**Lizards from suburban areas learn faster to stay safe**

Anuradha Batabyal\*, Maria Thaker

**SUPPLEMENTARY MATERIAL**

Habitat composition quantification

 We quantified habitat composition of rural and suburban study areas by calculating the percent area that is built-up by anthropogenic structures and non-built-up (vegetation or bare-ground). High resolution satellite images (200dpi) of 0.1 km2 each from 3 suburban and 3 rural sites were obtained for the years 2013 to 2018 using historical satellite imagery tool of Google Earth engine (Figure S1). Change in habitat composition was analysed across 6 years because the study species has an average lifespan of 1-2 years, and therefore expected to experience the change in habitat during their lifetime. We used the satellite image of the same sites across all years. Built-up and non-built-up areas were distinct in colour (built-up: white buildings and grey roads were combined, non-built-up: brown soil, rock, and green vegetation were combined), and thus we used colour as the criterion for extraction. We used Tin-Eye colour extraction software (TinEye: Image search and recognition company; url: https://labs.tineye.com/color/) to distinguish between built-up areas from non-built-up areas by extracting the dominant colour, assigning it with specific colour code (hexcode) and calculating the area covered by the specific colour or substrate composition (Figure S2).

 We tested whether habitat composition of built-up areas was different across suburban and rural study areas and across years by performing a generalised linear mixed effect model (glmmadmb, glmmADMB package in R) with percentage built-up area as the response, and habitat (rural and suburban) and year (2013-2018) as the fixed factors. We also included an interaction effect between habitat and year, and replicate sites within each habitat were added as the random effect. Variation due to replicates was negligible (SD<0.1). Since there was a significant interaction effect between habitat and year, we conducted a post-hoc analysis using lsmeans package and detailed results are reported in Results section of main paper (see Table S1 for detailed model parameters).



Figure S1. Map of the study area showing replicate suburban (U1, U2, U3) and rural sites (R1, R2, R3) in the greater Bangalore city region, India. An area of 0.1 km2 around each replicate site was used for habitat quantification.



Figure S2. Methodology for quantifying built-up and non-built-up area by colour extraction using TinEye colour extraction software

Table S1. Model parameters for habitat composition comparison

|  |
| --- |
| GLMM: Habitat compositionModel: Area built-up ~ Habitat\* Year+(1|site)AIC=229.3; Log-likelihood=-101.64 |
| Coefficient | Estimate | Std error | z | P |
| Intercept (rural-2013) | 0.3598  | 0.4824  | 0.75  | 0.4557  |
| Habitat (suburban) | 0.9321  | 0.6101  | 1.53  | 0.1265  |
| Year 2014 | -0.1503  | 0.7091  | -0.21  | 0.8322  |
| Year 2015 | 0.2091  | 0.6490  | 0.32  | 0.7473 |
| Year 2016 | 0.5563  | 0.6049  | 0.92  | 0.3578  |
| Year 2017 | 0.0455  | 0.6744  | 0.07  | 0.9463  |
| Year 2018 | 0.0890  | 0.6673  | 0.13  | 0.8940  |
| Suburban:2014 | 1.7242  | 0.8132  | 2.12  | 0.0340 \* |
| Suburban:2015 | 1.9635  | 0.7557  | 2.60  | 0.0094 \* |
| Suburban:2016 | 1.6531  | 0.7179  | 2.30  | 0.0213 \* |
| Suburban:2017 | 2.1660  | 0.7773  | 2.79  | 0.0053 \* |
| Suburban:2018 | 2.2244  | 0.7706  | 2.89  | 0.0039 \* |
| Random effect variance:  |
| Variance0.0003 | Std Dev0.01 |  |  |  |  |

Capture, housing and experimental setup for learning trials

Lizards were captured by noosing during the peak activity season (April-July 2016) and housed in the laboratory for up to 16-18 days for the learning trials. Only sexually mature males (SVL [mean ± SE] = 117.20 ± 1.87 mm) that were of similar age (in their first breeding season) were used. Each lizard was housed individually in a large rectangular enclosure (80 x 45 x 30cm) that also served as the testing enclosure. Enclosures were maintained indoors under ambient and visible spectrum white lamps for lighting. Each enclosure was lined with paper towels as a substratum, and were provided with a rock perch and a glass petri dish for water. Lizards were fed with live crickets before the start of trials each day. Each enclosure had a perch in the middle and two identical refuges constructed from PVC pipes that were cut longitudinally (15 cm long x 5 cm radius) and placed at the two farthest ends. Lizards from both suburban and rural areas showed similar levels of motivation to use the PVC pipes as refuges; there was no bias by the suburban lizards to use this artificial material with greater propensity compared to rural lizards (no significant difference in the latency to choose the first refuge by rural and suburban lizards during learning task; Wilcoxon U test: W=95, *P*=0.670).

