Cognitive performance and stress responsiveness in the Trinidadian Guppy – a multivariate approach

Pamela M. Prentice¹, Chloe Mnatzaganian¹, Thomas M. Houslay², Alex Thornton¹, Alastair J. Wilson¹

¹Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Cornwall, TR10 9FE, UK

²Department of Zoology, University of Cambridge, Cambridge, CB2 3EJ, UK

Abstract

Among-individual variation in cognitive performance has been recently demonstrated across a range of animal taxa. While this variation is a prerequisite for contemporary natural selection, it is also true that selection does not act on traits in isolation. Thus, the extent to which cognitive traits covary with other aspects of phenotype (e.g. personality traits) is expected to be an important factor in shaping evolutionary dynamics. Here we adopt a multivariate approach to test for spatial learning ability in a captive population of male Trinidadian guppies (*Poecilia reticulata*), and ask whether differences in cognitive performance are associated with (repeatable) differences in stress response behaviour. We focus on stress response for two reasons. First, functional links between cognitive traits and ‘stress coping style’ have been hypothesised. Second, individual-level studies of cognitive performance typically rely on multiple testing paradigms that may themselves be a stressor. Thus, there is a risk that variation in stress responsiveness is itself a cause of apparent, but artefactual variance in cognitive ability. Using a set of fish exposed repeatedly to two distinct spatial learning tasks (maze layouts), and an acute stress response test (open field trial), we find differences among-individuals in task performance that are repeatable within- and across maze layouts. On average performance improves with experience in the first maze, consistent with spatial learning, but not the second. In both mazes there is among-individual variation in the trajectory of mean performance with trial number suggesting individuals differing in ‘learning rate’. Acute stress response behaviour is repeatable but predicts neither average time to solve the maze nor learning rate. We thus find no support for among-individual correlation between acute stress response and cognitive performance. However, we highlight the possibility that cumulative, chronic stress effects may nonetheless cause observed declines in performance across repeats for some individuals (leading to lack of improvement in mean time to solve the second maze). If so, this may represent a pervasive but difficult challenge for our ability to robustly estimate learning rates in studies of animal cognition.

Introduction

Cognition, or “information processing”, is defined as the set of mechanisms by which animals acquire, process, store and use information from the environment (Shettleworth, 2010), and is vital for carrying out day-to-day behaviours needed for survival and reproduction. While differences in cognitive performance among-species have long been studied in comparative psychology (for a review see Healy 2019), a more recent focus in behavioural ecology has been the characterisation of among-individual variation within populations of non-human animals (Ashton et al., 2018; Boogert et al., 2018; Lucon-Xiccato & Bisazza, 2017a). This among-individual variation is interesting from an evolutionary perspective, as it is a pre-requisite for natural selection and genetic variation – both of which are fundamental for adaptive evolution to occur (Wilson et al. 2010). However, selection does not act on traits in isolation. Functional links between variation in cognitive performance and other aspects of behaviour (including, for example neophobia, boldness and stress responsiveness) have been hypothesised (Griffin et al. 2015; Medina-García et al. 2017; Quinn et al. 2012; Sweis et al. 2013). Robustly testing these relationships is often challenging, requiring multivariate data collection and analyses to detect and describe patterns of variation between associated traits at the appropriate level (e.g., among-individual and/or among genotype; Dingemanse & Dochtermann, 2013). Nonetheless, such efforts are important if we hope to understand the adaptive evolution of cognition in the context of the wider phenotype (Thornton & Wilson, 2015). Here we address this broad goal in the more specific context of testing hypothesised links between cognitive performance and a stress-response (Gibelli et al., 2019; Øverli et al., 2007) in Trinidadian guppies (*Poecilia reticulata*).

Quantifying patterns of among-individual variation in cognitive traits is still in its infancy (Boogert et al., 2018; Rowe & Healy, 2014; Thornton et al., 2014), and empirical studies therefore remain somewhat limited (but see Ashton et al., 2018; Lucon-Xiccato & Bisazza, 2017b; Niemelä et al., 2013; Shaw et al., 2019; Sonnenberg et al., 2019 for examples). However it is now abundantly clear that populations typically harbour high levels of among-individual variation in behavioural traits more generally (Dingemanse & Réale, 2005). Individual differences in (mean) behaviours, commonly referred to as personality, can manifest as, for instance, variation in aggressiveness or sociability

towards conspecifics, or differences in response when faced with predators or other sources of perceived risk (Bridger et al. 2015; Réale et al. 2007). Since, strong directional or stabilising selection is usually predicted to erode variation (Roff, 2002), it is widely hypothesised that variation in personality traits is maintained by fitness trade-offs of some kind (Dingemanse et al., 2004; Quinn et al., 2016). For example, bolder individuals may be better at acquiring resources to invest in life history traits (e.g., growth, reproduction) but their behaviour may also expose them to greater predation risk. In this way personalities can themselves be viewed as components of life history strategies, leading to an expectation that they will be correlated with – and trade-off against – other aspects of physiological, reproductive, and behavioural phenotype (Réale et al. 2010; Sih, Bell, & Johnson 2004; Wolf et al. 2007). Certainly, arguments that trade-offs can maintain variation in cognitive performance parallel explanations made for widespread presence of personality. These could be trade-offs among cognitive domains, or between, for instance an overall cognitive performance trait (‘general intelligence’ (Burkart et al., 2017; Galsworthy et al., 2005; Plomin & Spinath, 2002)) or other aspects of phenotype.

Variation in stress response provides one putative source of among-individual differences in both personality traits and cognitive performance (Gibelli et al., 2019; Raoult et al., 2017). The widely used concept of stress coping style model predicts that individuals will vary – both behaviourally and physiologically – along a proactive/ reactive continuum (Coppens et al., 2010; Koolhaas et al., 1999; Sih, Bell, & Johnson 2004). As originally posited, the model predicts proactive coping styles will express more ‘fight or flight’ type behavioural responses induced by adrenaline-response to stressors. At the other extreme, reactive coping styles will be more behaviourally ‘passive’ (e.g., freezing or hiding) and show high HPA(I) activity leading to cortisol response (Carere et al., 2014; Øverli et al., 2007). Various links to cognitive performance variation have been suggested. For instance, proactive styles are broadly thought to be associated with ‘bold’, exploratory, risk-taking personalities that may present with more opportunities to learn initially. Conversely, greater behaviourally flexibility associated with reactive coping styles (Coppens et al., 2010) may be important for tasks such as reversal learning, that require an ability to acquire (and use) new

information under changing environmental conditions (Griffin et al., 2015; Koolhaas et al., 1999; Sih & Del Giudice, 2012). More generally, sensitivity to external stressors or challenges could impact performance in cognitive assays if more stressed individuals are simply less motivated and/or are focused on sources of risk rather than environmental cues of rewards.

Although hypothesised links between stress responsiveness (or coping style) and cognitive performance seem intuitive, empirical evidence is still limited to a small number of studies (Bebus et al., 2016; Bensky et al., 2017; Brust & Guenther, 2017; Lukowiak et al., 2014; Mazza et al., 2018; Mesquita et al., 2015; van Horik et al., 2017). There are also contrasting studies in which either a weak or no relationship was detected (Carazo et al., 2014; Cole et al., 2011; Guillette et al., 2015). It is also possible that relationships are variable across different aspects of cognition. For instance in sailfin mollies (*Poecilia latipinna*), less anxious fish performed better in a discrimination learning task than highly anxious individuals, whereas the opposite was found in a reversal learning task. (Gibelli et al., 2019). Clearly, there is need for more empirical work before a clear picture of the complex relationship between variation in cognitive performance and stress responsiveness/coping style is understood. Here we address this broad goal by testing the hypothesis that individual differences in cognitive performance and stress responsiveness are correlated in male Trinidadian guppies (*Poecilia reticulata*).

