

- 27 MJW; Visualization: BS; Writing – original draft: BS; Writing – review and editing: BS,
28 DWAN, MJW.

29 **Abstract**

30 We tested learning and behavioral flexibility in family-living gidgee skinks (*Egernia stokesii*)
31 using a multi-stage visual discrimination task that included acquisition and reversal stages
32 using simple and compound stimuli composed of black shapes superimposed on a colored
33 background. We evaluated how lizards learn compound cues through a probe test. Lizards
34 showed behavioral flexibility through reversal learning using simple stimuli (only color or
35 shape). Our lizards used compound stimuli to learn a discrimination but had problems
36 reversing and generalizing across novel compound stimuli. In the probe test, lizards chose
37 the correct stimulus in a novel pairing with a distractor feature even without prior experience
38 with compound stimuli. Our results suggest that some lizards are likely able to attend
39 selectively to the relevant features of our compound stimuli while ignoring irrelevant features
40 instead of using the configuration of a cue card as a whole to learn to discriminate between
41 compound stimuli. We hope that our work will spark interest in further studies looking at how
42 lizards (and reptiles in general) learn to solve visual discrimination problems.

43

44 *Keywords:* cognition, executive functions, reptile, attentional set shifting, squamate

45

46 **Introduction**

47 Much discussion on how animals learn about stimuli and their value has focused around
48 attention. Learning the solution to, for example, a discrimination problem involves learning to
49 attend to the relevant stimulus features and learning which value of the feature predicts the
50 reward (Pearce & Mackintosh, 2010). This is especially important if the stimulus presented is
51 composed of multiple features, for example, different shapes and colors. Natural stimuli
52 rarely vary across a single feature. For example, ripe fruit can typically be found only at
53 specific times of the year, at a specific location; they have a distinct odor and often a distinct
54 color. Not all of these cues, however, are of similar importance, meaning that some predict
55 the ripeness of a fruit better than others. Selective attention ensures that the focus lies on
56 the stimuli that are most important in a given situation (interference control, Diamond, 2013).

57 The traditional study of selective attention has generated different hypotheses about
58 how animals acquire a discrimination. Animals might learn discontinuously, attending to only
59 one feature at a time until they select a stimulus feature that is associated with
60 reinforcement. Animals might learn continuously, being able to perceive all features of a
61 stimulus simultaneously, but gradually paying more attention to the feature that predicts
62 reward while learning to ignoring others (learnt irrelevance - learning that some stimulus
63 features can be ignored; Castro & Wasserman, 2016). Furthermore, attention might not be
64 given to the absolute properties of each stimulus but rather to distinguishing features
65 (Mackintosh, 1965; Zentall & Riley, 2000; Zentall, 2005). Some animals do indeed gradually
66 learn to pay attention to relevant features while ignoring irrelevant information in
67 simultaneous discrimination problems with compound stimuli (Castro & Wasserman, 2016;
68 Dopson, Esber & Pearce, 2010; Lawrence, 1949) and they do learn something about the
69 irrelevant features, although less so compared to features that are associated with a reward
70 (Mackintosh, 1965). Pretraining of the relevant feature as a single stimulus, however,
71 reduces the amount learnt about the irrelevant features in a compound cue (Mackintosh,
72 1965).

73 An individual's ability, or inability, to discriminate between multi-featured stimuli might
74 depend on how they categorize stimuli. In the psychology literature, two systems are
75 distinguished: implicit and explicit. In an implicit system of categorization, many features are
76 recognized in parallel and a stimulus is perceived holistically as a single entity;
77 discriminations between stimuli are achieved through associative mechanisms. In an explicit
78 system, attentional processes are focused on distinct stimulus features highlighting those
79 features that are relevant for reinforcement and a rule might be derived to solve a
80 discrimination problem. Humans, rhesus macaques (*Macaca mulatta*) and capuchin
81 monkeys (*Cebus apella*) are biased towards using the explicit system trying to find a single
82 feature (or rule) that permits correct categorization even if, per the experimental design, no
83 such single feature is available. Pigeons (*Columba livia*), in contrast, show no such bias
84 (Smith et al., 2012). Rule transfer tests suggest that humans and macaques have, and use,
85 both the explicit and implicit system of categorization (Casale, Roeder, & Ashby, 2012;
86 Smith et al. 2015). Pigeons, however, seem to use one system which is rule based and
87 associative (Qadri, Ashby, Smith, & Cook, 2019).

88 Attention is an important process involved in the detection of change and in exerting
89 behavioral flexibility (i.e., the ability to adapt behavior to changed conditions; Brown & Tait,
90 2015). In animals, behavioral flexibility is most often studied using discrimination reversal
91 learning tasks (e.g., single reversal: Leal & Powell, 2012; Tebbich, Sterelny, & Teschke,
92 2010; Tebbich & Teschke, 2014; serial reversal: Liu, Day, Summers, & Burmeister, 2016) or
93 attentional set-shifting tasks (e.g., Birrell & Brown, 2000; Bissonette, Lande, Martins, &
94 Powell, 2012; Dias, Robbins, & Roberts, 1996; Roberts, Robbins, & Everitt, 1988; Szabo,
95 Noble, Byrne, Tait, & Whiting, 2018; 2019a), amongst others (e.g., Nicolakakis, Sol, &
96 Lefebvre, 2003; Sol & Lefebvre, 2000; Sol, Timmermans, & Lefebvre, 2002; Tebbich,
97 Sterelny, & Teschke, 2010). To solve a reversal learning task, animals first need to
98 recognize the change in a stimulus-reward relationship to stop responding to the formerly
99 rewarded stimulus and shift behavior towards the formerly unrewarded stimulus (Dias et al.,
100 1996). In set-shifting tasks (e.g., Wisconsin card sorting task in human children, Berg, 1948;

101 or the intradimensional/extradimensional – ID/ED – attentional set-shifting task in animals,
102 Mackintosh & Little, 1969; Roberts et al., 1988; Szabo et al. 2018; 2019a), individuals need
103 to overcome an attentional bias towards one set of features (often referred to as belonging
104 to the same dimension; e.g., lines) and shift attention towards stimuli from within another set
105 of features (second dimension; e.g., shapes; Brown & Tait, 2015). Failure to do so might
106 either be caused by an inability to shift attention away from previously relevant features
107 (attentional perseveration) or the interference of learnt irrelevance (Castro & Wasserman,
108 2016). Furthermore, discrimination of compound stimuli (consisting of two or more features)
109 might be impeded if stimuli within the same feature set (dimension) are distant (e.g.,
110 wavelength if color is used) or if stimulus features are spatially separated (e.g., one feature
111 being intra-maze cues and the other extra-maze cues – cues outside the maze) (Trobalon,
112 Miguelez, McLaren & Mackintosh, 2003).

113 We have previously shown that reversal learning and intra-dimensional/ extra-
114 dimensional (ID/ED) attentional set-shifting tasks can be successfully applied in lizards
115 (Szabo et al., 2018; 2019a; Szabo & Whiting, 2020b). A growing body of literature is using
116 reversal learning to quantify lizard behavioral flexibility in the visual (e.g., Leal & Powell,
117 2012) and spatial domain (e.g., Noble, Carazo, & Whiting, 2012) and in relation to foraging
118 ecology (e.g., Day, Crews, & Wilczynski, 1999), incubation temperature (e.g., Clark, Amiel,
119 Shine, Noble, & Whiting, 2014) or habitat features (e.g., Batabyal & Thaker, 2019) although
120 a single reversal by itself might not be sufficient to demonstrate behavioral flexibility. Here,
121 we investigated the visual discrimination learning ability of the Australian gidgee skink
122 (*Egernia stokesii*), a family-living lizard. Gidgee skinks are a medium-sized species that
123 inhabit arid or semi-arid arboreal and rocky habitats in long-term, stable, kin-based, multi-
124 generational family groups comprising of a monogamous mating pair and their offspring
125 (Chapple, 2003; Cogger, 2014; Gardner, Bull, Cooper, & Duffield, 2001). They are part of
126 the *Egernia* group, a clade of Australian lizard species expressing varying degrees of
127 sociality (Chapple, 2003; Whiting & While, 2017). Gidgee skinks are omnivorous with a large
128 amount of plant material in their diet, they are active during the day, have a long life span of

129 about 25 years and take up to five years to reach sexual maturity (Chapple, 2003; Duffield &
130 Bull, 1998; 2002).

131 Our aim was to investigate, for the first time, learning and behavioral flexibility in this
132 social lizard. To this end, we tested lizards using a multi-stage visual two-choice
133 discrimination task which included multiple reversal stages. We presented lizards with cards
134 painted in different colors (background) containing different shapes superimposed on top
135 (compound stimulus with multiple features). Reinforcement was associated with one feature
136 set (e.g., one of two background colors) while the second set (e.g., two different shapes)
137 was irrelevant for reinforcement. We also investigated what lizards attend to when learning
138 compound stimuli. We hypothesized that lizards would either (1) use associative learning
139 perceiving each compound stimulus holistically by combining all stimulus features (e.g.,
140 color and shape) and remembering which two out of four cards were associated with a
141 reward; or (2) selectively attend to the stimulus features predicting reward (e.g., color) while
142 learning to ignore irrelevant information (e.g., all shapes).

