Exploration and spatial cognition show long-term repeatability but no heritability in the Aegean wall lizard

Reference:
De Meester Gilles, Pafilis Panayiotis, Vasilakis Gabriel, Van Damme Raoul.- Exploration and spatial cognition show long-term repeatability but no heritability in the Aegean wall lizard
Animal behaviour - ISSN 1095-8282 - 190(2022), p. 167-185
Full text (Publisher's DOI): https://doi.org/10.1016/J.ANBEHAV.2022.06.007
To cite this reference: https://hdl.handle.net/10067/1899710151162165141

Exploration and spatial cognition show long-term repeatability but no heritability in the Aegean wall lizard.

Gilles De Meester[1,2], Panayiotis Pafilis[2], Gabriel Vasilakis[2], & Raoul Van Damme[1]

[1] Department of Biology, Functional Morphology Group, University of Antwerp, Wilrijk, Belgium.

[2] Department of Biology, Section of Zoology and Marine Biology, National & Kapodistrian University of Athens, Athens, Greece.

Declaration of interest

The authors declare no conflict of interest.

Word count

18 250 words (including references), 6 figures, 4 tables, 5 tables in appendix and 2 figures in supplementary material.

[1] Corresponding author:

Gilles De Meester

Campus Drie Eiken, Building D – Room D.141

Universiteitsplein 1, 2610 Wilrijk, Belgium

email: Gilles.DeMeester@uantwerpen.be

telephone: +3232658916

ORCID ID: 0000-0001-7051-9957

ABSTRACT

Recently, biologists have become increasingly interested in cognitive variation among individuals, and how it relates to differences in fitness. However, very few studies so far have studied the long-term repeatability and heritability of cognitive performance in wild animals. This is nevertheless crucial information to fully understand the potential ecological and evolutionary impact of individual variation in cognitive performance. In 2019, we assessed exploration, problem-solving and spatial and reversal learning in 66 Aegean wall lizards (*Podarcis erhardii*), then released them in semi-natural enclosures consisting of either simple or complex habitat. One year later, we recaptured and retested the surviving lizards and their offspring to estimate the long-term repeatability and heritability of these behavioural and cognitive characteristics. We found that exploration and spatial learning were moderately repeatable, but reversal learning only marginally and learning flexibility and problem-solving not at all. Reversal learning ability declined over time in lizards kept in simple habitat, but not in those kept in complex habitats – suggesting habitat-dependent cognitive plasticity. To our knowledge, this is the first study demonstrating (long-term) consistency in cognitive traits within a non-avian reptile. The combination of modest repeatability and low heritability does suggest that within our study species, personality and cognitive variation among individuals and populations is mostly moulded by environmental effects.

Keywords: cognition, animal personality, *Podarcis*, behavioural repeatability, behavioural plasticity,

INTRODUCTION

The evolution of cognition, i.e. the acquisition, retention and use of environmental information (Dukas, 2004), is sometimes regarded as one of the most enigmatic topics within the study of biology (Thornton, Isden, & Madden, 2014). The benefits of cognition seem obvious: learning and problem-solving equip animals with the necessary behavioural flexibility to deal with changing environmental conditions (Sol, 2009). However, cognition is also costly, as it requires energetically expensive neural processes and tissues (Aiello & Wheeler, 1995; Buechel et al., 2018). Recently, the field of cognitive ecology has started to adopt an individual-based approach to identify the drivers of cognitive evolution (Boogert et al., 2018). Individuals can differ remarkably in their cognitive abilities, and such variation can provide the raw material for natural and sexual selection to act on. Hence, a small, but growing, number of studies have tried to relate individual variation in cognition to individual differences in fitness (reviewed in Boogert et al., 2018; Morand-Ferron, 2017; Morand-Ferron, Cole, & Quinn, 2016). Although valuable and informative, these studies have nonetheless been criticized for several reasons. To establish that a cognitive trait evolves by natural selection, one should not only prove its link with survival and/or reproduction, but also show that cognitive variation among individuals is consistent (i.e. repeatable) and heritable (Boogert et al., 2018; Cauchoix & Chaine, 2016; Griffin, Guillette, & Healy, 2015; Morand-Ferron, 2017; Thornton et al., 2014). Yet, few studies have verified these assumptions in non-human animals (but see references below).

Repeatability (R) is an estimate of how much of the phenotypic variation in a population can be explained by consistent differences among individuals (Bell, Hankison, & Laskowski, 2009; Boake, 1989; Nakagawa & Schielzeth, 2010). Behavioural repeatability does not necessarily exclude plasticity at the individual level. For instance, repeatability of cognitive performance would be demonstrated if over multiple repetitions of a learning task the relative order of fast to slow learners remains stable, even if intrinsic (e.g. age) or extrinsic (e.g. season) conditions influence individual performance (Griffin et al., 2015). Animals can either be retested on the same, slightly altered, task (temporal repeatability) or can be subjected to different protocols designed to measure the same cognitive ability (contextual repeatability) (Cauchoix et al., 2018).

Measuring the repeatability of cognitive performance, or behaviour in general, is deemed essential for various reasons. Firstly, repeatability sets the upper limit to heritability (h^2) (Boake, 1989; but see Dohm, 2002), and as such determines whether and how fast a trait may respond to selection (Boake, 1989; Croston et al., 2015; Morand-Ferron et al., 2016; Troisi et al., 2021). Secondly, measuring whether differences in cognitive traits are consistent is needed to understand the ecological and evolutionary relevance of their relation with a multitude of other biological traits (Soha et al., 2019), such as life-history (Cole et al., 2012), secondary sexual traits (Alvarez-Quintero, Velando, & Kim, 2021) or personality (consistent interindividual differences in behaviour across time and context, Réale et al., 2007). There is currently a strong interest in exploring how personality and cognition covary (Dougherty & Guillette, 2018). However, interpretation of such relationships (or the lack thereof) often assumes that cognition is repeatable as well, yet this is rarely verified (Griffin et al., 2015; Sommer-Trembo & Plath, 2018). Last but not least, many authors have pointed out that performance on a cognitive task can be influenced by other, non-cognitive, factors, such as distraction, motivation or previous experience (Morand-Ferron et al., 2016; Rowe & Healy, 2014). Repeated measurements are thus necessary to validate whether we are accurately quantifying cognitive variation.

While measuring repeatability of non-cognitive personality traits has almost become standard procedure in behavioural research (Bell et al., 2009), studies assessing the repeatability and consistency of animal cognition are much rarer in comparison (but see e.g. Ashton et al., 2018; Brust & Guenther, 2017; Cauchoix et al., 2018; Cole, Cram, & Quinn, 2011; Cooke et al., 2021; Gibelli & Dubois, 2017; Langley et al., 2018; Mason et al., 2021; Reichert et al., 2020; Rodríguez & Gloudeman, 2011; Schuster, Carl, & Foerster, 2017a; Schuster et al., 2017b; Shaw, 2017; Shaw et al., 2019; Soha et al., 2019; Sommer-Trembo & Plath, 2018; Sorato et al., 2018; Tello-Ramos et al., 2018; Troisi et al., 2021). A recent meta-analysis by Cauchoix et al. (2018) reported low to moderate values for repeatability of cognitive performance (temporal: R = 0.18 – 0.28, contextual: R = 0.20 – 0.27), albeit this was based on a small number of (mostly unpublished) datasets ($N_{temporal}$ = 22 studies, $N_{contextual}$ = 27 studies). Nevertheless, most of these past studies had three major limitations.

Firstly, repeatability of cognition, and personality, is commonly measured on relative short timescales, e.g. days or weeks in between repeated tests (estimates from > 1 year: 9 % in Bell et al., 2009; 31 % in Cauchoix et al., 2018 albeit only five species). Within a short timeframe, individuals are more likely to be tested under similar intrinsic and extrinsic conditions, leading to an inflation of repeatability estimates and potentially pseudo-repeatability (Dingemanse & Dochtermann, 2013). Over a longer period, plastic responses to differential environmental or developmental alterations may decrease the repeatability of the behavioural traits under study (Bell et al., 2009). Secondly, behavioural repeatability is frequently tested in animals raised and/or kept in controlled lab conditions. These will experience less environmental variation than their wild counterparts, which may result in biased repeatability estimates not representative for natural populations (Archard & Braithwaite, 2010; Stamps & Groothuis, 2010). Hence, verifying how consistent interindividual differences in cognition and personality are over longer timescales in natural conditions is critical information when trying to understand the role of such variation in ecological and evolutionary processes, but such data is largely lacking (but see e.g. Carlson, Tetzlaff, & Rutz, 2020; Debeffe et al., 2015; Payne et al., 2021 for personality and e.g. Ashton et al., 2018; Cole et al., 2011; Shaw, 2017; Tello-Ramos et al., 2018 for cognition). Lastly, to our knowledge, no study to date investigated the long-term repeatability of cognition and personality within the same (wild) study system, despite the growing evidence that both aspects of behaviour are closely linked (Dougherty & Guillette, 2018).

Repeatability is often used as an estimation of heritability (Boake, 1989; but see Dohm, 2002), here defined in its narrow sense as the proportion of phenotypic variation in a population that can be explained by additive genetic effects (Falconer & Mackay, 1996). An alternative approach, however, is to measure behaviour of both parents and offspring and employ modern statistical methods to determine the amount of additive genetic variation (Colby, Kimock, & Higham, 2021; de Villemeuril, 2012). This also allows to determine the relative contribution of genetic versus permanent environmental effects in shaping phenotypic variation, which is key to understanding how a trait evolves (Croston et al., 2015). Although evidence for a genetic basis of cognition has been inferred by artificial selection studies (e.g. in guppies: Buechel et al., 2018; in parasitoid wasps: Liefting et al., 2018; in fruit flies: Mery & Kawecki, 2002), common garden experiments (e.g. in black-capped chickadees: Roth, LaDage, & Pravosudov, 2010) and genome-wide association studies (e.g. among dog breeds: Gnanadesikan et al., 2020), actual heritability estimations for cognition are rare. Cognitive abilities tend to be heritable in humans ($h^2 = 0.26 – 0.85$), primates ($h^2 = 0.21 – 0.91$) and laboratory mice ($h^2 = 0.21$-0.50) (reviewed in Croston et al., 2015) but data on non-traditional study taxa are scarce (Croston et al., 2015; but see: Carrete et al., 2017; Langley et al., 2020; Quinn et al., 2016; Smith, Philips, & Reichard, 2015; Vardi et al., 2020). The few data available often come from laboratory populations (Croston et al., 2015), and thus heritability estimates may have been biased due to founder effects, inbreeding and artificial selection (Langley et al., 2018; Sorato et al., 2018; but see Dochtermann et al., 2019) and to all individuals being raised under the same standardized conditions (Croston et al., 2015; Smith et al., 2015; Vardi et al., 2020). More research on the heritability of cognitive traits in wild populations is needed (but see: Carrete et al., 2017; Quinn et al., 2016), in order to advance our understanding of their evolution in nature.

We aimed to test the long-term repeatability and heritability of spatial cognition, problem-solving and exploration within the Aegean wall lizard (*Podarcis erhardii* Bedriaga 1882). In 2019, we measured personality and cognition in 66 individual lizards and released them in semi-natural enclosures for a survival experiment. After one year, surviving lizards and their offspring were recaptured and re-subjected to the same behavioural assays. We specifically addressed some of the aforementioned limitations of previous research by 1) using a non-traditional study organism: to our best knowledge this is the first study to measure repeatability of cognition in a non-avian reptile, 2) measuring behavioural repeatability over a sufficient long timescale (20 % of this species' average lifespan, Valakos 1990), 3) keeping and raising lizards in semi-natural environments and 4) studying both personality and cognition. This way, we hope to gain more ecologically relevant insights regarding the repeatability and heritability of cognition within this species. In addition, our lizards were kept in two contrasting habitats (either structural simple or complex). Habitat complexity is known to affect spatial learning within this species (De Meester, Pafilis & Van Damme, 2022), and by keeping adults and juveniles in two different environments, we hope to learn more regarding the role of selection versus plasticity in shaping such variation.

MATERIAL AND METHODS

Study species and overall experimental design

The Aegean wall lizard is a medium-sized (40 -75 mm) insectivorous lizard, which is widespread across the Southern Balkans (Lymberakis et al., 2018; Valakos et al., 2008). It can be considered an ecological generalist, as demonstrated by its broad habitat use (e.g. Mediterranean scrublands, open sand dunes, urban habitats, ... - Lymberakis et al., 2018; Valakos et al., 2008) and its dietary flexibility (arthropods, snails, eggs, fruits and occassionally conspecifics - Adamopoulou, Valakos, & Pafilis, 1999; Brock, Donihue, & Pafilis, 2014; Madden & Brock, 2018).

The initial batch of lizards was collected in May 2019 on Naxos Island (Cyclades, Greece), at five locations that differed in structural habitat complexity: two “complex” sites (Eggares: 37°07'49.1"N, 25°26'18.9"E and Rachi Polichnitou: 37°00'53.0"N, 25°24'10.7"E), covered in dense phrygana and maquis vegetation, dry stone walls and rock outcrops and three “simple” sites (Manto: 37°05'22.0"N, 25°21'42.1"E, Grotta: 37°06'41.8"N, 25°23'09.8"E and Alyko: 36°58'45.3"N, 25°23'21.0"E) that were characterized by small patches of vegetation in an overall open landscape. Seventy-one lizards were captured (by lasso, hand or pitfall) and transported to the National and Kapodistrian University of Athens (NKUA) for housing and behavioural experiments (see later). Lizards were kept in cotton bags in a cold cool box to reduce stress during transportation (Heathcote et al., 2014). Five of these lizards died in captivity. The behavioural data of the adult lizards collected in 2019 was previously reported in De Meester et al. (2022), as part of a larger study on whether variation in personality and cognition across populations of lizards could be explained by differences in structural habitat complexity, but was re-used here to specifically test the long-term repeatability, plasticity and heritability of personality and cognition. Upon completion of the experiments, the remaining 66 adults (N_{female} = 32, N_{male} = 34) were then released in four semi-natural enclosures on Naxos, in order to follow up their survival and reproduction over the course of one year. Lizards were released in July 2019 and recaptured in July 2020. We then transported the survivors (N_{female} = 22, N_{male} = 21) back to the NKUA and repeated a subset of the original behavioural experiments with them. The 2019 protocols were followed as closely as possible in order to measure the long-term temporal (rather than contextual) repeatability of personality and cognition. In addition, we collected and tested new individuals found within the enclosures, both juveniles (N = 43) and adults (N_{female} = 9, N_{male} = 12) for heritability estimations. These new adults were unmarked, and thus did not belong to the previous released batch. At the moment of capture, it was still unclear whether these 'unknown adults' were intruders or the full-grown offspring from a previous batch of lizards. All data collected in 2020, on both adults and juveniles, was new and has not been published before.

Housing

Lizards were housed at the animal facilities of the NKUA, individually in plastic terraria (adults: 22 x 20 x 17 cm, juveniles: 18 x 9 x 13 cm l x w x h). Terraria contained a water bowl, sand and stone bricks (adults) or coconut fibre and a plastic refuge (juveniles), and were placed underneath 60 W incandescent lamps for thermoregulation. Animals were fed three times per week: adults with mealworms (*Tenebrio molitor*) and juveniles with either maggots (larva of a calliphorid fly) or small mealworms. Prey items were always dusted with Terravit Powder (JBL, GmbH & Co. KG). Room temperature was maintained around 28 ± 2 °C.

Behavioural experiments

Initially, in 2019, lizards were subjected to four cognitive tasks and three separate personality assays (described in De Meester et al., 2022). Due to time constraints, we only retested the surviving adults on a subset of these tasks in 2020: an exploration assay, a spatial + reversal learning task and a problem-solving task (in this order). Their offspring were submitted to the same tests, excluding the problem-solving task.