The location of safe and unsafe refuges (right or left side of the tank) was counter balanced for all lizards and no effect of side was found to affect the choice of refuges (Chi-square test of proportion for side bias shows no significant difference between right or left refuge during the 1st trial: *χ*2=0.33, *P*=0.855). Similarly, we found no colour/cue preference by lizards during the start of the experiment (Chi-square test of proportion for colour bias shows no significant difference between safe-colour associated refuge and unsafe refuge during the 1st trial: *χ*2=0.83, *P*=0.361).

Before the start of the trials, we measured the mass (g) and SVL (mm) for all lizards which we used to calculate body condition as scaled mass index (as per Peig and Green, 2009).

Scaled mass index= Mi \* (Lo/Li)^ bSMA

where Mi and Li are the body mass and the SVL of individual i respectively; bSMA is the scaling exponent estimated by the SMA regression of M on L; Lo is the mean SVL value for the study population. The scaling exponent bSMA was calculate indirectly by dividing the slope from an OLS regression (bOLS) by the Pearson’s correlation coefficient r (as per Peig and Green, 2009).

Learning criterion

Both the learning and reversal learning trials involved simulating a predator attack and recording the refuge choice of lizards. If a lizard was found inside the “unsafe refuge” at the beginning of a trial, the refuge was lifted up and once the lizard came back out onto the perch, the trial began (lizards were found inside unsafe refuge at the start of trials for n=7 trials during learning task and n=18 trials during reversal learning task). Alternately, if a lizard was already found inside the designated “safe refuge” at the start of a trial, it was considered to have chosen the safe refuge and the latency to choose that safe refuge was recorded as zero seconds. This criterion of learning is biologically relevant as these agamid lizards prefer to remain inside refuges during the night, and for general safety. However, to ensure that this choice did not inflate the results, we conducted the statistical analyses with (n = 607 trials for learning and n = 560 for reversal learning) and without these trials (n = 549 trials for learning and n = 530 for reversal learning). When excluding the trials in which individuals were already inside safe refuge, we found that the latency to select the safe refuge during the learning (z=-73.38, *P*<0.001) and reversal learning (z=-36.72, *P*<0.001) tasks decreased with the number of trials. Furthermore, suburban lizards were significantly faster than rural lizards in choosing the safe refuge (Task1: z=-2.85, *P* =0.004; Task2: z=-4.24, *P*<0.001). Thus, there were no statistical differences in the final conclusion with or without these trials; hence, the results reported in the main manuscript include all trial data. The criterion for learning was standardised at five consecutive correct choices, but we continued the testing trials for all lizards for 20 trials. Notably, all but three lizards exhibited learning based on the 5-consecutive-correct-choices criterion continued to choose the correct refuge until the end of the 20 trials. The three individuals that failed to choose the safe refuge on the 6th trial proceeded to choose the safe refuge on all subsequent trials till the 20th trial. We therefore included these 3 individuals in the reversal learning task.

Statistical analysis of behavioural data

 We first compared the proportion of individuals from suburban and rural habitats that showed learning or reversal learning using chi square tests for equality of proportion with continuity correction. We then tested whether the latency to select the safe refuge significantly decreased with time as this pattern would indicate that lizards were learning. We chose latency to select the safe refuge as our response because number of incorrect choices and latency to choose the safe refuge were tightly correlated (Pearson’s correlation coefficient = 0.726, t=26.00, *P* < 0.001, 95% CI: 0.68, 0.76). We used a generalised linear mixed model (GLMM, glmmADMB package in R) with latency to select safe refuge as the response variable (Poisson error distribution) and trial number, habitat (suburban and rural) and body condition as fixed factors, with individual ID and replicate site as random effects. We also ran a two sample T -test to compare body condition across our two populations. To supplement these analyses, we ran a generalized linear mixed model (GLMM) using correct and incorrect choices (‘1’ or ‘0’) across all trial data with a binomial probability distribution (i.e. binomial error – logit link) and included trial number, habitat (suburban and rural) and body condition as fixed factors, with individual ID and replicate site as random effects. We ran two GLMMs, one each for learning and reversal learning trials (see results in Table S2). We accounted for individual lizards by including lizard ID as a random effect in all our models but we could only run a random intercept model as the random slope model had poor model convergence. To account for this we re-ran our model using generalized estimating equations (GEEs) and included an AR1 correlation structure to test whether temporal correlation affected our estimates. This gave similar results to the broader GLMMs and thus we present results from the random intercept model with latency to choose the safe refuge as the response variable in our main manuscript. The major conclusion about learning responses remained same whether we use latency to select safe refuge or correct vs incorrect choices as the response variable in the statistical analyses.