The guppy is a freshwater poeciliid fish that is widely used as a model in behavioural and evolutionary ecology. Methods for assaying among-individual ‘personality’ variation are well established in this species generally (Burns & Rodd, 2008; White et al., 2016), while guppies have also been used in cognitive studies targeting colour and shape discrimination learning (Lucon-Xiccato & Bisazza, 2014, 2016), numerical learning (Kotrschal et al., 2013), and spatial learning (Kotrschal et al., 2015; Lucon-Xiccato & Bisazza, 2017b, 2017c). Here, we investigate the relationship between behavioural stress response and performance in a spatial learning task in which male guppies repeatedly navigated a maze to access females as a reward. The cognitive task was repeated using a second, differently structure maze in order that we could assess not just variation in learning within a single spatial context, but also ask whether – for instance – individuals displaying greater performance

in trials using the first maze subsequently also performed better in the second. In the wild, male guppies usually utilize large home ranges during mate search and foraging (Croft et al. 2003), and as such spatial learning is expected to be an ecologically relevant trait (Brown & Braithwaite, 2005). For our measure of stress responsiveness, we utilise ‘Open Field Trials’ (OFT). Widely used across species as a paradigm for characterising behavioural differences related to exploration, activity, and ‘shy-bold’ type variation (Bell et al., 2009; Gosling, 2001), previous studies on this captive population of guppies have highlighted its utility for assaying behavioural stress response (see e.g., Prentice et al 2020). Observed behaviours expressed in the OFT are both repeatable and plastic with respect to experimentally-manipulated stressor severity (specifically perceived predation risk) (Houslay et al., 2018). We also know from pedigree-based quantitative genetic studies that individual (mean) behaviours and their predictability (defined as within-individual variance) are heritable (Prentice et al., 2020; White et al., 2019; White & Wilson, 2019). Furthermore, there is some evidence of genetic integration between OFT behaviour and cortisol expression, strengthening the view that the OFT provides an appropriate assay of behavioural stress response; (Houslay et al., 2019).

In what follows we: i) test for evidence of learning in naïve guppies repeatedly exposed to a spatial learning task (maze), ii) ask whether individuals differ in cognitive performance across repeated trials and if so; iii) whether performance in the first maze predicts performance in a second spatial context (i.e. reconfigured maze). We predict that time to complete the mazes (our proxy of cognitive performance) will, on average, improve with experience consistent with spatial “learning”, but that individuals will consistently differ in cognitive performance within each maze. We also predict that individual performance in the first maze will be positively correlated with performance in the second, consistent with stable differences in cognitive ability. Finally, iv) we test the hypothesis that individual differences in cognitive performance will be associated with differences in stress responsiveness. However, in the current absence of specific models, we make no *a priori* predictions about the sign of this relationship.

Methods

Study site and housing

All behavioural assays were carried out on guppies from a captive population (derived from wild fish collected in the Aripo River, Trinidad in 2008) housed at the University of Exeter's Penryn campus. Adult males ($n = 64$) were randomly sampled from the stock population, and housed in groups of 8 in separate home tanks (15 l, $18.5 \times 37 \times 22$ cm) maintained at 23–24°C on a 12:12 light/dark cycle. The tanks shared a recirculating sump water supply which underwent a 25% water change once per week. All fish were fed to satiation twice daily on commercial flake food and live brine shrimp (*Artemia salina*) to control as much as possible for energetic and nutritional states prior to testing. We elected to focus on males only for several reasons. First, pilot studies showed a high occurrence of 'freezing' behaviour in females (relative to males) when introduced to the maze. While freezing can be a component of the behavioural stress response (Houslay et al., 2018), we considered that frequent occurrence during the cognitive assay would complicate data interpretation. Second, males show consistent sexual reproductive motivation towards females (Burns & Rodd, 2008), enabling the use of females as a 'reward' for males solving the maze (Kotrschal et al., 2015). Third, male guppies exhibit distinctive markings and colouration on body and fins. By recording and sketching these for each fish we were able identify individuals within groups without the need to subject individuals to invasive tagging.

Ethics

This work was conducted under the auspices of the Animals (Scientific Procedures Act) out with approval of the University of Exeter research ethics committee, under licence from the Home Office (UK) (Licence Number PPL30/3256). Experimental procedures and behavioural assays were developed in accordance with the principles of the three R's and ASAB guidelines (Buchanan et al., 2020) for use of animals. All periods of handling and emersion were kept to a minimum and only fish deemed healthy and exhibiting normal behaviour were used in trials. At the end of the experiment,

fish were returned to a designated ‘retirement’ tank (containing females as well as males) and not used in any further experiments.

Overview of behavioural testing scheme

We used a repeated measures approach to test for among-individual (co)variation in spatial learning performance and stress responsiveness. Spatial learning was first assessed by repeatedly trialling individuals in a maze apparatus (Maze A, Figure 1). Each individual fish was tested once per day for 11 consecutive days with improvement in time to complete the maze interpreted as ‘learning’. This is consistent with previous studies using either time to complete an objective or to perform a particular task to investigate variation in cognitive performance among-individuals (Guillette et al., 2015; Lucon-Xiccato & Bisazza, 2016; Mazza et al., 2018; Zidar et al., 2018). We acknowledge that this interpretation strictly requires the implicit assumption that the contribution of any other factors to among-individual variation (e.g., motivation, energetic state, experience previous to the experiment; Rowe & Healy 2014) is negligible relative to differential cognitive performance. We attempted to mitigate against other sources of among-individual variation as far as possible using standardised housing and husbandry conditions. Following completion of spatial learning trials using Maze A, individuals were tested for stress responsiveness three times each over a three-week period using Open Field Trials (OFT) with a mean (range) of 4 (1-5) days between successive trials. Finally, fish were retested in a second maze (Maze B) with a different layout, and repeat trials conducted (as before) one per day for 11 consecutive days. Thus, in total, the design called for all individuals to complete 22 spatial learning trials, 11 on each of two different maze layouts (distributed across two different mazes) and three OFT over a total testing period of 43 days. Note that the sample size declined slightly across the experiment as (i) a few mortalities occurred naturally within the testing period and, (ii) we proactively ‘retired’ any fish not deemed to be feeding well and behaving normally in their home tanks as a precaution against cumulative adverse effects. Thus 63 fish experienced Maze A, which declined to n=60 at trial 11 and OFT testing. Five fish were then removed prior to experiencing Maze B (n= 55 at trial 1 and n=53 at trial 11).

Spatial Learning Trials

In order to facilitate more rapid data collection, a single aquarium (25 x 45 x 25cm) was divided into two, with each half containing an identical version of maze A (A1, A2). Two replicates of maze B were similarly constructed (Figure 1). This allowed two fish to be tested concurrently during trials. Each maze consisted of 6 opaque Perspex panels (8 cm), spaced 5cm apart (Fig 1). A visually transparent perforated panel at one end of each maze was used to separate a small holding area (12.5 x 10 x 25 cm) contain two adult females selected haphazardly from stock. During trials the experimental maze tanks were lit from below by one fluorescent lamp and filled to a depth of 8 cm with room temperature water (approx. 23-24 °C). The water was taken from the same recirculating system used to house the male groups and was changed between each housing group (i.e. after every 4 runs with two fish trialled per run). Stimulus females were also changed at the same time.

