores, indicating that faster learners were not the most flexible. These results reveal how little we know about behavioral flexibility, and provide an immense opportunity for future research to explore how individuals and species can use behavior to react to changing environments.

1   **Behavioral flexibility and problem solving in an invasive bird**

2   Corina J. Logan*

3   SAGE Center for the Study of the Mind, University of California Santa Barbara, Santa Barbara,

4   CA 93103 USA

5   *Current address: Department of Zoology, University of Cambridge, Cambridge CB2 3EJ

6   United Kingdom

7   E-mail: cl417@cam.ac.uk, phone: +44 7868 357 294, fax: +44 1223 336 676

9   **BACKGROUND**

10   Behavioral flexibility, defined here as the ability to change preferences when circumstances

11   change based on learning from previous experience or using causal knowledge, is frequently

12   implicated as a key factor involved in problem solving success and adapting behavior to

13   changing environments (e.g., Lefebvre et al. 1997, Griffin & Guez 2014, Buckner 2015, Chow et

14   al. 2016). Those individuals or species that are more behaviorally flexible are predicted to learn

15   faster and more efficiently, and rely on more learning strategies to solve problems (Griffin &

16   Guez 2014). Testing behavioral flexibility experimentally requires individuals to change their

17   behavior in response to changes in the task.

18   A common way to experimentally test behavioral flexibility uses a reversal learning

19   paradigm where individuals first learn to prefer a particular option, and then, once proficient, the

20   reward contingencies are altered such that previously correct choices are now unsuccessful,

21   therefore the individual must learn to change its preference (e.g., Bond et al. 2007, Tebbich et al.

22   2010, Ghahremani et al. 2010, Buckner 2013). The few investigations that have examined the

23   relationship between behavioral flexibility and problem solving abilities produced mixed results.

24   Two studies investigating behavioral flexibility and problem solving speed found that faster

25   learners were slower to reverse their preferences (invasive Indian mynas: Griffin et al. 2013,

26   threatened Florida scrub-jays: Bebus et al. 2016). Another study found no correlation between

27   reversal learning speed and problem solving speed or ability (color and shape discrimination,

28   spatial memory, and motor skills; spotted bowerbirds: Isden et al. 2013). Reversal learning speed

29   is thought to positively correlate with inhibition: When the task changes subjects must inhibit the

30   previously learned behavior to be able to learn the new behavior (Manrique et al. 2013, Griffin &

31   Guez 2014, Liu et al. 2016). However, this idea is challenged by an experiment in rats that were

32   genetically modified to increase inhibition (Homberg et al. 2007). Knock out rats with improved

33   inhibition showed no difference in their reversal learning speed from non-modified rats

34   (Homberg et al. 2007). This suggests that behavioral flexibility may rely more on individuals

35   continuing to sample their environment rather than simply inhibiting a response when a behavior

36   is no longer rewarded. This variety of contrasting results indicates how little is known about

37   behavioral flexibility in terms of how it relates to problem solving abilities and cognition.

38   Griffin and Guez (2014) propose that behavioral flexibility is a multi-faceted trait: some

39   aspects are measurable in problem solving tasks while other aspects are measurable in other

40   contexts, therefore individuals might exhibit flexibility in some contexts but not others.

41   Behavioral flexibility is usually studied in relation to problem solving speed (Griffin et al. 2013,

42   Bebus et al. 2016), not problem solving success, and, while behavioral flexibility is also tested in

43   non-reversal learning paradigms (e.g., multi-access box: Auersperg et al. 2011, Manrique et al.

44   2013; episodic-like memory and future planning: Clayton & Dickinson 1998, Dally et al. 2006,

45   Raby et al. 2007), it is generally tested in only one context per study. Therefore, our

46   understanding of the mechanisms underlying behavioral flexibility is lacking.

47   To begin to address these gaps, I investigated behavioral flexibility in one of the most

48   invasive species in North America, the great-tailed grackle (*Quiscalus mexicanus*, family

49   Icteridae, hereafter referred to as grackles; Peer 2011). Species that rapidly adapt to novel

50   environments are presumed to require the ability to behaviorally respond to changing

51   circumstances within the course of their lifetime (Sol & Lefebvre 2000, Sol et al. 2002, Sol et al.

52   2005, Sol et al. 2007), thus many invasive species are likely candidates for possessing behavioral

53   flexibility. I investigated whether grackles are behaviorally flexible and efficient problem

54   solvers, whether they vary in behavioral flexibility across contexts, whether flexibility correlates

55   with problem solving ability and speed, and whether individuals that are more flexible use more

56   learning strategies.

57   I tested behavioral flexibility in two contexts by measuring preferences (due to learning,

58   attending to function, or previous experience) and then requiring individuals to change

59   preferences after modifying the task. A color association task (context 1) involved a gold tube

60   and a silver tube placed on the table at the same time and with one of the tubes containing hidden

61   food. Individuals learned to associate food with first the gold tube (learning speed; Experiment

62   1) and then the silver tube (a modified version of reversal learning in that there was only 1

63   reversal; Experiment 2). I used this task to compare the speed with which grackles learn and

64   reverse preferences compared with other species, and to examine which learning strategies

65   grackles use to become proficient. Economics theory calls this type of problem a contextual,

66   binary multi-armed bandit and predicts two different learning strategies as solutions to the

67   problem (McInerney 2010). These learning strategies involve a trade off between an exploration

68   phase and an exploitation phase. The pattern of the trade off indicates which learning strategy

69   was used. For example, a short exploration phase before switching almost exclusively to the

70   exploitation phase (marked by significantly more correct choices) indicates one learning strategy,

71   whereas a long exploration phase before eventually choosing significantly more correct choices

72   (exploitation phase) indicates a different learning strategy.

73   The Aesop's Fable paradigm (context 2) examines problem solving ability and involves

74   food floating in a partially filled water tube, which is solved by inserting objects into the tube to

75   raise the water level and bring the food within reach. It has been used to explore the cognitive

76   abilities underlying problem solving in rooks (Bird & Emery 2009), Eurasian jays (Cheke et al.

77   2011), humans (Cheke et al. 2012), New Caledonian crows (Taylor et al. 2011, Jelbert et al.

78   2014, Logan et al. 2014), and Western scrub-jays (Logan et al. 2016). While great-tailed grackles

79   are not reported to use tools (grackles are not listed in Lefebvre et al. 2002), non-tool using

80   species have successfully participated in the Aesop's Fable tests (rooks: Bird & Emery 2009,

81   Eurasian jays: Cheke et al. 2011, and Western scrub-jays: Logan et al. 2016) as well as other

82   tool-using tests (e.g., the trap tube; rooks: Seed et al. 2006, Tebbich et al. 2007), therefore I

83   expect grackles to be capable of performing these experiments. I compared grackle problem

84   solving performance with previously tested species to determine whether grackles are efficient

85   problem solvers. I modified the standard Aesop's Fable experiments to test behavioral flexibility

86   by assessing whether they prefer to drop the more functional heavy objects in a Heavy vs. Light

87   experiment, and whether they change these preferences in a follow up experiment where the

88   heavy objects become non-functional (Heavy vs. Light Magic).

89

90   **GENERAL METHODS**

91   **Ethics**

92 This research was carried out in accordance with permits from the U.S. Fish and Wildlife Service

93 (scientific collecting permit number MB76700A), California Department of Fish and Wildlife

94 (scientific collecting permit number SC-12306), U.S. Geological Survey Bird Banding

95 Laboratory (federal bird banding permit number 23872), and the Institutional Animal Care and

96 Use Committee at the University of California Santa Barbara (IACUC protocol numbers 860 and

97 860.1).

98

99 **Subjects and Study Site**

100 Eight wild adult great-tailed grackles (4 females and 4 males) were caught using a walk-in baited

101 trap measuring 0.61m high by 0.61m wide by 1.22m long (design from Overington et al. 2011).

102 Birds were caught (and tested) in two batches: batch 1 at the Andree Clark Bird Refuge (4 birds

103 [Tequila, Margarita, Cerveza, and Michelada] in September 2014, released in December) and

104 batch 2 at East Beach Park (4 birds [Refresco, Horchata, Batido, and Jugo] in January 2015,

105 released in March) in Santa Barbara, California. They were housed individually in aviaries

106 measuring 183cm high by 119cm wide by 236cm long at the University of California Santa

107 Barbara for 2-3 months while participating in the experiments in this study. Grackles were given

108 water *ad libitum* and unrestricted amounts of food (Mazuri Small Bird Food) for at least 20 hrs

109 per day, with their main diet being removed for up to 4 hrs on testing days while they

110 participated in experiments and received peanuts or bread when successful. Grackles were aged

111 by plumage and eye color and sexed by plumage and weight following Pyle (2001). Biometrics,

112 blood, and feathers were collected at the beginning and end of their time in the aviary. Their

113 weights were measured at least once per month, first at the time of trapping using a balancing

114   scale, and subsequently by placing a kitchen scale covered with food in their aviary and

115   recording their weight when they jumped onto the scale to eat.

116

117   **Experimental Set Up**

118   Apparatuses were placed on top of rolling tables (60cm wide by 39cm long) and rolled into each

119   individual's aviary for testing sessions, which lasted up to approximately 20min. If habituation to

120   an apparatus was needed, as indicated by a bird's unwillingness to immediately approach and eat

121   from it (this was the case for most birds with the stone dropping training apparatus), it was

122   placed in their aviary overnight and they were fed from it. If an apparatus had parts that would

123   allow a bird to learn how the task worked, these parts were taped over to prevent learning. If a

124   grackle approached an apparatus and ate from it without hesitating, it was considered habituated.

125   If re-habituation was needed, as indicated by an unwillingness to approach the apparatus after the

126   training or experiment began, the habituation process was repeated. Color tubes were baited with

127   peanut pieces and/or bread. Water tubes were baited with 1/16 of a peanut attached to a small

128   piece of cork with a tie wrap for buoyancy (hereafter referred to as a peanut float). The area

129   around the top of the water tube (the standing platform) was also sometimes baited with smaller

130   peanut pieces and bread crumbs, and more peanut floats could be added to the inside of the water

131   tube to encourage the bird to interact with the task. If more than one peanut float was in the tube,

132   the bird was given the opportunity, after retrieving the first peanut float, to insert more objects

133   into the tube to retrieve the other peanut floats. If a bird started to lose motivation (e.g., refuse to

134   come to the table or interact with the apparatus) for participating in a task because they were

135   unsuccessful (as in Heavy vs. Light Magic), I baited the standing platform between trials to

136   reward their participation and keep them interested in finishing the experiment. A trial was

137   terminated when the bird solved the task or did not interact with the apparatus for at least 1 min.

138   All water tube experiments (3-6) consisted of 20 trials per bird and were recorded with a Nikon

139   D5100 camera on a tripod placed inside the aviary. Experiments were given in the following

140   order: 1, 3, 4, 5, 6, 2, which was the same for all birds (except Tequila who began, but did not

141   complete, a sand vs. water experiment between Experiments 1 and 3). Grackles took 1-7 days to

142   complete an experiment, which could have spanned the course of up to 19 days. Birds were

143   tested with the experimenter just outside the aviary door and in full sight of the grackle. Grackles

144   are already habituated to humans because they forage in an urban setting, often coming within a

145   meter of the nearest humans. They quickly habituated to the testing set up and their first 5 days

146   in the aviary were spent habituating them to a human presence just outside their aviary door.