Experiments were conducted from May until July in 2019, and from August until September in 2020, and were performed between 10:00 and 19:00. Each individual received 20 – 30 minutes basking time underneath a 100 W heat bulb prior to transferring them to experimental arenas, in order to achieve sufficiently high body temperatures. All experiments were filmed using a digital camera (JVC Everio GZ-HM400) or a GoPro (Grundig HD 720P). Experimental equipment (e.g. refuges, novel objects, etc.) was cleaned between trials with 70 % alcohol and water (Vicente & Halloy, 2017).

Exploratory behaviour

Exploratory behaviour is the tendency of an individual to gather new environmental information (Verbeek, Drent, & Wiepkema, 1994). It facilitates the discovery and exploitation of novel habitats and resources, but may be costly due to e.g. an increased risk of predation and parasite infection (Bajer et al., 2015; Baxter-Gilbert, Riley, & Whiting, 2019). Aegean wall lizards on Naxos experience strong seasonal fluctuations in food availability (De Meester et al., 2021), during which they may benefit from more exploratory behaviour to find alternative food sources. Exploration was tested using a classical novel arena test (cfr. Carazo et al., 2014; Damas-Moreira et al., 2019; De Meester et al., 2022; McEvoy et al., 2015). Two distinct novel arenas (60 x 60 x 30 cm l x w x h, either plywood or sand substrate), which contained four identical objects (either pinecones or stones, one in each corner) and four plastic refuges (either black or white) were used (Supplementary Figure S1a-b). We specifically used two different arenas to avoid a confounding effect of habituation (McEvoy et al., 2015). A lizard was put in the centre of the arena on a platform underneath an opaque container. After three minutes, the container was lifted and the animal was free to explore the arena for ten minutes. On the videos, we divided the arena in four equal quadrants and scored the following behaviours (starting from the moment the lizard left the platform): the latency until the first transition from one quadrant to another, total number of transitions between quadrants, number of times it investigated an object or refuge (by touching it with the snout or front legs), number of times it entered a refuge, the total time spent within refuges and the latency to explore all four quadrants of the arena. Lizards were tested once in each novel arena (random order – but 2019 adults retained the same order in 2020) with at least one day in between trials.

Spatial and reversal learning

Spatial cognition is the capacity of an animal to learn and remember the location of resources in its environment and is thus deemed a key aspect of an individual's fitness (Dukas, 2004; Tello-Ramos et al., 2018). Small lizards, including *P. erhardii*, tend to escape towards a refuge (e.g. a crevice in a rock wall, underneath a log, etc) when approached by a predator. This antipredator strategy is likely to be more successful if lizards are capable of remembering the location of adequate safe hiding spots (Noble, Carazo, & Whiting, 2012). Field observations suggest that lacertid lizards often repeatedly flee towards the same shelter, thus implying an important role of spatial memory (Font, 2019)

We tested spatial learning in our lizards using an ecological relevant antipredator task, in which subjects needed to learn the difference between a safe and unsafe refuge in order to make a successful escape (following the methodology of Carazo et al., 2014; De Meester et al., 2021; Font, 2019; Noble et al., 2012; Vardi et al., 2020). Lizards were tested in a large arena (60 x 60 x 30 cm) with two identical refuges (black plastic cups) opposite of each other. In and around the arena, visual spatial cues were presented to allow orientation (Supplementary Figure S1c-e). At the start of each trial, a lizard was positioned in the centre of the arena underneath a transparent cover for two minutes. After lifting the cover, we 'attacked' the lizard by poking and chasing it with a paintbrush. We always tried to attack the animal from straight above in order to avoid influencing in which direction it fled. Subjects needed to escape either to the left or the right refuge (relative to the observer, counterbalanced among original populations). Entering the 'safe' refuge resulted in two minutes of undisturbed rest, after which the individual was brought back to its terrarium. Entering the 'unsafe' refuge was penalized by continuing the predator attack. Trials ended when the lizard entered the safe refuge or until 120 s had passed (after which it was captured and allowed to rest in the correct refuge for two more minutes). Trials were limited to 120 s to reduce the amount of stress inflicted on the animal and avoid exhaustion. Each lizard received three trials per day for five consecutive days, and per trial we noted how many times the lizard entered the unsafe refuge ('errors').

Lizards living in a dynamic environment need to keep track of changes in their environment, update their memories frequently, and adjust their behaviour accordingly (Noble et al., 2012). Replacing old obsolete with new information requires cognitive flexibility, which is commonly measured using a reversal learning task (Audet & Lefebvre, 2017; Noble et al., 2012; Tebbich & Teschke, 2014). Following previous reversal learning protocols (e.g. Bebus et al., 2016; Boussard et al., 2020; Madden et al., 2018; Mason et al., 2021; van Horik et al., 2018), we changed the identity of the refuges after a standardized number of trials (15) for each individual: safe became unsafe and vice versa. Lizards received fifteen more trials over five days in order to relearn the location of the safe refuge. Adults retested in 2020 started the spatial learning phase with the same refuge designated as safe as in 2019 to make performances in both years comparable.

Escape Box

Problem-solving requires animals to express a new behaviour or apply an old behaviour in a novel context, and is therefore considered an indicator of behavioural flexibility (Griffin & Guez, 2014; Tebbich & Teschke, 2014; but see Audet & Lefebvre, 2017). Lizards with better problem-solving skills may increase their foraging efficiency, e.g. by being better at extracting hidden or difficult prey (Cooper et al., 2019). We tested problem-solving using an escape box task (Supplementary Figure S1f-g). A lizard was locked in a transparent Plexiglas box (17.4 x 17.4 x 6.5 cm l x w x h) which was placed opposite of a pile of stones underneath a 60 W heat bulb inside a larger arena. Hence, the lizard needed to escape from the box in order to get access to this hiding/basking spot. This was possible by performing a novel motor act i.e. sliding open a white opaque door (3.2 x 2.4 cm l x h). The door was already slightly opened and contained grooves to allow manipulation. Lizards received a single 30 min trial, and we recorded the time it took each lizard to escape. If lizards did not escape, we assigned them the maximum time as score.

In 2019, lizards were tested in two batches, either before or after the spatial cognition protocol for logistic reasons. In 2020, lizards were either tested on the same day as the spatial cognition protocol or later. Juveniles were not tested in the escape box as they proved to be unable to move the door in a few pilot trials. The video of one lizard in 2020 was lost.

Semi-natural enclosures

Upon completion of the experiments in 2019, 66 adult lizards were released in four semi-natural enclosures on Naxos as part of a survival experiment. Prior to release, each individual was photographed and individually marked by toe clipping for the purpose of identifying them upon recapture. Toe-clipping is a standard procedure to allow individual recognition in reptilian studies and is generally considered to have little to no negative effects either short- or long-term (Langkilde & Shine, 2006; reviewed in Perry et al., 2011). We removed maximum two toes per lizard (depending on already missing toes upon capture). In addition, we took small tail clips (± 1 cm) from each individual for parentage assignment (cfr. Huyghe et al., 2009; Huyghe et al., 2010). Tissues were stored in 96 % ethanol at 4 °C.

The four enclosures were located on a private field on Naxos, consisting of abandoned agricultural terraces with dense Mediterranean maquis and phrygana. Each enclosure was rectangularly shaped and constructed by fencing in approximately 100 m² of land with smooth metal fences (70 cm height and 30 cm deep). In two enclosures, we pruned the vegetation in order to mimic an open environment comparable to the structural simple sites where lizards were captured. The two other enclosures mimicked the more complex habitats (see Figure 1 and Supplementary Figure S2). Within each enclosure, we placed four piles of stones (± 30 cm high) for basking and shelter. We released 16 – 17 individuals in each enclosure based on their overall performance on the spatial cognition task (as to have more or less comparable numbers of good and bad learners in each enclosure) with approximately an equal number of a) males and females and b) lizards originating from complex and simple habitats (see Figure 1).

Terrestrial predators (e.g. snakes) were removed from the enclosures, although Megarian banded centipedes (*Scolopendra cingulata*), which are capable of catching and consuming *P. erhardii* (Deimezis-Tsikoutas, Kapsalas, & Pafilis, 2020), could not be entirely eliminated, and on one occasion a brown rat (*Rattus norvegicus*) intruded a complex enclosure, but was quickly removed. We also attempted to scare off avian predators by suspending reflective disks above the enclosures. Enclosures were normally checked biweekly by volunteers, who also placed leftover fruits and vegetables in each enclosure to attract flying insects. Vegetation around the enclosures was routinely cut down twice a year (early spring and mid-summer). Unfortunately, due to the Covid-19 pandemic, the enclosures were only irregularly checked throughout spring 2020, and early-spring maintenance had to be postponed until summer. Although this resulted in both enclosures growing denser vegetation, their relative differences in complexity did not change.

Sixty-six known lizards were released in the enclosures in July 2019. In July 2020, we recaptured every lizard in the enclosures during a seven-day recapture session. Survivors were identified based on toe clips and photographs of dorsal patterns. In total, we collected 43 survivors (59 – 71 mm snout-vent length SVL), 45 juveniles (29 – 37 mm) and 21 unknown adults (54 – 69 mm) from the enclosures (see Figure 1 for sample sizes per enclosure), all of which were then transported to the NKUA for subsequent personality and cognition tests (see above). We were unable to estimate the exact age of our animals at the time of capture, although the presence of umbilical scars in juveniles indicated that they had hatched relatively recent. Upon completion of the behavioural experiments in 2020, we also collected tail tissues for genetic analyses from unknown adults and juveniles (same protocol as described above for the adults in 2019).

Parentage assignment

Genetic analyses were conducted based on the protocol of Huyghe et al. (2010). DNA was extracted by placing ± 2 mm^3 of tail tissue in Chelex extraction buffer (consisting of 0.2 mL 10 % Chelex, 20 µL 1 % SDS and 2 µL of 20 mg/mL proteinase K), which was then put inside a stirring incubator (Eppendorf, thermomixer comfort), initially at 65°C for 60 minutes followed by 95°C for 15 minutes (modified from Small et al., 1998).

Parentage was assessed using microsatellite genotypes from nine loci that have been successfully used in congeneric species before (Pmeli-02, Pmeli-04, Pmeli-07, Pmeli-13, Pmeli-14 and Pmeli-19 from Huyghe et al., 2009; B3, B4 and B6 from Nembrini & Oppliger, 2003). For each DNA-sample, we prepared three different primer mixtures, each of them containing fluorescently labelled primers for three of the nine loci. Next, we mixed 1.25 µL of each primer mix with 6.25 µL Qiagen multiplex PCR master mix 2x and 3.5 µL water, to which we then added 1.5 µL of DNA extract. Mixtures were centrifuged and placed in the thermocycler (Biometra, T-professional thermocycler) for PCR amplification. PCR conditions were as follows: 15 minutes of denaturation at 95°C, 30 cycles of 30 s denaturation at 72°C, 90 s of annealing at 57°C or 60°C (Huyghe et al., 2009; dependent on the primers, see Nembrini & Oppliger, 2003) and 60 s of extension at 72°C. This was followed by another 30 minutes of extension at 60°C. Success of the PCR was then visually checked via gel-electrophoresis. After appropriate dilution, successful PCR-products were sent to an external lab (X) for microsatellite detection on an AB 3130XL Genetic Analyser (Life Technologies, Carlsbad, CA, USA).

Microsatellite data was first processed in the Geneious Prime software version 2021.0.3. (http://www.geneious.com/) for loci identification and then in Cervus version 3.0.7. (Kalinowski, Taper, & Marshall, 2007) for parentage assignment. We conducted separate parentage analyses for each enclosure, and used the unknown adults both as potential offspring of the 2019 adults and as potential parents of the juveniles. Proportion of mistyped loci was set to 5 % and relaxed and strict (trio) confidence intervals were equal to 80 and 95 % respectively.

Statistical analyses

Statistical analyses were performed in R version 3.5.1. (R Core Team, 2018). Prior to analysis, data were transformed where necessary to meet model assumptions.

First, we reduced the number of variables measured in the exploration test using a principal component analysis (PCA) with the ‘princomp’ function (with a correlation matrix to standardize variables) (‘stats’ package, R Core Team, 2018). We retained the first two principal components as these both had an eigenvalue > 1 (Kaiser-Guttman criterion, Kaiser, 1991) (cfr. Petelle, Martin, & Blumstein, 2015; Thys, Pinxten, & Eens, 2021; Vanden Broecke et al., 2021).

Secondly, to verify that lizards did learn during our spatial + reversal learning task, we used two separate generalized mixed-effect models (GLMMs), with a negative binomial distribution ('glmer.nb' function in 'lmer4' package, Bates et al., 2015) for the spatial and reversal phase respectively. The number of errors per trial was fitted as response variable, while trial number, side of the safe refuge (left/right, relative to the observer) and lizard age group (adults 2019, adults 2020, unknown adults & juveniles) were included as predictors. To test whether learning was consistent across age groups and independent of the rewarded side, we added age*trial and safe side * trial interactions as well. Original population, batch and lizard ID (with a random intercept and slope for trial) were included as random effects. Model assumptions were checked using the ‘RVAideMemoire’ (Hervé, 2020) and ‘performance’ (Lüdecke et al., 2021) packages. Statistical significance of GLMMs was tested with Wald Chi-square tests using the ‘Anova’ function ('car' package, Fox & Weisberg, 2019) Interactions were removed if not-significant.

Next, we estimated long-term (temporal) repeatability in personality and cognition in the subset of lizards that were tested in both 2019 and 2020 ($N = 43$) using a series of linear mixed-effect models (LMMs) ('lmer' function, 'lmerTest' package, Kuznetsova, Brockhoff, & Christensen, 2017). For exploration, we used PC1 and PC2. Next, we used the mean number of errors made over 15 trials as scores for spatial (SL) and reversal learning (RL), as better learners should more quickly learn the location of the correct refuge and thus make fewer wrong entrees (cfr. Sonnenberg et al., 2019; Tello-Ramos et al., 2018). The mean number of errors over both stages of the task was used as an indicator of overall learning flexibility (FLEX score). For the repeatability of the escape box task, we used escape times (ESC time). Lower scores on the cognitive tasks generally reflect better cognitive performance (fewer errors, less time to escape). We started by fitting global full models including the following variables: year (2019/2020), original habitat (simple/complex), enclosure type (simple/complex), sex, SVL (scaled, as risk-taking behaviour may be size-dependent, e.g. Bajer et al., 2015), tail status (complete/damaged, as this can affect lizard behaviour, cfr. Michelangeli et al., 2020) and side of the safe refuge (left/right, for the spatial + reversal task only). We also included all two- and three-way interactions between year, original habitat and enclosure type to see whether cognitive performance and personality would change over time in a habitat-dependent way. The following random factors were added to the models: lizard ID, arena type (plywood/sand, only for exploration), original population, enclosure ID and batch (only for spatial cognition and escape box). Where necessary, the 'bobyqa' optimizer was used to ensure model convergence (Bates et al., 2015). Next, we adopted a model selection approach (Symonds & Moussalli, 2011). Starting from the full global model, we generated a set of candidate models with the 'dredge' function ('MuMIn' package, Barton, 2013). The top model with the lowest Akaike information criterion corrected for small sample sizes (AICc) was selected, as well as alternative candidate models within ≤ 2 AICc units from it (cfr. Gomes et al., 2020; Symonds & Moussalli, 2011). We then determined the relative importance (RI) of all predictors by calculating their summed Akaike weights over all candidate models. Variables with a RI ≥ 0.50 (Gardner et al., 2020; Gomes et al., 2020; Simpson & McGraw, 2018, 2019) were used to construct a final model to test which factors influenced personality and cognitive performance, and to calculate the (adjusted) repeatability with the 'rptR' package (Stoffel et al., 2017). We built a series of similar models to estimate the short-term repeatability of personality within each subgroup of lizards (all 2019 adults, surviving 2019 adults , 2020 adults, unknown adults and juveniles). Significance of predictors is based on F-tests calculated using Kenward-Roger Degrees of Freedom Approximation ('anova' function, 'stats' package).