At each trial, two males were individually netted from their home tank and quickly identified from natural markings. Each was randomly allocated to one of the two maze replicates and carefully placed within a perforated plastic tube in the ‘start’ zone (Figure 1). They were given 60 s to acclimate before the plastic tubes were removed. A Sunkwang C160 video camera mounted above the tank allowed the fish to be observed without disturbance. Tracking software (<http://www.biobserve.com>) was then used to determine the *start latency* as the (post-acclimation) time taken before a fish started the maze by leaving the ‘start’ zone, and *maze time* as the latency from starting to completing the maze (with completion defined as reaching the ‘end’ zone; Figure 1). On reaching the ‘end’ zone individuals were given 60 s undisturbed visual access to the females before an opaque plastic sheet was inserted to obstruct females from view. Following the 60 s reward period, fish were netted and returned to the home tank. To ensure standardized exposure to the reward stimulus, individuals that did not complete the maze within 480 s post-acclimation period (irrespective of whether they had started) were gently guided through the maze to the end zone using a net behind them and then experienced 60 s visual access to the females. Following the 60 s reward

period, fish were netted and returned to the home tank. These fish were assigned a right censored value of 480 second for *maze time*.

Figure 1. Aerial view of the maze designs used in the experiments (A and B), each tank was split into two identical mazes (1 and 2).

Open Field Trial (OFT)

OFTs to characterise stress responsiveness closely followed the protocol described in White et al. (2016). For each trial, a single individual was netted from the home tank, quickly identified and introduced gently into the centre of an open arena (a 30 × 20 cm tank filled to 5 cm water placed on a lightbox). A cardboard screen was around the tank prevented visual disturbance and a Sunkwang C160 video camera mounted above the arena again allowed movement to be tracked. Following a 30 s acclimation period, individuals' movements were tracked for 4 minutes and 30 s to determine *track length* (total distance swum (cm)) and *area covered* (percent of tank area covered). These two observed behaviours which are known to be repeatable and heritable in this population (Houslay et al., 2018; White et al., 2019; White & Wilson, 2019), were used to calculate the derived trait of *relative*

area following Houslay et al. (2019). *Relative area* is the observed area covered in the trial minus the expected area covered under a simulated ‘random swim’ of length equal to the observed track length (see Houslay et al. (2019) for further detail on simulations). Low values of *relative area* result from a ‘flight type’ behavioural stress response in which individuals swim rapidly (yielding a high track length) but exhibit thigmotaxis (staying close to the walls and seeking escape from the arena) and thus cover relatively little of the arena area. In contrast low values of *relative area* correspond to efficient exploration (i.e. a high proportion of the arena covered given distance swum), by putatively less stressed fish.

Statistical Analysis

Data from both types of behavioural assay were analysed using univariate and multivariate linear mixed effect models fitted by REML (restricted maximum likelihood) using ASReml within R (<http://www.vsni.com>) (Gilmour et al., 2009). By including individual identity as a random effect in these models we test for and characterise among-individual (co)variation. Traits were mean centred and scaled to standard deviation units to ease interpretation of results and facilitate convergence of multivariate models. For *maze time* we did this using the overall mean and standard deviation of observations from both mazes in order to preserve any meaningful differences in performance between A and B. With traits in standard deviation units (sdu), estimates of among-individual variance (V_{ind}) can be interpreted as repeatabilities (i.e. proportion of the observed phenotypic variance explained by among-individual differences). However, we also calculate estimates of adjusted repeatability (R), the proportion of phenotypic variance explained by consistent among-individual differences, after controlling for fixed effects on the mean (Nakagawa & Schielzeth, 2010). Thus $R = V_{ind} / (V_{ind} + V_R)$ where V_R is the residual (within-individual) variance estimated from each model. The significance of random effects was tested using likelihood ratio tests (LRT), while fixed effects (included in the various models as described below) were tested using conditional F-statistics. All models assumed Gaussian error structures, an assumption that was deemed acceptable based on visual inspection of the model residuals.

273

274 *Univariate analyses of maze performance and spatial learning*

275 We use *maze time* as our observed measure of performance. Here we describe in full the
 276 univariate analysis of data collected in maze A (subsequently *maze time_A*). Identical procedures were
 277 then applied to data from maze B. First, we visualised the distribution of *maze time_A* across repeat
 278 using box plots and also plotted the proportion of mazes completed as a function of repeat to see if a
 279 pattern of increasing average performance (i.e. decreasing *maze time* and/or increasing proportion of
 280 successful completion) was immediately apparent. Next a series of three nested models with identical
 281 fixed effects but differing random effect structure were fitted to the centred and scaled *maze time_A*
 282 data. All models included a fixed effect of *trial number* (the cumulative number of trials experienced
 283 by an individual, treated as a continuous variable), allowing us to test for improvement in the mean
 284 (indicative of learning). Additional fixed effects were included as statistical controls for potential
 285 sources of variance not relevant to hypotheses being tested here. These included time of day (in
 286 minutes after 9 am), maze replicate (as a factor denoting position 1 or 2 in maze tank), and order
 287 caught from the home tank. The latter was to account for any cumulative disturbance effect of
 288 removing fish sequentially from the home tank and/or build-up of chemical cues in the maze between
 289 water changes.

290 The first model contained no random effects, while the second contained a random intercept
 291 of individual identity. Likelihood ratio test (LRT) comparison of these models was conducted to test
 292 the hypothesis that individuals differ in their average performance (*maze time_A*) across the 11 repeats,
 293 and we estimated the (adjusted) repeatability of performance under the second model. For the LRT we
 294 assume twice the difference in model log-likelihoods is distributed as a 50:50 mix of X^2_1 and X^2_0
 295 following Stram & Lee (1994). The third model was a first order random regression (i.e. a random
 296 slope and intercept model) in which each individual's deviation from the fixed effect mean *maze time*
 297 can change as a linear function of *trial number* (1-11). Variation in random slopes means that there is
 298 among-individual variation around the mean *maze time_A* - *trial number* relationship. Thus, LRT

comparison of the second and third models thus provides a test for among-individual variation in learning rate. This comparison is conducted assuming the test statistics is distributed as X^2_2 since the third model has two extra parameters (a slope variance and a slope-intercept covariance). Note that among-individual variance in slopes cannot be scaled to a repeatability as within individual variance in slope is not estimable (using data from a single maze; see below). Nor is its magnitude directly comparable to random intercept variance since slopes and intercepts are in different units. However, under the third model, among-individual variance in learning (slope) means that among-individual variance *maze time_A* changes with *trial number* (Supplemental Materials Figure 1). Thus, to understand the biological effect size of estimated variance in slopes, we use the third model to predict among-individual variance (V_{ind}) and adjusted repeatability (R) of *maze time_A* at both initial (trial 1) and final (trial 11) performance (following e.g., Nussey et al. (2007); see Supplemental Materials Table 3 for didactic explanation and corresponding code). We note that among-individual variation at final performance has been used to infer differences in cognitive ability in studies adopting similar repeated measures designs (e.g. Langley et al. 2020) and so also has a useful biological interpretation here.