147

148   **Experimenters**

149   My research assistant, Luisa Bergeron, and I conducted Experiments 1-2; I conducted

150   Experiments 3-6; and my assistants (Luisa Bergeron, Alexis Breen, Michelle Gertsvolf, Christin

151   Palmstrom, and Linnea Palmstrom) and I conducted the stone dropping training.

152

153   **Statistical Analyses**

154   *Color Association Tests*

155   Two analyses were performed on the color association data (Experiments 1 and 2). First, a bird

156   was considered to pass this test if it chose correctly at least 17 out of the most recent 20 trials

157   (with a minimum of 8 or 9 correct choices out of 10 on the two most recent sets of 10; binomial

158   test: $p=0.003$ for 17/20). Once the bird reached proficiency using this analysis, their individual

159   learning strategy was identified using a contextual, binary multi-armed bandit (see McInerney

160   2010 for a review). It was contextual in that the subject was allowed to make only one choice per

161   trial, and binary because there were two options on the table, one containing a reward and the

162   other containing no reward. I categorized grackle learning strategies by matching them to the two

163   known approximate strategies of the contextual, binary multi-armed bandit: epsilon-first and

164   epsilon-decreasing (McInerney 2010). The following equations refer to the different phases

165   involved in each strategy:

166   *Equation 1 (exploration phase):* ϵN

167   *Equation 2 (exploitation phase):* $(1-\epsilon) N$

168   N is the number of trials given, and epsilon, ϵ, represents the subject's uncertainty about the

169   location of the reward, starting at complete uncertainty ($\epsilon=1$) at the beginning of the experiment

170   and decreasing rapidly as individuals gain experience with the task (exploration phase where the

171   rewarded color is chosen below or at chance levels) and switch to the exploitative phase (the

172   rewarded color is chosen significantly above chance levels). Because the grackles needed to

173   learn the rules of the task, they necessarily had an exploration phase. The epsilon-first strategy

174   involves an exploration phase followed by an entirely exploitative phase. The optimal strategy

175   overall would be to explore one color in the first trial and the other color in the second trial, and

176   then switch to an exploitative strategy (choose the rewarded color significantly above chance

177   levels). In this case there would be no pattern in the choices in the exploration phase because it

178   would consist of sampling each color only once. In the epsilon-decreasing strategy, birds would

179   start by making some incorrect choices and then increase their choice of gold gradually as their

180   uncertainty decreases until they choose the rewarded color significantly above chance levels. In

181   this case, a linear pattern emerges during the exploration phase.

182   To determine whether faster learners are also more flexible, I used a Spearman's rank

183   correlation test to examine whether learning speed (number of trials to learn a preference)

184   correlated with reversal learning scores (number of trials to reverse a preference minus the

185   number of trials to learn the preference) in the color association test.

186

187   *Aesop's Fable tests*

188   To make the water tube experimental results comparable with previous studies, I used two-tailed

189   binomial tests to determine whether each grackle chose particular objects or tubes at random

190   chance (null hypothesis: $p \geq 0.05$) or significantly above chance (alternative hypothesis: $p < 0.05$).

191   The Bonferroni-Holm correction was applied to p-values within each experiment to correct for

192   an increase in false positive results that could arise from conducting multiple tests on the same

193   dataset.

194   Generalized linear mixed models (GLMMs) were used to determine whether birds

195   preferred particular objects or tubes (response variable: correct/more correct choice or

196   incorrect/less correct choice) in a water tube experiment and whether the trial number or bird

197   influenced choices (explanatory variables: experiment, trial number, bird), and to control for the

198   non-independence of multiple choices per trial (random factor: choice number). I used minimal

199   belief priors (V=1, nu=0) and fixed the variance component to one (fix=1) because the

200   measurement error variance was known, as is standard when choices are binary (Hadfield 2010).

201   I ensured that the Markov chain for this test model converged by manipulating the number of

202   iterations (nitt=150000 for the null model, nitt=600000 for the test model), the number of

203   iterations that must occur before samples are stored (burnin=30000), and the intervals the

204   Markov chain stores (thin=300) until successive samples were independent as indicated by low

205   (<0.1) correlations (autocorr function, MCMCglmm package: Hadfield 2014a,b) and there were

206 no trends when visually inspecting the time series of the Markov chain (function:

207 plot(testmodel$Sol); Hadfield 2014a,b). I compared this test model to a null model where I

208 removed all explanatory factors and set it to 1.

209   I determined whether the test model was likely given the data, relative to the null model,

210 by using Akaike weights (range: 0-1, all model weights sum to 1; Akaike 1981; Weights

211 function, MuMIn package: Bates et al. 2011). The Akaike weight indicates the "relative

212 likelihood of the model given the data" (Burnham and Anderson 2002, p. xxiii) and models with

213 Akaike weights greater than 0.89 are considered reliable models because they are highly likely

214 given the data (Burnham and Anderson 2002). The test model was highly likely given the data

215 (Akaike weight=1.00) and the null model was not (Akaike weight=3.4e-30), therefore the test

216 model was highly likely given the data. To investigate the potential effects of season or order of

217 testing, I conducted a GLMM to determine whether the batch to which the bird belonged

218 (explanatory variable: batch=1 or 2) influenced their test performance (response variable: correct

219 or incorrect choice) while controlling for the non-independence of multiple choices per trial

220 (random factor: choice number). The null model was highly likely given the data (Akaike

221 weight=0.94), while the batch model was not (Akaike weight=0.06), indicating that batch did not

222 influence test performance. GLMMs were carried out in R v3.2.1 (R Core Team 2016) using the

223 MCMCglmm function (MCMCglmm package, Hadfield 2014a) with a binomial distribution

224 (called categorical in MCMCglmm) and logit link.

225

226 *Cross-Context Analysis*

227 To determine whether those grackles that were more behaviorally flexible in the water tube

228 context (yes or no) were also more behaviorally flexible in the color association context (the

229   number of trials to meet criterion in experiment 2 minus the number of trials to meet criterion in

230   experiment 1), I used a Spearman's rank correlation test. Margarita and Cerveza were included,

231   even though their attraction to the magnet in Experiment 4 biased their results, to increase the

232   sample size of individuals that participated in both contexts to 5.

233

234   **Data Availability**

235   The   data   are   available   at   the   KNB   Data   Repository:

236   https://knb.ecoinformatics.org/#view/corina_logan.15.6 (Logan 2015).

237

238   **Video**

239   Watch video clips showing examples of each experiment at: https://youtu.be/GhR6fGG1yc4.

240

241   **CONTEXT 1: COLOR ASSOCIATION TESTS**

242   **Experiment 1: Color Association Test (learning speed)**

243   *Experiment 1: Methods*

244   To assess how many trials it takes a grackle to form an association between food and color, they

245   were given a gold and a silver tube with food (peanut pieces or bread) always hidden in the gold

246   tube (Logan et al. 2014 & 2016). Grackles were first trained on a blue tube where they learned to

247   search for hidden food. Each color tube set up consisted of a PVC tube (outer diameter 26mm,

248   inner diameter 19mm) mounted on two pieces of plywood glued together at a right angle (whole

249   apparatus measuring 50mm wide by 50mm tall by 67mm deep). Each tube was placed at

250   opposite ends of a table with the tube openings facing the side walls so the bird could not see

251   which tube contained the food. Tubes were pseudorandomized for side and the left tube was

252   always placed first, followed by the right to avoid behavioral cueing. Pseudorandomization

253   consisted of alternating location for the first two trials of a session and then keeping the same

254   color on the same side for at most 2 consecutive trials thereafter. Each trial consisted of placing

255   the tubes on the table, and then the bird had the opportunity to choose one tube by looking into it

256   (and eating from it if it chose the gold tube). Once the bird chose, the trial ended by removing the

257   tubes.

258

259   *Experiment 1: Results*

260   All grackles learned to associate the gold tube with a food reward by reaching criterion (at least

261   17/20 correct trials) in 20-40 trials (Table 1). Using the binary multi-armed bandit analysis to

262   assess which learning strategies birds used, Refresco used the epsilon-first strategy because he

263   first explored (i.e., made unsuccessful and/or successful choices) and then exploited (i.e., was

264   successful) every trial thereafter: He explored in his first trial (he failed by choosing silver) and

265   then always chose gold after that (Figure 1). The rest of the grackles used the epsilon-decreasing

266   strategy by exploring more at the beginning and gradually increasing their success until they

267   performed above chance levels by the end of the experiment (Figure 1). Horchata and Jugo had

268   exceptions to this strategy: Horchata started a second exploration phase at the end of her

269   experiment, and Jugo's pattern of exploration did not linearly increase in his first several trials

270   (though it did thereafter). In the first part of Jugo's experiment, he did not appear to follow any

271   particular rules during his learning phase such as 'always choose the left side' or 'always

272   alternate sides', therefore it is unknown how this part of his exploration phase should be

273   classified according to economics theory.

274

275   *Experiment 1: Discussion*

276   Grackles were fast to learn an initial preference in the color association task (average 31 trials).

277   Their performance is similar to Western scrub-jays (Logan et al. 2016), 3 species of Darwin's

278   finches (Tebbich et al. 2010), and pigeons (Lissek et al. 2002) who learned in an average of

279   between 40-56 trials using a similar experimental design and passing criterion. These species are

280   faster than Pinyon jays, Clark's nutcrackers, a different group of Western scrub-jays (Bond et al.

281   2007), and Indian mynas (Griffin et al. 2013) who learned on average between 122-280 trials and

282   also had a similar passing criterion and experimental design. The Indian myna's design was

283   different, however they are included for comparison because they are one of the only other

284   species that has been tested in a similar behavioral flexibility-problem solving context. If

285   grackles trade off learning speed for inhibition (implicated in behavioral flexibility; Manrique et

286   al. 2013, Griffin & Guez 2014, Liu et al. 2016) or continuing to sample their environment, as

287   Indian mynas and Florida scrub jays appear to (Griffin et al. 2013, Bebus et al. 2016), then those

288   grackles that are faster learners will not be the most flexible (e.g., the fastest to reverse their

289   initially learned preference). However, if these traits are independent, then I expect grackle

290   learning speed not to correlate with flexibility as in spotted bowerbirds (Isden et al. 2013).

291

292   **Experiment 2: Color Association Reversal (learning speed)**

293   *Experiment 2: Methods*

294   The methods were the same as in the color association test (Experiment 1), except the food was

295   always placed in the silver tube rather than the gold tube, thus forcing the bird to reverse their

296   preference to consistently obtain the food. Because many other experiments (3-6) occurred

297   between Experiments 1 and 2, I first checked whether the grackles remembered Experiment 1

298   before moving them to Experiment 2. If they were successful in 9 or 10 out of their first 10 trials,

299   indicating that they remembered that the food was always in the gold tube, then they moved onto

300   reversal learning with the food always in the silver tube. If they were not successful in their first

301   10 trials, then they were given a refresher on Experiment 1 until they re-passed the original

302   criterion before moving onto reversal learning.

303

304   *Experiment 2: Results*

305   For their Experiment 1 refresher, Margarita, Cerveza, Michelada, and Jugo remembered that

306   food was always in the gold tube because they passed in their first 10 trials (Table 1, Figure 2).

307   Tequila, Horchata, Refresco, and Batido needed to re-achieve proficiency on Experiment 1,

308   requiring 30-80 trials before moving onto Experiment 2 (Table 1). Their re-learning patterns

309   followed the epsilon-decreasing strategy that all birds used before, except for Refresco who used

310   the epsilon-first strategy the first time and switched to the epsilon-decreasing strategy for the

311   refresher (Figure 2).

312   Most grackles reversed their color association, indicating they were behaviorally flexible.

313   Seven out of 8 grackles met the reversal learning success criterion (17 correct choices out of the

314   most recent 20 trials) in 70-130 trials (Table 1), but Batido stopped participating before reaching

315   criterion (Figure 3). All birds used the epsilon-decreasing strategy, but they were slower to learn

316   to reverse their previously learned preference than they were to initially learn the preference, and

317   many continued to explore throughout the experiment (Figure 3).

318   Faster learners were not more flexible. In the color association test, learning speed

319   (number of trials to learn a preference) did not significantly correlate with reversal learning

320   scores (number of trials to reverse a preference minus the number of trials to learn the

321   preference; Spearman's rank correlation test: S=84.14, p=0.25, rho=-0.50, n=7).