143

144 **Methods**

145 ***Study animals***

146 We hand-captured 22 wild, adult *Egernia stokesii* (snout-vent-length > 145 mm; Chapple,
147 2003), of undetermined sex, around Fowlers Gap Arid Zone Research Station (-31.086972
148 S, 141.704836 E), New South Wales, Australia. Three were captured during March 2018
149 and 19 during November 2018. Within one week of capture, individuals were transported to
150 Macquarie University by car in cloth bags within a cooler box and transferred into individual
151 plastic tubs (683 L x 447 W x 385 H mm).

152

153 ***Housing***

154 Lizards were housed in a temperature-controlled environment ($24 \pm 2^\circ\text{C}$ SD), relative
155 humidity between 30-60% and a light cycle of 12 h (06:00 – 18:00 h). Snout-vent-length

156 (SVL), total length (TL) and head width (HW) of all lizards were determined on the day of
157 capture and one week before the start of testing with the addition of lizard weight. We
158 installed heat cords underneath the enclosures to increase temperature on one side to up to
159 33°C ($\pm 2^\circ\text{C}$ SD) and iButtons (Thermochron iButton model DS1921) recorded temperature
160 hourly within enclosures. Each enclosure was lined with paper and equipped with a refuge
161 for shelter, a water bowl (heavy, poly resin reptile water bowls made to look like rock, 130 L
162 x 110 W x 40 H mm) and two wooden ramps.

163

164 ***Husbandry***

165 When not being tested, lizards were fed three times a week. On Monday, Wednesday and
166 Friday they were fed an assortment of small cut fruit and vegetables. Additionally, on Friday
167 lizards were fed 3-4 adult crickets (powdered with aristopet Repti-vite and URS Ultimate
168 Calcium). During experiments, lizards were fed only on Fridays because they were
169 consuming fruit-based baby food (0.08 ± 0.01 g; HeinzTM, various flavors) as positive
170 reinforcement during trials for the rest of the week. During regular feeding, vegetables and
171 fruit were presented in green dishes (60 mm diameter), while within trials reinforcement was
172 presented in 55 mm diameter petri dishes, with the outside covered in black electrical tape.
173 Lizards had *ad libitum* access to water.

174

175 ***Habituation and pre-training***

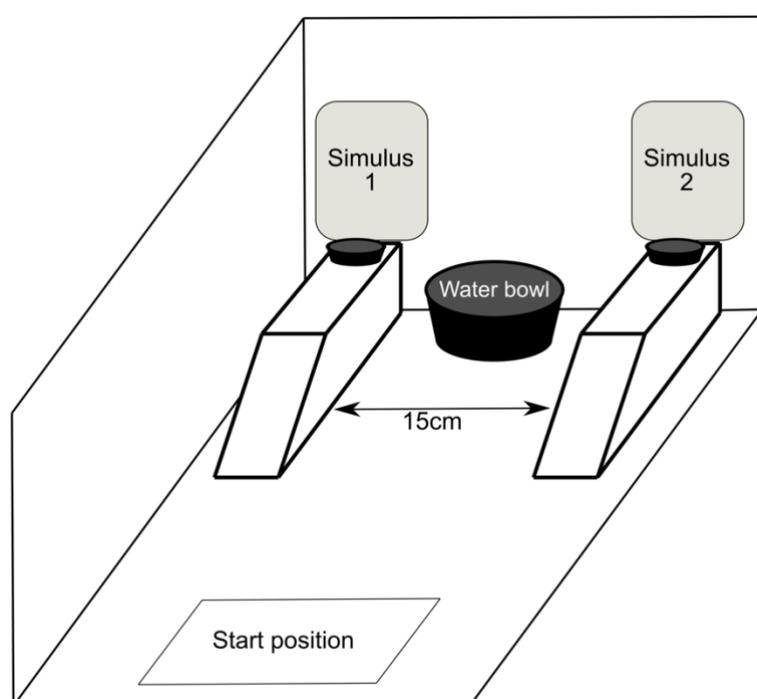
176 Prior to the study, all lizards were feeding consistently and had habituated to captivity over
177 the course of at least one month (three lizards had been in captivity for nine months before
178 testing began; electronic supplementary material Table S1).

179 Pre-training was conducted one week prior to testing and lasted for a maximum of
180 two weeks (30 trials). During phase 1 (N = 5 trials), one baited food dish was presented in
181 front of the water bowl for 1.5 h with the refuge removed. Individuals moved on to phase 2 if
182 they ate the reward in at least 4/5 (80%) trials. Phase 2 lasted at least 10 trials but a

183 maximum of 25 trials. One baited food dish was presented on top of a ramp, for 1.5 h and
184 the refuge was removed. An individual moved on to the first testing stage if it ate the reward
185 in at least 80% of phase 2 trials. For each lizard, two food dishes were available and used in
186 a random order throughout pre-training to ensure that lizards had distributed their own scent
187 evenly on both dishes before the first test stage. The side (ramp) on which a food dish was
188 presented in phase 2 was randomized but counter-balanced (each dish appeared an equal
189 number of times on the left and right ramp).

190 To prevent stress-induced learning impairment (Langkilde & Shine, 2006), animals
191 were kept and tested in their home enclosures throughout the experiment. Two lizards did
192 not reach criterion in pre-training phase 2 and were therefore removed from the experiment
193 due to low food motivation (electronic supplementary material Table S1). The three lizards
194 captured during March 2018 had pre-experience with one cognitive assay testing their motor
195 response inhibition (Szabo, Hofer, & Whiting, 2020a) whereas the other 19 animals were
196 experimentally naïve.

197



198 **Figure 1.** Schematic set-up used during experiments. Cards depicting stimuli were
199 presented on top of wooden ramps at one end of the enclosure. Food dishes were

200 presented right in front of the cards elevated on the ramps. The water bowl was present
201 during trials between ramps that were approximately 15 cm apart. Lizards started each trial
202 from the starting position, on the warm side of the enclosure (heated by heat cord to $33 \pm$
203 2°C SD) opposite the set-up.

204

205 ***Setup and procedure***

206 The study was conducted from December 2018 to March 2019 (three trials/day, five
207 days/week). Subjects were randomly divided into two groups: one group was tested on a
208 shape discrimination (Stimulus group shape, $N = 11$), the other group was tested on a color
209 discrimination (stimulus group color, $N = 10$). Average snout-vent-length (SVL; a commonly
210 used measure for lizard size) of groups was equated as best as possible (mean $\text{SVL}_{\text{color}} =$
211 179.2 ± 17.0 SD; mean $\text{SVL}_{\text{shape}} = 178.9 \pm 15.5$ SD).

212 To start a trial, the refuge was slowly placed on top of an animal and both were
213 slowly moved to the start position at one end of the enclosure opposite to the ramps, stimuli
214 and dishes (Figure 1). We ensured that the refuge was covering the head and most of the
215 body to prevent the animals from seeing any unconsciously given cues from the
216 experimenter while setting up a trial. Next, stimulus cards were fixed to the back wall of the
217 enclosure on top of the ramps with Bostik Blu-Tack adhesive putty. This prevented cards
218 from falling off the ramps during trials. Thereafter, the already-baited (2 ± 0.3 g of reward
219 each) and closed (with fine PVC coated Polyester insect screen, Cyclone Screening Pet
220 Mesh) food dishes were placed on top of the ramps directly in front of and as close as
221 possible to the cue cards. The dish in front of the incorrect stimulus was completely covered
222 with mesh blocking access to the reward. A hole in the mesh on the dish in front of the
223 correct stimulus made the reward accessible. Using mesh allowed food odor to be
224 detectable from both dishes controlling for the use of food chemicals to solve the tasks. By
225 using ramps we ensured that the stimulus cards were visible from any location in the
226 enclosure and that animals were unable to see into the dishes and determine which one
227 provided food before making a choice (gidgee skinks are saxicolous and good climbers;

228 Chapple, 2003). Each individual received the same refuge, ramps, cards, and dishes for the
229 duration of the experiment.

230 A trial started after about one minute of acclimation with the removal of the refuge
231 exposing the lizard to the set-up. The order in which the subjects received the set-up and
232 started each trial were alternated throughout the course of the study to prevent order effects.
233 Trials lasted for 1.5 h each. Trial length was chosen to give lizards ample opportunity to
234 make a choice (individuals can take over an hour to start moving). At the end of the trial we
235 returned the refuge to the enclosure and removed dishes and cards. Between trials, in the
236 40 minute inter-trial interval, both dishes were cleaned and refilled with fresh baby food
237 (making sure that both dishes were touched to assure even odor distribution). Animals were
238 tested between 07:30-13:30 h. We left lizards undisturbed during trials to minimize stress
239 and videotaped (H.264 Digital Video Recorder, 3-Axis Day & Night Dome Cameras) each
240 trial to be scored later.