Univariate analysis of relative area

To verify our expectation that individuals would show consistent differences in stress responsiveness, we fit a simple random intercepts model to (scaled and centred) *relative area*. This model included fixed effects of trial number (1-3), and time of day (in minutes after 9 am in which each trial took place) as well as a random effect of individual identity as. Adjusted repeatability (R) of *relative area* was calculated and the significance of among individual variance tested by LRT comparison to a simplified model with no random effect (assuming the test statistic was distributed as a 50:50 mix of X^2_1 and X^2_0 as above).

Multivariate modelling of Maze A, Maze B and OFT data combined

Finally, to test the predicted correlation structure between cognitive performance and stress responsiveness, we formulated a trivariate mixed model in which the three response variables were *maze time_A*, *maze time_B* and *relative area*. Fixed effects were exactly as described above on all three traits. Random effects were also as described above (i.e. individual level random intercepts and slopes for *maze time_A* and *maze time_B* but a random intercept only for *relative area*) but the multivariate formulation allowed us to estimate the full 5x5 among-individual covariance matrix (**ID**) among these effects. Since each observation of a fish provided data on a single trait only, residual covariances among traits were fixed to zero. After fitting the model, we compared it to a simplified fit in which all among-trait covariance elements in **ID** were constrained to zero. This provides a global test of individual covariance between traits. We then scaled estimated pairwise covariances in **ID** to their corresponding correlations for easier interpretation (noting for a pair of effects x, y the correlation $r_{xy} = \text{COV}_{xy} / (\text{V}_x \text{V}_y)^{0.5}$). This allowed us to scrutinise the correlation structure between stress responsiveness and cognitive performance in both mazes A and B, using both final performance and learning rate (i.e. random regression slope) as measures of cognition. Additionally, it allowed us to estimate the individual level correlation in cognitive performance measures (final *maze time* performance, learning) across mazes. These are not strictly equivalent to individual repeatabilities of cognitive performance measures across mazes (as opposed to individual repeatability of *maze time* across trials within mazes) because estimates could be negative. However, they can be readily interpreted in those terms; a strong positive correlation between, for example, individual *learning* in maze A and maze B means this latent variable is highly repeatable across mazes. Conversely, a negative correlation means that individuals learning faster in maze A tend to learn more slowly in maze B (and *vice versa*).

Results

Performance in Maze A

Plots of the raw data suggest that average time to complete Maze A decreases across trials, and that the success rate (proportion of individuals completing the tasks within the 480 s) tends to increase (Figure 2). These patterns are qualitatively consistent with expectations if (average) performance improves as a consequence of learning. The mixed model analysis of *maze time_A* confirms statistical support for this with a significant negative effect of trial repeat number (based on the full random slope and intercept model; coefficient = -0.043 (0.014) sdu, $F_{1,59,8} = 10.140$, $P = 0.003$). This effect size equates to an estimated decrease of 91.9 seconds in average *maze time* over the 11 trials. Other fixed effects of order caught and maze position were non-significant (see Supplementary Material Table S1). Likelihood ratio tests (LRT) confirmed among-individual variation in *maze time_A* (comparison of null and random intercept models; $\chi^2_{0,1} = 155$, $P \leq 0.001$). Under the random intercept model, repeatability of *maze time_A* conditional on fixed effects was estimated as $R_A = 0.343$ (0.05).

LRT comparison of the random intercept and first order random regression models showed the latter to be a significantly better fit to the data ($\chi^2_2 = 25.0$, $P \leq 0.001$). This comparison provides evidence for among-individual variance in the rate of change of *maze time_A* across repeated trials (interpretable, with caveats discussed below, as variation in rate of learning). Among-individual variance in intercepts (int) and slope (slp) were estimated as $V_{ind_{int}} = 0.394$ (0.102) and $V_{ind_{slp}} = 0.006$ (0.002) respectively while the among-individual intercept–slope correlation was estimated as ($r_{ind_{int},ind_{slp}} = -0.489$ (0.147)). Biological interpretation of these parameters is not completely straightforward. Given the scaling of *trial number* in the random effect structure of the model (see Supplemental Materials Table S3) $V_{ind_{int}}$ is interpretable as among individual variance in *maze time_A* at first trial. While slope variance is in different units and thus not of directly comparable magnitude, variation in slopes actually means that among-individual variance in the observed trait (V_{ind} for *maze time_A*) changes with trial repeat number. Here the random regression model predicts values of

$V_{ind_{A1}} = 0.394$ (0.102), and $V_{ind_{A11}} = 0.542$ (0.131)) at first and last trial in maze A respectively, suggesting more among individual variation in performance at the end of trials than at the beginning. The corresponding predictions of repeatability at first and last observed trial are $R_{A1} = 0.431$ (0.070) and $R_{A11} = 0.511$ (0.067). The negative intercept-slope correlation ($r_{ind_{A.int,A.slp}} = -0.489$ (0.147), $\chi^2_2 = 6.182$, $P = 0.045$), means that individuals with higher intercepts (high *maze time_A* at trial 1, tended to have lower slopes (i.e., more negative, indicative of faster learning). These patterns are represented visually in Figure 3, which shows the individual reaction norms predicted from the best linear unbiased predictions (BLUPs) of random intercept and slope for each fish (following e.g., Houslay & Wilson (2017)).

Figure 2. Plots of raw data of *maze time* across both maze designs. Boxplots (a) and (c) show the data distributions for time to complete Maze A and Maze B respectively across the 11 trials. Black boxes display data of all individuals and grey boxes represent only those individuals that successfully

completed the task within 480 s. Horizontal lines within box correspond to behavioural medians, box boundaries correspond to first and third quartiles. When present, whiskers correspond to 10th and 90th percentiles, and points correspond to outliers. Plots (b) and (c) represent mean and standard errors for time to complete Maze A and B respectively. Colours represent the same groups; black error bars represent mean and standard errors of *maze time* for all individuals, and grey represent only those individuals that successfully completed the maze in the allocated time.

Performance in Maze B

In contrast to Maze A, plotting *maze time_B* data reveals no clear increase in performance (i.e. decrease in time) across trials. Furthermore, there is actually a trend towards fewer individuals successfully completing the task (Figure 2). However, we note that if the censored data points are excluded to leave only successfully complete trials, there is a decreasing trend in *maze time_B* with trial number. The mixed model analysis (which uses data from all trials) confirms the lack of improvement in the mean *maze time_B*, with a (non-significant) positive estimate of the trial repeat number effect (from random slope and intercept model; coefficient = 0.014 (0.014), $F_{1,538,2} = 1.193$, $P = 0.301$). Effects of order caught and maze position were not significant (Table S2). Likelihood ratio tests (LRT) between the univariate random intercept model and the null model with no random effect, shows the presence of significant among-individual variation for *maze time_B* ($\chi^2_{0,1} = 182.041$ $P \leq 0.001$), with a corresponding repeatability estimate of $R_B = 0.401$ (0.055). The random slope model was a significantly better fit again ($\chi^2_2 = 9.995$ $P = 0.007$) providing evidence of among-individual variation in the performance-trial number relationship. Among-individual variance in intercepts (int) and slope (slp) were estimated as 0.472 (0.130) and 0.004 (0.002) respectively. These estimates mean predicted values of $V_{ind_{B1}} = 0.472$ (0.130) and $V_{ind_{B11}} = 0.635$ (0.162)) which correspond to repeatabilities of $R_{BI} = 0.439$ (0.074) and $R_{BII} = 0.512$ (0.071). Given that there is no (significant) effect of trial number on mean *maze time_B* the presence of among-individual variance in slope suggest that some individuals are improving (consistent with learning) while for others performance is tending to get worse across repeats in Maze B. Furthermore, the among-individual intercept–slope correlation was non-significant as ($r_{ind_{B.int,B.slp}} =$

-0.302 (0.214), $\chi^2_2 = 1.476$, $P = 0.478$). The predicted patterns are again represented visually by plotting the individual reaction norms (Figure 3).