Narrow-sense heritability (h^2) was estimated by employing Bayesian mixed-effect animal models with the 'MCMCglmm' package (Hadfield, 2010). Animal models use both phenotypic data (here behavioural measurements) and pedigree information (based on our paternity assignment) to calculate the additive genetic variance of a given trait (σ^2_a). For these analyses, we used the complete dataset of all lizards that were released and captured in the enclosures, including the repeated measures of the 2020 adults. All response variables were z-transformed, but given the strong side bias (see results), SL and RL scores were z-transformed per rewarded side, to make cognitive performance among individuals comparable (cfr. Guillette et al., 2009). We once again used a model selection approach, starting from full global models with the following predictors: enclosure type (simple/complex, as structural complexity of the environment may affect behavioural development of individuals e.g. Spence, Magurran, & Smith, 2011), age group (adults 2019/adults unknown/juvenile, to account for age-differences in mean levels of behaviour, e.g. Rohrer et al., 2020), SVL (standardized per age group), tail status, and an age group*enclosure type interaction. We included the following random effects: lizard ID linked to the pedigree (σ^2_a, additive genetic variance), lizard ID independent of pedigree (to account for repeated measurements and permanent environmental effects), enclosure ID, novel arena (exploration data only), batch (SL + RL + FLEX only) and mother ID (to avoid that maternal effects would inflate h^2). We calculated a dominance matrix using the 'nadiv' package (Wolak, 2012) and implemented this as an additional random factor in the MCMCglmms to account for (genetic) dominance effects. Heritability was calculated from the final models as σ^2_a / σ^2_p with σ^2_p being the total phenotypic variance (de Villemeuril, 2012). Each animal model was initially run for 1 000 000 iterations, with a burn-in period of 100 0000 iterations and a 200 iteration thinning interval, and a parameter expanded prior (v = 1, nu = 1, alpha.mu = 0, v.mu = 1) (de Villemeuril, 2012). We checked convergence of chains and autocorrelation of all models, and in case of high autocorrelation (>0.10) we increased the number of iterations, the burn-in period and/or the iteration thinning interval. One juvenile was removed from the animal models due to missing data.

For both the (G)LMMs and the MCMCglmms, post-hoc multiple comparisons between different levels of a significant fixed factor and/or different slopes were performed with the 'emmeans' and 'emtrends' functions respectively using Tukey's method (Lenth et al., 2019).

Ethical note

All experiments and procedures were approved by the Ethical Committee of the University of Antwerp (file ID: 2017-67) and the Greek Ministry of Environment and Energy (permit IDs: 7ZΠP4653Π8-E76, ΨH424653Π8-ΩY2 and 69I44653Π8-ΔΣ1). Experiments and procedures were conducted in accordance with national legislation and adhered to the ASASB/ABS guidelines for the use of animals in behavioural research and teaching. Animals were checked daily while in captivity to monitor their welfare. Adult lizards were released at the initial site of capture, juveniles and intruders were retained for a follow-up experiment. Five lizards died in captivity in 2019, one lizard in 2020, and another one escaped from captivity in 2020.

RESULTS

We identified parents of 37 (82 %) juveniles. We were unable to identify the parents of any of the unknown adults (neither when matched with the current or previous batch of released lizards), suggesting that these lizards originated from the surrounding field and somehow managed to get into the enclosures. Four of these 'intruders' interbred with known adults and sired/birthed five of the juveniles in our dataset. Hence, their data was retained for the analyses. From the 66 lizards initially released, 14 males (42 %) and 17 females (52 %) reproduced (parent of at least one juvenile). Within that subset, males fathered on average 2.50 ± 0.28 (SE) juveniles (max. 6) and females birthed on average 2.00 ± 0.32 young (max. 5). Of the 'intruders', only two males (17 %) and two females (22 %) reproduced. Each male fathered a single juvenile, while the females birthed one and two juveniles respectively.

Descriptive statistics for all behavioural tests are given for the subset of adults tested in both years and per age group in the Appendix (Table A1 and A2).

Exploration PCA

The results of the PCA on the explorative behaviours are summarized in Table 1. The first principal component (eigenvalue = 1.76) explained 44.33 % of the total variance. Higher scores on PC1 corresponded to a higher number of transitions in the arena, more frequently touching objects, more refuges entered and more time spent in them, a lower latency to explore the entire arena and to enter the first refuge, and thus to overall more explorative behaviour. PC2 (eigenvalue = 1.29) explained 23.60 % of the total variance and represented a contrast between faster exploration (lower latency to first transition, more transitions, lower latency to explore the whole arena) versus more time spent inside the refuges (Table 1).

Exploration – short term repeatability within each group

Composition and output of the final models are reported in Table 2, as well as adjusted and unadjusted (short-term) repeatability estimates within each age group.

The short-term repeatability of exploration PC1 showed considerable variation across groups, e.g. being moderately high in 2020 adults but almost non-existent in intruders, 2019 adults and juveniles (see Table 2). Exploration PC1 increased with size in 2020 adults (est = 0.612 ± 0.224, $F_{1,52}$ = 6.465, p = 0.014) and tended to be lower in 2019 adults with an intact tail (intact: 0.843 ± 0.133, damaged: 1.732 ± 0.216, $F_{1,64}$ = 3.064, p = 0.085) and male 2020 adults (females: -0.024 ± 0.227, males: -0.405 ± 0.312, $F_{1,37}$ = 4.071, p = 0.051).

Exploration PC2 was moderately repeatable within intruders but not in 2020 adults or juveniles. 2019 adults showed significant repeatability for PC2 in the entire dataset, but not in the subgroup of survivors (Table 2). Exploration PC2 was higher in 2019 adults (both complete dataset and survivors only) with an intact tail (intact: 0.290 ± 0.107, damaged: -1.069 ± 0.362, $F_{1,61}$ = 8.731, p = 0.004), females in 2020 (females: 0.630 ± 0.210, males: -0.098 ± 0.190, $F_{1,35}$ = 5.970, p = 0.020; Figure 2), and tended to be higher in larger juveniles (0.242 ± 0.123, $F_{1,40}$ = 3.3776, p = 0.059) and 2019 adults originating from a simple habitat (complex: -0.194 ± 0.136; simple: 0.546 ± 0.151; $F_{1,3}$ = 6.134, p = 0.094).

Exploration – long term repeatability

Final models for the long-term repeatability of exploration are given in Table 3. PC1 scores differed significantly between years ($F_{1,152}$ = 41.171, p < 0.001), with lizards having lower scores in 2020 than 2019 (682 ± 450 % decrease, t = -6.552, p < 0.001, Figure 3a), and larger lizards were more explorative (0.472 ± 0.150, $F_{1,134}$ = 8.812, p = 0.004). Sex and enclosure type did not affect exploration PC1 (all p > 0.10, Table 3). Interindividual variation in PC1 was moderately repeatable from 2019 to 2020 (R_{adj} = 0.280, CI = [0.091; 0.436], LRT: p < 0.001; Table 3).

Male and female lizards differed in their exploration PC2 scores ($F_{1,35}$ = 9.032, p = 0.005). Males obtained lower scores, meaning that they explored more slowly and spent more time hiding (females: 0.484 ± 0.134, males: -0.065 + 0.130; t = -3.021, p = 0.005; Figure 2). Original habitat type did not predict exploration PC2 scores ($F_{1,3}$ = 2.276, p = 0.237). No other variable or interaction had sufficient high importance to be included in the final model. Lizards did not show long-term consistency in interindividual variation in PC2 (R_{adj} = 0.060, CI = [0.000; 0.201], LRT: p = 0.178, Table 3).

Exploration – heritability

The selected animal models (MCMCglmm) are reported in Table 4. The final model for exploration PC1 included age group and SVL (Table 4), albeit the latter did not affect explorative behaviour (posterior mean + 95% credibility interval: 0.057 [-0.044; 0.160]). Juveniles had lower exploration scores than adults (post-hoc pairwise comparisons: adults '19 – juveniles: 1.135 [0.864; 1.423], adults '20 - juveniles: 0.494 [0.185; 0.783], intruders - juveniles: 0.590 [0.204; 0.972]). Adults in 2019 behaved more explorative than intruders (intruders - adults '19: -0.548 [-0.893; -0.209]) and 2020 adults (adults '19 – adults '20: 0.642 [0.409; 0.858]). (Figure 3b; Table A3). Heritability of exploration PC1 did not differ from zero (h^2 = 0.031, CI = [0; 0.110]), meaning that additive genetic variance contributed almost nothing to the observed phenotypic variation.

The best model explaining variation in exploration PC2 was the null model (Table 4), thus exploration PC2 was unaffected by age, enclosure type, SVL or tail status. Heritability for exploration PC2 scores was not different from zero either ($h^2 = 0.057$, CI = [0; 0.178]).

Spatial + reversal learning – within-year performance

Full results of the GLMMs on learning performance over time are given in Table A4 – A5 but summarized here. During the spatial learning task, lizards significantly decreased the number of errors they made over consecutive trials (-0.027 ± 0.008, $\chi^2 = 11.970$, df =1, $p < 0.001$) independent of safe side (safe side * trial: $\chi^2 = 2.063$, df =1, $p = 0.151$) and consistent across age groups (age * trial: $\chi^2 = 3.846$, df =3, $p = 0.279$) (Figure 4). Nevertheless, a side-bias was observed ($\chi^2 = 307.027$,df =1, $p < 0.001$) with lizards committing more errors when the safe refuge was on the right side of the arena (left: 0.230 ± 0.018, right: 1.253 ± 0.039, $z = -17.522$, $p < 0.001$). In addition, the age groups also differed in their overall number of errors ($\chi^2 = 28.202$, df =3, $p < 0.001$). Juveniles made fewer errors (0.519 ± 0.040) than 2019 adults (0.824 ± 0.041, $z = 5.204$, $p < 0.001$), 2020 adults (0.770 ± 0.050, $z = 3.615$, $p = 0.002$) and intruders (0.812 ± 0.061, $z = 3.497$, $p = 0.003$) (Table A5)

During the reversal phase, learning curves differed among age groups (age * trial: $\chi^2 = 10.387$, df =3, $p = 0.016$). Within each group, the number of errors decreased significantly with trial number (adults' 19: -0.027 ± 0.012, $z = -2.606$, $p = 0.009$; adults '20: -0.073 ± 0.014, $z = -2.767$, $p = 0.029$; intruders: -0.036 ± 0.020, $z = -2.319$, $p = 0.020$; juveniles: -0.070 ± 0.017, $z = -4.071$, $p < 0.001$; Figure 4), but adults learnt faster (steeper slope) in 2020 than 2019 ($z = -2.767$, $p = 0.029$) and juveniles tended to learn faster than their parents in 2019 ($z = 2.325$, $p = 0.092$) (Table A5). Similarly, lizards improved over time independent of whether the safe refuge was left or right in the arena, but did so faster in case of the former (left: -0.069 + 0.015, $z = -4.931$, $p < 0.001$; right: - 0.034 ± 0.009, $z = 4.007$, $p < 0.001$; safe side * trial: $\chi^2 = 4.577$, df =1, $p = 0.032$).

Spatial + reversal learning – repeatability

Adults did not differ in SL scores between both years ($F_{1,1} = 3.331$, $p = 0.317$), independent of either original habitat and/or enclosure (neither included in final model, Table 3). We did find evidence for a side bias in cognitive performance ($F_{1,37} = 106.93$, $p < 0.001$) with lizards making more errors if the safe refuge was on the right side of the arena (left: 0.305 ± 0.463, right: 1.376 ± 0.083, $t = 10.806$, $p < 0.001$), and there was a trend for lizards with an intact tail to make more errors (intact: 0.899 ± 0.084, damaged: 0.667 ± 0.168, $F_{1,63} = 3.602$, $p = 0.062$). Lizards showed relatively modest consistency in their spatial learning performance across years, even when adjusting for this side bias ($R_{adj} = 0.398$, CI = [0.124; 0.622], LRT: $p = 0.004$; Table 3).

In contrast, RL scores differed between years depending on the enclosure in which lizards were kept (enclosure*year: $F_{1,39} = 7.924$, $p = 0.008$). Nevertheless, post-hoc pairwise comparisons did not reveal any significant differences (all $p > 0.10$, see Table A5). Visual inspection of the data suggested that lizards kept in simple enclosures made more reversal errors in 2020 compared to 2019 (75 ± 34 % increase), which was less prominent in the complex enclosures (12 ± 17 % increase) (Figure 5). Once again, lizards made fewer errors if the safe refuge was on the left side of the arena (left: 0.292 ± 0.032, right: 1.338 ± 0.074, $F_{1,34} = 195.677$, $p < 0.001$). There was no overall effect of original habitat nor of sex, SVL or tail status, as these were either not included in the final model or non-significant (Table 3). Reversal learning, corrected for side bias and the enclosure*year interaction, showed moderate long-term repeatability, although this was only marginally significant ($R_{adj} = 0.251$, CI = [0.000; 0.545], LRT: $p = 0.061$) (Table 3).

None of the variables or their interactions explained variation in FLEX scores, as the null model was the best model (Table 3). Long-term repeatability of learning flexibility was low and not significant ($R = 0.192$, CI = [0.000; 0.460], LRT: $p = 0.105$). (Table 3).

Spatial + reversal learning – heritability

Variation in SL scores was explained by neither age, nor enclosure type, nor SVL as none of these variables had sufficient high importance (all $R < 0.50$) to be included in the final (Bayesian) animal model (Table 4). Tail status was included in the final model but did not affect SL score (0.346, CI = [-0.026; 0.780]). The heritability of spatial learning performance did not differ from zero ($h^2 = 0.054$, CI = [0; 0.175]).

The final animal model for RL scores included the enclosure type * age interaction. A series of post-hoc pairwise comparisons (see Table A3) revealed that: juveniles from simple enclosures made fewer errors than their parents (adults) in either 2019 or 2020 (adults$_{simple}$ '19 – juveniles$_{simple}$: 0.692 [0.014; 1.449]; adults '20$_{simple}$ – juveniles$_{simple}$: 1.187 [0.355; 1.946]). Juveniles from complex enclosures likewise performed better than their parents in either year (adults' 19$_{complex}$ – juveniles$_{complex}$: 0.961 [0.266; 1.698]; adults$_{complex}$ '20 – juveniles$_{complex}$: 0.833 [0.095; 1.530]) but also than the adults from the simple enclosures in 2020 (adults$_{simple}$ '20 – juveniles$_{complex}$: 1.271 [0.264; 2.437]) (Figure 6). Curiously, RL scores from juveniles from simple and complex enclosures did not differ from each other, and neither from the RL scores of 2019 adults in the opposite enclosure type (Table 4, Table A3; Figure 6). Heritability of reversal learning was also not different from zero (h^2 =0.074, CI = [0.000; 0.249]).

Learning flexibility was not predicted by any of the aforementioned variables (Table 4), and did not show significant heritability (h^2 = 0.053, CI = [0; 0.167]).

Escape box – repeatability

Most lizards succeeded in escaping from the box (2019: 34/41, 2020: 38/41). Neither year, original habitat, enclosure, SVL, sex or any of their interactions was included in the final model, and thus did not explain variation in escape times among individuals. Overall, long-term consistency of escape time was non-existent (R_{adj} = 0.000, CI = [0.000; 0.307], LRT: p = 1) (Table 3).