241 A choice was scored as correct if the lizard approached the dish in front of the
242 correct stimulus first and passed over it with its head. We assumed that if the head passed
243 over the dish the lizard was able to see into the chosen dish. A choice was scored as
244 incorrect if the lizard first approached the dish in front of the incorrect stimulus. We scored
245 trial latency (time from removal of the refuge to when the head passed over a dish for the
246 first time) and choice latency (from the first directed, uninterrupted forward movement of the
247 whole body to choice; an interruption is defined as no movement for 20 s or more). Animals
248 were able to visit both dishes multiple times within a trial being able to correct wrong choices
249 in each trial (self-correction). To determine the point of task acquisition we used a pre-
250 determined criterion of 6/6 or 7/8 correct choices in consecutive trials (based on Szabo et
251 al., 2018; 2019a). To prevent decreased motivation to participate due to prolonged exposure
252 to the same task, a maximum number of trials was given in each task. In stage 1 we allowed
253 70 trials while in all subsequent stages (2 to 6) we allowed only 60 trials. If acquisition was
254 not reached within the maximum number of trials a lizard was removed from the experiment
255 ('non-learner'). To ensure unbiased scoring of a lizard's choice behavior, a random subset of

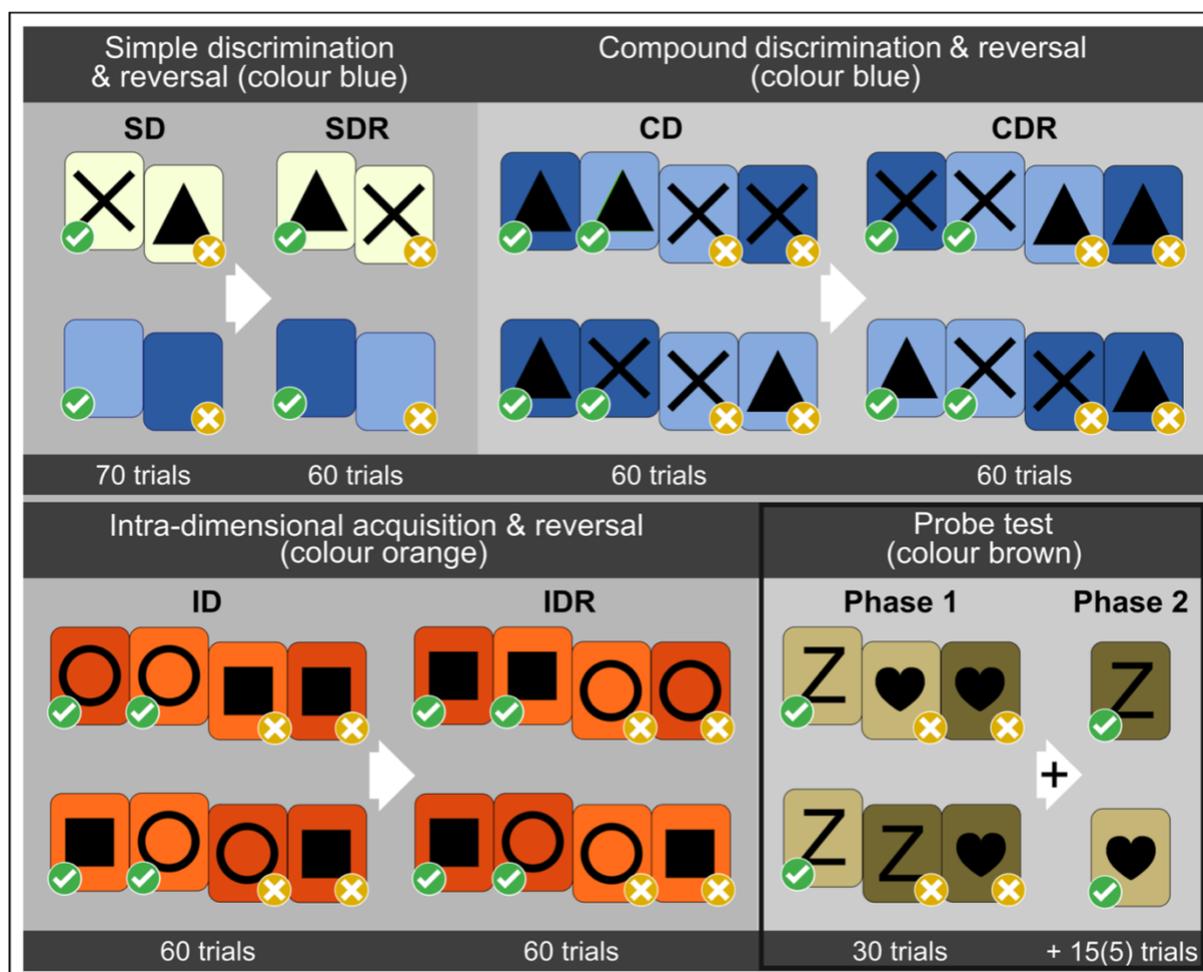
256 trials (14%) was scored by an independent observer. Cohen's kappa inter-observer
 257 reliability was calculated at 0.903 (high; Falissard, 2012).

258

259 **Stimulus cards**

260 Cards were made of pressed wood (rectangle coaster, Boyle Industries Pty Ltd, 113 L x 93
 261 W x 3 H mm), sprayed with spray paint and finally, shapes were drawn onto the sprayed
 262 cards (Figure 2). Color pairs were chosen to be easily discriminable based on the lizard
 263 perceptual system (Fleishman, Loew, & Whiting, 2011). During experiments, each stimulus
 264 (and dish) was never presented more than twice in a row on the same side and the left/right
 265 position of each feature within compound stimuli varied independently of each other. For
 266 more details on spray paint brand and color as well as shape measurements see Szabo et
 267 al., 2018.

268



269

270 **Figure 2.** *Sequence in which stimuli were presented.* Half of the lizards were assigned to
271 stimulus group 1 (top row) and learned to discriminate between shapes (shape as the
272 relevant feature). The other half were assigned to stimulus group 2 (bottom row) and started
273 learning to discriminate between colors (color as the relevant feature). The first couple of
274 stages are called simple discrimination (SD) and reversal (SDR) in which cards were made
275 up of only one feature (e.g., either sprayed in one color or a black shape on beige
276 background). One stimulus was associated with a reward (tick mark) while the other was not
277 associated with a reward (X). This was followed by two stages (compound discrimination –
278 CD, and reversal - CDR) in which a second, but irrelevant feature was added to the cue
279 cards (either a background color other than beige or a shape superimposed over the colored
280 background). Next, novel colors and shapes were used in the intra-dimensional acquisition
281 (ID) and reversal (IDR). Finally, in the probe test, again novel colors and shapes were
282 introduced, but only three possible combinations of stimuli were presented during phase 1.
283 After lizards reached the predetermined learning criterion, the fourth and novel combination
284 was introduced five times out of 15 in phase 2 to test if lizards had learnt to use compound
285 stimuli or had separated stimuli into feature sets (or dimensions) to make a successful
286 discrimination. Except for SD and the probe test, lizards received a maximum of 60 trials to
287 reach the learning criterion.

288

289 ***Test stages***

290 In stage 1, simple discrimination (SD), lizards were trained to discriminate between two
291 color/shape stimuli (single feature) and thereafter on a simple discrimination reversal (SDR)
292 of the same stimuli (Figure 2). Next, a second set of features was added to the cards: for
293 lizards trained on shape, background colors were added to the cards and for lizards trained
294 to discriminate between colors, shapes were superimposed on the background color (Figure
295 2). These irrelevant features were intended as a distractor. A compound discrimination
296 acquisition (CD) and a reversal (CDR) were tested. Lastly, lizards that learnt to discriminate
297 between these compound cues were presented with an intra-dimensional shift (Figure 2).

298 We presented lizards with novel stimuli (novel colors and shapes) for all features in an
299 acquisition (ID) and reversal stage (IDR). We had planned to also test our lizards on an
300 extra-dimensional shift (similar to Szabo et al., 2018; 2019a); however, none made it this far.

301 After lizards had finished the test stages they were tested using a probe test to
302 investigate if they had the ability to separate cards into distinct stimulus features when
303 learning to discriminate compound cues. Only lizards that had successfully reached the
304 learning criterion in at least one stage were tested on the probe test. The probe test was
305 similar to an intra-dimensional shift. We again introduced a novel set of shape and color
306 stimuli (Figure 2); however, the positive stimulus feature was paired with only one of the
307 distractor features (phase 1). For example, the reinforced shape 'Z' was paired with a light
308 brown background only while the non-reinforced heart-shape was paired with both a light
309 and dark brown background (Figure 2). Similar to the previous test stages, in phase 1,
310 lizards had to learn the discrimination. They could either learn the configuration of the whole
311 reinforced card (e.g., Z with a light brown background) or learn the distinct feature (e.g.,
312 shape Z) predicting reward. If a lizard reached the learning criterion (6/6 or 7/8) within 30
313 trials it was given 15 additional trials (phase 2). In trials 1, 3, 7, 13 and 15 of phase 2, we
314 presented the missing compound stimulus (e.g., 'Z' with the dark brown background). We
315 predicted that, if lizards had learnt to choose based on the configuration of cards, then the
316 introduced compound card would be novel and lizards would chose at random in trials 1, 3,
317 7, 13 and 15. Conversely, if they had learnt to choose based on the reward predicting
318 feature, the new distractor would have no effect on performance and lizards would choose
319 correctly in trials 1, 3, 7, 13 and 15.

320

321 ***Odor control***

322 Because we had cut a hole into only one of the mesh sheets covering the food dishes (which
323 might have changed food odor diffusion) we tested 10 randomly chosen lizards in an odor
324 control task (electronic supplementary material Table S1). We administered 10 trials each
325 after lizards had finished test trials. We used compound stimuli (stages CD and CDR) but

326 neither of the color or shape features, dish or side of stimulus/dish presentation predicted
327 reinforcement. Dishes and cards were thoroughly cleaned with detergent before testing. We
328 used a one-sample Wilcoxon test to compare lizards number of correct choices
329 (approaching the open dish) out of 10 trials to chance performance (= 5/10 trials). The
330 lizards choice did not deviate from chance (One sample Wilcox test, $V = 8.5$, median = 5; $p =$
331 0.746).