Figure 3. Spatial learning traits across Maze A and Maze B as a function of trial number, *maze time_A* (a), and *maze time_B* (b). Grey lines represent individual predicted reaction norms (BLUPs) from univariate random slope models for each trait. Coloured lines are used to illustrate reaction norms for a small random set of arbitrarily chosen individuals tested in both mazes. Black dashed line represents the trend in fixed effect mean *maze time* across repeat trials.

427

428 *Among-individual differences in OFT behaviour*

429 We found evidence of significant among-individual variation in *relative area*,
 430 (*repeatability*(with *SE*), $R = 0.465$ (0.089), $\chi^2_{0,1} = 20.421$, $P < 0.001$). This replicates previous
 431 findings in the same population (Prentice et al., 2020) though the current estimate of repeatability is
 432 somewhat higher, likely due to differences in study design (e.g. the current study used a shorter inter-
 433 observation interval and was limited to males only). Fixed effects from the OFT behaviour models are
 434 presented in the Supplementary Material (Table S2) for completeness, although are not directly
 435 relevant to our hypotheses in this study.

436 *Multivariate model*

437 The full multivariate model (**ID**) of *maze time_A*, *maze time_B* and *relative area* provides
 438 evidence of some significant among-individual covariance structure between observed traits
 439 (comparison of the full model to one in which all among-individual between trait covariances are
 440 fixed to zero; $\chi^2_8 = 44.094$, $P \leq 0.001$). Examination of the estimated covariances and correlations
 441 (Table 1) suggests this result is largely driven by a strong positive correlation between the individual
 442 intercepts for *maze time_A* and *maze time_B* ($r_{ind_{A,int},B,int} = 0.686$ (0.135)). In other words
 443 performance at first trial is positively correlated at the individual level across mazes (since
 444 $r_{ind_{A,int},B,int} = r_{ind_{A1},B1}$). Using the multivariate random regression model to predict the corresponding
 445 correlation at final trial (i.e. trial 11), performances across mazes yields an estimate (SE) of
 446 $r_{ind_{A11},B11} = 0.602$ (0.131). Thus, our results suggest strong positive among-individual correlation
 447 of performance as measured by maze time across trials and mazes. This is not only the case for first
 448 and last performance, but also for intermediate trial numbers as can be shown by transforming the **ID**
 449 estimate from the random regression model (as shown in Table1) to a ‘character state’ correlation
 450 matrix among the full set of trials and maze specific observations, and *relative area* (see
 451 Supplementary Material Table S3 for this matrix and an explanation of the transformation).

However, returning to a reaction norm interpretation of results, we do not find evidence that reaction norm slopes (i.e. putative rates of learning) are correlated across mazes. While the multivariate model corroborates the presence of among-individual slope variance in mazes A and B, the correlation between them was only weakly positive and non-significant ($r_{ind_{A.slp,B.slp}} = 0.216$ (0.266); Table 1). Nor do we find statistical support for among-individual correlation between maze performance intercepts or slopes (for either maze) and relative area.

Table 1. Among individual variance–covariance–correlation matrix from the final trivariate model of *maze time_A*, *relative area* and *maze time_B*. Variances are shown on the diagonal (dark grey shading), with covariances below and correlations above. Light grey shading denotes within trait covariance/correlation estimates (i.e. between reaction norm intercepts and slopes). Standard errors are shown in parentheses and bold font denotes nominally significant pairwise estimates assuming approximate 95% CI of $\pm 1.96SE$).

	<i>Maze time_A</i>		<i>Relative area</i>	<i>Maze time_B</i>	
	<i>intercept_A</i>	<i>slope_A</i>		<i>intercept_B</i>	<i>slope_B</i>
<i>intercept_A</i>	0.436 (0.113)	-0.489 (0.147)	0.286 (0.175)	0.686 (0.135)	-0.129 (0.254)
<i>slope_A</i>	-0.027 (0.013)	0.007 (0.003)	-0.075 (0.202)	-0.006 (0.209)	0.216 (0.266)
<i>Relative area</i>	0.127 (0.084)	-0.004 (0.011)	0.451 (0.118)	0.024 (0.125)	0.377 (0.231)
<i>intercept_B</i>	0.299 (0.091)	-0.003 (0.012)	0.011 (0.085)	0.437 (0.119)	-0.309 (0.212)
<i>slope_B</i>	-0.005 (0.010)	0.001 (0.001)	0.016 (0.011)	-0.013 (0.012)	0.004 (0.002)

Discussion

Here, we show evidence of among-individual differences in performance – measured as time to complete a maze – in guppies exposed to a spatial learning test paradigm. Performance of individuals is repeatable both within, and across, the two spatial learning tasks (i.e. mazes) presented. However, the question of whether there is robust evidence of learning, on average or by individual fish, is somewhat less clear cut. In particular, in the first maze used (A) we find evidence of improvement in mean performance consistent with learning (on average). We also find among-individual variation in this rate of improvement, and so – putatively their rate of learning. However, the same fish exposed to maze B show (on average), no increase in performance across successive trials. We found among-individual correlation structure between performances (i.e. time in the maze) but not learning (i.e. rate of improvement) across the 2 spatial learning tasks. We did not however find any significant association between individual differences in maze performance (or learning) and repeatable stress responsiveness as measured in the open field trials. In what follows we describe each of these findings in more detail and discuss them in the wider context of the cognitive literature.

The data from Maze A show that on average, time to complete the maze improves across repeated trials. This improvement suggests that spatial learning is occurring in the guppies, a finding consistent with previous studies of this species (Fong et al., 2019; Kotrschal et al., 2015; Lucon-Xiccato & Bisazza, 2017c). We also see evidence of consistent, repeatable differences among-individuals in performance in Maze A. This is shown in our reaction norm models as significant among-individual variance in intercept, which strictly represents performance at first trial. However, using among-individual variation in intercepts and slope to predict the corresponding variance at, and correlation among-, all trials (see Supplemental Materials Table 3 for derivation and presentation of these estimates) reveals that in fact individual performance is positively correlated across all trials from 1 to 11. In simple terms, fish that are faster than average at completing Maze A in their first trial, tend to be faster than average across all subsequent trials too. Predicted repeatability of *maze time* is moderately high relative to many behavioural studies (e.g., 43% at trial 1, 51% at trial 11) but broadly comparable to estimates reported from similar assays designed to test cognitive variation; see

Cauchoix, Hermer, Chaine, & Morand-Ferron (2017) for an overview. We note that a contributing factor is likely to be short inter-observation period (here 24 hrs) typical of cognitive studies, since behavioural repeatabilities generally declines as this increases (Boulton et al., 2014).