322

323   *Experiment 2: Discussion*

324   Grackles demonstrated behavioral flexibility in the color association task by quickly reversing

325   their initially learned preference (average 91 trials). Their performance was similar to 3 species

326   of Darwin's finches who reversed in an average of 76-95 trials (Tebbich et al. 2010). Darwin's

327   finches and grackles reversed more quickly than pigeons (Lissek et al. 2002), Pinyon jays,

328   Clark's nutcrackers, Western scrub-jays (Bond et al. 2007), and Indian mynas (Griffin et al.

329   2013) who learned on average between 142-380 trials.

330   Faster learners were not less flexible (e.g., slower to reverse a preference), which is

331   predicted if learning speed trades off with inhibition or continuing to sample the environment,

332   and was found in Indian mynas and Florida scrub-jays (Manrique et al. 2013, Griffin et al. 2013,

333   Griffin & Guez 2014, Bebus et al. 2016, Liu et al. 2016). There was no correlation between

334   grackle learning speed and behavioral flexibility, similar to results found in spotted bowerbirds

335   that were tested in the wild using a different experimental design and passing criterion (6

336   consecutive correct choices before trying an incorrect choice in 2 consecutive sessions; Isden et

337   al. 2013). Regardless of methodological differences, such unpredictable variation in behavioral

338   flexibility across species suggests that the underlying mechanisms are poorly understood and

339   require further investigation.

340

341   **CONTEXT 2: AESOP'S FABLE TESTS**

342   **Context 2: General Methods and Results**

343   *Spontaneous Stone Dropping*

344   **Methods:** Birds were given two sequential 5 min trials with the stone dropping training

345   apparatus and two stones to see whether they would spontaneously drop stones down tubes. The

346   stone dropping training apparatus was a clear acrylic box with a tube on top. The box contained

347   out of reach food on top of a platform that was obtainable by dropping a stone into the top of the

348   tube, which, when contacting the platform, forced the magnet holding it up to release the

349   platform (design as in Bird and Emery 2009 with the following tube dimensions: 90mm tall,

350   outer diameter=50mm, inner diameter=37 or 44mm). The food then fell from the platform to the

351   table. At the end of the first 5 min trial, the stones were moved to different locations on the table

352   and on the wooden blocks. The blocks made it easier to access the top of the tube.

353   **Results:** No grackle spontaneously dropped stones down the tube of the platform

354   apparatus, indicating that this was not a behavior that was easily innovated. Therefore, they all

355   underwent stone dropping training to allow them to participate in experiments 3-6.

356

357   *Stone Dropping Training*

358   **Methods:** Those birds that did not spontaneously drop stones down the tube on the stone

359   dropping training apparatus were trained to push or drop stones down tubes using this same

360   apparatus (Figure 4). Birds were given two stones and went from accidentally dropping stones

361   down the tube as they pulled at food under the stones, which were balanced on the edge of the

362   tube opening, to pushing or dropping stones into the tube from anywhere near the apparatus.

363   Once the bird proficiently pushed or dropped stones into the apparatus 30 times, they moved on

364   to obtain their reachable distance on a water tube.

365   Stone pushing/dropping proficiency was defined as consistently directing the stone to

366   tube opening from anywhere on the standing platform at the top of the apparatus. Not all motions

367   had to be in the direction of the tube opening because some grackles preferred to move the stone

368   to a particular location on the ramp (which may initially be in the opposite direction from the

369   tube) and push or drop it in from there, or push the stone in shorter, angular strokes. It was

370   permissible for a bird to throw one of the stones off the side of the apparatus (which occurred

371   sporadically in training and experiments) as long as they proficiently put the other stone in the

372   tube.

373   The training procedure was modified from Logan et al. (2014) to allow stone pushing

374   from a clear cast acrylic ramp placed on top of the tube rather than stone dropping by picking up

375   the stone from the table and putting it into the tube without a ramp (Figure 4). The modification

376   was necessary because grackles seem to form associations between the stones and the top of the

377   tube, the stones and the table where the food comes out, and the stones falling in one direction

378   only: down. When I placed the stones below the level of the top of the tube to try to train them to

379   pick the stones up and put them in the top of the tube, the grackles took the stones and dropped

380   them off the side of the apparatus or table, often placing them on the table and then looking at

381   where the platform should have fallen open. Placing the standing platform at the tops of the

382   water tubes for the experiments was implemented to mitigate this limitation. Once this change

383   was made, it was no longer necessary to train the grackles to pick up and drop the stones because

384   pushing them into the tube sufficed and required less training.

385   Similar to Western scrub-jays (Logan et al. 2016), the grackles inserted objects while

386   standing at the top of the tube rather than standing on the ground. The different standing position

387   should not influence their perception of the objects as they were inserted into the tube because

388   their heads were always over the top of the tube at the time of insertion, regardless of where they

389   were standing.

390   **Results:** Most grackles learned to push stones into a tube on the platform apparatus in

391   135-362 trials (Table 2), however Michelada showed a neophobic reaction (refused to approach

392   and interact with the apparatus even after repeated re-habituation attempts) to the stone falling

393   down the tube and did not habituate to this event, and Jugo learned too slowly to become

394   proficient by the time he needed to be released. Therefore, Michelada and Jugo were excluded

395   from the stone dropping experiments.

396

397   **Reachable Distance**

398   To determine how high to set the water levels in the Aesop's Fable experiments, a bird's

399   reachable distance was obtained. Food was placed on cotton inside a resealable plastic bag,

400   which was stuffed inside a standard water tube (a clear acrylic tube [170mm tall, outer

401   diameter=51mm, inner diameter=38mm] super glued to a clear acrylic base [300x300x3mm]) to

402   obtain the reachable distance without giving the bird experience with water. The food was first

403   placed within reach and then lowered into the tube in 1cm increments until the bird could not

404   reach it. The lowest height the bird could still reach was considered its reachable distance and

405   water levels in subsequent experiments were set to allow the desired number of objects to bring

406   the food within reach.

407

408   **An Additional Experiment that Only Tequila Began: Water vs. Sand**

409   **Methods:** To determine whether grackles can discriminate between functional and non-

410   functional substrates, two standard water tubes were placed on the table: one partially filled with

411   water (functional) and the other partially filled with sand (non-functional) to equal levels

412   (pseudorandomized for side; similar to Jelbert et al. 2014, Logan et al. 2014, Logan et al. 2016).

413   Before the experiment began, birds were habituated to the water and sand tubes for 10 trials by

414   taping over the openings at the top and placing food (peanut pieces) on the tops and at the bases

415   of each tube (there was no food inside the tubes). A trial continued until the bird ate all four food

416   pieces. The first tube from which food was taken was recorded and used as an indicator of a

417   potential preference. Preferences were discouraged by placing relatively more bait on the non-

418   preferred tube in the next trial. During the 20-trial experiment, four stones were placed in pairs

419   on the standing platforms at the top of each tube, and birds could insert them into the functional

420   water tube or non-functional sand-tube.

421   **Results:** Tequila participated in two trials, inserting 1 stone into the water tube in trial 1,

422   and 1 stone into first the water tube and then 1 stone in the sand tube in trial 2. However, he

423   started to refuse to participate in stone dropping all together, therefore I eliminated this

424   experiment and implemented a water tube proficiency assessment to re-motivate him to

425   participate in stone dropping experiments.

426

427   **Water Tube Proficiency Assessment**

428   **Methods:** Upon the implementation of the water tube proficiency assessment for Tequila after

429   the Water vs. Sand experiment, I required this assessment for all of the other birds to ensure a

430   more similar level of water tube experience and to ensure they were able to transfer their stone

431   dropping skills from a platform apparatus to a water tube context. This was likely not necessary

432   for the other grackles though because Tequila transferred his stone dropping skills directly from

433   the platform apparatus to a water tube. Grackles were given a standard tube partially filled with

434   water with a peanut float and four stones (9-14g, each displaces 5-6mm water), which they could

435   drop into the tube to raise the water level and consequently reach the food. Once a bird

436   accomplished 30 consecutive proficient trials, they moved on to Experiment 3. Proficiency was

437   defined as in the stone dropping training section above.

438   **Results:** Most grackles immediately applied their stone dropping skills to a water tube

439   context as indicated by their proficiency on their first trial (Cerveza, Margarita, Refresco, and

440   Batido). Horchata was proficient by her second trial. After Tequila's refusal to insert stones into

441   tubes following his Water vs. Sand experiment, he needed 76 trials of water tube proficiency

442   assessment to complete stone insertion proficiency again.

443

444   **Accidental Object Insertions**

445   Because objects were placed near the top of the tube to allow birds to push objects into the tube,

446   it was also possible to accidentally push or kick an object into the tube. Accidental insertions

447   were noted (see Supplementary Tables S1-S3) and included in analyses because birds could learn

448   about the affordances of the task if an object fell into the water, regardless of whether it was

449   chosen or accidental. Some trials were allowed to consist of only an accidental insertion or

450   insertions because the bird was losing motivation (e.g., refusing to come to the table or interact

451   with the apparatus) and would not have finished the trial otherwise. Counting these as trials errs

452   on the conservative side because throwing the data out and not counting it as a trial removes the

453   ability to account for learning in analyses.

454

455   **Experiment 3: Heavy vs. Light**

456   *Experiment 3: Methods*

457 I modified the Aesop's Fable paradigm to test behavioral flexibility by requiring birds to change

458 preferences using four experiments involving two preference changes, similar to reversal

459 learning experiments. In Experiment 3 (Heavy vs. Light), grackles were given heavy and light

460 objects with the former being twice as functional as the latter, therefore grackles should prefer to

461 insert heavy objects if they attend to the functional properties of the task or if they can quickly

462 form an association between the more functional option and success. However, unlike in most

463 previous experiments (e.g., Cheke et al. 2011, Taylor et al. 2011, Jelbert et al. 2014, Logan et al.

464 2014, but see Logan et al. 2016), the light objects sank rather than floated, thus if enough were

465 inserted, the food could be reached. I made this modification so that in Experiment 4 (Heavy vs.

466 Light Magic) when the heavy objects became non-functional by sticking to a magnet placed

467 inside the tube above the water, the light objects would now be the functional option because

468 they could fall past the magnet into the water. Individuals that prefer heavy objects or have no

469 preference in the Heavy vs. Light experiment should change their preference in the Heavy vs.

470 Light Magic experiment to preferring neither object or light objects. This would indicate that

471 their preferences are sensitive to changing contexts.

472 In the Heavy vs. Light experiment, one standard water tube was presented with 4 heavy

473 (steel rod wrapped in fimo clay, weight=10g, each displaces 2-3mm of water) and 4 light (plastic

474 tube partially filled with fimo clay, weight=2g, each displaces 1-1.5mm of water) objects placed

475 in pseudorandomized (as explained for the color association test) pairs near the top of the tube

476 (both objects were 21-24mm long and 8mm in diameter; Figure 5A). Heavy objects had a larger

477 volume (1,056-1,207mm^3) and displaced 0.5-2mm more water than light objects (volume

478 roughly 500mm^3), which had a hollow end. Thus the heavy objects were more functional than

479 the light objects, but importantly, both objects were functional. Each bird had three opportunities

480   to interact with the objects before the experiment began: one heavy and one light object was

481   placed on the table (pseudorandomized for side) with food underneath and on top of each object.

482   The object that was first touched was recorded and a trial continued until the bird interacted with

483   both objects. If one object was preferred (as indicated by approaching it first 2-3 times), then

484   more food was placed on the other object to try to eliminate any object preference before the

485   experiment began. Each grackle needed to choose each object type first at least once during this

486   process, which resulted in Horchata receiving four interactions and Batido five to ensure a lack

487   of preference. After object interaction trials, each bird was given the 20-trial experiment.