DISCUSSION

In the last few years, a growing number of studies has focused on interindividual variation in cognition. Despite this interest, information on the long-term consistency of such individual differences, as well as on their heritability, is still lacking. Here, we report moderate repeatability in explorative behaviour (PC1) and spatial learning in Aegean wall lizards kept in semi-natural conditions for one year (20 % of their average lifespan). In contrast, reversal learning was only marginally repeatable, and showed habitat-dependent plasticity. Problem-solving and learning flexibility were not repeatable. Last, heritability estimates were not different from zero for any of the traits.

Exploration

Our lizards displayed repeatable individual variation in exploration PC1 (general exploratory behaviour) across years ($R_{adj} = 0.280$). This result corroborates previous findings that personality variation can be consistent over long and considerable portions of a species life (eastern box turtles: Carlson et al., 2020; roe deer: Debeffe et al., 2015; sleepy lizards: Payne et al., 2021; zebra fishes: Thomson et al., 2020; European starlings: Thys et al., 2017a; zebra finches: Wuerz & Krüger, 2014), although not always (collared flycatchers: Garamszegi et al., 2015). In contrast, exploration PC1 did not show significant heritability ($h^2 = 0.031$). Explorative behaviour is generally found to be moderately heritable (Dochtermann et al., 2019) albeit this varies greatly among studies ($R^2 = 0.019 - 0.25$ in European green lizards: Bajer et al., 2015; $h^2 = 0.22 - 0.37$ in great tits: Dingemanse et al., 2002; $h^2 = 0.355 - 0.362$ in yellow-bellied marmots: Petelle et al., 2015; $h^2 = 0.08$ in red squirrels: Taylor et al., 2012).

Thus, the consistent individual variation in exploration PC1 could not be explained by additive genetic differences among individuals. We should, however, take into account that our sample size (37 juveniles, 16 fathers, 19 mothers) was relatively small compared to former studies on heritability (median N = 336, range = 6 - 11 092, only 14 % with $N < 100$ in the meta-analysis of Dochtermann et al., 2019) Hence, it is not impossible that additive genetic variance does contribute to behavioural variation in *P. erhardii*, but we were simply unable to detect it (Blanckenhorn & Perner, 1994). Nonetheless, the low genetic variation in our lizards could also be due to going through a genetic bottleneck when introduced in our enclosures (Carrete et al., 2017) or could be a consequence of strong directional selection on explorative behaviour in the past (Boake, 1989; Falconer & Mackay, 1996; Wheelwright, Keller, & Postma, 2014). Large seasonal fluctuations in precipitation and accordingly arthropod abundances on Naxos (Adamopoulou et al., 1999; De Meester et al., 2021; Karamaouna, 1987; Parashi, 1988) may exert strong selection on explorative behaviour within Aegean wall lizards if it facilitates the discovery and acquisition resource acquisition (Bajer et al., 2015; Baxter-Gilbert et al., 2019). However, we did observe negative selection on exploration in a previous batch of lizards from 2018 to 2019, but not in the current batch (De Meester et al., submitted). Ideally, we should thus verify the heritability of personality (and cognition) in completely natural populations.

Regardless of the reasons, low heritability (if accurate) but moderate repeatability does imply that personality variation in *P. erhardii* mostly arises due to strong environmental effects (Petelle et al., 2015; Quinn et al., 2016; Vardi et al., 2020). This is further supported by the extremely low short-term repeatability of exploration PC1 within juveniles (R = 0.005). In hindsight, juveniles were captured and transferred to captive lab conditions too soon after hatching (as indicated by the presence of umbilical scars) and thus effectively grew up in the same standardized environment. A lack of genetic differences plus little divergence in personal experiences may explain their low behavioural repeatability (Archard & Braithwaite, 2010; Stamps & Groothuis, 2010). Short-term repeatability is slightly higher (but not significant) in 2019 adults (R_{adj} = 0.079 - 0.085) and moderate in 2020 adults (R_{adj} = 0.333), giving additional support for the hypothesis that personality variation develops over time. Behavioural repeatability is often predicted to change with age, although in which direction is highly debated (Carlson et al., 2020). Both within- and among- individual variance in a population can increase or decrease over time due to a multitude of processes (overview in Carlson et al., 2020), including selection (Bell et al., 2009), divergence in personal experiences (Stamps & Groothuis, 2010), state-behaviour feedback loops (Kok et al., 2019; Sih et al., 2015), canalization (Kok et al., 2019), changes in the costs of behavioural flexibility (Polverino et al., 2016) or in the developmental dynamics of the physiological mechanisms underlying behaviour (Bell et al., 2009; Stamps & Groothuis, 2010). Such changes are not necessarily monotonic over time (Thys et al., 2021). A valuable follow-up experiment would be to measure personality multiple times across the lifetime of the same cohort of lizards starting from birth, to test more explicitly how and when personality variation develops in this species.

Following up on this, we did find evidence for changes in (average) explorative behaviour with age. Adult lizards behaved more explorative in 2019 than 2020, which could simply reflect senescence (Brommer & Class, 2015). In addition, juveniles had lower PC1 scores than adults, which is in line with the idea that younger animals should behave more cautious to allow future reproduction, while adults should take more risks to increase current reproduction (Schuster et al., 2017a; Wolf et al., 2007). Nevertheless, we should note that all lizards tested in 2020 (intruders included) behaved less explorative than the 2019 adults. Lizards were tested in May and August during 2019 and 2020 respectively, thus seasonal fluctuations in behaviour (Aragón, López, & Martín, 2001; Jenssen, Greenberg, & Hovde, 1995; Kerr & Bull, 2006) may explain the differences between years. Indeed, Naxian *P. erhardii* become less active as ambient temperatures rise during summer (Catsadorakis, 1984). Alternatively, restricted space use and physical activity within the enclosures compared to a natural environment could also have led to a plastic decrease in explorative behaviour over time in every group (Oosthuizen et al., 2013).

In sharp contrast, individual differences in exploration PC2 (fast exploration versus hiding) were not consistent across years (R = 0.060), nor did they show significant heritability ($h^2 = 0.057$). Interestingly, exploration PC2 showed considerable short-term repeatability within the complete dataset of 2019 adults ($R_{adj} = 0.211$), but not within the subset of survivors in either 2019 ($R_{adj} = 0.130$) or 2020 ($R_{adj} = 0.162$). Lower repeatability among survivors may be a consequence of strong directional selection on exploration PC2 (Bell et al., 2009). Indeed, female PC2 scores were higher in the survivors than in the complete batch of released adults (survivors: 0.339 ± 0.167; all: 0.212 ± 0.142) while the opposite occurred in males (survivors: -0.031 ± 0.180; all: 0.165 ± 0.160). Interestingly, male and female survivors differed in PC2-scores in 2020 but not 2019. This implies that sex-dependent plasticity also occurred across years. Male and female lizard can indeed differ in how their behaviour changes over the breeding season (Aragón et al., 2001). Sex-dependent selection and plasticity would have respectively decreased inter-individual and increased within-individual variance (Carlson et al., 2020), and thus both contributed to overall lower behavioural repeatability of PC2 on the long-term.

Cognition

Adult lizards showed moderate repeatability in spatial learning performance across years ($R_{adj} = 0.398$). Our study hence adds to a small body of evidence that individual variation in spatial learning abilities can be repeatable over longer timescales (pheasants: Langley et al., 2018; Eurasian harvest mice: Schuster et al., 2017b; mountain chickadees: Tello-Ramos et al., 2018; but see Soha et al., 2019 on song sparrows). To our best knowledge, this is the first study demonstrating cognitive repeatability (either short- or long-term) in a non-avian reptile. On the other hand, heritability for spatial learning was not different from zero ($h^2 = 0.054$). While heritability estimates for spatial learning vary greatly across literature ($h^2 = 0.27$ in chimpanzees: Hopkins, Russell, & Schaeffer, 2014; $h^2 = 0.09 - 0.23$ in pheasants:

Langley et al., 2020; $h^2 = 0.27$ in rose bitterlings: Smith et al., 2015), our results are in line with the only other study investigating heritability of (spatial) cognition in lizards (no significant mother-offspring regression in delicate skinks: Vardi et al., 2020).

Whether his low heritability is a consequence of directional selection, founder effects or too low sample sizes can unfortunately not be verified with our current dataset. It would not be unreasonable to expect selection for spatial learning in *P. erhardii*, as it may contribute to successfully escaping predators (Font, 2019) and remembering the location of resources during periods of food scarcity (De Meester et al., 2021). We did indeed observe selection on spatial learning in our enclosures, although in the opposite direction and only in females (De Meester et al., submitted).

We previously reported differences in spatial learning performance between lizards originating from structural simple and complex habitats (De Meester et al., 2022). Assuming that spatial learning is not heritable, then such intraspecific variation would be entirely due to plasticity (Morand-Ferron et al., 2016), as also hypothesized for the lizards in Vardi et al. (2020). Indeed, being reared in structural complex environments has a positive effect on brain (size) and cognitive development in fish and lizards (Carbia & Brown, 2019; LaDage et al., 2016; Spence et al., 2011; Vardi et al., 2020). Our juveniles made fewer errors during the spatial learning compared to adults, which indicates that spatial cognition may indeed be plastic in *P. erhardii*. Higher learning abilities in juveniles could be a consequence of the higher need for behavioural plasticity in early life (Fischer et al., 2014; Szabo et al., 2019) or of cognitive decline with age (Bonte, Kemp, & Fagot, 2014). Alternatively, juvenile lizards should be more motivated to find the safe refuge due to an higher vulnerability to predation (Martín & López, 1995). Interestingly, in contrast to reversal learning, spatial learning performance did not show habitat-dependent plastic changes across years. This implies that if variation in spatial learning is caused by permanent environmental effects, such effects may be limited to a critical period during early life. It could thus be an interesting follow-up experiment to test the cognitive performance of newly born lizards, release them in our enclosures, and follow up their cognitive development in both habitat types.

Demonstrating that individual variation in learning is repeatable validates that we are truly measuring cognitive variation (Ashton et al., 2018; Thornton et al., 2014) and helps us to understand its ecological and evolutionary relevance (Boake, 1989; Morand-Ferron et al., 2016; Soha et al., 2019). Nevertheless, we should be aware of the possibility of pseudo-repeatability (Cooke et al., 2021; Dingemanse & Dochtermann, 2013; Mason et al., 2021), i.e. behavioural repeatability could be a consequence of consistent differences in other non-cognitive variables among individuals. For example, Cooke et al. (2021) demonstrated that problem-solving performance in great tits (*Parus major*) was highly repeatable, until corrected for hunger motivation and experience. Nonetheless, the long time-interval between repeated tests should have drastically reduced the chances of pseudo-repeatability (Niemelä & Dingemanse, 2017). Spatial learning was also unaffected by lizard personality in this dataset (De Meester et al., 2022), and tail status was corrected for, thus it is also unlikely that individuals simply differed consistently in their willingness to seek shelter. Biases for certain stimuli, such as a colour (Mason et al., 2021) or a side preference (our results) could also increase repeatability estimates if test subjects differ consistently in whether they are trained to pick the preferred or unpreferred cue. However, learning performance remained significantly repeatable even after adjusting for the side bias of our lizards. Lastly, behavioural repeatability could as well be influenced by memories from a previous testing round (Griffin et al., 2015). Yet, if lizards remembered the location of the safe refuge from the previous year, they should have made fewer errors or learnt faster in 2020, which was not the case. Nonetheless, it would be good to test the contextual repeatability of spatial learning in *P. erhardii* as well. Using various tasks aimed at measuring the same cognitive ability, e.g. training lizards to locate food or mates instead of shelter, or testing spatial learning at different scales, may help to minimize the influence of pseudo-repeatability and memory (Brust & Guenther, 2017; Cauchoix et al., 2018; Griffin et al., 2015; Troisi et al., 2021).

Next, we found that reversal learning was only marginally repeatable (R_{adj} = 0.251) and learning flexibility not at all (R = 0.192), and that neither showed significant heritability (h^2_{RL} = 0.074, h^2_{FLEX} = 0.053). Previous studies reported reversal learning to be both repeatable (song sparrows: Soha et al., 2019) and not repeatable (mountain chickadees: Tello-Ramos et al., 2018) , while overall being modestly heritable (R^2 = 0.31 among 51 strains of lab mice: Laughlin et al., 2011; h^2 = 0.26 in red junglefowl: Sorato et al., 2018). The low repeatability of reversal learning and learning flexibility is in sharp contrast with the rather high repeatability of spatial learning. A similar result was obtained for wild mountain chickadees by Tello-Ramos et al. (2018). One possible explanation may be that cognitive flexibility is more plastic and sensitive to environmental changes (Tello-Ramos et al., 2018). Indeed, lizards kept in simple enclosures seemingly made more errors during the reversal in 2020 than in 2019. If individuals within a group change their behaviour inconsistently from each other, due to differential personal experiences, then behavioural repeatability is indeed expected to decrease (Brommer & Class, 2015).

Changes in reversal learning performance may be a consequence of deviations in neurogenesis rates, a process known to be stimulated by spatial complexity and impaired by structural simplicity, even in adults (Dunlap, 2016; Kempermann, Kuhn, & Gaga, 1997; LaDage et al., 2013). Neurogenesis facilitates reversal learning but importantly, appears to be less relevant for the initial acquisition of information (Burghardt et al., 2012; Kalm et al., 2013; Swan et al., 2014). In addition, stress is known to down-regulate neurogenesis (Mirescu & Gould, 2006). Lizards in the simple open enclosures may have experienced more stress, due to e.g. feeling more exposed to aerial predators or more intense competition for the fewer resources. Thus, stress and habitat simplicity may have inhibited the rate of neurogenesis, leading to reduced reversal learning in lizards kept in simple enclosures. Importantly, the fact that changes in neurogenesis are not expected to influence the capacity to learn an initial (spatial) association may explain why habitat complexity did not lead to differential changes in spatial learning performance.

The rate of neurogenesis is also often believed to decline with age (Amrein et al., 2004; Molowny, Nacher, & López-García, 1995), which possibly explains why juvenile lizards showed better reversal learning than adults. Yet, strangely enough, juveniles only outperformed the adults in their own enclosure, but did not differ from adults in the opposite enclosure type (with the exception of juveniles from complex enclosures making fewer errors than 2020 adults in simple enclosures). Why these age-differences seem habitat-specific is unclear to us, especially given that 2019 adults were tested prior to release into the enclosures.

Finally, problem-solving ability, here measured with an escape box task, showed the lowest repeatability (R = 0) of all cognitive traits. Long term consistency of problem-solving is very rarely tested, and previous studies have demonstrated both low (R = 0.002 - 0.02 in North Island robins, Shaw, 2017) and relative high temporal repeatability (R = 0.27 - 0.54 in great tits: Cole et al., 2011). Cauchoix et al. (2018) found that (contextual) repeatability of cognition was significantly lower for latency measures, such as our escape times, likely due to ceiling or floor effects. Among-individual variation may be lowered because all failing individuals were assigned the same maximum score, or because most lizards solved the task within a comparable short time due to its apparent ease.

Problem-solving assays have been criticized, as it is often unclear whether individual variation in performance truly reflects cognitive variation or is due to non-cognitive differences (e.g. hunger, motivation, …) among test subjects (Audet & Lefebvre, 2017; Morand-Ferron et al., 2016; Shaw, 2017). Especially when only measured once, the outcome of a problem-solving task can be strongly influenced by intrinsic and extrinsic conditions (Cauchoix et al., 2018). The fact that escape times were not repeatable in our study seems to validate such concerns, and illustrates the danger of linking performance in a (single) problem-solving task to e.g. personality, life-history or fitness without any information regarding its repeatability. Following the suggestion of Thornton et al. (2014), problem-solving should have been tested over multiple trials within each year, and measure the repeatability of lizards' improvement or the mean solving time (Cauchoix et al., 2018).