332

333 ***Statistical analyses***

334 We analyzed general learning performance within each stage using Bayesian generalized
335 linear mixed effects models (GLMM; Hadfield, 2010). We only looked at the first three stages
336 because our sample sizes were too low in later stages ($N = 1$ in CDR and ID, none of the
337 lizards learnt during IDR). Criterion trials (last six to eight trials in which the learning criterion
338 of 6/6 or 7/8 consecutive correct choices was reached) were included in the analyses. We
339 were also interested in comparing learning performance between stimulus groups and
340 capture locations. Because gidgee skinks live in multi-generational family groups, we used
341 capture location as a proxy for group identity to investigate if groups differed in their learning
342 ability. We compared groups only if data on a minimum sample of three animals were
343 available in each group (Forstmeier, Wagenmakers, & Parker, 2017). Therefore, we
344 allocated lizards captured within the same rock outcrop into the same group to increase
345 sample size. This was warranted because gidgee skinks live in large, stable, kin-based
346 family groups comprised of a monogamous pair and multiple generations of offspring that
347 exhibit low rates of dispersal (Duffield & Bull, 2002; Gardner et al., 2001; Gardner, Bull, &
348 Cooper, 2002) and high site fidelity (Gardner, Bull, Fenner, Murray, & Donnellan, 2007;
349 Pearson, Godfrey, Bull, & Gardner, 2016). Due to small sample sizes, stimulus groups were
350 compared only within the first stage and certain capture locations had to be excluded from
351 the analysis (due to low sample size); location was therefore only compared in the first two
352 stages.

353 To investigate possible differences in learning in stage 1 (SD), our model included
354 choice (1 – correct, 0 – incorrect) as the response variable and trial, stimulus group and their
355 interaction as the fixed effects. Because certain capture locations had to be excluded
356 because of low sample size, we ran an extra model to compare performance between
357 capture locations to avoid limiting our analysis of general learning performance and stimulus
358 group due to removing data. We did not correct alpha levels but instead present estimates
359 and confidence intervals generated from our models (Nakagawa, 2004). The model
360 investigating differences between capture locations included choice (1 – correct, 0 –
361 incorrect) as the response variable and trial, capture location and their interaction as the
362 fixed effects. Similarly, in stage 2 (SDR) we ran a model to look at general learning
363 performance (sample size for stimulus groups was too small to run an analysis) with choice
364 (1 – correct, 0 – incorrect) as the response variable and trial as the fixed effect and a second
365 model (same as above) for capture location. Finally, sample sizes were too small to
366 compare stimulus groups or capture locations in stage 3. Therefore, we only ran a model to
367 look at general learning performance using choice (1 – correct, 0 – incorrect) as the
368 response variable and trial as the only fixed effect.

369 We also made two comparisons between stages: (1) to compare acquisition (SD) to
370 reversal (SDR) performance and (2) to compare learning between stages in which simple
371 stimuli (SD) were used to when distractor features (CD) were added (compound stimuli). No
372 further comparisons were made due to limitations in sample size. We focused only on data
373 from the respective stages of interest and used Bayesian GLMMs with choice (1 – correct, 0
374 – incorrect) as the response variable and trial, stage and their interaction as the fixed effects.
375 Furthermore, we compared the number of correct/incorrect choices performed in the last 10
376 trials of stage 2 (SDR) to the number of correct/incorrect choices made in the first 10 trials of
377 stage 3 (CD) using the Fisher's exact test of independence to further investigate how the
378 addition of an irrelevant feature affected lizard performance. Additionally, we used a
379 Bayesian GLMM to investigate if lizards showed a training effect (improvement across
380 stages) after being repeatedly tested on similar problems. The model was based on the

381 whole dataset and choice (1 – correct, 0 – incorrect) was used as the response variable and
382 trial, stage, and their interactions as the fixed effects. To account for changes in
383 performance across stages, we used a random effect of trial, stage, and their interaction,
384 nested in animal identity. Due to our limited dataset we did not estimate all covariates
385 (compared to all other Bayesian models for which all covariates were estimated) and z-
386 transformed trial to ensure good model conversion. All other Bayesian models included a
387 random effect of trial nested in individual identity to account for differences in intercepts and
388 slopes (respectively) and to account for autocorrelation between successive choices.
389 Furthermore, for all models, we confirmed that no auto-correlation (correlation between lags
390 < 0.1 ; Hadfield, 2010) occurred in all Bayesian models, that sufficient mixing (by visually
391 inspecting plots of MCMC chains; Hadfield, 2010) took place and that the Markov chain was
392 run for long enough (Heidelberg and Welch diagnostic tests; Hadfield, 2010). All statistical
393 analyses were run in R version 3.5.3 (R Development Core Team, 2008).

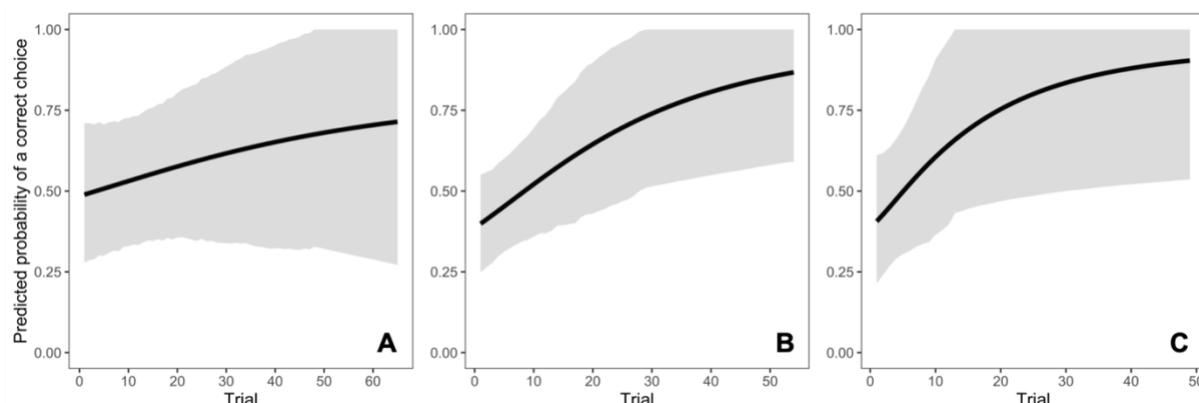
394 We also investigated if errors, food motivation during test trials, a side bias or body
395 condition differed between successful learners and non-learners (removed due to not
396 reaching the learning criterion). For details on these analyses, see the electronic
397 supplementary material. No statistical analyses were conducted on the data collected from
398 the probe test because only two lizards reached the learning criterion in phase 1 and were
399 tested in phase 2 of the probe test. All datasets generated and the code used for analyses in
400 the current study are available at the Open Science Framework (link for review purposes:
401 https://osf.io/q2uf6/?view_only=f766af725ccf4241829a122fa559428d).

402

403 ***Ethics***

404 The experiments performed in this study were strictly non-invasive observations of behavior
405 and followed the guidelines laid out by the Association for the Study of Animal
406 Behavior/Animal Behavior Society for the treatment of animals in behavioral research and
407 Teaching (2018). Experimental procedures were approved by the Macquarie University

408 Animal Ethics Committee (ARA # 2013/031). Animal collection from the wild was approved
409 by the New South Wales National Parks and Wildlife Service (OEH; License #SL101972).
410



411
412 **Figure 3.** Change in the predicted probability of making a correct choice from trial 1 to trial *n*
413 (last trial is dependent on when lizards reach the learning criterion) for the first three stages.
414 (A) Learning curve of the 12 lizards tested in the simple discrimination (stage 1 - SD). (B)
415 Learning curve of the 8 lizards tested in the simple discrimination reversal (stage 2 - SDR).
416 (C) Learning curve of the 6 lizards tested in the compound discrimination (stage 3 - CD).
417 Results shown only include lizards that reached the learning criterion. 95% credible intervals
418 are shown in grey.

419 420 **Results**

421 Overall, 12/22 lizards (55%) reached the learning criterion during the simple discrimination
422 (SD); 8/12 (67%) reached criterion during the reversal (SDR), and six (75%) also learnt to
423 discriminate between compound stimuli (CD). However, only one lizard managed to reach
424 the learning criterion during the reversal of the compound discrimination (CDR) and the
425 intra-dimensional acquisition (ID) when novel colors and shapes were introduced but
426 subsequently failed in the intra-dimensional reversal (IDR). Lizards that were excluded did
427 make more errors compared to lizards that reached criterion; non-learners did not show
428 lower food motivation during test trials and neither a side bias or body condition explained
429 learning performance (for details see the electronic supplementary material).