488

489   *Experiment 3: Results*

490   Grackles varied in how efficiently they solved problems. Four grackles (Tequila, Margarita,

491   Batido, and Refresco) were 3.4-5.2 times more likely to choose heavy objects rather than the less

492   functional light objects, while two grackles (Cerveza and Horchata) had no preference (they were

493   0.6-1.4 times more likely to choose heavy objects; see Table 2 for binomial test results and Table

494   3 for GLMM results). Cerveza and Horchata's performances improved across trials: they were

495   3.9-4.4 times more likely to choose heavy objects as trial number increased, indicating that they

496   learned through trial and error that the heavy objects were more functional (Table 3). The other

497   grackles' performances did not improve with increasing trial number, indicating that they might

498   have been biased toward heavy objects from the beginning of the experiment or perhaps they

499   used prior knowledge to solve the task (e.g., causal knowledge about functional differences

500   regarding weight or volume; Table 3). Horchata was not motivated to participate in the water

501   tube experiments: she required bait between almost all trials to get her to continue to interact

502   with the apparatus, which might have influenced her lack of success. All choices in all trials for

503   all birds are presented Supplementary Table S1.

504

505   *Experiment 3: Discussion*

506   Despite not being a tool-using species, grackles performed well in the tool-using Aesop's Fable

507   paradigm object discrimination tests. Four out of 6 grackles discriminated between object types

508   as indicated by their preference for inserting heavy objects significantly more than light objects.

509   Their object discrimination performance is similar to that in other successful species where

510   individuals preferred to insert heavy objects that sank rather than light objects that floated and

511   thus were not functional at all (Cheke et al. 2011 and 2012, Taylor et al. 2011, Jelbert et al. 2014,

512   Logan et al. 2014). This is in contrast to 4-year-old children who performed poorly by having no

513   object preference (Cheke et al. 2012) and Western scrub-jays who successfully obtained the food

514   but did not discriminate between object types (Logan et al. 2016). Perhaps individuals who

515   discriminated between object types did so because they discriminated between the causal

516   properties of the objects, and thus used causal cognition to solve this task. However, other

517   explanations cannot be ruled out yet: they may have had an initial preference for heavy objects

518   (object bias hypothesis), they might have noticed that inserting a heavy object brings the food

519   closer to the top of the tube than inserting a light object (perceptual-motor feedback hypothesis),

520   or they may have associated retrieving food with the heavy objects (Jelbert et al. 2015).

521   To begin to address the object bias hypothesis, grackles had a modified version of Heavy

522   vs. Light where the light objects, rather than floating and being non-functional, displaced about

523   half the amount of water as the heavy objects (as in Logan et al. 2016). In previous studies

524   testing sinking vs. floating, all participating individuals preferred the sinking objects: 2/2

525   Eurasian jays (Cheke et al. 2011), 4/4 New Caledonian crows (Taylor et al. 2011), 6/6 New

526   Caledonian crows (Jelbert et al. 2014), 6/6 New Caledonian crows (Logan et al. 2014), and

527   children age 5 and over (Cheke et al. 2012). Additionally, in a similar Aesop's Fable experiment

528   using objects of the same weight but different volumes (Solid vs. Hollow), all individuals chose

529   the larger volume significantly more: 5/5 New Caledonian crows (Jelbert et al. 2014) and 6/6

530   New Caledonian crows (Logan et al. 2014). This led to the alternative hypothesis that these

531   individuals, rather than attending to function, had a bias toward heavy objects because they were

532   more similar to familiar objects in the wild than the light objects (Logan et al. 2014, Jelbert et al.

533   2015). Making both objects functional, but to different degrees, allowed me to partially test the

534   object bias hypothesis: Rather than making a dichotomous choice where the heavy object was the

535   only functional option, birds could either have an object bias and/or attend more closely to the

536   functional differences in the properties of the objects and choose the heavy object more, or

537   exhibit no object bias by having no object preference and still succeed. Indeed, Western scrub-

538   jays that participated in this modified Heavy vs. Light experiment showed no object preferences,

539   indicating that birds might not have a general bias toward heavy objects (Logan et al. 2016). Two

540   grackles had no object preference and successfully retrieved the reward, which further supports

541   the notion that, at the species level, object biases are not the default. The object bias hypothesis

542   cannot be ruled out for the 4 grackles that preferred to insert heavy objects when both objects

543   were functional. The heavy objects were approximately the same weight as the stones used in

544   stone dropping training, therefore individuals that preferred heavy objects may have relied on an

545   association between weight and function instead of, or in addition to, any similarity the heavy

546   objects might have had to objects in the wild.

547   Three of the 4 grackles that preferred heavy objects did not show a learning effect across

548   the 20 trials in this experiment, indicating that they relied on prior information about the world to

549   solve this task, which suggests that they had an object bias or they may have used causal

550   cognition. Their performance suggests that they did not simply associate the heavy objects with

551   reaching the food because both object types could result in a reward.

552

553   **Experiment 4: Heavy vs. Light Magic**

554   *Experiment 4: Methods*

555   The set up was the same as in Experiment 3, except there were magnets (2 super magnets on the

556   outside and 3 on the inside of the tube) attached to the tube above the water level such that the

557   heavy objects would stick to the magnets and not displace water, while the light objects could

558   fall past the magnets into the water, thus being the functional choice (Figure 5B). Birds were

559   given 3 heavy and 3 light objects, placed in pseudorandomized pairs near the top of the tube, and

560   20 trials were conducted.

561

562   *Experiment 4: Results*

563   Two grackles exhibited behavioral flexibility by changing their preference from Experiment 3.

564   Tequila and Refresco changed from choosing significantly more heavy objects in Experiment 3

565   to having no object preferences in this experiment (binomial test: p=0.60 and 1.00, respectively),

566   whereas Batido continued to prefer the non-functional heavy objects (binomial test: p=0.02) and

567   Horchata continued to have no object preference (binomial test: p=1.00; see Table 2 for binomial

568   test results, Table 3 for GLMM results, and Supplementary Table S2 for all choices made by all

569   birds). Margarita continued to prefer heavy items (binomial test: p=0.02) and Cerveza went from

570   having no preference to choosing significantly more non-functional heavy items (binomial test:

571   p=0.02) because they exhibited an intense interest in the magnet (Table 2; see a video clip at:

572   https://youtu.be/GhR6fGG1yc4). They repeatedly stuck heavy objects to the magnet and

573   attempted to pull them off, requiring almost no rewards between trials for participating, which

574   indicated a high degree of motivation (motivation that rapidly decreases if they fail experiments).

575   The experiment did not have the intended effect on their behavior. Tequila gave up after 17

576   trials, refusing to drop either type of object into the tube, indicating he may have inhibited his

577   heavy preference while at the same time not persisting with the light objects. Tequila and

578   Refresco's performance improved with trial number, indicating that they learned through trial

579   and error about which object was functional (Table 3). The other grackle's performances did not

580   change or they decreased with increasing trial number, indicating that they did not learn about

581   which object was functional (Table 3). Even though Tequila and Refresco did not learn to prefer

582   light objects in the amount of trials given, they did exhibit flexibility in that they changed their

583   preferences from heavy in the previous experiment to having no preference in this experiment.

584

585   *Experiment 4: Discussion*

586   Behavioral flexibility was exhibited by grackles because they changed their preferences when the

587   task changed. When the heavy objects in the Heavy vs. Light Magic experiment were no longer

588   functional because they stuck to a magnet, 2 grackles changed from having preferred heavy

589   objects when they were functional in Heavy vs. Light to having no object preference in the

590   Magic experiment. This demonstrates either an attention to the functional properties of objects or

591   that some grackles who previously might have been biased toward heavy objects were able to

592   counteract this bias in light of new associative information (i.e., associating light objects with the

593   previously rewarded movement of an object falling into the water). Either way, their behavior

594   changed when circumstances changed.

595   No grackle completely switched their preference to the light objects, which may have

596   been due to the difficult design of the apparatus, both theoretically and physically. Theoretically,

597   this experiment is similar to the U-tube Aesop's Fable experiment where birds that rely on causal

598   cues to solve these kinds of tasks should fail because no causal information is available about

599   how the apparatus works (Cheke et al. 2011, Jelbert et al. 2014, Logan et al. 2014). Instead, birds

600   must rely solely on associative cues to solve the task (e.g., dropping stones into the blue tube

601   results in a reward in its adjacent tube). If birds generally solve these kinds of problems

602   associatively, then they should learn this association at a similar rate to other Aesop's Fable

603   experiments. However, if birds rely to some degree on causal cues to solve these problems, then

604   their learning speed should be impeded. The grackle's lack of learning to associate light objects

605   with a food reward may have been partially due to an attendance to causal cues when solving

606   these tasks. The apparatus was difficult physically as well: if one heavy item was inserted into

607   the tube, it stuck to the magnet and blocked access to the food regardless of how many light

608   objects were inserted. Thus, grackles had to inhibit inserting any heavy objects to solve this

609   problem, which made the task difficult. Tequila and Refresco's performance is consistent with

610   previous interpretations of Eurasian jay behavior on the U-tube apparatus: they were slow to

611   learn from associative cues when only associative information was available, therefore they

612   likely rely to a degree on causal cues when solving these tasks (Cheke et al. 2011). However, in

613   the case of the grackles, I cannot rule out that the physical difficulty of the apparatus was the

614   cause of their slow learning, which, if addressed, might have resulted in their quickly learning

615   this task using associative cues.

616

617 **Experiment 5: Narrow vs. Wide Equal Water Levels**

618 *Experiment 5: Methods*

619 Experiments 5 and 6 followed the same methods used for New Caledonian crows (Logan et al.

620 2014). To solve Experiment 5, objects must be inserted into a narrow (functional) rather than a

621 wide (non-functional) tube when water levels are equal in both tubes. In Experiment 6, the

622 narrow tube becomes non-functional because the water level is too low, therefore birds must

623 change their preference to inserting objects into the functional wide tube or to having no

624 preference (as long as they successfully reach the food in most trials) to demonstrate behavioral

625 flexibility.

626   To determine whether birds attend to volume differences in Experiment 5, a wide tube

627 and a narrow tube with equal water levels were presented with four objects made out of fimo

628 clay (30x10x5mm, 3-4g, each object displaced 1-2mm in the wide tube and 5-6mm in the narrow

629 tube; Logan et al. 2014; Figure 6). Two objects were placed on the platform near the narrow tube

630 opening and two objects on the platform near the wide tube opening. The objects were functional

631 only if dropped into the narrow tube because the water levels were set such that dropping all of

632 the objects into the wide tube would not bring the floating food within reach. However, dropping

633 1-2 objects into the narrow tube would raise the water level enough to reach the food. Both tubes

634 were 170mm tall with 3mm thick lids that constricted the opening to 25mm in diameter to

635 equalize the bird's access to the inside of each tube, and super glued to a clear acrylic base

636 (300x300x3mm). The wide tube (outer diameter=57mm, inner diameter=48mm,

637 volume=307,625mm³) was roughly equally larger than the standard water tube (dimensions

638 above, volume=192,800mm³) as the narrow tube was smaller (outer diameter=38mm, inner

639   diameter=25mm, volume=83,449mm^3). The position of the tubes was pseudorandomized for side

640   to ensure that tube choices were not based on a side bias, and 20 trials were conducted. Before

641   the experiment began, each bird had three opportunities to interact with the object, as in

642   Experiment 3, only here it was simply to habituate them to the clay object (one object type) and

643   not to train the birds not to prefer one object type over another.

644

645   *Experiment 5: Results*

646   Grackles did not discriminate between water volumes, indicating they were not efficient problem

647   solvers in this context. All three grackles that participated in this experiment displayed no

648   preference for dropping objects into the functional narrow tube or the non-functional wide tube

649   (binomial test: p=1.00 for every bird, Table 2; see Table 3 for GLMM results, and

650   Supplementary Table S3 for all choices by all birds). None of the grackles' performances

651   improved with trial number, indicating that they did not learn to distinguish which tube was

652   functional (Table 3). Batido appeared to rely on the strategy of dropping all objects into both

653   tubes regardless of which tube he received a reward from, although in trial 12, he picked up the

654   objects from the wide tube area and dropped them into the narrow tube even though he was

655   trained to only push stones, not drop them (Supplementary Table S3).