CONCLUSION

Very few studies so far have tested the long-term consistency and heritability of personality and cognition, especially so for wild animals, despite the fact that this information is crucial to understand the potential evolutionary and ecological impact of such behavioural variation (Boogert et al., 2018; Cauchoix & Chaine, 2016; Dukas, 2004; Griffin et al., 2015; Morand-Ferron et al., 2016). Our study showed that individual differences in some, but not all, aspects of exploration and cognitive performance were consistent in semi-wild Aegean wall lizards across years, but neither cognition nor personality were heritable.

The low heritability estimates would imply that all of our behavioural traits have very little evolutionary potential, even if selection would act upon them, although this needs to be verified in natural populations. Our results do suggest that both cognition and personality within Aegean wall lizards are at least partially plastic, changing with age, depending on both sex and habitat complexity. Our study thus illustrates that long-term studies on the repeatability of cognition in wild animals can advance our understanding of the role of both genetic and environmental factors in shaping cognitive variation.

REFERENCES

Adamopoulou, C., Valakos, E. D., & Pafilis, P. (1999). Summer diet of *Podarcis milensis, P. gaigeae* and *P. erhardii* (Sauria: Lacertidae). *Bonner Zoologische Beitrage, 48*, 275-282.

Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Current Anthropology, 36*, 199-221. doi: 10.1086/204350

Alvarez-Quintero, N., Velando, A., & Kim, S. Y. (2021). Smart mating: the cognitive ability of females influences their preference for male cognitive ability. *Behavioral Ecology, 32*(5), 803-813. doi: 10.1093/beheco/arab052

Amrein, I., Slomianka, L., Poletaeva, II, Bologova, N. V., & Lipp, H. P. (2004). Marked species and age-dependent differences in cell proliferation and neurogenesis in the hippocampus of wild-living rodents. *Hippocampus, 14*(8), 1000-1010. doi: 10.1002/hipo.20018

Aragón, P., López, P., & Martín, J. (2001). Seasonal changes in activity and spatial and social relationships of the Iberian rock lizard, *Lacerta monticola*. *Canadian Journal of Zoology, 79*(11), 1965-1971. doi: 10.1139/cjz-79-11-1965

Archard, G. A., & Braithwaite, V. A. (2010). The importance of wild populations in studies of animal temperament. *Journal of Zoology, 281*, 149-160. doi: 10.1111/j.1469-7998.2010.00714.x

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature, 554*(7692), 364-367. doi: 10.1038/nature25503

Audet, J.-N., & Lefebvre, L. (2017). What's flexible in behavioral flexibility? *Behavioral Ecology, 28*(4), 943-947. doi: 10.1093/beheco/arx007

Bajer, K., Horvath, G., Molnar, O., Torok, J., Garamszegi, L. Z., & Herczeg, G. (2015). European green lizard (*Lacerta viridis*) personalities: Linking behavioural types to ecologically relevant traits at different ontogenetic stages. *Behavioural Processes, 111*, 67-74. doi: 10.1016/j.beproc.2014.11.020

Barton, K. (2013). MuMIn: multi-model inference. R Package version 1.43.17 from http://CRAN.R-project.org/package=MuMIn

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*, 1-48. doi: 10.18637/jss.v067.i01

Baxter-Gilbert, J., Riley, J. L., & Whiting, M. J. (2019). Bold New World: urbanization promotes an innate behavioral trait in a lizard. *Behavioral Ecology and Sociobiology, 73*(8). doi: 10.1007/s00265-019-2713-9

Bebus, S. E., Small, T. W., Jones, B. C., Elderbrock, E. K., & Schoech, S. J. (2016). Associative learning is inversely related to reversal learning and varies with nestling corticosterone exposure. *Animal Behaviour, 111*, 251-260. doi: 10.1016/j.anbehav.2015.10.027

Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: a meta-analysis. *Animal Behaviour, 77*(4), 771-783. doi: 10.1016/j.anbehav.2008.12.022

Blanckenhorn, W. U., & Perner, D. (1994). Heritability and repeatability of behavioural attributes aflecting foraging success and fitness in water striders. *Animal Behaviour, 48*, 169-176.

Boake, C. R. B. (1989). Repeatability - its role in evolutionary studies of mating behavior. *Evolutionary Ecology, 3*, 173-182.

Bonte, E., Kemp, C., & Fagot, J. (2014). Age effects on transfer index performance and executive control in baboons (*Papio papio*). *Frontiers in Psychology, 5*, 188. doi: 10.3389/fpsyg.2014.00188

Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1756). doi: 10.1098/rstb.2017.0280

Boon, A. K., Reale, D., & Boutin, S. (2007). The interaction between personality, offspring fitness and food abundance in North American red squirrels. *Ecology Letters, 10*(11), 1094-1104. doi: 10.1111/j.1461-0248.2007.01106.x

Boussard, A., Buechel, S. D., Amcoff, M., Kotrschal, A., & Kolm, N. (2020). Brain size does not predict learning strategies in a serial reversal learning test. *Journal of Experimental Biology, 223*(Pt 15). doi: 10.1242/jeb.224741

Brock, K., Donihue, C. M., & Pafilis, P. (2014). New records of frugivory and ovophagy in *Podarcis* (Lacertidae) lizards from East Mediterranean Islands. *North-Western Journal of Zoology, 10*(1), 223-225.

Brommer, J. E., & Class, B. (2015). The importance of genotype-by-age interactions for the development of repeatable behavior and correlated behaviors over lifetime. *Frontiers in Zoology, 12*(Suppl 1), S2. doi: 10.1186/1742-9994-12-S1-S2

Brust, V., & Guenther, A. (2017). Stability of the guinea pigs personality - cognition - linkage over time. *Behavioural Processes, 134*, 4-11. doi: 10.1016/j.beproc.2016.06.009

Buechel, S. D., Boussard, A., Kotrschal, A., van der Bijl, W., & Kolm, N. (2018). Brain size affects performance in a reversal-learning test. *Proceedings of the Royal Society B: Biological Sciences, 285*(1871). doi: 10.1098/rspb.2017.2031

Burghardt, N. S., Park, E. H., Hen, R., & Fenton, A. A. (2012). Adult-born hippocampal neurons promote cognitive flexibility in mice. *Hippocampus, 22*(9), 1795-1808. doi: 10.1002/hipo.22013

Carazo, P., Noble, D. W., Chandrasoma, D., & Whiting, M. J. (2014). Sex and boldness explain individual differences in spatial learning in a lizard. *Proceedings of the Royal Society B: Biological Sciences, 281*(1782), 20133275. doi: 10.1098/rspb.2013.3275

Carbia, P. S., & Brown, C. (2019). Environmental enrichment influences spatial learning ability in captive-reared intertidal gobies (*Bathygobius cocosensis*). *Animal Cognition, 22*(1), 89-98. doi: 10.1007/s10071-018-1225-8

Carlson, B. E., Tetzlaff, S. J., & Rutz, C. (2020). Long-term behavioral repeatability in wild adult and captive juvenile turtles (*Terrapene carolina*): Implications for personality development. *Ethology, 126*(6), 668-678. doi: 10.1111/eth.13024

Carrete, M., Centeno-Cuadros, A., Méndez, M., Agudo, R., & Donázar, J. A. (2017). Low heritability in tool use skills in a wild vulture population. *Animal Behaviour, 129*, 127-131. doi: 10.1016/j.anbehav.2017.05.015

Catsadorakis, G. (1984). Aspects of ecology and behaviour in the lizard *Podarcis erhardii*. *Biologia Gallo-Hellenica, 11*, 99-110.

Cauchoix, M., & Chaine, A. S. (2016). How can we study the evolution of animal minds? *Frontiers in Psychology, 7*, 358. doi: 10.3389/fpsyg.2016.00358

Cauchoix, M., Chow, P. K. Y., van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason, G., . . . Morand-Ferron, J. (2018). The repeatability of cognitive performance: a meta-analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1756). doi: 10.1098/rstb.2017.0281

Colby, A. E., Kimock, C. M., & Higham, J. P. (2021). Endocranial volume is variable and heritable, but not related to fitness, in a free-ranging primate. *Scientfic Reports, 11*(1), 4235. doi: 10.1038/s41598-021-81265-w

Cole, E. F., Cram, D. L., & Quinn, J. L. (2011). Individual variation in spontaneous problem-solving performance among wild great tits. *Animal Behaviour, 81*(2), 491-498. doi: 10.1016/j.anbehav.2010.11.025

Cole, E. F., Morand-Ferron, J., Hinks, A. E., & Quinn, J. L. (2012). Cognitive ability influences reproductive life history variation in the wild. *Current Biology, 22*(19), 1808-1812. doi: 10.1016/j.cub.2012.07.051

Cooke, A. C., Davidson, G. L., van Oers, K., & Quinn, J. L. (2021). Motivation, accuracy and positive feedback through experience explain innovative problem solving and its repeatability. *Animal Behaviour, 174*, 249-261. doi: 10.1016/j.anbehav.2021.01.024

Cooper, T., Liew, A., Andrle, G., Cafritz, E., Dallas, H., Niesen, T., . . . Mendelson, J. (2019). Latency in problem solving as evidence for learning in varanid and helodermatid lizards, with comments on foraging techniques. *Copeia, 107*(1). doi: 10.1643/ch-18-119

Croston, R., Branch, C. L., Kozlovsky, D. Y., Dukas, R., & Pravosudov, V. V. (2015). Heritability and the evolution of cognitive traits. *Behavioral Ecology, 26*(6), 1447-1459. doi: 10.1093/beheco/arv088

Damas-Moreira, I., Riley, J. L., Harris, D. J., & Whiting, M. J. (2019). Can behaviour explain invasion success? A comparison between sympatric invasive and native lizards. *Animal Behaviour, 151*, 195-202. doi: 10.1016/j.anbehav.2019.03.008

Dammhahn, M. (2012). Are personality differences in a small iteroparous mammal maintained by a life-history trade-off? *Proceedings of the Royal Society B: Biological Sciences, 279*(1738), 2645-2651. doi: 10.1098/rspb.2012.0212

De Meester, G., Pafilis, P., & Van Damme, R. (2022). Bold and bright: shy and supple? The effect of habitat type on personality-cognition covariance in the Aegean wall lizard (*Podarcis erhardii*). *Animal Cognition*. doi: 10.1007/s10071-021-01587-0

De Meester, G., Sfendouraki-Basakarou, A., Pafilis, P., & Van Damme, R. (2021). Dealing with the unexpected: the effect of environmental variability on behavioural flexibility in a Mediterranean lizard. *Behaviour*, 1-31. doi: 10.1163/1568539X-bja10088

de Villemeuril, P. (2012). Estimation of a biological trait heritability using the animal model Retrieved 28-06-2021, from http://devillemereuil.legtux.org/wp-content/uploads/2012/12/tuto_en.pdf

Debeffe, L., Lemaître, J. F., Bergvall, U. A., Hewison, A. J. M., Gaillard, J. M., Morellet, N., . . . Kjellander, P. (2015). Short- and long-term repeatability of docility in the roe deer: sex and age matter. *Animal Behaviour, 109*, 53-63. doi: 10.1016/j.anbehav.2015.08.003

Deimezis-Tsikoutas, A., Kapsalas, G., & Pafilis, P. (2020). A rare case of saurophagy by *Scolopendra cingulata* (Chilopoda: Scolopendridae) in the central Aegean Archipelago: a role for insularity? *Zoology and Ecology, 30*(30), 48-51. doi: 10.35513/21658005.2020.1.6

Dingemanse, N. J., Both, C., Drent, P. J., van Oers, K., & van Noordwijk, A. J. (2002). Repeatability and heritability of exploratory behaviour in great tits from the wild. *Animal Behaviour, 64*(6), 929-938. doi: 10.1006/anbe.2002.2006

Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in behaviour: mixed-effect modelling approaches. *Journal of Animal Ecology, 82*(1), 39-54. doi: 10.1111/1365-2656.12013

Dochtermann, N. A., Schwab, T., Anderson Berdal, M., Dalos, J., & Royaute, R. (2019). The heritability of behavior: A meta-analysis. *Journal of Heredity, 110*(4), 403-410. doi: 10.1093/jhered/esz023

Dohm, M. R. (2002). Repeatability estimates do not always set an upper limit to heritability. *Functional Ecology, 16*, 273-280.

Dougherty, L. R., & Guillette, L. M. (2018). Linking personality and cognition: a meta-analysis. *Philosophical Transactions of the Royal Society B Biological Sciences, 373*(1756). doi: 10.1098/rstb.2017.0282

Dukas, R. (2004). Evolutionary biology of animal cognition. *Annual Review of Ecology, Evolution, and Systematics, 35*, 347-374. doi: 10.2307/annurev.ecolsys.35.112202.30000014

Dunlap, K. D. (2016). Fish neurogenesis in context: Assessing environmental influences on brain plasticity within a highly labile physiology and morphology. *Brain Behavior and Evolution, 87*(3), 156-166. doi: 10.1159/000446907

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. New York: Longman.

Fischer, B., van Doorn, G. S., Dieckmann, U., & Taborsky, B. (2014). The evolution of age-dependent plasticity. *The American Naturalist, 183*(1), 108-125. doi: 10.1086/674008

Font, E. (2019). Rapid learning of a spatial memory task in a lacertid lizard (*Podarcis liolepis*). *Behavioural Processes, 169*, 103963. doi: 10.1016/j.beproc.2019.103963

Fox, J., & Weisberg, S. (2019). *An {R} Companion to Applied Regression*. Thousand Oaks, CA.: Sage.

Garamszegi, L. Z., Marko, G., Szasz, E., Zsebok, S., Azcarate, M., Herczeg, G., & Torok, J. (2015). Among-year variation in the repeatability, within- and between-individual, and phenotypic correlations of behaviors in a natural population. *Behavioral Ecology and Sociobiology, 69*(12), 2005-2017. doi: 10.1007/s00265-015-2012-z

Gardner, K. M., Mennill, D. J., Newman, A. E. M., & Doucet, S. M. (2020). Social and physiological drivers of rapid colour change in a tropical toad. *General and Comparative Endocrinology, 285*, 113292. doi: 10.1016/j.ygcen.2019.113292

Gibelli, J., & Dubois, F. (2017). Does personality affect the ability of individuals to track and respond to changing conditions? *Behavioral Ecology, 28*(1), 101-107. doi: 10.1093/beheco/arw137

Gnanadesikan, G. E., Hare, B., Snyder-Mackler, N., Call, J., Kaminski, J., Miklosi, A., & MacLean, E. L. (2020). Breed differences in dog cognition associated with brain-expressed genes and neurological functions. *Integrative and Comparative Biology, 60*(4), 976-990. doi: 10.1093/icb/icaa112

Gomes, A. C. R., Guerra, S., Silva, P. A., Marques, C. I., Trigo, S., Boogert, N. J., & Cardoso, G. C. (2020). Proactive common waxbills make fewer mistakes in a cognitive assay, the detour-reaching task. *Behavioral Ecology and Sociobiology, 74*, 1-15. doi: 10.1007/s00265-020-2809-2

Griffin, A. S., & Guez, D. (2014). Innovation and problem solving: a review of common mechanisms. *Behavioural Processes, 109 Pt B*, 121-134. doi: 10.1016/j.beproc.2014.08.027

Griffin, A. S., Guillette, L. M., & Healy, S. D. (2015). Cognition and personality: an analysis of an emerging field. *Trends in Ecology & Evolution, 30*(4), 207-214. doi: 10.1016/j.tree.2015.01.012

Guillette, L. M., Reddon, A. R., Hurd, P. L., & Sturdy, C. B. (2009). Exploration of a novel space is associated with individual differences in learning speed in black-capped chickadees, *Poecile atricapillus*. *Behavioural Processes, 82*(3), 265-270. doi: 10.1016/j.beproc.2009.07.005

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software, 33*, 1-22. doi: 10.18637/jss.v033.i02

Heathcote, R. J. P., Bell, E., d'Ettorre, P., While, G. M., & Uller, T. (2014). The scent of sun worhip: basking experience alters scent mark composition in male lizards. *Behavioural Ecology and Sociobiology, 68*, 861-870. doi: 10.1007/s00265-014-1700-4

Hervé, M. (2020). RVAideMemoire: Testing and Plotting Procedures for Biostatistics. R package version 0.9-75.