430 In the simple discrimination (SD), the probability of making a correct choice
431 increased from 0.489 in trial 1 to 0.715 in trial 65 (Figure 3 A). Our lizards improved by 46%
432 across 65 trials which was not significant (GLMM, $N = 12$, $\text{trial} = 0.019$, $CI_{\text{low}} = -0.028$, $CI_{\text{up}} =$
433 0.068 , $p = 0.362$). During the reversal of the simple discrimination (SDR), in which a change
434 in the stimulus-reward relationship occurred, the probability of lizards making a correct
435 choice increased from 0.399 in trial 1 to 0.867 in trial 54 (Figure 3 B). Our lizards improved
436 118% across 54 trials which was marginally significant (GLMM, $N = 8$, $\text{trial} = 0.057$, $CI_{\text{low}} = -$
437 0.008 , $CI_{\text{up}} = 0.127$, $p = 0.043$). After the introduction of an irrelevant distractor the probability
438 of lizards choosing correctly increased from 0.406 in trial 1 to 0.904 in trial 49 (Figure 3 C).
439 Our lizards improved by 123% across 49 trials which was not significant albeit we had a
440 small sample size at this stage (GLMM, $N = 6$, $\text{trial} = 0.098$, $CI_{\text{low}} = -0.032$, $CI_{\text{up}} = 0.282$, $p =$
441 0.083). Finally, lizards captured at different locations did not differ in their learning
442 performance (GLMM, $p > 0.05$; electronic supplementary material Table S2), and neither did
443 stimulus groups (GLMM, $SG = 0.237$, $CI_{\text{low}} = -0.882$, $CI_{\text{up}} = 1.399$, $p = 0.674$; $\text{trial}_{SG} = 0.002$,
444 $CI_{\text{low}} = -0.063$, $CI_{\text{up}} = 0.067$, $p = 0.981$).

445 The comparison between the acquisition of the discrimination (SD) to the reversal
446 (SDR) revealed no difference between stages (GLMM, $\text{stage} = -0.099$, $CI_{\text{low}} = -0.387$, $CI_{\text{up}} =$
447 0.191 , $p = 0.506$; $\text{trial}_{\text{stage}} = 0.005$, $CI_{\text{low}} = -0.010$, $CI_{\text{up}} = 0.020$, $p = 0.493$). Adding irrelevant
448 features to the cue cards did not affect lizards' performance (GLMM, $\text{stage} = -0.078$, $CI_{\text{low}} = -$
449 0.470 , $CI_{\text{up}} = 0.326$, $p = 0.696$; $\text{trial}_{\text{stage}} = 0.007$, $CI_{\text{low}} = -0.012$, $CI_{\text{up}} = 0.024$, $p = 0.463$;
450 Fisher's exact test of independence, $p > 0.05$; electronic supplementary material Table S3).
451 Furthermore, improvement when solving similar problems across stages was non-significant
452 (GLMM, $\text{SDR}_{\text{learning rate}} = 0.180$, $CI_{\text{low}} = -0.276$, $CI_{\text{up}} = 0.665$, $p = 0.436$; $\text{CD}_{\text{learning rate}} = 0.296$,
453 $CI_{\text{low}} = -0.340$, $CI_{\text{up}} = 0.933$, $p = 0.341$).

454 Of the 12 lizards tested in the probe test, only two reached the predetermined
455 learning criterion in phase 1 within the 30 trials given in this stage. Lizard STOK4 (stimulus
456 group shape) reached the criterion in 10 trials and STOK16 (stimulus group color) in 6 trials.
457 Within phase 2, STOK4 chose correctly in 60% (6/10) of phase 1 trials and 80% (4/5) of

458 trials in which the novel card was used. STOK16 chose correctly in 40% (4/10) of phase 1
459 trials and 80% (4/5) in those trials in which the novel card was used.

460

461 **Discussion**

462 Gidgee skinks showed evidence of behavioral flexibility by completing a minimum of one
463 reversal and learnt a discrimination using compound stimuli. However, only one lizard learnt
464 a second reversal and an intra-dimensional acquisition, although it failed at the third (intra-
465 dimensional) reversal. Because gidgee skinks live in family groups which are genetically
466 distinct from other groups (Gardner et al., 2001), we also hypothesized that groups might
467 differ in cognitive performance. Capture location did, however, not predict differences in
468 performance, at least not in this task.

469 We were also interested in what stimuli lizards relied upon while learning to
470 discriminate compound cues; a factor largely ignored in lizard learning studies. Lizards
471 selected the card showing the novel pairing between the relevant and irrelevant feature in
472 four out of five trials. Some studies have implemented probe tests similar to what we have
473 tested here, to evaluate what stimuli animals attend to during ID/ED attentional set-shifting
474 tasks (Dias et al., 1996; Dias, Robbins, & Roberts, 1997; Roberts et al., 1988). In these
475 studies animals were probe tested after they had finished their first intra-dimensional
476 acquisition. During the probe test, the relevant (reinforced) features were kept the same as
477 in the previous stage, but the irrelevant features were replaced with novel exemplars.
478 Common marmosets (*Cullithrix jacchus*) do not change performance compared to the
479 preceding ID stage, showing that they were unaffected by the change in the irrelevant
480 features (Dias et al., 1996; 1997; Roberts et al., 1988). The two lizards that were tested on
481 phase 2 probe trials both showed greater accuracy (higher proportion of correct choices) in
482 probe trials as compared to trials in which phase 1 cards were presented, which suggests
483 that they might have learnt to ignore irrelevant information (learnt irrelevance), as did
484 marmosets. The added irrelevant information had previously been associated with the

485 incorrect stimulus. Had lizards attended to the irrelevant features they might have mistakenly
486 valued it as incorrect based on previous experience. One alternative explanation for our
487 lizards' good performance during the five probe trials could be that they were attracted to the
488 novel stimulus combination. However, trials were 1.5 hours long increasing the likelihood
489 that lizards had habituated to the sight of, or familiarized themselves with, the novel feature
490 combination after experience in the first trial. Nonetheless, novelty of the stimulus
491 combination could have caused our result. Testing lizards using a similar approach such as
492 the one used in common marmosets (Dias et al., 1996; 1997; Roberts et al., 1988) would
493 give further insights into whether novelty affected our lizards' performance in probe trials.
494 Another alternative explanation for our result could be that lizards had learnt which stimuli
495 were incorrect based on their experience in test stages and phase 1 of the probe test.
496 Studies using a midsession reversal approach, in which the reversal of a simultaneous two-
497 choice discrimination takes place in the middle of each session, have shown that rats
498 develop a win-stay/ loose-shift strategy to reverse effectively in each session while pigeons
499 use timing instead of the result of their choice (McMillan & Roberts, 2012; Rayburn-Reeves,
500 Stagner, Kirk, & Zentall, 2013). Consequently, rats eventually only make one mistake per
501 session (first reversal trial) while pigeons make a number of anticipatory errors (before the
502 reversal occurs) as well as perseverative errors (after the reversal occurs) (McMillan &
503 Roberts, 2012; Rayburn-Reeves et al., 2013). Other studies suggest that the errors made by
504 pigeons are caused by problems learning to inhibit responding. Pigeons learn which stimuli
505 to reject instead of which stimuli to choose (McMillan, Sturdy, & Spetch, 2015; Zentall,
506 Halloran, & Peng, 2020). Because lizards are more closely related to birds than mammals
507 (Alföldi et al., 2011), it would therefore not be surprising to find that lizards also rely more on
508 information about which stimuli to reject rather than which stimuli to choose. If this is the
509 case, then it is not surprising that lizards had no problem choosing the novel, correct
510 stimulus combination, because the information needed for lizards to make the appropriate
511 response (which stimuli to reject) in probe trials had not changed. Importantly, only one of
512 the two successful individuals had previous experience with compound cues, the other lizard

513 had only been tested with single feature cards. Furthermore, one lizard was tested with color
514 as the relevant feature and the second with shape indicating that both colors and shapes
515 were of similar discriminability to lizards. This is also confirmed by our analysis showing no
516 difference in performance between stimulus groups. With data on only two lizards, however,
517 we are unable to determine exactly how lizards learn compound cues. Furthermore, their low
518 accuracy in trials with phase 1 stimuli indicates that these two lizards had not fully learnt the
519 compound discrimination which might be connected with our learning criterion (discussed
520 below). More tests are certainly needed to better understand the processes involved when
521 lizards learn compound stimuli and this is likely to be a highly rewarding line of study in the
522 future.

523 Our results show that gidgee skinks are able to learn discriminations based on visual
524 cues and that they possess some reversal learning ability. However, our results also
525 highlight some potential issues with applying our learning criterion. Lizards performance had
526 not plateaued at the point when they were considered to have learnt the discrimination,
527 especially given the very slow increase in the probability of making a correct choice that we
528 observed (Figure 3). Consequently, they may have been moved to the next stage too early.
529 Using a set learning criterion (fixed number of correct trials) might only be sufficient after
530 running pilots to determine its appropriateness for a given species. Our methods can,
531 however, be improved by giving animals more than 2-3 trials a day and testing until a certain
532 performance threshold is reached across multiple testing days (e.g., 80-90% correct across
533 2-3 days; e.g., Reiner & Powers, 1978; Roberts et al., 1988). If only a limited number of trials
534 can be given within a day, we suggest giving animals multiple post criterion trials to evaluate
535 the stability of performance after the criterion is reached (e.g., Szabo, Noble, & Whiting,
536 2019b). Applying set learning criteria to decide on when to advance animals to more
537 advanced tasks is common in studies across vertebrates (e.g., Ashton, Ridley, Edwards, &
538 Thornton, 2018; Batabyal & Thaker, 2019; Day et al., 1999; Logan, 2016; Roberts et al.,
539 1988; Tebbich et al., 2010) but further study is needed to determine the effectiveness of
540 these procedures in lizards.