Accepting that improvement across repeated trials can be interpreted as learning (caveats to this are discussed below), our random regression model also provides evidence for among-individual variation in spatial learning in Maze A. Usefully, our modelling strategy allowed all observations to contribute to estimating variance in the latent cognitive trait (learning) while avoiding statistically problematic ‘two-step’ analysis (Houslay & Wilson, 2017). Although this strategy is now widely used in studies of behavioural plasticity, it has not yet been widely adopted by researchers focussing specifically on animal cognition (but see e.g., Langley et al., 2020). In addition to finding variance in slopes (learning), we estimated a negative among-individual intercept-slope correlation using the Maze A data; individuals with higher intercepts (i.e. *maze time* at first trial) tend to have lower (more negative) slopes. While it is therefore the case that those fish performing poorly initially exhibit higher rates of learning, it is also true - as noted above - that individual performance (*maze time*) is positively correlated across trials 1-11. These two results are entirely compatible because differences in learning (slope) are not sufficiently pronounced that initially poor performing (but fast learning) fish will generally ‘overtake’ initially good performing (but slow learning) individuals in expected time to complete the maze by trial 11. We cannot comment on what fitness consequences, if any, the variation detected here would have in wild fish. Nonetheless, this finding does highlight a danger with the general presumption that cognitive abilities are under positive selection. Here, if we assumed that fitness benefits were accrued by rapidly achieving a spatial task (e.g. locating a resource) regardless of mechanism, it would be the slower learners that were advantaged. Thus, while it is tempting to assume fast learners will achieve better outcomes, they may sometimes simply be those with the ‘most room for improvement’.

Thus, findings from Maze A are consistent with our initial predictions that time to complete the maze would improve (on average) with experience due to spatial learning, but that individuals would also vary in both performance (*maze time*) and learning (rate of change in performance with

experience). We also found that individuals that were quicker (over all trials) to complete Maze A, tended to be quicker (over all trials) to complete Maze B. While this could be attributable to cognitive differences, there are certainly other possibilities. For instance more explorative and/or less neophobic individuals may be generally faster at solving tasks (Boogert et al., 2006; Bousquet et al., 2015; Zidar et al., 2018). Similarly there could be among-individual variation in perceived cue salience (Meyer et al., 2012), individual physiology (Bókony et al., 2014), or motivation (van Horik & Madden, 2016). Regardless, of these unknowns, an important difference between Maze A and Maze B was that we found no evidence of learning on average in the latter. In fact, for Maze B the mean *maze time* actually increased slightly, though not significantly, across trials. Despite this, patterns of individual variation around the mean trajectory were largely similar to those found in Maze A. Thus, there is among-individual variation in intercept (*maze time* at trial 1) and also in slope. Given that there is no (significant) change in mean performance, but there is significant variation in slopes, we conclude that some individuals are improving (learning) in Maze B while others are getting worse with experience. We also note that, as in Maze A, slope variance is present, but not sufficiently high to break down the positive correlation structure of individual performance (*maze time*) across trials 1-11.

Although we did not formally test for differences in average slope between maze A and B, we note that approximate 95% confidence intervals do not overlap (estimated as $\text{coefficient} \pm 1.96\text{SE}$). Several possibilities may explain the finding of spatial learning on average in A but not B. First, the results from maze A may be a false positive (Fraser et al., 2018; Sterne & Smith, 2001). However coinciding with previous studies which show this species is capable of learning an initial spatial learning task (Fong et al., 2019; Kotrschal et al., 2015; Lucon-Xiccato & Bisazza, 2017b), we assume this unlikely. Second, it may be that the layout of maze B was, in some unknown way, more challenging to learn. This could certainly be true if, for instance learning to navigate a new maze following the acquisition of a previously learnt layout, poses a more challenging task such that more trials would be required to detect improvement. There is some evidence for such effects in guppies. For instance, Lucon-Xiccato & Bisazza (2014) found that on average guppies took 14.61 trials to

learn a reversed colour cue association, while Fong et al., (2019) found that on average, 15.30 trials were required for guppies to learn a reversed maze layout.

A third, possible explanation is that learning does lead to gains in maze B performance, but that these are being masked at the level of the sample mean by concurrent changes in aspects of average individual ‘state’ that reduces cognitive performance and/or motivation. One plausible hypothesis is that chronic stress responses arise cumulatively from repeated capture and handling necessitated by the experimental design (Huntingford et al., 2006; Warren & Callaghan, 1976; Wong, Dykstra, Campbell, & Earley, 2008). If so, this could negatively impact affected individuals and offset expected improvements in mean performance across trial number. Presently we cannot directly test this possibility, and variation in susceptibility to chronic stress response is not well understood. Nonetheless, our experiment does confirm repeatable among-individual variation ($R = 0.465$ (0.089) in *relative area* covered in the OFT, used here as a measure of acute behavioural stress response. This replicates previous results using independent data sets of fish from the same captive population (Houslay et al., 2019; Prentice et al., 2020; White et al., 2016). Acute stressor exposure has been shown to affect cognitive performance in spatial learning tasks in both mammals and fish (Gaikwad et al., 2011; R. Y. Wong et al., 2019). At the individual level, there is also evidence to suggest short-term measures of acute stress responses can predict longer term organismal performance under chronic and/or repeated stressor exposure (Øverli et al., 2007; Salak-Johnson & Mcglone, 2007; Segerstrom & Miller, 2004).

Here, our modelling approach did not provide compelling statistical support for strong relationships between *relative area* and either initial *maze time*, or learning (i.e. improvement in *maze time*) in either Maze A or B. However, the estimated correlation between *relative area* and *maze time* in Maze B actually rises to $r_{ind} = 0.336$ (0.169) by trial 11 and thus approaching nominal significance at $\alpha = 0.05$ (assuming a lower 95% CI of $r_{ind} - 1.96SE$). To explore this further we conducted a *post hoc* likelihood ratio test comparison of a bivariate model of *maze time_B* and *relative area*; all effects as described for the trivariate model earlier) to the corresponding model fit where among-individual covariances between *relative area* and *maze time_B* (intercept and slope) were constrained to zero. This

confirmed no significant covariance between *relative area* and *maze time_B* ($\chi^2_2 = 3.098$, $P = 0.212$). Thus, we do not find statistical support for the prediction, made under the stress coping style model, that (acute) stress responsiveness will (co)vary with cognitive performance (Coppens et al., 2010; Griffin et al., 2015; Sih & Del Giudice, 2012). Nonetheless, the possibility that chronic stress negatively impacts apparent learning cannot be completely excluded here. Empiricists rightly seek to minimise the possibility of stress confounding conclusions from cognitive studies. However, we suggest the assumption that individuals remain (equally) ‘unstressed’ over experimental periods requiring repeated observations (and often repeated capture and or social isolation) is difficult to validate in practice.

In summary, here we have evidence of consistent differences among-individuals in spatial task performance in the guppy *P. reticulata*. Individual performance is repeatable across trials within- and between two different spatial tasks (i.e. maze layouts). This among-individual variation in performance may well be mediated by cognitive factors but differences in ‘personality’ (e.g. neophobia, exploratory tendency) may also contribute. We also find evidence of improved performance with experience, consistent with spatial learning. In both tasks variation around the trajectory of mean performance across trial number was present. While this means individuals can be considered as differing in ‘spatial learning rate’ it is important to note that performance declines for some individuals, especially in the second maze where there was no improvement in average time across 11 trials. We show here that an individual’s (repeatable) behavioural response to an acute stress stimulus does not predict either average performance in the maze or learning rate. However, we suggest the possibility that cumulative, chronic stress effects may contribute to declining performance (or reduced improvement) in our study. If individuals generally differ in susceptibility to chronic stress, this may represent a widespread but currently poorly acknowledged challenge for characterisation of cognitive variation in animal studies.