Hopkins, W. D., Russell, J. L., & Schaeffer, J. (2014). Chimpanzee intelligence is heritable. *Current Biology, 24*(14), 1649-1652. doi: 10.1016/j.cub.2014.05.076

Huyghe, K., Breugelmans, K., Small, M., Tadic, Z., R, V. A. N. D., Vanhooydonck, B., & Backeljau, T. (2009). Characterization of polymorphic microsatellite markers in the Dalmatian wall lizard

Podarcis melisellensis (Squamata: Lacertidae). *Molecular Ecology Resources, 9*(1), 299-301. doi: 10.1111/j.1755-0998.2008.02381.x

Huyghe, K., Small, M., Vanhooydonck, B., Herrel, A., Tadic, Z., Van Damme, R., & Backeljau, T. (2010). Genetic divergence among sympatric colour morphs of the Dalmatian wall lizard (*Podarcis melisellensis*). *Genetica, 138*(3), 387-393. doi: 10.1007/s10709-010-9435-2

Jenssen, T. A., Greenberg, N., & Hovde, K. A. (1995). Behavioral profile of free-ranging male lizards, *Anolis carolinensis,* across breeding and post-breeding seasons. *Herpetological Monographs, 9*, 41-62. doi: 10.2307/1466995

Kaiser, H. F. (1991). Coefficient alpha for a principal component and the Kaiser-Guttman rule. *Psychological Reports, 68*, 855-858. doi: 10.2466/pr0.1991.68.3.855

Kalinowski, S., Taper, M., & Marshall, T. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology, 16*, 1099-1106. doi: 10.1111/j.1365-294X.2007.03089.x

Kalm, M., Karlsson, N., Nilsson, M. K., & Blomgren, K. (2013). Loss of hippocampal neurogenesis, increased novelty-induced activity, decreased home cage activity, and impaired reversal learning one year after irradiation of the young mouse brain. *Experimental Neurology, 247*, 402-409. doi: 10.1016/j.expneurol.2013.01.006

Karamaouna, M. (1987). *Ecology of millipedes in Mediterranean coniferous ecosystems of southern Greece.* PhD-thesis, University of Athens.

Kempermann, G., Kuhn, H. G., & Gaga, F. H. (1997). More hippocampal neurons in adult mice living in an enriched environment. *Nature, 386*, 493-495. doi: 10.1038/386493a0

Kerr, G. D., & Bull, C. M. (2006). Movement patterns in the monogamous sleepy lizard (*Tiliqua rugosa*): effects of gender, drought, time of year and time of day. *Journal of Zoology, 269*(2), 137-147. doi: 10.1111/j.1469-7998.2006.00091.x

Kok, E. M. A., Burant, J. B., Dekinga, A., Manche, P., Saintonge, D., Piersma, T., & Mathot, K. J. (2019). Within-individual canalization contributes to age-related increases in trait repeatability: A longitudinal experiment in red knots. *The American Naturalist, 194*(4), 455-469. doi: 10.1086/704593

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13), 1-26. doi: 10.18637/jss

LaDage, L. D., Maged, R. M., Forney, M. V., Roth, T. C., 2nd, Sinervo, B., & Pravosudov, V. V. (2013). Interaction between territoriality, spatial environment, and hippocampal neurogenesis in male side-blotched lizards. *Behavioral Neuroscience, 127*(4), 555-565. doi: 10.1037/a0032852

LaDage, L. D., Roth, T. C., Sinervo, B., & Pravosudov, V. V. (2016). Environmental experiences influence cortical volume in territorial and nonterritorial side-blotched lizards, Uta stansburiana. *Animal Behaviour, 115*, 11-18. doi: 10.1016/j.anbehav.2016.01.029

Langkilde, T., & Shine, R. (2006). How much stress do researchers inflict on their study animals? A case study using a scincid lizard, *Eulamprus heatwolei*. *Journal of Experimental Biology, 209*(Pt 6), 1035-1043. doi: 10.1242/jeb.02112

Langley, E. J. G., Adams, G., Beardsworth, C. E., Dawson, D. A., Laker, P. R., van Horik, J. O., . . . Madden, J. R. (2020). Heritability and correlations among learning and inhibitory control traits. *Behavioral Ecology, 31*(3), 798-806. doi: 10.1093/beheco/araa029

Langley, E. J. G., van Horik, J. O., Whiteside, M. A., Beardsworth, C. E., & Madden, J. R. (2018). The relationship between social rank and spatial learning in pheasants, *Phasianus colchicus:* cause or consequence? *PeerJ, 6*, e5738. doi: 10.7717/peerj.5738

Laughlin, R. E., Grant, T. L., Williams, R. W., & Jentsch, J. D. (2011). Genetic dissection of behavioral flexibility: reversal learning in mice. *Biological Psychiatry, 69*(11), 1109-1116. doi: 10.1016/j.biopsych.2011.01.014

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means, from https://github.com/rvlenth/emmeans

Liefting, M., Hoedjes, K. M., Le Lann, C., Smid, H. M., & Ellers, J. (2018). Selection for associative learning of color stimuli reveals correlated evolution of this learning ability across multiple stimuli and rewards. *Evolution, 72*, 1449-1459. doi: 10.1111/evo.13498

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software, 6*, 3139. doi: 10.21105/joss.

Lymberakis, P., Pafilis, P., Poulakakis, N., Sotiropoulos, K., & Valakos, E. D. (2018). The Amphibians and Reptiles of the Aegean Sea. In S. Sfenthourakis, P. Pafilis, A. Parmakelis, N. Poulakakis & K. A. Triantis (Eds.), *Biogeography and Biodiversity of the Aegean. In honour of Prof. Moysis Mylonas.* (pp. 169-189). Nicosia, Cyprus: Broken Hill Publishers Ltd.

Madden, I., & Brock, K. M. (2018). An extreme record of cannibalism in *Podarcis erhardii mykonensis* (Reptilia: Lacertidae) from Siros island, Cyclades, Greece. *Herpetology Notes, 11*, 291-292. doi: 10.1111/1365-2435.12550

Madden, J. R., Langley, E. J. G., Whiteside, M. A., Beardsworth, C. E., & van Horik, J. O. (2018). The quick are the dead: pheasants that are slow to reverse a learned association survive for longer in the wild. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1756). doi: 10.1098/rstb.2017.0297

Martín, J., & López, P. (1995). Escape behaviour of juvenile *Psammodromus algirus* lizards: constraint of or compensation for limitations in body size? *Behaviour, 132*, 181-192.

Mason, D., Zajitschek, S., Anwer, H., O'Dea, R. E., Hesselson, D., & Nakagawa, S. (2021). Low repeatability of aversive learning in zebrafish (*Danio rerio*). *Journal of Experimental Biology, 224*, jeb240846. doi: 10.1101/2020.11.16.385930

McEvoy, J., While, G. M., Sinn, D. L., Carver, S., & Wapstra, E. (2015). Behavioural syndromes and structural and temporal consistency of behavioural traits in a social lizard. *Journal of Zoology, 296*(1), 58-66. doi: 10.1111/jzo.12217

Mery, F., & Kawecki, T. J. (2002). Experimental evolution of learning ability in fruit flies. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 14274-14279. doi: 10.1073/pnas.222371199

Michelangeli, M., Melki-Wegner, B., Laskowski, K., Wong, B. B. M., & Chapple, D. G. (2020). Impacts of caudal autotomy on personality. *Animal Behaviour, 162*, 67-78. doi: 10.1016/j.anbehav.2020.02.001

Mirescu, C., & Gould, E. (2006). Stress and adult neurogenesis. *Hippocampus, 16*(3), 233-238. doi: 10.1002/hipo.20155

Molowny, A., Nacher, J., & López-García, C. (1995). Reactive neurogenesis during regeneration of the lesioned medial cerebral cortex of lizards. *Neuroscience, 68*, 823-836.

Morand-Ferron, J. (2017). Why learn? The adaptive value of associative learning in wild populations. *Current Opinion in Behavioral Sciences, 16*, 73-79. doi: 10.1016/j.cobeha.2017.03.008

Morand-Ferron, J., Cole, E. F., & Quinn, J. L. (2016). Studying the evolutionary ecology of cognition in the wild: a review of practical and conceptual challenges. *Biological Reviews, 91*(2), 367-389. doi: 10.1111/brv.12174

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews, 85*(4), 935-956. doi: 10.1111/j.1469-185X.2010.00141.x

Nembrini, M., & Oppliger, A. (2003). Characterization of microsatellite loci in the wall lizard *Podarcis muralis* (Sauria: Lacertidae). *Molecular Ecology Notes, 3*, 123-124. doi: 10.1046/j.1471-8286

Niemelä, P. T., & Dingemanse, N. J. (2017). Individual versus pseudo-repeatability in behaviour: Lessons from translocation experiments in a wild insect. *Journal of Experimental Biology, 86*(5), 1033-1043. doi: 10.1111/1365-2656.12688

Noble, D. W., Carazo, P., & Whiting, M. J. (2012). Learning outdoors: male lizards show flexible spatial learning under semi-natural conditions. *Biology Letters, 8*(6), 946-948. doi: 10.1098/rsbl.2012.0813

Oosthuizen, M. K., Scheibler, A. G., Bennett, N. C., & Amrein, I. (2013). Effects of laboratory housing on exploratory behaviour, novelty discrimination and spatial reference memory in a subterranean, solitary rodent, the Cape mole-rat (*Georychus capensis*). *PLoS One, 8*(9), e75863. doi: 10.1371/journal.pone.0075863

Parashi, L. M. (1988). *Study of spiders in maquis ecosystems of southern Greece (mainland- insular)* PhD-thesis, University of Athens.

Payne, E., Sinn, D. L., Spiegel, O., Leu, S. T., Gardner, M. G., Godfrey, S. S., . . . Sih, A. (2021). Consistent after all: behavioural repeatability in a long-lived lizard across a 6-year field study. *Animal Behaviour, 174*, 263-277. doi: 10.1016/j.anbehav.2021.01.025

Perry, G., Wallace, M. C., Perry, D., Curzer, H., & Muhlberger, P. (2011). Toe clipping of amphibians and reptiles: science, ethics and the law. *Journal of Herpetology, 45*, 547-555. doi: 10.2307/41415330

Petelle, M. B., Martin, J. G., & Blumstein, D. T. (2015). Heritability and genetic correlations of personality traits in a wild population of yellow-bellied marmots (*Marmota flaviventris*). *Journal of Evolutionary Biology, 28*(10), 1840-1848. doi: 10.1111/jeb.12700

Polverino, G., Ruberto, T., Staaks, G., & Mehner, T. (2016). Tank size alters mean behaviours and individual rank orders in personality traits of fish depending on their life stage. *Animal Behaviour, 115*, 127-135. doi: 10.1016/j.anbehav.2016.03.013

Quinn, J. L., Cole, E. F., Reed, T. E., & Morand-Ferron, J. (2016). Environmental and genetic determinants of innovativeness in a natural population of birds. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1690). doi: 10.1098/rstb.2015.0184

R_Core_Team. (2018). A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Réale, D., Reader, S. M., Sol, D., McDougall, P. T., & Dingemanse, N. J. (2007). Integrating animal temperament within ecology and evolution. *Biological Reviews, 82*(2), 291-318. doi: 10.1111/j.1469-185X.2007.00010.x

Reichert, M. S., Crofts, S. J., Davidson, G. L., Firth, J. A., Kulahci, I. G., & Quinn, J. L. (2020). Multiple factors affect discrimination learning performance, but not between-individual variation, in wild mixed-species flocks of birds. *Royal Society Open Science, 7*(4), 192107. doi: 10.1098/rsos.192107

Rodríguez, R. L., & Gloudeman, M. D. (2011). Estimating the repeatability of memories of captured prey formed by *Frontinella communis* spiders (Araneae: Linyphiidae). *Animal Cognition, 14*, 675-682. doi: 10.1007/s10071-011-0402-9

Roth, T. C., LaDage, L. D., & Pravosudov, V. V. (2010). Learning capabilities enhanced in harsh environments: a common garden approach. *Proceedings of the Royal Society B: Biological Sciences, 277*(1697), 3187-3193. doi: 10.1098/rspb.2010.0630

Rowe, C., & Healy, S. D. (2014). Measuring variation in cognition. *Behavioral Ecology, 25*(6), 1287-1292. doi: 10.1093/beheco/aru090

Schuster, A. C., Carl, T., & Foerster, K. (2017a). Repeatability and consistency of individual behaviour in juvenile and adult Eurasian harvest mice. *Naturwissenschaften, 104*(3-4), 10. doi: 10.1007/s00114-017-1430-3

Schuster, A. C., Zimmermann, U., Hauer, C., & Foerster, K. (2017b). A behavioural syndrome, but less evidence for a relationship with cognitive traits in a spatial orientation context. *Frontiers in Zoology, 14*, 19. doi: 10.1186/s12983-017-0204-2

Shaw, R. C. (2017). Testing cognition in the wild: factors affecting performance and individual consistency in two measures of avian cognition. *Behavioural Processes, 134*, 31-36. doi: 10.1016/j.beproc.2016.06.004

Shaw, R. C., MacKinlay, R. D., Clayton, N. S., & Burns, K. C. (2019). Memory performance influences male reproductive success in a wild bird. *Current Biology, 29*(9), 1498-1502 e1493. doi: 10.1016/j.cub.2019.03.027

Sih, A., Mathot, K. J., Moiron, M., Montiglio, P. O., Wolf, M., & Dingemanse, N. J. (2015). Animal personality and state-behaviour feedbacks: a review and guide for empiricists. *Trends in Ecology & Evolution, 30*(1), 50-60. doi: 10.1016/j.tree.2014.11.004

Simpson, R. K., & McGraw, K. J. (2018). It's not just what you have, but how you use it: solar-positional and behavioural effects on hummingbird colour appearance during courtship. *Ecology Letters, 21*(9), 1413-1422. doi: 10.1111/ele.13125

Simpson, R. K., & McGraw, K. J. (2019). Experimental trait mismatches uncover specificity of evolutionary links between multiple signaling traits and their interactions in hummingbirds. *Evolution, 73*(3), 436-451. doi: 10.1111/evo.13662

Small, M. P., Beacham, R. E., Withler, R. E., & Nelson, R. J. (1998). Discriminating coho salmon (*Oncorhynchus kisutch*) populations within the Fraser River, British Columbia, using microsatellite DNA markers. *Molecular Ecology, 7*, 141-155. doi: 10.1046/j.1365-294x.1998.00324.x

Smith, C., Philips, A., & Reichard, M. (2015). Cognitive ability is heritable and predicts the success of an alternative mating tactic. *Proceedings of the Royal Society B: Biological Sciences, 282*(1809), 20151046. doi: 10.1098/rspb.2015.1046