656   Some grackles did not initially transfer from dropping previous object types to dropping

657   the clay objects used in this experiment. It appeared as though they were trying to solve the

658   problem, but did not perceive the clay objects as being the kind of object one would drop into a

659   water tube. In these cases, additional training was implemented using a single standard water

660   tube and a mixture of clay objects and stones until the bird was willing to drop objects into the

661   tube even if they consisted only of clay objects. Cerveza transferred to dropping clay objects

662   after 4 training trials, but Tequila and Margarita were excluded from this experiment because

663   they did not transfer to dropping clay objects into tubes. After 14 training trials on a regular

664   water tube with stones and clay objects available to Tequila, it was clear that it would take many

665   more training trials than there was time for because his motivation was greatly diminished.

666   Margarita refused to participate in the training trials. Horchata was also excluded from this

667   experiment because she refused to interact with the objects.

668

669   *Experiment 5: Discussion*

670   Grackles did not discriminate between water volumes in the Narrow vs. Wide equal water level

671   experiment. Perhaps their skill in water displacement experiments is limited to objects, however

672   more experiments involving object and tube properties would need to be conducted to confirm

673   this.

674

675   **Experiment 6: Narrow vs. Wide Unequal Water Levels**

676   *Experiment 6: Methods*

677   Those grackles that passed Experiment 5 continued to this experiment to determine whether their

678   tube choices adjusted to changing circumstances. This experiment was the same as Experiment

679   5, except the water level in the narrow tube was lowered to 5cm from the table, thus making the

680   food unreachable even if all objects were dropped into this tube (as in Logan et al. 2014). The

681   water level in the wide tube was raised such that the bird could reach the food in 1-4 object

682   drops, and 20 trials were conducted.

683

684   *Experiment 6: Results*

685 No grackle passed Experiment 5, indicating they were not sensitive to the differences in water

686 volumes; therefore they were not given Experiment 6, which would have investigated their

687 behavioral flexibility in this context.

688

689 *Experiment 6: Discussion*

690 While grackles were not given experiment 6, it is interesting to compare their behavioral

691 flexibility in the Heavy vs. Light experiments with New Caledonian crow performances in the

692 Narrow vs. Wide experiments because crows previously showed behavioral flexibility in this

693 context (New Caledonian crows were not given the Heavy vs. Light Magic experiment because it

694 was not invented yet, which is why I cannot directly compare their behavioral flexibility

695 performances). Four out of 6 New Caledonian crows preferred to drop objects into the functional

696 narrow tube rather than the non-functional wide tube. When the wide tube became the functional

697 option, 3 New Caledonian crows changed their preference to the wide tube and 1 changed to no

698 preference (Logan et al. 2014). Grackle performances in the Heavy vs. Light experiments were

699 similar to the New Caledonian crow that changed from narrow to no preference in the Narrow

700 vs. Wide experiments.

701

702 **Context 2: First Choices on First Trials**

703 First trial performance on tasks that can involve more than trial and error learning is considered

704 an important feature when evaluating cognitive performance because it gives an indication of

705 what the individuals might infer about the task. All six grackles chose the more functional heavy

706 objects as their first choice in their first trial in Heavy vs. Light, which indicates that they

707 preferred the heavy objects from the very beginning of the experiment (Supplementary Table

708   S1). Five out of six grackles chose the non-functional heavy objects in Heavy vs. Light Magic

709   (Supplementary Table S2), which is not surprising given that they had learned to prefer heavy

710   objects in the previous experiment and had likely never interacted with a magnet before,

711   therefore they should have had no reason to have a prior understanding of how the Magic

712   experiment worked. Two out of three grackles chose the functional narrow tube in Narrow vs.

713   Wide with equal water levels, indicating no group-level initial preference for a particular tube

714   (Supplementary Table S3).

715

716   **Context 2: Did choice number influence the results?**

717   Individuals could learn how the water tube tasks worked with each choice they made, potentially

718   making each choice dependent on previous choices. Multiple choices could be made per trial;

719   therefore I analyzed how independent choice number was. Choice number was modeled as a

720   random factor in the GLMM and did not influence the results, indicating that choices appear

721   independent of each other (Table 3). This result combined with the result that some individuals

722   increased their success across the 20-trials, indicates that learning occurred at the trial level, not

723   at the level of individual choices. Perhaps learning was linked to the receipt of a food reward,

724   which was obtained near the end of a trial, rather than the movement of the food in the tube,

725   which could occur with each choice.

726

727   **Did behavioral flexibility correlate across contexts?**

728   Those grackles that were more behaviorally flexible in the Aesop's Fable context (i.e., changed

729   preferences) were not more behaviorally flexible in the color association context (i.e., faster to

730   reverse a previously learned preference; Spearman's rank correlation test: S=28.89, p=0.45,

731   rho=-0.44, n=5; Table 4).

732

733   **GENERAL DISCUSSION**

734   Grackles demonstrated behavioral flexibility in two contexts, the Aesop's Fable paradigm and a

735   color association test. Contrary to predictions, behavioral flexibility did not correlate across

736   contexts: Those grackles that exhibited flexibility in the Aesop's Fable paradigm were not the

737   fastest to reverse a previously learned color preference in the color association test. Four out of 6

738   grackles exhibited efficient problem solving abilities, similar to other successful species, by

739   choosing heavy objects significantly more in the Heavy vs. Light experiment. Problem solving

740   efficiency did not appear to be directly linked with behavioral flexibility because not all grackles

741   (2/4) that were efficient problem solvers changed their preference in the Aesop's Fable

742   paradigm. Problem solving speed (Experiment 1) also did not significantly correlate with

743   reversal learning scores (Experiment 2), indicating that faster learners were not the most flexible.

744   The lack of correlations within and across contexts for behavioral flexibility and for

745   behavioral flexibility and problem solving supports the hypothesis that behavioral flexibility is

746   not only an independent source of variation distinct from problem solving ability and speed

747   (Cole et al. 2011), but also independent across contexts (Griffin & Guez 2014). This finding

748   suggests that investigations of behavioral flexibility should conduct multiple tests of behavioral

749   flexibility in different contexts to more fully understand how it relates to itself, cognition, and

750   other variables of interest (Griffin and Guez 2014). Maintaining many independent sources of

751   variation in a population could be useful for successfully adapting to new environments. Western

752   bluebirds were found to rely on a range of dispersal strategies that already existed in their

753   population when colonizing a new habitat (Duckworth 2008). Thus, the more sources of

754   variation a population can access, the more likely it is that at least some of this variation will suit

755   the new situation and allow the species to adapt.

756   One potential explanation for why individuals varied in behavioral flexibility across

757   contexts regards the type of cognition used in each context. Causal cognition and/or trial and

758   error learning based on multiple cues (e.g., object type, movement of the food with each object

759   drop) could be used to solve the Aesop's Fable tasks (Experiments 3-6), whereas only trial and

760   error learning based on one cue (i.e., color) could be used to solve the color association tasks

761   (Experiments 1-2). If individuals varied in the kinds of cues they attended to this might have

762   caused the differences in performance across contexts. Such individual variation in attention to

763   particular cues was found in Eurasian jays (Cheke et al. 2011). Behavioral flexibility will need to

764   be tested in more contexts to determine whether individual differences are due to differential cue

765   use or different contexts.

766   A higher number of learning strategies in the color association tests did not necessarily

767   indicate flexibility in the color association context or the Aesop's Fable context. Refresco was

768   one of the two behaviorally flexible individuals in the water tube experiments (3-4), and about

769   average in reversing a color preference (Experiment 2; Table 4). He was also the only grackle to

770   use more than one learning strategy in the color association experiments: he used the epsilon-first

771   strategy to sample the environment once before arriving at the correct solution and then he

772   stayed with the correct choice for the rest of Experiment 1. He then switched his learning

773   strategy to epsilon-decreasing for his color association refresher and for reversal learning

774   (Experiment 2), which is the same strategy the rest of the birds used in Experiments 1 and 2.

775   Individuals using the epsilon-decreasing strategy sample the environment extensively before

776   consistently making the correct choice. Because there was almost no individual variation in

777   learning strategies it is difficult to understand how this trait covaries with behavioral flexibility.

778   However, it suggests that a variety of learning strategies is not required for a large amount of

779   variation in behavioral flexibility, problem solving ability, and problem solving speed to exist in

780   a population.

781   That behavioral flexibility did not correlate across contexts or with problem solving

782   ability (Experiment 3) or speed (Experiment 1) reveals how little we know about behavioral

783   flexibility, and provides an immense opportunity for future research to explore how individuals

784   and species can use behavior to react to changing environments.

785

786   **ACKNOWLEDGEMENTS**

787   I am grateful to Luisa Bergeron, Christin Palmstrom, Linnea Palmstrom, and Michelle Gertsvolf

788   for trapping and aviary assistance; Alexis Breen for helping to train Jugo; Brigit Harvey for stone

789   dropping training consultations; Steve Rothstein for scouting grackles and for use of the aviaries;

790   Joe Jablonski and David Bothman for making the apparatuses; Jill Zachary and Kathy Frye at

791   Santa Barbara City Parks and Recreation for use of the Andree Clark Bird Refuge and East

792   Beach Park; Estelle Sandhaus and Chris Briggs at the Santa Barbara Zoo for access to wild

793   grackles; Karrie Black for managing purchasing and grants; Alex Thornton for logistical advice;

794   Dieter Lukas for conceptual input and analysis feedback, Alecia Carter for manuscript

795   restructuring advice, and the rest of LARG for feedback; Krista Fahy at the Santa Barbara

796   Museum of Natural History for helping with permit applications; Kristine Johnson, Sarah

797   Overington, Julie Morand-Ferron, and Neeltje Boogert for grackle advice and trap plans and

798   manuscript feedback; Manny Garcia for veterinary and permit support; Mary Hunsicker and

799   Bertrand Lemasson for assistance with making the trap; Margaret Tarampi, Eric Egenolf,

800   Rebecca Schaefer, and Sam Franklin for brainstorming object designs; Will Hoppitt for GLMM

801   effect size interpretation assistance; and Irina Mikhalevich, Ljerka Ostojić, Jennifer Vonk, Zoe

802   Johnson-Ulrich, and an anonymous reviewer for manuscript feedback.

803

804   **REFERENCES**

805   **Akaike H. 1981.** Likelihood of a model and information criteria. *Journal of*

806   *Econometrics* **16**:3-14 doi:10.1016/0304-4076(81)90071-3

807   **Auersperg, AMI, von Bayern AMP, Gajdon GK, Huber L, Kacelnik, A. 2011.**

808   Flexibility in problem solving and tool use of kea and New Caledonian crows in a

809   multi access box paradigm. *PloS One* **6**:e20231. doi:10.1371/journal.pone.0020231

810   **Bates D, Maechler M, Bolker B. 2011.** lme4: Linear mixed-effects models using S4 classes. R

811   package version 0.999375-42. http://CRAN.R-project.org/package=lme4. Accessed 12

812   January 2016

813   **Bebus, SE, Small TW, Jones, BC, Elderbrock EK, Schoech SJ. 2016.** Associative learning is

814   inversely related to reversal learning and varies with nestling corticosterone exposure.

815   *Animal Behaviour* **111**:251-260 doi: 10.1016/j.anbehav.2015.10.027

816   **Bird CD, Emery NJ. 2009.** Rooks use stones to raise the water level to reach a floating worm.

817   *Current Biology* **19**:1410-1414 doi:10.1016/j.cub.2009.07.033

818   **Bond AB, Kamil AC, Balda RP. 2007.** Serial reversal learning and the evolution of behavioural

819   flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus*,

820   *Nucifraga columbiana, Aphelocoma californica*). *Journal of Comparative Psychology*

821   **121**:372 doi:10.1037/0735-7036.121.4.372

822   **Buckner C. 2013.** A property cluster theory of cognition. *Philosophical Psychology* **28**:307-336

823   doi: 10.1080/09515089.2013.843274

824   **Burnham KP, Anderson DR. 2002.** *Model selection and multimodel inference: a practical*

825   *information- theoretic approach*, 2nd edn. New York, NY: Springer.