Table S2. Model parameters for learning and reversal learning tasks

|  |
| --- |
| **GLMM**: **Task 1 Learning**Model1: latency to choose safe refuge~ habitat + trial number + body condition + (1|ID) + (1|site)AIC=6615; Log-likelihood=-3301.10 |
| Coefficient | Estimate | Std error | z | P |
| Intercept (rural) | 2.72170  | 0.25502  | 10.67  | <2e-16 |
| Habitat (suburban) | -0.43030  | 0.18027  | -2.39  | 0.017\* |
| Trial number | -0.06889  | 0.00304  | -22.68  | <2e-16\* |
| Body condition | 0.00155  | 0.00265  | 0.58  | 0.559  |
| Random effect variance: |  |  |  |  |
| IDSite |  | Variance0.220.05 | Std Dev0.470.23 |  |  |
| Model2: choice (safe/unsafe) ~ habitat + trial number + body condition+ (1|ID) + (1|site)AIC=432.2; Log-likelihood=-210.089 |
| Coefficient | Estimate | Std error | z | P |
| Intercept (rural) | -0.519953  | 0.534010  | -0.97  | 0.3302  |
| Habitat (suburban) | 1.30442  | 0.409130  | 3.19  | 0.0014 \* |
| Trial number | 0.198059  | 0.027189  | 7.28  | 3.2e-13 \* |
| Body condition | 0.000531  | 0.005134  | 0.10  | 0.9177  |
| Random effect variance: |  |  |  |  |
| IDSite |  | Variance0.370.06 | Std Dev0.610.25 |  |  |
| **GEE: Task 1 Learning**Model1: latency to choose safe refuge~ habitat + trial number + body condition + (1|ID) |
| Coefficient | Estimate | Std error | Wald | P |
| Intercept (rural) | 3.479816  | 0.266510 | 170.49  | < 2e-16 \* |
| Habitat (suburban) | -0.624663  | 0.174991  | 12.74 | 0.0003\* |
| Trial number | -0.102906  | 0.010253 | 100.74 | < 2e-16 \* |
| Body condition | 0.001558  | 0.002756  | 0.32 | 0.5718  |
| Estimated correlation parameter: alpha |
| Estimate0.42 | Std Err0.07 |  |  |  |  |
|  |
| **GLMM**: **Task 2 Reversal Learning**Model 3: latency to choose safe refuge~ habitat + trial number + body condition+ (1|ID) + (1|site)AIC= 5390.9; Log-likelihood=-2689.44 |
| Coefficient | Estimate | Std error | z | P |
| Intercept (rural) | 3.43006  | 0.23944  | 14.33  | < 2e-16 |
| Habitat (suburban) | -0.76745  | 0.22381  | -3.43  | 0.0006\* |
| Trial number | -0.08736  | 0.00278  | -31.41  | < 2e-16\* |
| Body condition | 0.00354  | 0.00208  | 1.71  | 0.0881 |
| Random effect variance: |  |  |  |  |
| IDSite | Variance0.120.04 | Std Dev0.350.20 |  |  |
| Model 4: choice (safe/unsafe) ~ habitat + trial number + body condition+ (1|ID) + (1|site)AIC=545.1; Log-likelihood=-267.53 |
| Coefficient | Estimate | Std error | z | P |
| Intercept (rural) | -0.96958  | 0.48848  | -1.98  | 0.047 |
| Habitat (suburban) | 0.75006  | 0.31531  | 2.38  | 0.017 \*  |
| Trial number | 0.21504  | 0.02294 | 9.38  | <2e-16 \* |
| Body condition | -0.00569  | 0.00464  | -1.23  | 0.220  |
| Random effect variance: |  |  |  |  |
| IDSite |  | Variance0.327.104e-05 | Std Dev0.570.008 |  |  |
| **GEE**: Task 2 Reversal LearningModel 3: latency to safe refuge~ habitat + trial number + body condition+ (1|ID) |
| Coefficient | Estimate | Std error | Wald | P |
| Intercept (rural) | 3.65846  | 0.24388 | 225.02  | < 2e-16 |
| Habitat (suburban) | -0.52825  | 0.17039  | 9.61  | 0.0019\* |
| Trial number | -0.05109  | 0.01140  | 20.07  | 7.5e-06 \* |
| Body condition | -0.00128  | 0.00250  | 0.26  | 0.6091  |
| Estimated correlation parameter: alpha |
| Estimate0.84 | Std Err0.02 |  |  |  |  |