Soha, J. A., Peters, S., Anderson, R. C., Searcy, W. A., & Nowicki, S. (2019). Performance on tests of cognitive ability is not repeatable across years in a songbird. *Animal Behaviour, 158*, 281-288. doi: 10.1016/j.anbehav.2019.09.020

Sol, D. (2009). Revisiting the cognitive buffer hypothesis for the evolution of large brains. *Biology Letters, 5*(1), 130-133. doi: 10.1098/rsbl.2008.0621

Sommer-Trembo, C., & Plath, M. (2018). Consistent individual differences in associative learning speed are not linked to boldness in female Atlantic mollies. *Animal Cognition, 21*(5), 661-670. doi: 10.1007/s10071-018-1201-3

Sonnenberg, B. R., Branch, C. L., Pitera, A. M., Bridge, E., & Pravosudov, V. V. (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. *Current Biology, 29*(4), 670-676 e673. doi: 10.1016/j.cub.2019.01.006

Sorato, E., Zidar, J., Garnham, L., Wilson, A., & Lovlie, H. (2018). Heritabilities and co-variation among cognitive traits in red junglefowl. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1756). doi: 10.1098/rstb.2017.0285

Spence, R., Magurran, A. E., & Smith, C. (2011). Spatial cognition in zebrafish: the role of strain and rearing environment. *Animal Cognition, 14*(4), 607-612. doi: 10.1007/s10071-011-0391-8

Stamps, J. A., & Groothuis, T. G. (2010). The development of animal personality: relevance, concepts and perspectives. *Biological Reviews, 85*(2), 301-325. doi: 10.1111/j.1469-185X.2009.00103.x

Stoffel, M. A., Nakagawa, S., Schielzeth, H., & Goslee, S. (2017). rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution, 8*(11), 1639-1644. doi: 10.1111/2041-210x.12797

Swan, A. A., Clutton, J. E., Chary, P. K., Cook, S. G., Liu, G. G., & Drew, M. R. (2014). Characterization of the role of adult neurogenesis in touch-screen discrimination learning. *Hippocampus, 24*(12), 1581-1591. doi: 10.1002/hipo.22337

Symonds, M. R. E., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology, 65*, 13-21. doi: 10.1007/s00265-010-1037-6

Szabo, B., Noble, D. W. A., Byrne, R. W., Tait, D. S., & Whiting, M. J. (2019). Precocial juvenile lizards show adult level learning and behavioural flexibility. *Animal Behaviour, 154*, 75-84. doi: 10.1016/j.anbehav.2019.06.003

Taylor, R. W., Boon, A. K., Dantzer, B., Reale, D., Humphries, M. M., Boutin, S., . . . McAdam, A. G. (2012). Low heritabilities, but genetic and maternal correlations between red squirrel behaviours. *Journal of Animal Ecology, 25*(4), 614-624. doi: 10.1111/j.1420-9101.2012.02456.x

Tebbich, S., & Teschke, I. (2014). Coping with uncertainty: woodpecker finches (*Cactospiza pallida*) from an unpredictable habitat are more flexible than birds from a stable habitat. *PLoS One, 9*(3), e91718. doi: 10.1371/journal.pone.0091718

Tello-Ramos, M. C., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2018). Memory in wild mountain chickadees from different elevations: comparing first-year birds with older survivors. *Animal Behaviour, 137*, 149-160. doi: 10.1016/j.anbehav.2017.12.019

Thomson, H. R., Lamb, S. D., Besson, A. A., Johnson, S. L., & Wright, J. (2020). Long-term repeatability of behaviours in zebrafish (*Danio rerio*). *Ethology, 126*(8), 803-811. doi: 10.1111/eth.13038

Thornton, A., Isden, J., & Madden, J. R. (2014). Toward wild psychometrics: linking individual cognitive differences to fitness. *Behavioral Ecology, 25*(6), 1299-1301. doi: 10.1093/beheco/aru095

Thys, B., Eens, M., Aerts, S., Delory, A., Iserbyt, A., & Pinxten, R. (2017a). Exploration and sociability in a highly gregarious bird are repeatable across seasons and in the long term but are unrelated. *Animal Behaviour, 123*, 339-348. doi: 10.1016/j.anbehav.2016.11.014

Thys, B., Pinxten, R., & Eens, M. (2021). Long-term repeatability and age-related plasticity of female behaviour in a free-living passerine. *Animal Behaviour, 172*, 45-54. doi: 10.1016/j.anbehav.2020.11.021

Thys, B., Pinxten, R., Raap, T., De Meester, G., Rivera-Gutierrez, H. F., & Eens, M. (2017b). The female perspective of personality in a wild songbird: Repeatable aggressiveness relates to exploration behaviour. *Scientific Reports, 7*(1), 7656. doi: 10.1038/s41598-017-08001-1

Troisi, C. A., Cooke, A. C., Davidson, G. L., de la Hera, I., Reichert, M. S., & Quinn, J. L. (2021). No evidence for cross-contextual consistency in spatial cognition or behavioral flexibility in a passerine. *Animal Behavior and Cognition, 8*(3), 446-461. doi: 10.26451/abc.08.03.08.2021

Valakos, E. D. (1990). *The ecology of the lizard Podarcis erhardii (Bedriaga, 1882) (Sauria: Lacertidae) in a typical insular ecosystem on Naxos Isl.* Ph. D. Ph. D. Thesis, University of Athens.

Valakos, E. D., Pafilis, P., Sotiropoulos, K., Lymberakis, P., Maragou, P., & Foufopoulos, J. (2008). *The Amphibians and Reptiles of Greece*. Frankfurt am Main: Chimaira.

van Horik, J. O., Langley, E. J. G., Whiteside, M. A., Laker, P. R., & Madden, J. R. (2018). Intra-individual variation in performance on novel variants of similar tasks influences single factor explanations of general cognitive processes. *Royal Society Open Science, 5*(7), 171919. doi: 10.1098/rsos.171919

Vanden Broecke, B., Bongers, A., Mnyone, L., Matthysen, E., & Leirs, H. (2021). Nonlinear maternal effects on personality in a rodent species with fluctuating densities. *Current Zoology, 67*(1), 1-9. doi: 10.1093/cz/zoaa032

Vardi, R., Goulet, C. T., Matthews, G., Berger-Tal, O., Wong, B. B. M., & Chapple, D. G. (2020). Spatial learning in captive and wild-born lizards: heritability and environmental effects. *Behavioral Ecology and Sociobiology, 74*(2). doi: 10.1007/s00265-020-2805-6

Verbeek, M. E. M., Drent, P. J., & Wiepkema, P. R. (1994). Consistent individual differences in early exploratory behaviour of male great tits. *Animal Behaviour, 48*, 1113-1121.

Vicente, N. S., & Halloy, M. (2017). Interaction between visual and chemical cues in a *Liolaemus* lizard: a multimodal approach. *Zoology, 125*, 24-28. doi: 10.1016/j.zool.2017.07.006

Wheelwright, N. T., Keller, L. F., & Postma, E. (2014). The effect of trait type and strength of selection on heritability and evolvability in an island bird population. *Evolution, 68*(11), 3325-3336. doi: 10.1111/evo.12499

Wolak, M. (2012). nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. *Methods in Ecology and Evolution, 3*, 792-796. doi: 10.1111/j.2041-210X.2012.00213.x

Wolf, M., van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs favour the evolution of animal personalities. *Nature, 447*(7144), 581-584. doi: 10.1038/nature05835

Wuerz, Y., & Krüger, O. (2014). Personality over ontogeny in zebra finches: long-term repeatable traits but unstable behavioural syndromes. *Frontiers in Zoology, 12*(Suppl 1), S9.

TABLES

Table 1. Principal Component Analysis of exploration behaviours. Loadings with an absolute value > 0.3 (bold) were considered to substantially contribute to a principal component (cfr. Boon, Reale, & Boutin, 2007; Dammhahn, 2012; Thys et al., 2017b).

	Comp 1	**Comp 2**	**Comp 3**
Eigenvalue	1.76	1.29	1.00
% variance	44.33	23.60	14.16
First transition	-0.130	**-0.411**	**0.739**
# transitions	**0.422**	**0.357**	0.268
Latency to explore all quadrants	**-0.347**	**-0.474**	0.156
# touches	**0.335**	0.287	**0.556**
# refuges entered	**0.462**	**-0.339**	
Latency to enter first refuge	**-0.432**	0.286	0.177
Time spent hiding	**0.413**	**-0.448**	-0.122

Table 2. Overview of the final models and their results for the short-term (within-year) repeatability of exploration within each age group. Models were constructed based on a model selection approach (see main text), using predictors with an relative importance (RI) ≥ 0.50. Repeatability (R) was calculated using the 'rptR' package in R (Stoffel et al., 2017). Both the adjusted and unadjusted repeatability are given, with their 95% confidence interval (square brackets). Their significance was tested using a log-likelihood ratio test. For the meaning of the exploration PCs, we refer to Table 1. Statistical significance is reported as: '°' $p < 0.10$, '*' $p < 0.05$, '**' $p < 0.01$, '***' $p < 0.001$.

Personality Trait	Age group	*N*	RI	Confounding factors	*F*-stats	*P*
Exploration PC1	Adults '19 (Box-cox λ = 1.4)	66	0.62	Tail	$F_{1,64} = 3.064$	0.085 °
				R_{adj}	0.085 [0.000; 0.310]	0.242
				R	0.101 [0.000; 0.314]	0.202
	Adults '19 (survivors only) (Box-cox λ = 1.3)	43	/	/	/	/
				R_{adj}	/	/
				R	0.079 [0.000; 0.370]	0.303
	Adults '20	43	0.79	Sex	$F_{1,37} = 4.071$	0.051 °
			1.00	**SVL**	$F_{1,52} = 6.465$	**0.014 ***
				R_{adj}	**0.333** [0.087; 0.602]	**0.007 ****
				R	**0.449** [0.167; 0.652]	**< 0.001 *****
	Intruders	21	/	/	/	/
				R_{adj}	/	/
				R	0.010 [0.000; 0.393]	0.480
	Juveniles	44	/	/	/	/
				R_{adj}	/	/
				R	0.005 [0.000; 0.237]	0.480
Exploration PC2	Adults '19 (Box-cox λ = 0.8)	66	1.00	Habitat	$F_{1,3} = 6.134$	0.094 °
			1.00	**Tail**	$F_{1,61} = 8.731$	**0.004 ****
				R_{adj}	**0.211** [0.000; 0.436]	**0.034 ***
				R	**0.239** [0.022; 0.433]	**0.011 ***
	Adults '19 (survivors only)	43	1.00	Habitat	$F_{1,3} = 4.183$	0.145
			1.00	**Tail**	$F_{1,38} = 6.215$	**0.017 ***
				R_{adj}	0.130 [0.000; 0.410]	0.197
				R	0.169 [0.000; 0.430]	0.104
	Adults '20	43	1.00	**Sex**	$F_{1,35} = 5.970$	**0.020 ***
				R_{adj}	0.162 [0.000; 0.445]	0.136
				R	0.227 [0.000; 0.465]	0.064 °
	Intruders	21	0.71	Tail	$F_{1,31} = 2.679$	0.111
				R_{adj}	**0.448** [0.038; 0.722]	**0.018 ***
				R	**0.386** [0.000; 0.686]	**0.030 ***
	Juveniles	44	0.68	SVL	$F_{1,40} = 3.776$	0.059 °
				R_{adj}	0.075 [0.000; 0.335]	0.280
				R	0.093 [0.000; 0.352]	0.228

Table 3. Overview of the final models and their results for the long-term (across-year) repeatability of exploration and cognition. Models were constructed based on a model selection approach (see main text), using predictors with an relative importance (RI) ≥ 0.50. Repeatability (R) was calculated using the 'rptR' package in R (Stoffel et al., 2017). Both the adjusted and unadjusted repeatability are given, with their 95% confidence interval (square brackets). Their significance was tested using a log-likelihood ratio test. For the meaning of the exploration PCs, we refer to Table 1. Statistical significance is reported as: '°' $p < 0.10$, '*' $p < 0.05$, '**' $p < 0.01$, '***' $p < 0.001$.

Personality/Cognition Trait	***N***	**RI**	**Confounding factors**	***F*-stats**	***P***
Exploration PC1	43	0.67	Enclosure	$F_{1,2} = 2.047$	0.284
		0.52	Sex	$F_{1,38} = 2.409$	0.129
		1.00	**Year**	$F_{1,152} = 41.171$	**< 0.001 *****
		1.00	**SVL**	$F_{1,134} = 8.812$	**0.004 ****
			R_{adj}	**0.280** [0.091; 0.436]	**< 0.001 *****
			R	**0.188** [0.021; 0.335]	**0.003 ****
Exploration PC2	43	0.83	Habitat	$F_{1,3} = 2.276$	0.237
		1.00	**Sex**	$F_{1,35} = 9.032$	**0.005 ****
			R_{adj}	0.060 [0.000; 0.201]	0.178
			R	**0.117** [0.000; 0.249]	**0.039 ***
SL Score (log)	42	0.77	Year	$F_{1,1} = 3.331$	0.317
		1.00	**Safe side**	$F_{1,37} = 106.93$	**< 0.001 *****
		0.86	*Tail*	$F_{1,63} = 3.602$	0.062 °
			R_{adj}	**0.398** [0.124; 0.622]	**0.004**
			R	**0.786** [0.515; 0.868]	**< 0.001 *****
RL Score (log)	42	1.00	Habitat	$F_{1,2} = 4.932$	0.141
		0.64	Enclosure	$F_{1,2} = 0.242$	0.672
		0.58	Sex	$F_{1,33} = 2.798$	0.104
		0.64	Year	$F_{1,2} = 2.006$	0.324
		1.00	**Safe side**	$F_{1,34} = 195.677$	**< 0.001 *****
		0.64	**Enclosure * Year**	$F_{1,39} = 7.924$	**0.008 ****
			R_{adj}	0.251 [0.000; 0.545]	0.061 °
			R	**0.805** [0.530; 0.874]	**< 0.001 *****
Flex Score	42	/	/	/	/
			R_{adj}	/	/
			R	0.192 [0.000; 0.460]	0.105
ESC Time (Box-cox $\lambda = 0.2$)	41	/	/	/	/
			R_{adj}	/	/
			R	0.000 [0.000; 0.307]	1

Table 4. Overview of the final animal models (MCMCglmm) and their results for the heritability of exploration and cognition. Models were constructed based on a model selection approach (see main text), using predictors with an relative importance (RI) ≥ 0.50. Posterior means + 95% credible intervals (between square brackets) are reported. Predictors were considered to be important if the 95 % credible interval did not overlap with zero (bold). Heritability was calculated from both the final and null models. Higher exploration scores correspond to more explorative behaviour, while higher scores for spatial learning (SL), reversal learning (RL) and learning flexibility (FLEX) reflect more errors and thus worse cognitive performance.