826   **Cheke LG, Bird CD, Clayton NS. 2011.** Tool-use and instrumental learning in the Eurasian jay

827   (*Garrulus glandarius*). *Animal Cognition* **14**:441-455 doi:10.1007/s10071-011-0379-4

828   **Cheke LG, Loissel E, Clayton NS. 2012.** How do children solve Aesop's Fable? *PLoS ONE*

829   **7**:e40574 doi:10.1371/journal.pone.0040574

830   **Chow PKY, Lea SE, Leaver LA. 2016.** How practice makes perfect: the role of persistence,

831   flexibility and learning in problem-solving efficiency. *Animal Behaviour* **112**:273-283

832   doi: 10.1016/j.anbehav.2015.11.014

833   **Clayton NS, Dickinson A. 1998.** Episodic-like memory during cache recovery by scrub jays.

834   *Nature* **395**:272-274 doi:10.1038/26216

835   **Dally JM, Emery NJ, Clayton NS. 2006.** Food-caching western scrub-jays keep track of who

836   was watching when. *Science* **312**:1662-1665 doi:10.1126/science.1126539

837   **Duckworth RA. 2008.** Adaptive dispersal strategies and the dynamics of a range expansion. *The*

838   *American Naturalist* **172**:S4-S17 doi:10.1086/588289

839   **Ghahremani DG, Monterosso J, Jentsch JD, Bilder RM, Poldrack RA. 2010.** Neural

840   components underlying behavioural flexibility in human reversal learning. *Cerebral*

841   *Cortex* **20**:1843-1852 doi:10.1093/cercor/bhp247

842   **Griffin AS, Guez D. 2014.** Innovation and problem solving: a review of common mechanisms.

843   *Behavioural Processes* **109**:121-134 doi:10.1016/j.beproc.2014.08.027

844 **Griffin AS, Guez D, Lermite F, Patience M. 2013.** Tracking changing environments:

845   Innovators are fast, but not flexible learners. *PLoS ONE* **8**:e84907

846   doi:10.1371/journal.pone.0084907

847 **Hadfield J. 2010.** MCMCglmm: Markov chain Monte Carlo methods for generalised linear

848   mixed models.

849   http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=E395FF7367BF580D916E465F

850   55302C60?doi=10.1.1.160.5098&rep=rep1&type=pdf Accessed 7 May 2015

851 **Hadfield J. 2014a.** MCMCglmm: MCMC generalised linear mixed models. R package.

852   https://cran.r-project.org/web/packages/MCMCglmm/index.html Accessed 30 July 2015

853 **Hadfield J. 2014b.** MCMCglmm course notes. http://cran.r-

854   project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf. Accessed 7 May

855   2015

856 **Isden J, Panayi C, Dingle C, Madden J. 2013.** Performance in cognitive and problem-solving

857   tasks in male spotted bowerbirds does not correlate with mating success. *Animal*

858   *Behaviour* **86**:829-838. doi:10.1016/j.anbehav.2013.07.024

859 **Jelbert SA, Taylor AH, Cheke LG, Clayton NS, Gray, RD. 2014.** Using the Aesop's fable

860   paradigm to investigate causal understanding of water displacement by New Caledonian

861   crows. *PLoS ONE* **9**:e92895 doi:10.1371/journal.pone.0092895

862 **Jelbert SA, Taylor AH, Gray RD. 2015.** Investigating animal cognition with the Aesop's Fable

863   paradigm: Current understanding and future directions. *Communicative & Integrative*

864   *Biology* **8**:e1035846 doi:10.1080/19420889.2015.1035846

865 **Lefebvre L, Whittle P, Lascaris E, Finkelstein A. 1997.** Feeding innovations and forebrain

866   size in birds. *Animal Behaviour* **53**:549-560 doi:10.1006/anbe.1996.0330

867   **Lefebvre L, Nicolakakis N, Boire D. 2002**. Tools and brains in birds. *Behaviour* **139**:939-973

868   doi:10.1163/156853902320387918

869   **Lissek S, Diekamp B, Güntürkün O. 2002.** Impaired learning of a colour reversal task after

870   NMDA receptor blockade in the pigeon (*Columba livia*) associative fore-brain

871   (Neostriatum Caudolaterale). *Behavioral Neurosciences* **116**:523-529 doi:10.1037//0735-

872   7044.116.4.523

873   **Liu Y, Day LB, Summers K, Burmeister SS. 2016.** Learning to learn: advanced behavioural

874   flexibility in a poison frog. *Animal Behaviour* **111**:167-172.

875   doi:10.1016/j.anbehav.2015.10.018

876   **Logan C. 2015.** Great-tailed grackle behavioral flexibility and problem solving experiments,

877   Santa Barbara, CA USA 2014-2015. KNB Data Repository.

878   https://knb.ecoinformatics.org/#view/corina_logan.15.6

879   **Logan CJ, Harvey B, Schlinger BA, Rensel M. 2016.** Western scrub-jays do not appear to

880   attend to functionality in Aesop's Fable experiments. *PeerJ* **4**:e1707 doi:

881   10.7717/peerj.1707

882   **Logan CJ, Jelbert SA, Breen AJ, Gray RD, Taylor AH. 2014.** Modifications to the Aesop's

883   Fable paradigm change performances in New Caledonian crows. *PLoS ONE* **9**:e103049

884   doi:10.1371/journal.pone.0103049

885   **Manrique HM, Völter CJ, Call J. 2013.** Repeated innovation in great apes. *Animal Behaviour*

886   **85**:195-202 doi:10.1016/j.anbehav.2012.10.026

887   **McInerney RE. 2010.** *Multi-Armed Bandit Bayesian Decision Making*. University of Oxford,

888   Oxford: Technical Report.

889   **Overington SE, Cauchard L, Cote KA, Lefebvre L. 2011.** Innovative foraging behaviour in

890   birds: What characterizes an innovator? *Behavioural Processes* **87**:274-285

891   doi:10.1016/j.beproc.2011.06.002

892   **Peer BD. 2011.** Invasion of the emperor's grackle. *Ardeola* **58**:405-409

893   doi:10.13157/arla.58.2.2011.405

894   **Pyle P. 2001.** *Identification Guide to North American Birds Part 1*. Ann Arbor, MI: Sheridan

895   Books, Inc.

896   **R Core Team. 2016.** *R: A language and environment for statistical computing*. R Foundation for

897   Statistical Computing, Vienna, Austria. https://www.R-project.org.

898   **Raby CR, Alexis DM, Dickinson A, Clayton NS. 2007.** Planning for the future by western

899   scrub-jays. *Nature* **445**:919-921 doi:10.1038/nature05575

900   **Seed AM, Tebbich S, Emery NJ, Clayton NS. 2006.** Investigating physical cognition in rooks,

901   *Corvus frugilegus*. *Current Biology* **16**:697-701 doi:10.1016/j.cub.2006.02.066

902   **Sol D, Duncan RP, Blackburn TM, Cassey P, Lefebvre L. 2005.** Big brains, enhanced

903   cognition, and response of birds to novel environments. *Proceedings of the National*

904   *Academy of Sciences* **102**:5460-5465 doi:10.1073□pnas.0408145102

905   **Sol D, Lefebvre L. 2000.** Behavioural flexibility predicts invasion success in birds introduced to

906   New Zealand. *Oikos* **90**:599-605 doi:10.1034/j.1600-0706.2000.900317.x

907   **Sol D, Székely T, Liker A, Lefebvre L. 2007.** Big-brained birds survive better in nature.

908   *Proceedings of the Royal Society of London B* **274**:763-769 doi:10.1098/rspb.2006.3765

909   **Sol D, Timmermans S, Lefebvre L. 2002.** Behavioural flexibility and invasion success in birds.

910   *Animal Behaviour* **63**:495-502 doi:10.1006/anbe.2001.1953

911   **Taylor AH, Elliffe DM, Hunt GR, Emery NJ, Clayton NS, Gray RD. 2011.** New Caledonian

912   crows learn the functional properties of novel tool types. *PLoS ONE* **6**:e26887

913   doi:10.1371/journal.pone.0026887

914   **Tebbich S, Seed AM, Emery NJ, Clayton NS. 2007.** Non-tool-using rooks, *Corvus frugilegus*,

915   solve the trap-tube problem. *Animal Cognition* **10:**225-231 doi: 10.1007/s10071-006-