References

- 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.
<https://doi.org/10.1038/nature25503>
- 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. <https://doi.org/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. <https://doi.org/10.1016/j.anbehav.2008.12.022>
- Bensky, M. K., Paitz, R., Pereira, L., & Bell, A. M. (2017). Testing the predictions of coping styles theory in threespined sticklebacks. *Behavioural Processes*, 136, 1–10.
<https://doi.org/10.1016/j.beproc.2016.12.011>
- Bókony, V., Lendvai, Á. Z., Vágási, C. I., Pătraș, L., Pap, P. L., Németh, J., Vincze, E., Papp, S., Preiszner, B., Seress, G., & Liker, A. (2014). Necessity or capacity? Physiological state predicts problem-solving performance in house sparrows. *Behavioral Ecology*, 25(1), 124–135.
<https://doi.org/10.1093/beheco/art094>
- Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. In *Philosophical Transactions of the Royal Society B: Biological Sciences* 373 (1756), 20170280. <https://doi.org/10.1098/rstb.2017.0280>
- Boogert, N. J., Reader, S. M., & Laland, K. N. (2006). The relation between social rank, neophobia and individual learning in starlings. *Animal Behaviour*, 72(6), 1229–1239.
<https://doi.org/10.1016/j.anbehav.2006.02.021>
- Boulton, K., Grimmer, A. J., Rosenthal, G. G., Walling, C. A., & Wilson, A. J. (2014). How stable are personalities? A multivariate view of behavioural variation over long and short timescales in the

621 sheepshead swordtail, *Xiphophorus birchmanni*. *Behavioral Ecology and Sociobiology*, 68(5),
622 791–803. <https://doi.org/10.1007/s00265-014-1692-0>

623 Bousquet, C. a. H., Petit, O., Arrivé, M., Robin, J.-P., & Sueur, C. (2015). Personality tests predict
624 responses to a spatial-learning task in mallards, *Anas platyrhynchos*. *Animal Behaviour*, 110,
625 145–154. <https://doi.org/10.1016/j.anbehav.2015.09.024>

626 Bridger, D., Bonner, S. J., & Briffa, M. (2015). Individual quality and personality: bolder males are
627 less fecund in the hermit crab *Pagurus bernhardus*. *Proc Biol Sci*, 282(1803), 20142492.
628 <https://doi.org/10.1098/rspb.2014.2492>

629 Brown, C., & Braithwaite, V. A. (2005). Effects of predation pressure on the cognitive ability of the
630 poeciliid *Brachyrhaphis episcopi*. *Behavioral Ecology*, 16(2), 482–487.
631 <https://doi.org/10.1093/beheco/ari016>

632 Brust, V., & Guenther, A. (2017). Stability of the guinea pigs personality – cognition – linkage over
633 time. In *Behavioural Processes* 134, 4–11. <https://doi.org/10.1016/j.beproc.2016.06.009>

634 Buchanan, K., Perera, T. B. de, ... C. C.-A., & 2012, U. (2020). Guidelines for the treatment of
635 animals in behavioural research and teaching. *Animal Behaviour*, 159, I–XI.
636 <https://doi.org/10.1016/j.anbehav.2019.11.002>

637 Burkart, J. M., Schubiger, M. N., & Van Schaik, C. P. (2017). The evolution of general intelligence.
638 *Behavioral and Brain Sciences*, 40, e195. <https://doi.org/10.1017/S0140525X16000959>

639 Burns, J. G., & Rodd, F. H. (2008). Hastiness, brain size and predation regime affect the performance
640 of wild guppies in a spatial memory task. *Animal Behaviour*, 76(3), 911–922.
641 <https://doi.org/10.1016/j.anbehav.2008.02.017>

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

- 645 Carere, C., Maestripieri, D., Caramaschi, D., Carere, C., Sgoifo, A., & Koolhaas, J. M. (2014).
646 Neuroendocrine and Autonomic Correlates of Animal Personalities. In *Animal Personalities* (pp.
647 353–377). <https://doi.org/10.7208/chicago/9780226922065.003.0013>
- 648 Cauchoix, M., Hermer, E., Chaine, A. S., & Morand-Ferron, J. (2017). Cognition in the field:
649 Comparison of reversal learning performance in captive and wild passerines. *Scientific Reports*,
650 7(1), 12945. <https://doi.org/10.1038/s41598-017-13179-5>
- 651 Cole, E. F., Cram, D. L., & Quinn, J. L. (2011). Individual variation in spontaneous problem-solving
652 performance among wild great tits. *Animal Behaviour*, 81(2), 491–498.
653 <https://doi.org/10.1016/j.anbehav.2010.11.025>
- 654 Coppens, C. M., de Boer, S. F., & Koolhaas, J. M. (2010). Coping styles and behavioural flexibility:
655 towards underlying mechanisms. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1560), 4021–4028. <https://doi.org/10.1098/rstb.2010.0217>
- 657 Dingemanse, N. J., Both, C., Drent, P. J., & Tinbergen, J. M. (2004). Fitness consequences of avian
658 personalities in a fluctuating environment. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 271(1541), 847–852.
- 660 Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in behaviour:
661 Mixed-effect modelling approaches. *Journal of Animal Ecology*, 82(1), 39–54.
662 <https://doi.org/10.1111/1365-2656.12013>
- 663 Dingemanse, N. J., & Réale, D. (2005). Natural selection and animal personality. *Behaviour*, 142(9–
664 10), 1159–1184.
- 665 Fong, S., Buechel, S. D., Boussard, A., Kotrschal, A., & Kolm, N. (2019). Plastic changes in brain
666 morphology in relation to learning and environmental enrichment in the guppy (*Poecilia*
667 *reticulata*). *Journal of Experimental Biology*, 222(10). <https://doi.org/10.1242/jeb.200402>
- 668 Fraser, H., Parker, T., Nakagawa, S., Barnett, A., & Fidler, F. (2018). Questionable research practices