Personality/Cognitive trait	RI	Confounding factors	Posterior mean + CI
Exploration PC1	/	Intercept	-0.035 [-1.616; 1.268]
	1.00	**Age (Adult '19)**	**0.549 [0.209; 0.893]**
		Age (Adult '20)	-0.093 [-0.450; 0.282]
		Age (Juv)	**-0.587 [-0.972; -0.204]**
	0.65	SVL	0.057 [-0.044; 0.160]
		h^2	0.031 [0.000; 0.110]
		$h^2_{null\ model}$	0.027 [0.000; 0.092]
Exploration PC2	/	Intercept	-0.001 [-1.376; 1.581]
		h^2	/
		$h^2_{null\ model}$	0.057 [0.000; 0.178]
SL Score	/	Intercept	-0.378 [-1.046. 0.307]
(log)	0.72	Tail (intact)	0.346 [-0.026. 0.780]
		h^2	0.054 [0.000; 0.175]
		$h^2_{null\ model}$	0.056 [0.000; 0.184]
RL Score	/	Intercept	-0.077 [-0.870; 0.617]
(log)	0.65	Enclosure (Simple)	0.558 [-0.828; 2.003]
	0.65	Age (Adult '19)	0.367 [-0.368; 1.000]
		Age (Adult '20)	0.245 [-0.358; 0.808]
		Age (Juvenile)	-0.586 [-1.366; 0.200]
	0.65	Enclosure (Simple)*Age (Adult '19)	-0.748 [-2.022; 0.588]
		Enclosure (Simple)*Age (Adult '20)	-0.122 [-1.487; 1.227]
		Enclosure (Simple)*Age (Juveniles)	-0.476 [-1.842; 0.968]
		h^2	0.074 [0.000; 0.249]
		$h^2_{null\ model}$	0.063 [0.000; 0.218]
Flex Score	/	Intercept	-0.103 [-0.701. 0.623]
		h^2	/
		$h^2_{null\ model}$	0.053 [0.000; 0.167]

FIGURE CAPTIONS

Figure 1. Graphical representation of the four enclosures (structural simple on top, structural complex bottom). Per enclosure, sample sizes per group and per sex are given. Small letters next to the numbers represent whether the lizards originated from a complex (c) or simple (s) habitat.

Figure 2. Exploration PC2 scores for surviving adults that were tested in both year ($N_{female} = 22$, $N_{male} = 21$). Orange boxplots represent exploration PC2 scores in 2019, light grey boxplots visualize PC2 scores when retested in 2020 and dark grey boxplots are the pooled data over both years (long-term repeatability – LTR). Statistical significant differences are indicted as follows: '°' $p < 0.10$, '*' $p < 0.05$, '**' $p < 0.01$, '***' $p < 0.001$. Higher scores represent lizards that are faster in exploring a novel arena and spent less time hiding.

Figure 3. A) boxplots representing the exploration PC1 scores for adult Aegean wall lizards tested in both 2019 and 2020 ($N = 43$). Statistical significant differences are indicted as follows: '°' $p < 0.10$, '*' $p < 0.05$, '**' $p < 0.01$, '***' $p < 0.001$. B) boxplots with the exploration PC1 scores per age group ($N_{adults\ '19} = 66$, $N_{adults\ '20} = 43$, $N_{intruders} = 21$, $N_{juveniles} = 45$). Age groups were considered different from each other if the 95 % credibility interval of their difference (as obtained from a MCMCglmm) did not overlap with zero, which is indicated with an '*'. In both graphs, higher scores represent more explorative behaviour, but see Table 1 for a more detailed explanation of the PC scores.

Figure 4. Performance of lizards (number of errors made) over consecutive trials in the spatial and reversal learning task. Significant regressions are indicated by a solid line, grey areas visualize the standard errors. Sample sizes are as follows: $N_{adults\ '19} = 66$, $N_{adults\ '20} = 42$, $N_{intruders} = 21$, $N_{juveniles} = 44$.

Figure 5. Boxplots visualising the reversal learning (RL) scores per year and per enclosure type. Higher scores indicate that lizards made more errors and thus correspond to lower cognitive performance. The same individual lizards were tested both in 2019 (orange) and 2020 (grey) after spending one year in semi-natural enclosures mimicking either a complex or simple habitat. Albeit a significant interaction was found between enclosure type and year, post-hoc pairwise comparisons did not reveal any significant differences among groups. Sample sizes were as follows: $N_{complex} = 25$, $N_{simple} = 17$.

Figure 6. Boxplots with the reversal learning (RL) scores per age group in enclosures with a complex habitat (left: $N_{adults\ '19} = 33$, $N_{adults\ '20} = 25$, $N_{intruders} = 18$, $N_{juveniles} = 21$) and simple habitat (right: $N_{adults\ '19} = 33$, $N_{adults\ '20} = 17$, $N_{intruders} = 3$, $N_{juveniles} = 22$). Age groups were considered different from each other if the 95 % credibility interval of their difference (as obtained from a MCMCglmm) did not overlap with zero, which is indicated with an '*'. In both graphs, higher scores represent more errors and thus worse cognitive performance.

APPENDIX

Table A1. Performance (mean ± SE) on the exploration test and cognitive tasks for lizards that were tested in both 2019 and 2020, given per original habitat and enclosure type. For the meaning of the exploration PCs, we refer to Table 1 in main text.

	Complex habitat		**Simple habitat**	
Enclosure	Complex	Simple	Complex	Simple
Exploration PC1	*N* = 13	*N* = 9	*N* = 13	*N* = 8
2019	1.20 ± 0.31	0.32 ± 0.38	1.25 ± 0.25	0.76 ± 0.23
2020	-0.14 ± 0.38	-0.58 ± 0.49	0.07 ± 0.31	-0.37 ± 0.37
Exploration PC2	*N* = 13	*N* = 9	*N* = 13	*N* = 8
2019	-0.18 ± 0.25	-0.21 ± 0.20	0.49 ± 0.23	0.59 ± 0.25
2020	0.19 ± 0.27	0.25 ± 0.25	0.55 ± 0.28	-0.02 ± 0.39
# Spatial Errors	*N* = 12	*N* = 9	*N* = 13	*N* = 8
2019	0.80 ± 0.20	0.79 ± 0.19	1.12 ± 0.23	1.03 ± 0.29
2020	0.58 ± 0.15	0.85 ± 0.29	0.97 ± 0.19	0.76 ± 0.23
# Reversal Errors	*N* = 12	*N* = 9	*N* = 13	*N* = 8
2019	1.0 ± 0.20	0.79 ± 0.26	0.61 ± 0.16	0.61 ± 0.18
2020	0.82 ± 0.18	1.13 ± 0.23	0.63 ± 0.15	0.79 ± 0.25
# Flexibility Errors	*N* = 12	*N* = 9	*N* = 13	*N* = 8
2019	0.90 ± 0.07	0.79 ± 0.09	0.85 ± 0.07	0.82 ± 0.11
2020	0.70 ± 0.06	0.99 ± 0.13	0.80 ± 0.05	0.78 ± 0.10
# Escaped from Box	*N* = 11	*N* = 9	*N* = 13	*N* = 8
2019	9	7	12	6
2020	11	8	11	8
Escape Time (s)	*N* = 11	*N* = 9	*N* = 13	*N* = 8
2019	885 ± 179	772 ± 243	717 ± 169	804 ± 228
2020	492 ± 93	767 ± 182	809 ± 150	546 ± 130

Table A2. Performance (mean ± SE) on the exploration test and cognitive tasks for each age group. For the meaning of exploration PCs, we refer to Table 1 in main text.

Enclosure	Adults 2019	Intruders	Juveniles
Exploration PC1	$N = 66$	$N = 21$	$N = 45$
	0.91 ± 0.13	0.02 ± 0.22	-1.14 ± 0.17
Exploration PC2	$N = 66$	$N = 21$	$N = 45$
	0.19 ± 0.11	-0.00 ± 0.19	-0.54 ± 0.12
# Spatial Errors	$N = 66$	$N = 21$	$N = 44$
	0.82 ± 0.08	0.57 ± 0.07	0.55 ± 0.09
# Reversal Errors	$N = 66$	$N = 21$	$N = 44$
	0.85 ± 0.08	0.70 ± 0.10	0.53 ± 0.08
# Flexibility Errors	$N = 66$	$N = 21$	$N = 44$
	0.83 ± 0.03	0.64 ± 0.05	0.54 ± 0.04

Table A3. Results of the post-hoc pairwise comparisons on the differences in exploration PC2 and RL scores between different age groups. Data was analysed using a MCMCglmm and pairwise comparisons were conducted using the 'emmeans' function in R (Lenth et al., 2019). For each pairwise comparison, the estimated difference + 95 % credible interval (CI, between brackets) is given. Bold indicates that the CI did not overlap with zero and the groups thus differed from each other.

Model	Predictor	Groups	Estimate + CI
Exploration PC2	Age	**Intruders - Adults '19**	**-0.548 [-0.893; -0.209]**
		Intruders – Adults '20	0.096 [-0.282; 0.450]
		Intruders – Juveniles	**0.590 [0.204; 0.972]**
		Adults '19 – Adults '20	**0.642 [0.409; 0.858]**
		Adults '19 – Juveniles	**1.135 [0.864; 1.423]**
		Adults '20 - Juveniles	**0.494 [0.185; 0.783]**
RL Scores	Enclosure * Age	Intruders$_{complex}$ – Intruders$_{simple}$	-0.565[-1.966; 0.790]
		Intruders$_{complex}$ – Adults$_{complex}$ '19	-0.370 [-1.055; 0.293]
		Intruders$_{complex}$ – Adults$_{simple}$ '19	-0.188 [-1.131; 0.766]
		Intruders$_{complex}$ – Adults$_{complex}$ '20	-0.248 [-0.833; 0.328]
		Intruders$_{complex}$ – Adults$_{simple}$ '20	-0.692 [-1.609; 0.236]
		Intruders$_{complex}$ – Juveniles$_{complex}$	0.582 [-0.169; 1.370]
		Intruders$_{complex}$ – Juveniles$_{simple}$	0.494 [-0.529; 1.606]
		Intruders$_{simple}$ – Adults$_{complex}$ '19	0.200 [-1.138; 1.729]
		Intruders$_{simple}$ – Adults$_{simple}$ '19	0.373[-0.885; 1.581]
		Intruders$_{simple}$ – Adults$_{complex}$ '20	0.330 [-1.040; 1.680]
		Intruders$_{simple}$ – Adults$_{simple}$ '20	-0.124 [-1.251; 1.173]
		Intruders$_{simple}$ – Juveniles$_{complex}$	1.160 [-0.381; 2.612]
		Intruders$_{simple}$ – Juveniles$_{simple}$	1.066 [-0.252; 2.394]
		Adults$_{complex}$ '19– Adults$_{simple}$ '19	0.180 [-0.634; 0.977]
		Adults$_{complex}$ '19– Adults$_{complex}$ '20	0.126 [-0.401; 0.740]
		Adults$_{complex}$ '19– Adults$_{simple}$ '20	-0.319 [-1.304; 0.648]
		Adults$_{complex}$ '19– Juveniles$_{complex}$	**0.961 [0.266; 1.698]**
		Adults$_{complex}$ '19– Juveniles$_{simple}$	0.856 [-0.187; 1.947]
		Adults$_{simple}$ '19– Adults$_{complex}$ '20	-0.057 [-0.977; 0.865]
		Adults$_{simple}$ '19– Adults$_{simple}$ '20	-0.503 [-1.070; 0.173]
		Adults$_{simple}$ '19– Juveniles$_{complex}$	0.779 [-0.140; 1.861]
		Adults$_{simple}$ '19– Juveniles$_{simple}$	**0.692 [0.014; 1.449]**
		Adults$_{complex}$ '20– Adults$_{simple}$ '20	-0.438 [-1.348; 0.455]
		Adults$_{complex}$ '20– Juveniles$_{complex}$	**0.833 [0.095; 1.530]**
		Adults$_{complex}$ '20– Juveniles$_{simple}$	0.743 [-0.259; 1.840]
		Adults$_{simple}$ '20– Juveniles$_{complex}$	**1.271 [0.264; 2.437]**
		Adults$_{simple}$ '20– Juveniles$_{simple}$	**1.187 [0.355; 1.946]**
		Juveniles$_{complex}$ – Juveniles$_{simple}$	-0.088 [-1.076; 0.834]

Table A4. Full model outcome of the GLMMs testing the performance of lizards over consecutive trials during the spatial and reversal learning task. If an interaction was non-significant, it was removed and main effects would be reported from an main-effect model only. Significance levels are indicated as follows: : '°' $p < 0.10$, '*' $p < 0.05$, '**' $p < 0.01$, '***' $p < 0.001$.

Response variable	**Predictor**	**Wald-stats**	***P***
Spatial learning	**Trial**	$\chi^2 = 11.970$, df =1	**<0.001 *****
	Safe Side	$\chi^2 = 307.027$,df =1	**<0.001 *****
	Group	$\chi^2 = 28.202$, df =3	**<0.001 *****
	Trial * Group	$\chi^2 = 3.846$, df =3	0.279
	Trial * Safe side	$\chi^2 = 2.063$, df =1	0.151
Reversal learning	**Trial**	$\chi^2 = 24.311$, df =1	**< 0.001 *****
	Safe Side	$\chi^2 = 6.001$,df =1	**0.014 ***
	Group	$\chi^2 = 10.395$, df =3	**0.015 ***
	Trial * Group	$\chi^2 = 10.387$, df =3	**0.016 ***
	Trial * Safe Side	$\chi^2 = 4.577$, df =1	**0.032 ***

Table A5. Post-hoc multiple pairwise comparisons for the (G)LMMs, using Tukey's method with the 'emmeans' and 'emtrends' functions in R (Lenth et al., 2019). Significance levels are indicated as follows: : '°' $p < 0.10$, '*' $p < 0.05$, '**' $p < 0.01$, '***' $p < 0.001$.

Model	Predictor	Groups	Ratio + SE	Stats
Spatial learning (per trial)	Age group	Adult 20 – Adult 19	0.848 ± 0.104	$z = -1.340$, $p = 0.537$
		Adult 20 – Intruder	0.925 ± 0.151	$z = -0.478$, $p = 0.964$
		Adult 20 – Juvenile	1.669 ± 0.237	**$z = 3.615$, $p = 0.002$** ***
		Adult 19 – Intruder	1.090 ± 0.165	$z = 0.572$, $p = 0.940$
		Adult 19 – Juvenile	1.968 ± 0.256	**$z = 5.204$, $p < 0.001$** ***
		Intruder – Juvenile	1.805 ± 0.305	**$z = 3.497$, $p = 0.003$** ***
Reversal learning (per trial)	Age * trial	**$\text{slope}_{\text{Adult 20}}$ – $\text{slope}_{\text{Adult 19}}$**	-0.046 ± 0.017	**$z = -2.767$, $p = 0.029$ ***
		$\text{slope}_{\text{Adult 20}}$ – $\text{slopeI}_{\text{ntruder}}$	-0.037 ± 0.023	$z = -1.619$, $p = 0.368$
		$\text{slope}_{\text{Adult 20}}$ – $\text{slope}_{\text{Juvenile}}$	-0.003 ± 0.021	$z = -0.146$, $p = 0.999$
		$\text{slope}_{\text{Adult 19}}$ – $\text{slope}_{\text{Intruder}}$	0.009 ± 0.022	$z = 0.418$, $p = 0.976$
		$\text{slope}_{\text{Adult 19}}$ - $\text{slope}_{\text{Juvenile}}$	0.043 ± 0.019	$z = 2.325$, $p = 0.092$ °
		$\text{slope}_{\text{Intruder}}$ - $\text{slope}_{\text{Juvenile}}$	0.034 ± 0.024	$z = 1.410$, $p = 0.493$
RL Scores LTR	Enclosure * Year	Complex 19 – Simple 19	1.112 ± 0.057	$t = 2.081$, $p = 0.277$
		Complex 19 – Complex 20	1.029 ± 0.048	$t = 0.614$, $p = 0.921$
		Complex 19 – Simple 20	0.964 ± 0.055	$t = -0.647$, $p = 0.911$
		Simple 19 – Complex 20	0.926 ± 0.053	$t = -1.344$, $p = 0.592$
		Simple 19 – Simple 20	0.867 ± 0.047	$t = -2.653$, $p = 0.151$
		Complex 20 – Simple 20	0.936 ± 0.048	$t = -1.293$, $p = 0.605$