541 Previously, we used the same methodology to test set shifting in two other, relatively
542 closely related lizard species, the tree skink (*Egernia striolata*; Szabo et al., 2018) and the
543 eastern blue-tongue lizard (*Tiliqua s. scincoides*; Szabo et al., 2019a). Both tree skinks and
544 blue-tongue lizards were able to learn multiple acquisition and reversal stages including an
545 intra- and extradimensional shift. Lizards started each stage choosing correctly at or below
546 chance and gradually improved until reaching the learning criterion. Tree skinks and blue-
547 tongue lizards showed significant evidence of learning at least in the early stages. Similar to
548 the current study, however, sample sizes decreased across stages and so did the power to
549 detect learning. Furthermore, in tree skinks, blue-tongue lizards and gidgee skinks, we did
550 not find differences in stage comparisons or training effects across stages using the same
551 exemplars of relevant stimulus features (e.g., X and triangle across the first four stages;
552 Figure 2). It is possible that our methodology is not well suited to test reversal learning, habit
553 formation (training effect) and behavioral flexibility (reversal learning and set shifting) in
554 lizards and is therefore not sensitive enough to detect changes in learning rate. Simply
555 increasing statistical power by testing higher numbers of individuals, however, might lead to
556 statistical results showing the expected differences between test stages.

557 Large individual differences in performance might have factored into our inability to
558 detect differences between stages. Unfortunately, our study precluded investigating among-
559 individual differences in learning in detail given our sample size (van de Pol, 2012).
560 Individual variation might, however, be caused by cognitive as well as non-cognitive factors
561 such as food motivation, neophobia or behavioral type (e.g., Boogert, Madden, Morand-
562 Ferron, & Thornton, 2018; Dougherty & Guillette, 2018). Larger samples of independent
563 animals will be necessary to statistically explore among-individual differences, and how
564 cognitive and non-cognitive factors impact learning (van de Pol, 2012). Looking at individual
565 differences in learning is beginning to receive greater attention across taxa, but less so in
566 reptiles. Nonetheless, recent research is starting to close this gap (e.g., Carazo, Noble,
567 Chandrasoma, & Whiting, 2014; Chung et al., 2017; Goulet et al., 2018; Munch et al., 2018;
568 Noble, Byrne, & Whiting, 2014).

569 One striking difference with our two previous studies is the high dropout rate of
570 gidgee skinks from the beginning of the first stage. We investigated possible causes
571 analyzing error rates, motivation (eating the reward), body condition and the influence of a
572 side bias. We found that animals removed as non-learners made more errors but did not
573 differ from learners in motivation, side bias or body condition (electronic supplementary
574 material). One factor causing lowered motivation to participate might have been the chosen
575 reward (baby food), where a different reward might have been more suitable. We did not test
576 food preference *a priori* in these lizards because we did not expect baby food to be less
577 desirable than other foods. Nonetheless, it could have factored into the high dropout rate.
578 Moreover, although we found similar patterns of performance as in previous studies (Szabo
579 et al., 2018; 2019a), our methodology might have been less well-suited to test gidgee skinks
580 compared to tree skinks and blue-tongue lizards. We previously found similarly high dropout
581 rates in sleepy lizards (*Tiliqua r. asper*), also a member of the Egernia group, learning to
582 discriminate simpler stimuli (Szabo & Whiting, 2020b) although many of these were captive-
583 bred individuals which may have contributed to their high dropout. With data on only four
584 species we are unable to make distinct conclusions on the underlying cause of the
585 differences between species. Using a different methodology to test visual discrimination
586 learning in these species could help establish whether gidgee skinks and sleepy lizards are
587 poor at visual discrimination learning compared to other relatively closely related lizard
588 species. Additionally, the different species tested might simply differ in, for example, how
589 long they need to be habituated to captivity and testing procedures. Furthermore, they might
590 differ in food motivation or neophobia and might even have different preferences for visual
591 stimuli, all of which might have attributed to the species differences.

592 One goal of testing gidgee skinks to discriminate simple and compound stimuli was
593 to compare their performance to that of tree skinks and the eastern blue-tongue lizards, all of
594 which differ in sociality. Tree skinks live in small groups (Chapple, 2003) while blue-tongue
595 lizards are solitary (Koenig, Shine, & Shea, 2001). Set-shifting tasks test for flexibility in
596 behavior and attention, and we hypothesized that such cognitive skills might be enhanced in

597 social species because of their need to identify, track, and remember individuals and their
598 relationships within their social group ('Social Intelligence Hypothesis'; Byrne & Whiten 1988;
599 Humphrey, 1976). In other vertebrates, group size has been linked to increased cognitive
600 ability including associative and reversal learning (e.g., Ashton et al., 2018). Evidence shows
601 that gidgee skinks can discriminate between group and non-group members (Bull, Griffin,
602 Lanham, & Johnston, 2000) and that mothers recognize their offspring (Main & Bull, 1996)
603 based on chemical cues. Whether visual cues are involved in individual recognition is
604 unknown. We can therefore not exclude the possibility that flexible use of visual information
605 is important for group cohesion. Although we did not compare learning performance directly
606 across species, behavioral flexibility and learning does not seem to be enhanced in group-
607 living lizards (i.e., gidgee skinks). In the future, testing social cognitive abilities (e.g., social
608 learning), which are more closely linked to sociality, will help us understand how cognition
609 might relate to group size in lizards. However, the social life of lizards might not pose the
610 same selective pressures to enhance cognition such as in other vertebrates (Byrne & Whiten
611 1988; Humphrey, 1976; Jolly, 1966). Furthermore, testing a broader range of lizard species
612 could prove informative with respect to alternative hypotheses such as the 'Ecological
613 Intelligence Hypothesis' which links enhanced cognition to a species' ecology (e.g., Byrne,
614 1997; Gibson, 1986; Rosati, 2007).

615 Taken together, this is the third study looking at discrimination learning of multi-
616 featured cues in lizards. We show that some gidgee skinks might be able to differentiate
617 feature sets when tested with compound stimuli without needing extensive pre-experience
618 with such cues, but further testing is needed to confirm our hypothesis. Such investigations
619 are important for evaluating how animals learn and what cognitive processes they use to
620 solve complex tasks. Understanding reptile cognition more broadly will help us better
621 understand the evolution of cognition in vertebrates and whether animals with different
622 evolutionary history have evolved learning processes convergently.

623

624 **Acknowledgements**

625 We thank Isabel Damas, Bruno Pleno, Faustine Degottex and Victor Frichot for their help
626 catching skinks and Jonathan Ogle for blind coding. This project was funded by the
627 Australian Society of Herpetologists (Student research grant to Birgit Szabo) and Macquarie
628 University (iMQRes scholarship and Rice Memorial Field Research Award received by Birgit
629 Szabo).

630

631 **References**

- 632 Alföldi, J., Di Palma, F., Grabherr, M., Williams, C., Kong, L., Mauceli, E., Russell, P., Lowe,
633 C. B., Glor, R. E., Jaffe, J. D., Ray, D. A., Boissinot, S., Shedlock, A. M., Botka, C.,
634 Castoe, T. A., Colbourne, J. K. Fujita, M. K., Moreno, R. G., ten Hallers, B. F.,
635 Haussier, D., Heger, A., Heiman, D. Janes, D. E., Johnson, J., de Jong, P.J.,
636 Koriabine, M. Y., Lara, M., Novick, P. A., Organ, C. L., Peach, S. E., Poe, S., Pollock,
637 D. D., de Queiroz, K., Sanger, T., Searle, S., Smith, J. D., Smith, Z., Swofford, R.,
638 Turner-Maier, J., Wade, J., Young, S., Zadissa, A., Edwards, S. V., Glenn, T. C.,
639 Schneider, C. J., Losos, J. B., Lander, E. S., Breen, M., Ponting, C. P. & Lindblad-
640 Toh, K. (2011). The genome of the green anole lizard and a comparative analysis
641 with birds and mammals. *Nature*, *477*, 587-591.
- 642 Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is
643 linked to group size and affects fitness in Australian magpies. *Nature*, *554*, 364-370.
- 644 Batabyal, A., & Thaker, M. (2019). Lizards from suburban areas learn faster to stay safe.
645 *Biology Letters*, *15*(2), Article 20190009. doi:10.1098/rsbl.2019.0009
- 646 Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal*
647 *of General Psychology*, *39*(1), 15–22.
- 648 Birrell, J. M., & Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional
649 set-shifting in the rat. *Journal of Neuroscience*, *20*(11), 4320-4324.
- 650 Bissonette, G. B., Lande, M. D., Martins, G. J., & Powell, E. M. (2012). Versatility of the
651 mouse reversal/set-shifting test: Effects of topiramate and sex. *Physiology &*
652 *Behavior*, *107*(5), 781-786
- 653 Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and
654 understanding individual differences in cognition. *Philosophical Transactions of the*
655 *Royal Society of London B Biological Sciences*, *373*(1756), Article 20170280.
656 doi:10.1098/rstb.2017.0280