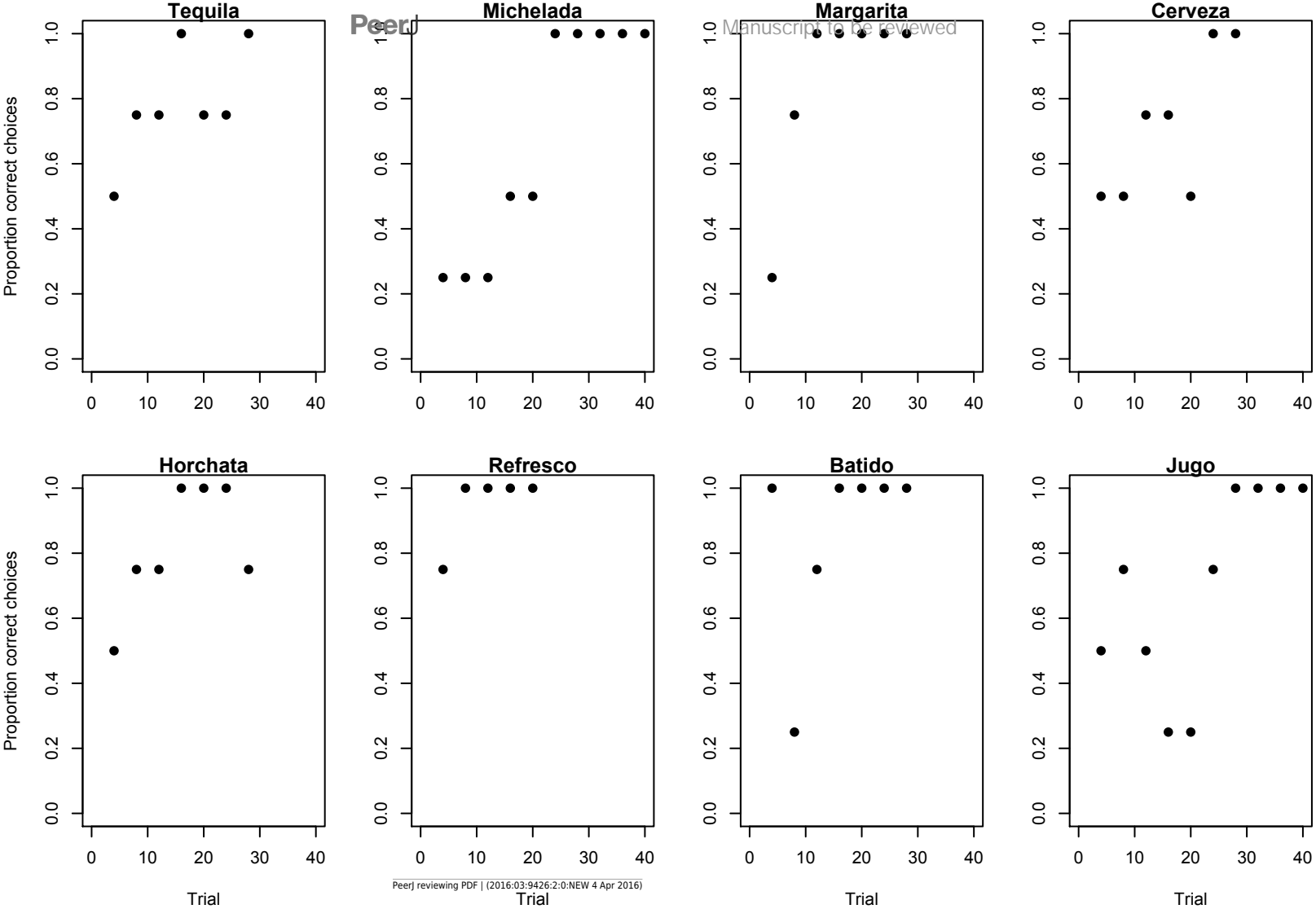
916   0061-4

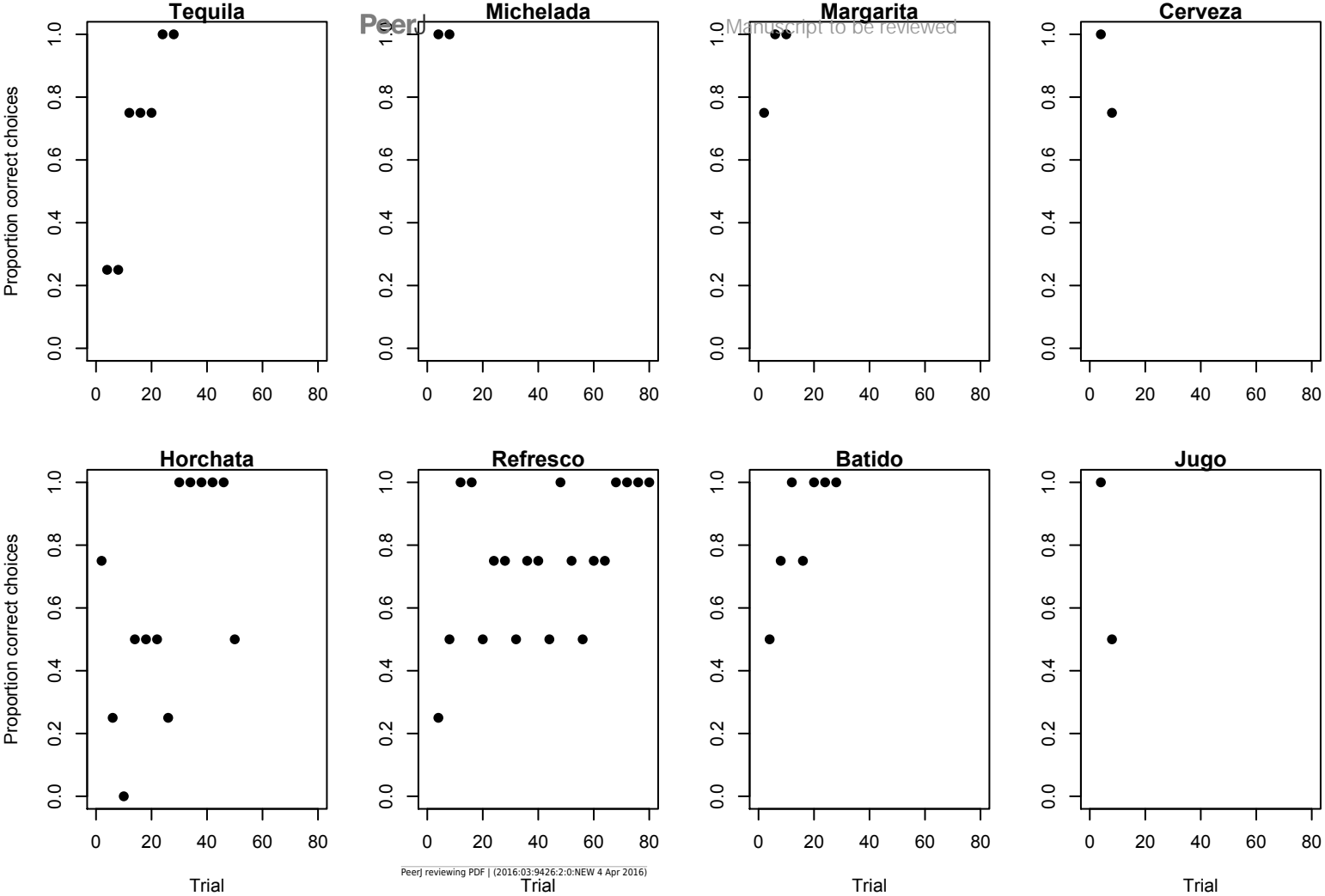
917   **Tebbich S, Sterelny K, Teschke I. 2010.** The tale of the finch: Adaptive radiation and

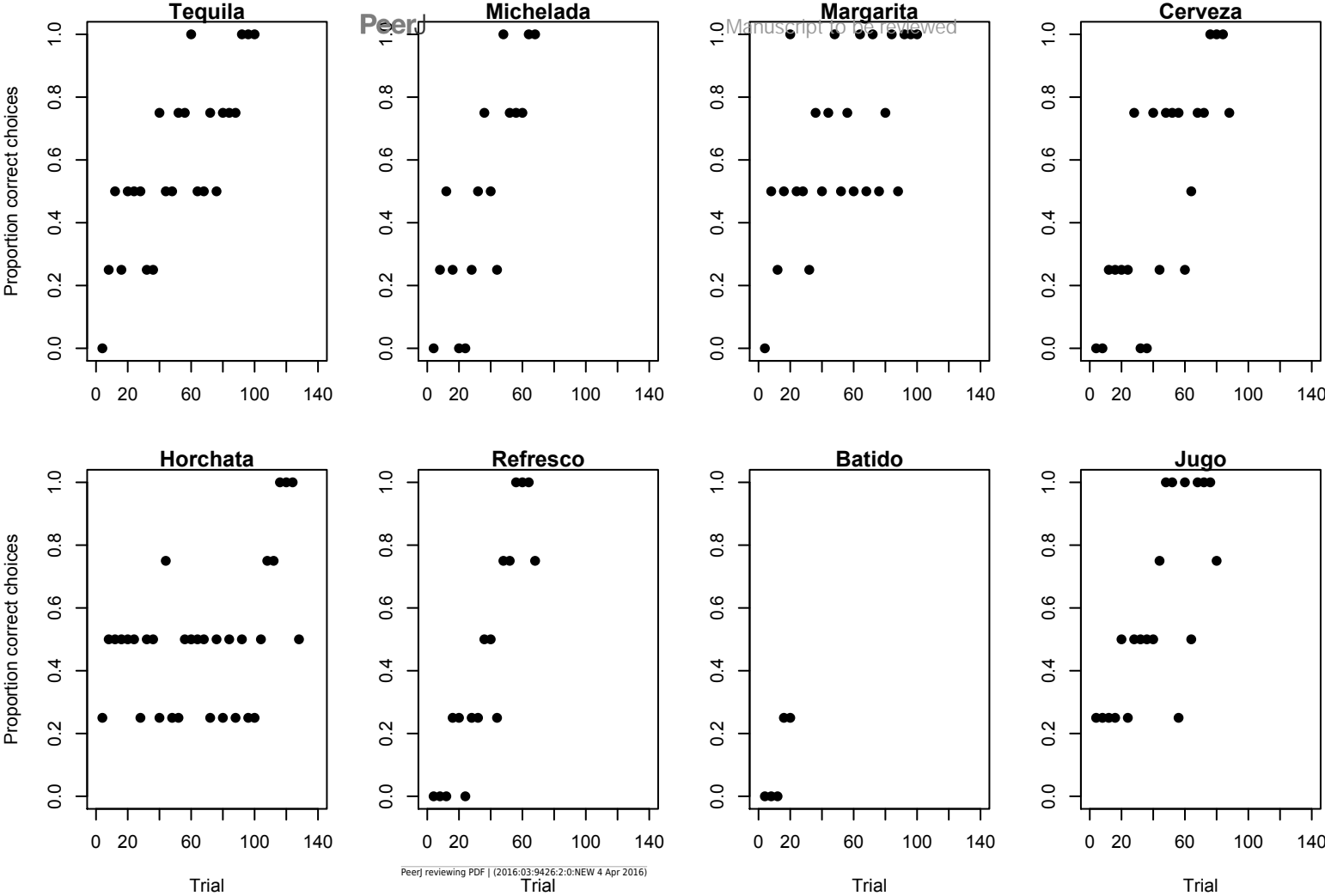
918   behavioural flexibility. *Philosophical Transactions of the Royal Society B* **365**:1099-1109

919   doi:10.1098/rstb.2009.0291

Learning strategies used during Experiment 1

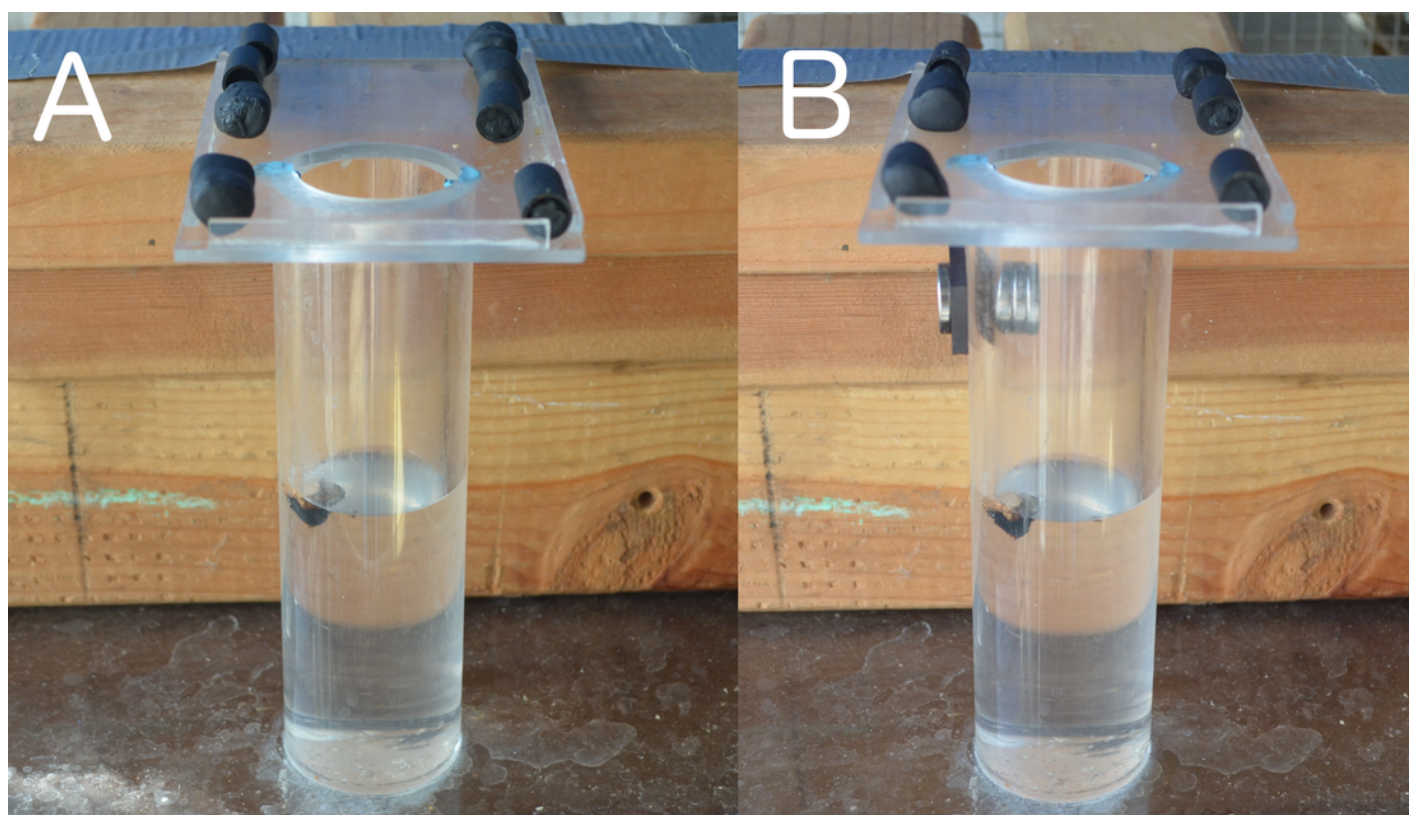
Learning strategies employed by grackles when learning to associate the gold tube with food in Experiment 1 as shown by the proportion of correct choices (non-overlapping sliding window of 4-trial bins) across the number of trials required to reach the criterion of 17/20 correct choices.