- 657 Brown, V. J., & Tait, D. S. (2015). Behavioral flexibility: attentional shifting, rule switching
658 and response reversal. In I. P. Stolerman, & L. H. Price (Eds.), *Encyclopedia of*
659 *Psychopharmacology* (pp. 264-269). Springer-Verlag.
- 660 Bull, C. M., Griffin, C. L., Lanham, E. J., & Johnston, G. R. (2000). Recognition of
661 peromones from group members in a gregarious lizard, *Egernis stokesii*. *Journal of*
662 *Herpetology*, 34(1), 92-99.
- 663 Byrne, R.W. (1997). The technical intelligence hypothesis: an additional evolutionary
664 stimulus to intelligence? In A. W. Whiten, & R. W. Byrne (Eds.), *Machiavellian*
665 *intelligence II: Extensions and evaluations* (pp. 289-311). Cambridge University
666 Press.
- 667 Byrne, R. W., & Whiten, A. W. (1988). *Machiavellian intelligence: Social expertise and the*
668 *evolution of intellect in monkeys, apes, and humans*. Clarendon Press/Oxford
669 University Press.
- 670 Carazo, P., Noble, D. W. A., Chandrasoma, D., & Whiting, M. J. (2014). Sex and boldness
671 explain individual differences in spatial learning in a lizard. *Proceedings of the Royal*
672 *Society B: Biological Sciences*, 281(1782), Article 20133275.
673 doi:papers3://publication/doi/10.1098/rspb.2013.3275
- 674 Casale, M. B., Roeder, J. L., & Ashby, F. G. (2012). Analogical transfer in perceptual
675 categorization. *Memory & Cognition*, 40(3), 434-449. [https://doi.org/10.3758/s13421-](https://doi.org/10.3758/s13421-011-0154-4)
676 [011-0154-4](https://doi.org/10.3758/s13421-011-0154-4).
- 677 Castro, L., & Wasserman, E. A. (2016). Attentional shifts in categorization learning:
678 Perseveration but not learned irrelevance. *Behavioural Processes*, 123, 63-73.
679 doi:10.1016/j.beproc.2015.11.001
- 680 Chapple, D. G. (2003). Ecology, life-history, and behavior in the Australian scincid genus
681 *Egernia*, with comments on the evolution of complex sociality in lizards.
682 *Herpetological Monographs*, 17, 145-180.

- 683 Chung, M., Goulet, C. T., Michelangeli, M., Melki-Wegner, B., Wong, B. B. M., & Chapple, D.
684 G. (2017). Does personality influence learning? A case study in an invasive lizard.
685 *Oecologia*, 185(4), 641-651. doi:10.1007/s00442-017-3975-4
- 686 Clark, B. F., Amiel, J. J., Shine, R., Noble, D. W. A., & Whiting, M. J. (2014). Colour
687 discrimination and associative learning in hatchling lizards incubated at 'hot' and
688 'cold' temperatures. *Behavioral Ecology and Sociobiology*, 68(2), 239-247.
689 doi:10.1007/s00265-013-1639-x
- 690 Cogger, H. G. (2014). *Reptiles and amphibians of Australia* (7th ed.). CSIRO PUBLISHING,
691 Victoria.
- 692 Day, L. B., Crews, D., & Wilczynski, W. (1999). Spatial and reversal learning in congeneric
693 lizards with different foraging strategies. *Animal Behaviour*, 57, 393-407.
694 doi:10.1006/anbe.1998.1007
- 695 Diamond, A. (2013). Executive functions. *Annual Reviews in Psychology*, 64, 135-168.
696 doi:10.1146/annurev-psych-113011-143750
- 697 Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Primate analogue of the Wisconsin card
698 sorting test: Effects of excitotoxic lesions of the prefrontal cortex in the marmoset.
699 *Behavioral Neuroscience*, 110(5), 872-886.
- 700 Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control
701 within prefrontal cortex with an analog of the Wisconsin card sort test- Restriction to
702 novel situations and independence from "on-line" processing. *The Journal of*
703 *Neuroscience*, 17(23), 9285-9297.
- 704 Dopson, J. C., Esber, G. R., & Pearce, J. M. (2010). Differences in the associability of
705 relevant and irrelevant stimuli. *Journal of Experimental Psychology Animal Behavior*
706 *Processes*, 36(2), 258-267. doi:10.1037/a0016588
- 707 Dougherty, L. R., & Guillette, L. M. (2018). Linking personality and cognition: a meta-
708 analysis. *Philosophical Transactions of the Royal Society of London B Biological*
709 *Sciences*, 373(1756), Article 20170282. doi:10.1098/rstb.2017.0282

- 710 Duffield, G. A., & Bull, C. M. (1998). Seasonal and ontogenetic changes in the diet of the
711 Australian skink *Egernia stokesii*. *Herpetologica*, 54, 414-419.
- 712 Duffield, G. A., & Bull, C. M. (2002). Stable social aggregations in an Australian lizard,
713 *Egernia stokesii*. *Naturwissenschaften*, 89, 424-427. doi:10.1007/s00114-002-0346-7
- 714 Falissard, B. (2012). *psy: Various procedures used in psychometry*. R package version 1.1.
715 <https://CRAN.R-project.org/package=psy>
- 716 Fleishman, L. J., Loew, E. R., & Whiting, M. J. (2011). High sensitivity to short wavelengths
717 in a lizard and implications for understanding the evolution of visual systems in
718 lizards. *Proceedings of the Royal Society B: Biological Sciences*, 278(1720), 2891-
719 2899. <https://doi.org/10.1098/rspb.2011.0118>.
- 720 Forstmeier, W., Wagenmakers, E. J., & Parker, T. H. (2017). Detecting and avoiding likely
721 false-positive findings - a practical guide. *Biological Reviews of the Cambridge*
722 *Philosophical Society*, 92(4), 1941-1968. doi:10.1111/brv.12315
- 723 Gardner, M. G., Bull, C. M., & Cooper, S. J. (2002). High levels of genetic monogamy in the
724 group-living Australian lizard *Egernia stokesii*. *Molecular Ecology*, 11, 1787-1794.
- 725 Gardner, M. G., Bull, C. M., Cooper, S. J., & Duffield, G. A. (2001). Genetic evidence for a
726 family structure in stable social aggregations of the Australian lizard *Egernia stokesii*.
727 *Molecular Ecology*, 10, 175-183.
- 728 Gardner, M. G., Bull, C. M., Fenner, A., Murray, K., & Donnellan, S. C. (2007). Consistent
729 social structure within aggregations of the Australian lizard, *Egernia stokesii* across
730 seven disconnected rocky outcrops. *Journal of Ethology*, 25, 263-270.
- 731 Gibson, K. R. (1986). Cognition, brain size, and the extraction of embedded food resources.
732 In J. G. Else, & P. C. Lee (Eds.), *Primate Ontogeny, Cognition, and Social Behaviour*
733 (pp. 93-105). Cambridge University Press.
- 734 Goulet, C. T., Michelangeli, M., Chung, M., Riley, J. L., Wong, B. B. M., Thompson, M. B., &
735 Chapple, D. G. (2018). Evaluating cognition and thermal physiology as components
736 of the pace-of-life syndrome. *Evolutionary Ecology*, 32(5), 469-488.
737 doi:10.1007/s10682-018-9948-1

- 738 Guidelines for the treatment of animals in behavioural research and teaching. (2018). *Animal*
739 *Behaviour*, 135, I-X. [doi:10.1016/j.anbehav.2017.10.001](https://doi.org/10.1016/j.anbehav.2017.10.001)
- 740 Hadfield, J. D. (2010). MCMC methods for multi-response generalised linear mixed models:
741 The MCMCglmm R package. *Journal of Statistical Software*, 33, 1-22.
742 [doi:10.18637/jss.v033.i02](https://doi.org/10.18637/jss.v033.i02)
- 743 Humphrey, N. (1976). The social function of intellect. In P. P. G. Bateson, & R. A. Hinde
744 (Eds.), *Growing points in ethology* (pp. 303-317). Cambridge University Press.
- 745 Jolly, A. (1966). Lemur social behavior and primate intelligence. *American Association for*
746 *the Advancement of Science*, 153(3735), 501-506.
- 747 Koenig, J., Shine, R., & Shea, G. (2001). The ecology of an Australian reptile icon: How do
748 blue-tongued lizards (*Tiliqua scincoides*) survive in suburbia? *Wildlife Research*,
749 28(3), 215-227.
- 750 Langkilde, T., & Shine, R. (2006). How much stress do researchers inflict on their study
751 animals? A case study using a scincid lizard, *Eulamprus heatwolei*. *Journal of*
752 *Experimental Biology*, 209, 1035-1043. [doi:10.1242/jeb.02112](https://doi.org/10.1242/jeb.02112)
- 753 Lawrence, D. H. (1949). Acquired distinctiveness of cues: I. Transfer between
754 discriminations on the basis of familiarity with the stimulus. *Journal of Experimental*
755 *Psychology*, 39, 770-784.
- 756 Leal, M., & Powell, B. J. (2012). Behavioural flexibility and problem-solving in a tropical
757 lizard. *Biology Letters*, 8(1), 28-30. [doi:10.1098/rsbl.2011.0480](https://doi.org/10.1098/rsbl.2011.0480)
- 758 Liu, Y., Day, L. B., Summers, K., & Burmeister, S. S. (2016). Learning to learn: advanced
759 behavioural flexibility in a poison frog. *Animal Behaviour*, 111, 167-172.
760 [doi:10.1016/j.anbehav.2015.10.018](https://doi.org/10.1016/j.anbehav.2015.10.018)
- 761 Logan, C. J. (2016). Behavioral flexibility in an invasive bird is independent of other
762 behaviors. *PeerJ*, 4, Article e2215. [doi:10.7717/peerj.2215](https://doi.org/10.7717/peerj.2215)
- 763 Mackintosh, N. J. (1965). Selective attention in animal discrimination learning. *Psychological*
764 *Bulletin*, 64, 124-150.