Learning strategies used during the Experiment 1 refresher

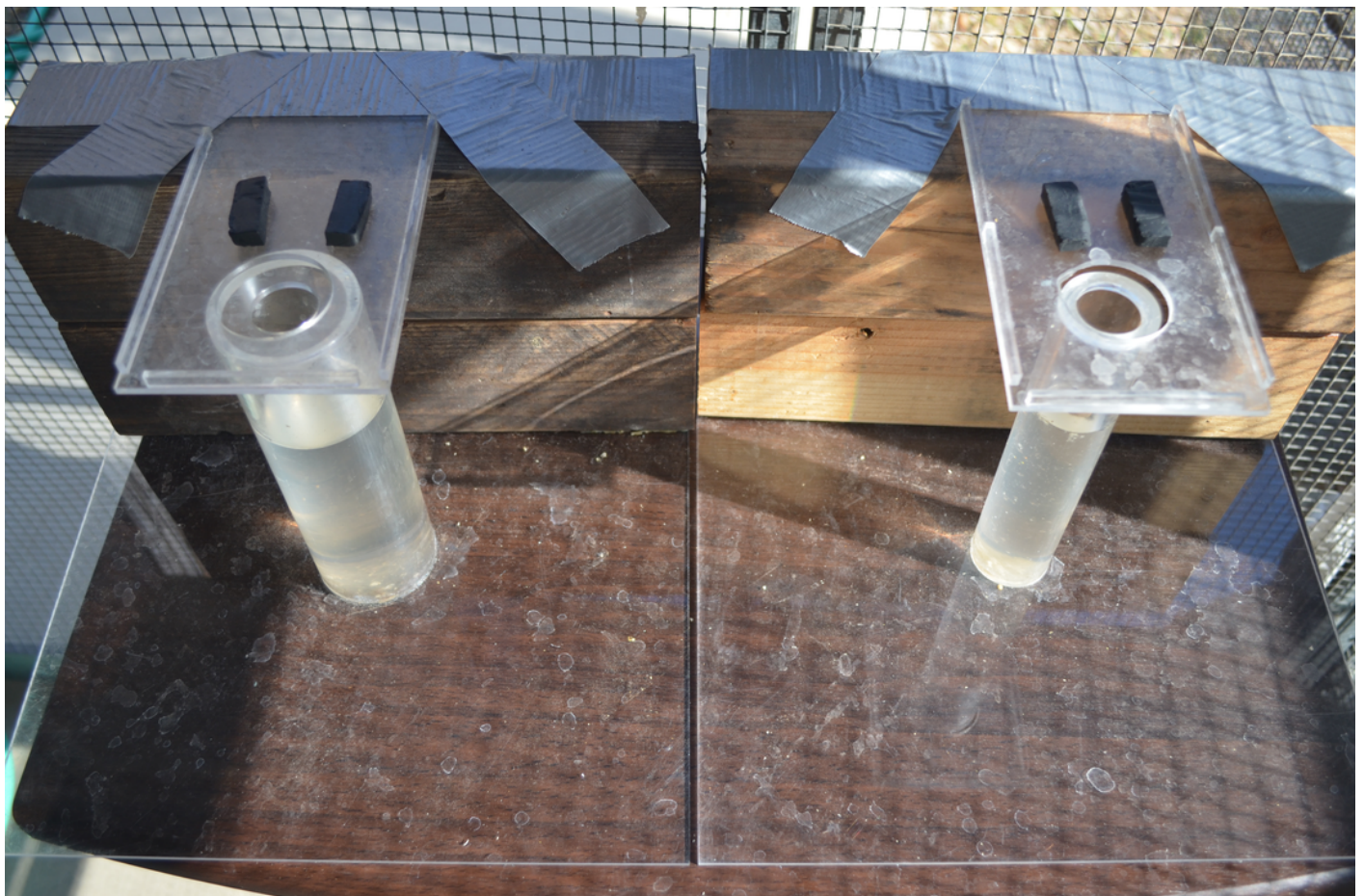
Learning strategies employed by grackles when checking whether they remember that the gold tube contained the food as shown by the proportion of correct choices (non-overlapping sliding window of 4-trial bins) across the number of trials required to reach the criterion of 17/20 correct choices.

Learning strategies used during reversal learning in Experiment 2

Learning strategies employed by grackles when learning to associate the silver tube with food, rather than their previously learned association between the gold tube and food, as shown by the proportion of correct choices (non-overlapping sliding window of 4-trial bins) across the number of trials required to reach the criterion of 17/20 correct choices.

Stone dropping training using a platform apparatus

Batido inserted a stone into the stone dropping training apparatus, which collapsed the platform and the peanut float fell out onto the table.

Heavy vs. Light (A) and Heavy vs. Light Magic (B)

Heavy objects are more functional than light objects in the Heavy vs. Light experiment (A), while the light objects are the only functional objects in the Heavy vs. Light Magic experiment (B).

Narrow vs. Wide

Dropping clay objects into the narrow tube in the Narrow vs. Wide equal water level experiment is the only way to reach the floating food.

Color association results

The number of trials needed to reach proficiency in Experiment 1 (learning speed), and Experiment 2 (reversal learning speed), and the number of trials needed to pass the Experiment 1 refresher (*=did not pass the refresher in 10 trials).

1 Table 1. Color association results: The number of trials needed to reach proficiency in

2 Experiment 1 (learning speed), and Experiment 2 (reversal learning speed), and the number of

3 trials needed to pass the Experiment 1 refresher (*=did not pass the refresher in 10 trials).

Bird	Sex	Experiment 1: Learning speed	Experiment 2: Reversal learning speed	Experiment 1 refresher
Tequila	M	30	100*	30
Margarita	F	30	100	10
Cerveza	F	30	90	10
Michelada	F	40	70	10
Horchata	F	30	130*	50
Refresco	M	20	70*	80
Batido	M	30	Incomplete*	30
Jugo	M	40	80	10

Performance per bird in the water tube experiments (3-6)

The number of stone dropping training trials needed to reach proficiency, and p-values from Bonferroni-Holm corrected (within experiment) binomial tests for each experiment (- = was not given this experiment). Note: Tequila was the first bird tested and I did not realise until after I trained him to pick up and drop the stones into the tube that I wanted to only train the other birds to push the stones into the tube to save training time. Therefore, the trial numbers for the other birds refer to proficiency to push objects into the tube, not pick up and drop them. Y=yellow, P=purple, B=blue, O=orange, R=red, G=green.

1 Table 2. Performance per bird per experiment: the number of stone dropping training trials

2 needed to reach proficiency, and the number of (more) functional choices/total number of

3 choices and p-values from Bonferroni-Holm corrected (within experiment) binomial tests for

4 each experiment (- = was not given this experiment). Note: Tequila was the first bird tested and I

5 did not realise until after I trained him to pick up and drop the stones into the tube that I wanted

6 to only train the other birds to push the stones into the tube to save training time. Therefore, the

7 trial numbers for the other birds refer to proficiency to push objects into the tube, not pick up and

8 drop them. Y=yellow, P=purple, B=blue, O=orange, R=red, G=green.

Bird (color rings)	Sex	Stone drop training trials	Heavy vs. Light	Heavy vs. Light Magic	Wide vs. Narrow Equal
Tequila (YP)	M	192 push / 233 drop	33/43 0.003 (heavy)	19/30 0.60	-
Margarita (PB)	F	362	41/49 0.00001 (heavy)	17/56 0.02 (heavy)	-
Cerveza (BO)	F	252	36/55 0.06	10/39 0.02 (heavy)	1.00
Michelada (OR)	F	-	-	-	-
Batido (OP)	M	179	38/51 0.002 (heavy)	28/37 0.02 (heavy)	1.00
Horchata (GR)	F	135	18/32 0.60	16/32 1.00	-
Refresco (PY)	M	204	46/67 0.009 (heavy)	17/35 1.00	1.00
Jugo (RB)	M	-	-	-	-

Examining test performance and learning effects in the water tube experiments (3-6)

Examining the influence of experiment, trial, and bird on test success (Test Performance) and whether success increased with trial number (Learning Effects), thus indicating a learning effect. GLMM: Choices Correct ~ Experiment*Trial*Bird, random = ~Choice Number. CI=credible intervals, Batido in italics indicates the intercept.

1 Table 3. Examining the influence of experiment, trial, and bird on test success (Test

2 Performance) and whether success increased with trial number (Learning Effects), thus

3 indicating a learning effect. GLMM: Choices Correct ~ Experiment*Trial*Bird, random =

4 ~Choice Number. CI=credible intervals, Batido in italics indicates the intercept.

	Test Performance			Learning Effects		
	Posterior Mean	**Lower 95% CI**	**Upper 95% CI**	**Posterior Mean**	**Lower 95% CI**	**Upper 95% CI**
Choice number	0.002	1.39E-16	0.002	-	-	-
Heavy vs. Light						
Batido	*1.36*	*-0.03*	*2.83*	-0.01	-0.13	0.12
Margarita	0.29	-2.31	2.83	0.04	-0.17	0.26
Cerveza	-1.92	-4.06	0.10	0.13	0.05	0.32
Horchata	-1.01	-3.14	1.14	0.01	-0.20	0.19
Refresco	-0.13	-2.07	2.13	-0.01	-0.19	0.16
Tequila	0.22	-1.96	2.47	-0.003	-0.20	0.22
Heavy vs. Light Magic						
Batido	-2.11	-4.66	-0.11	-0.06	-0.25	0.16
Margarita	-0.20	-3.63	3.14	-0.01	-0.32	0.28
Cerveza	2.13	-1.33	4.86	-0.18	-0.45	0.15
Horchata	2.60	-0.83	5.87	-0.03	-0.31	0.28
Refresco	-1.32	-4.73	1.92	0.33	0.03	0.62
Tequila	-1.05	-4.50	2.43	0.17	-0.17	0.50
Narrow vs. Wide						
Batido	-1.12	-3.10	0.88	0.01	-0.15	0.19
Cerveza	-0.21	-3.61	2.74	0.06	-0.20	0.33
Refresco	-0.48	-3.49	2.77	0.02	-0.23	0.25

Summary of results

The learning strategy or strategies employed by each bird as well as the number of trials to reach proficiency in Experiment 1, the number of trials needed to reverse their preference (Experiment 2) minus the number of trials needed to initially learn the preference (Experiment 1; a measure of behavioral flexibility), and whether they exhibited a preference change between Experiments 3 and 4. -=preferred to stick heavy objects to the magnet thus the experiment did not test what it was designed to test in these individuals, X=did not participate in this experiment.

1   Table 4. Summary of results: The learning strategy or strategies employed by each bird as well

2   as the number of trials to reach proficiency in Experiment 1, the number of trials needed to

3   reverse their preference (Experiment 2) minus the number of trials needed to initially learn the

4   preference (Experiment 1; a measure of behavioral flexibility), and whether they exhibited a

5   preference change between Experiments 3 and 4. (No)=preferred to stick heavy objects to the

6   magnet thus the experiment did not test what it was designed to test in these individuals however

7   they were included in the Spearman's rank test to increase the sample size, X=did not participate

8   in this experiment.

Bird	Sex	Learning strategy	Experiment 1: Learning speed	Behavioral flexibility (Exp 2 - Exp 1)	Behavioral flexibility (water tube)
Tequila	M	Epsilon-decreasing	30	70	Yes
Margarita	F	Epsilon-decreasing	30	70	(No)
Cerveza	F	Epsilon-decreasing	30	60	(No)
Horchata	F	Epsilon-decreasing	30	100	No
Refresco	M	Epsilon-first, then epsilon-decreasing	20	50	Yes
Batido	M	Epsilon-decreasing	30	-	No
Michelada	F	Epsilon-decreasing	40	30	X
Jugo	M	Epsilon-decreasing	40	40	X