- 765 Mackintosh, N. J., & Little, L. (1969). Intradimensional and extradimensional shift learning by
766 pigeons. *Psychonomic Science*, 14, 5-6.
- 767 McMillan, N., & Roberts, W. A. (2012). Pigeons make errors as a result of interval timing in a
768 visual, but not a visual-spatial, midsession reversal task. *Journal of Experimental*
769 *Psychology: Animal Behavior Processes*, 38, 440-445.
- 770 McMillan, N., Sturdy, C. B., & Spetch, M. L. (2015). When is a choice not a choice? Pigeons
771 fail to inhibit incorrect responses on a go/no-go midsession reversal task. *Journal of*
772 *Experimental Psychology: Animal Learning and Cognition*, 41, 255-265.
- 773 Main, A. R., & Bull, C. M. (1996). Mother–offspring recognition in two Australian lizards,
774 *Tiliqua rugosa* and *Egernia stokesii*. *Animal Behaviour*, 52, 193-200.
- 775 Munch, K. L., Noble, D. W. A., Botterill-James, T., Koolhof, I. S., Halliwell, B., Wapstra, E., &
776 While, G. M. (2018). Maternal effects impact decision-making in a viviparous lizard.
777 *Biology Letters*, 14(4), Article 20170556. doi:10.1098/rsbl.2017.0556
- 778 Nakagawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and
779 publication bias. *Behavioral Ecology*, 15(6), 1044-1045.
780 <https://doi.org/10.1093/beheco/arh107>
- 781 Nicolakakis, N., Sol, D., & Lefebvre, L. (2003). Behavioural flexibility predicts species
782 richness in birds, but not extinction risk. *Animal Behaviour*, 65(3), 445-452.
783 doi:10.1006/anbe.2003.2085
- 784 Noble, D. W. A., Byrne, R. W., & Whiting, M. J. (2014). Age-dependent social learning in a
785 lizard. *Biology Letters*, 10(7), Article 20140430. doi:10.1098/rsbl.2014.0430
- 786 Noble, D. W. A., Carazo, P., & Whiting, M. J. (2012). Learning outdoors: Male lizards show
787 flexible spatial learning under semi-natural conditions. *Biology Letters*, 8(6), 946-948.
788 doi:10.1098/rsbl.2012.0813
- 789 Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: a review and a possible
790 integration. In C. J. Mitchell, & M. E. Le Pelley (Eds.), *Attention and associative*
791 *learning: From brain to behaviour* (pp. 11-40). Oxford University Press.

- 792 Pearson, S. K., Godfrey, S. S., Bull, C. M., & Gardner, M. G. (2016). Larger lizards live
793 longer in the group-living *Egernia stokesii*. *Australian Journal of Zoology*, 64, 182-
794 191. <https://doi.org/10.1071/ZO16024>
- 795 Qadri, M. A. J., Ashby, F. G., Smith, J. D., & Cook, R. G. (2019). Testing analogical rule
796 transfer in pigeons (*Columba livia*). *Cognition*, 183, 256-268.
797 doi:10.1016/j.cognition.2018.11.011
- 798 R Development Core Team (2008). *R: A language and environment for statistical computing*.
799 R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL
800 <http://www.R-project.org>.
- 801 Rayburn-Reeves, R. M., Stagner, J. P., Kirk, C. R., & Zentall, T. R. (2013). Reversal learning
802 in rats (*Rattus norvegicus*) and pigeons (*Columba livia*): Qualitative differences in
803 behavioral flexibility. *Journal of Comparative Psychology*, 127, 202-211.
- 804 Reiner, A. J., & Powers, A. S. (1978). Intensity and pattern discrimination in turtles after
805 lesions of nucleus rotundus. *Journal of Comparative and Physiological Psychology*,
806 92(6), 1156-1168.
- 807 Roberts, A. C., Robbins, T. W., & Everitt, B. J. (1988). The effects of intradimensional and
808 extradimensional shifts on visual discrimination learning in humans and non-human
809 primates. *The Quarterly Journal of Experimental Psychology Section B*, 40(4), 321-
810 341.
- 811 Rosati, A. G. (2017). Foraging cognition: Reviving the ecological intelligence hypothesis.
812 *Trends in Cognitive Sciences*, 21, 691-702.
- 813 Smith, J. D., Berg, M. E., Cook, R. G., Murphy, M. S., Crossley, M. J., Boomer, J., Spiering,
814 B., Beran, M. J., Church, B. A., Ashby, F. G., & Grace, R. C. (2012). Implicit and
815 explicit categorization: a tale of four species. *Neuroscience and Biobehavioral*
816 *Reviews*, 36(10), 2355-2369. doi:10.1016/j.neubiorev.2012.09.003
- 817 Smith, J. D., Zakrzewski, A. C., Johnston, J. J. R., Roeder, J. L., Boomer, J., Ashby, F. G., &
818 Church, B. A. (2015). Generalization of category knowledge and dimensional
819 categorization in humans (*Homo sapiens*) and nonhuman primates (*Macaca*

- 820 *mulatta*). *Journal of Experimental Psychology: Animal Learning and Cognition*, 41(4),
821 322-335. <https://doi.org/10.1037/xan0000071>.
- 822 Sol, D., & Lefebvre, L. (2000). Behavioural flexibility predicts invasion success in birds
823 introduced to New Zealand. *Oikos*, 90(3), 599-605.
- 824 Sol, D., Timmermans, S., & Lefebvre, L. (2002). Behavioural flexibility and invasion success
825 in birds. *Animal Behaviour*, 63(3), 495-502. doi:10.1006/anbe.2001.1953
- 826 Szabo, B., Hoefler, S., & Whiting, M. J. (2020a). Are lizards capable of inhibitory control?
827 Performance on a semi-transparent version of the cylinder task in five species of
828 Australian skinks. *Behavioral Ecology and Sociobiology*, 74, Article 118.
829 doi:10.1007/s00265-020-02897-y
- 830 Szabo B., Noble, D. W. A., Byrne, R. W., Tait, D. S., Whiting, M. J. (2018). Subproblem
831 learning and reversal of a multidimensional visual cue in a lizard: evidence for
832 behavioural flexibility? *Animal Behaviour*, 144, 17-26. doi:
833 10.1016/j.anbehav.2018.07.018
- 834 Szabo, B., Noble, D. W. A., Byrne, R. W., Tait, D. S., Whiting, M. J. (2019a). Precocial
835 juvenile lizards show adult-level learning and behavioural flexibility. *Animal*
836 *Behaviour*, 154, 75-84. doi:10.1016/j.anbehav.2019.06.003
- 837 Szabo, B., Noble, D. W. A., & Whiting, M. J. (2019b). Context-specific response inhibition
838 and differential impact of a learning bias in a lizard. *Animal Cognition*, 22(3), 317-
839 329. doi:10.1007/s10071-019-01245-6
- 840 Szabo, B., & Whiting, M. J. (2020b). Do lizards have enhanced inhibition? A test in two
841 species differing in ecology and sociobiology. *Behavioural Processes*, 172, Article
842 104043. doi:10.1016/j.beproc.2020.104043
- 843 Tebbich, S., Sterelny, K., & Teschke, I. (2010). The tale of the finch: adaptive radiation and
844 behavioural flexibility. *Philosophical Transactions of the Royal Society of London B*
845 *Biological Sciences*, 365(1543), 1099-1109. doi:10.1098/rstb.2009.0291

- 846 Tebbich, S., & Teschke, I. (2014). Coping with uncertainty: woodpecker finches (*Cactospiza*
847 *pallida*) from an unpredictable habitat are more flexible than birds from a stable
848 habitat. *PLoS One*, 9(3), Article e91718. doi:10.1371/journal.pone.0091718
- 849 Trobalon, J. B., Miguelez, D., McLaren, I. P. L., & Mackintosh, N. J. (2003). Intradimensional
850 and extradimensional shifts in spatial learning. *Journal of Experimental Psychology:*
851 *Animal Behavior Processes*, 29(2), 143-152. doi:10.1037/0097-7403.29.2.143
- 852 van de Pol, M. (2012). Quantifying individual variation in reaction norms: how study design
853 affects the accuracy, precision and power of random regression models. *Methods in*
854 *Ecology and Evolution*, 3(2), 268-280. doi:10.1111/j.2041-210X.2011.00160.x
- 855 Whiting, M. J., & While, G. M. (2017). Sociality in Lizards. In D. R. Rubenstein, & P. Abbot
856 (Eds.), *Comparative Social Evolution* (pp. 390-426). Cambridge University Press.
- 857 Zentall, T. R. (2005). Selective and divided attention in animals. *Behavioural Processes*, 69,
858 1-15.
- 859 Zentall, T. R., Halloran, M., & Peng, D. (2020). Midsession reversal learning: Pigeons learn
860 what stimulus to avoid. *Journal of Experimental Psychology Animal Learning and*
861 *Cognition*, 46(2), 101-106. doi:10.1037/xan0000240
- 862 Zentall, T. R., & Riley, D. A. (2000). Selective attention in animal discrimination learning. *The*
863 *Journal of General Psychology*, 127(1), 45-66.