- 669 in ecology and evolution. *PLoS ONE*, 13(7), e0200303.
 670 <https://doi.org/10.1371/journal.pone.0200303>
- 671 Gaikwad, S., Stewart, A., Hart, P., Wong, K., Piet, V., Cachat, J., & Kalueff, A. V. (2011). Acute
 672 stress disrupts performance of zebrafish in the cued and spatial memory tests: The utility of fish
 673 models to study stress-memory interplay. *Behavioural Processes*, 87(2), 224–230.
 674 <https://doi.org/10.1016/j.beproc.2011.04.004>
- 675 Galsworthy, M. J., Paya-Cano, J. L., Liu, L., Monleón, S., Gregoryan, G., Fernandes, C., Schalkwyk,
 676 L. C., & Plomin, R. (2005). Assessing reliability, heritability and general cognitive ability in a
 677 battery of cognitive tasks for laboratory mice. *Behavior Genetics*, 35(5), 675–692.
 678 <https://doi.org/10.1007/s10519-005-3423-9>
- 679 Gibelli, J., Aubin-Horth, N., & Dubois, F. (2019). Individual differences in anxiety are related to
 680 differences in learning performance and cognitive style. *Animal Behaviour*, 157, 121–128.
 681 <https://doi.org/10.1016/j.anbehav.2019.09.008>
- 682 Gilmour, A. R., Gogel, B. J., Cullis, B. R., Thompson, R., & Butler, D. (2009). ASReml user guide
 683 release 3.0. *VSN International Ltd, Hemel Hempstead, UK*.
- 684 Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research?
 685 *Psychological Bulletin*, 127(1), 45–86. <https://doi.org/10.1037/0033-2909.127.1.45>
- 686 Griffin, A. S., Guillette, L. M., & Healy, S. D. (2015). Cognition and personality: An analysis of an
 687 emerging field. In *Trends in Ecology and Evolution* 30 (4), 207–214.
 688 <https://doi.org/10.1016/j.tree.2015.01.012>
- 689 Guillette, L. M., Hahn, A. H., Hoeschele, M., Przyslupski, A. M., & Sturdy, C. B. (2015). Individual
 690 differences in learning speed, performance accuracy and exploratory behaviour in black-capped
 691 chickadees. *Animal Cognition*, 18(1), 165–178. <https://doi.org/10.1007/s10071-014-0787-3>
- 692 Healy, S. D. (2019). The face of animal cognition. In *Integrative Zoology* 14 (2), 132–144. John

693 Wiley & Sons, Ltd. <https://doi.org/10.1111/1749-4877.12361>

694 Houslay, T. M., Earley, R. L., White, S. J., Lammers, W., Grimmer, A. J., Travers, L. M., Johnson, E.
695 L., Young, A. J., & Wilson, A. J. (2019). Genetic integration of the stress response. *BioRxiv*,
696 770586. <https://doi.org/10.1101/770586>

697 Houslay, T. M., Vierbuchen, M., Grimmer, A. J., Young, A. J., & Wilson, A. J. (2018). Testing the
698 stability of behavioural coping style across stress contexts in the Trinidadian guppy. *Functional*
699 *Ecology*, 32(2), 424–438. <https://doi.org/10.1111/1365-2435.12981>

700 Houslay, T. M., & Wilson, A. J. (2017). Avoiding the misuse of BLUP in behavioural ecology.
701 *Behavioral Ecology*, 28(4), 948–952. <https://doi.org/10.1093/beheco/arx023>

702 Huntingford, F. A., Adams, C., Braithwaite, V. A., Kadri, S., Pottinger, T. G., Sandøe, P., & Turnbull,
703 J. F. (2006). Current issues in fish welfare. In *Journal of Fish Biology* 68 (2), 332–372.
704 <https://doi.org/10.1111/j.0022-1112.2006.001046.x>

705 Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H.,
706 De Jong, I. C., Ruis, M. A. W., & Blokhuis, H. J. (1999). Coping styles in animals: current
707 status in behavior and stress-physiology. *Neuroscience & Biobehavioral Reviews*, 23(7), 925–
708 935. [https://doi.org/10.1016/S0149-7634\(99\)00026-3](https://doi.org/10.1016/S0149-7634(99)00026-3)

709 Kotrschal, A., Corral-Lopez, A., Amcoff, M., & Kolm, N. (2015). A larger brain confers a benefit in a
710 spatial mate search learning task in male guppies. *Behavioral Ecology*, 26(2), 527–532.
711 <https://doi.org/10.1093/beheco/aru227>

712 Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I., Immler, S.,
713 Maklakov, A. A., & Kolm, N. (2013). Artificial selection on relative brain size in the guppy
714 reveals costs and benefits of evolving a larger brain. *Current Biology*, 23(2), 168–171.

715 Langley, E. J. G., Adams, G., Beardsworth, C. E., Dawson, D. A., Laker, P. R., van Horik, J. O.,
716 Whiteside, M. A., Wilson, A. J., & Madden, J. R. (2020). Heritability and correlations among

717 learning and inhibitory control traits. *Behavioral Ecology*.
718 <https://doi.org/10.1093/beheco/araa029>

719 Lucon-Xiccato, T., & Bisazza, A. (2014). Discrimination reversal learning reveals greater female
720 behavioural flexibility in guppies. *Biology Letters*, 10(6), 20140206.

721 Lucon-Xiccato, T., & Bisazza, A. (2016). Male and female guppies differ in speed but not in accuracy
722 in visual discrimination learning. *Animal Cognition*, 19(4), 733–744.
723 <https://doi.org/10.1007/s10071-016-0969-2>

724 Lucon-Xiccato, T., & Bisazza, A. (2017a). Individual differences in cognition among teleost fishes. In
725 *Behavioural Processes* 141, 184–195. <https://doi.org/10.1016/j.beproc.2017.01.015>

726 Lucon-Xiccato, T., & Bisazza, A. (2017b). Sex differences in spatial abilities and cognitive flexibility
727 in the guppy. *Animal Behaviour*, 123, 53–60. <https://doi.org/10.1016/j.anbehav.2016.10.026>

728 Lucon-Xiccato, T., & Bisazza, A. (2017c). Complex maze learning by fish. *Animal Behaviour*, 125,
729 69–75. <https://doi.org/10.1016/j.anbehav.2016.12.022>

730 Lukowiak, K., Orr, M., de Caigny, P., Lukowiak, K. S., Rosenegger, D., Han, J. Il, & Dalesman, S.
731 (2014). Ecologically relevant stressors modify long-term memory formation in a model system.
732 *Behavioural Brain Research*, 214(1), 18–24. <https://doi.org/10.1016/j.bbr.2010.05.011>

733 Mazza, V., Eccard, J. A., Zaccaroni, M., Jacob, J., & Dammhahn, M. (2018). The fast and the
734 flexible: cognitive style drives individual variation in cognition in a small mammal. *Animal*
735 *Behaviour*, 137, 119–132. <https://doi.org/10.1016/j.anbehav.2018.01.011>

736 Mesquita, F. O., Borcato, F. L., & Huntingford, F. A. (2015). Cue-based and algorithmic learning in
737 common carp: A possible link to stress coping style. *Behavioural Processes*, 115, 25–29.
738 <https://doi.org/10.1016/j.beproc.2015.02.017>

739 Meyer, P. J., Lovic, V., Saunders, B. T., Yager, L. M., Flagel, S. B., Morrow, J. D., & Robinson, T. E.

(2012). Quantifying individual variation in the propensity to attribute incentive salience to reward cues. *PLoS ONE*, 7(6). <https://doi.org/10.1371/journal.pone.0038987>

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

Niemelä, P. T., Vainikka, A., Forsman, J. T., Loukola, O. J., & Kortet, R. (2013). How does variation in the environment and individual cognition explain the existence of consistent behavioral differences? *Ecology and Evolution*, 3(2), 457–464. <https://doi.org/10.1002/ece3.451>

Nussey, D. H., Wilson, A. J., & Brommer, J. E. (2007). The evolutionary ecology of individual phenotypic plasticity in wild populations. *Journal of Evolutionary Biology*, 20(3), 831–844. <https://doi.org/10.1111/j.1420-9101.2007.01300.x>

Øverli, Ø., Sørensen, C., Pulman, K. G. T., Pottinger, T. G., Korzan, W., Summers, C. H., & Nilsson, G. E. (2007). Evolutionary background for stress-coping styles: Relationships between physiological, behavioral, and cognitive traits in non-mammalian vertebrates. In *Neuroscience and Biobehavioral Reviews* 31(3), 396–412. Pergamon. <https://doi.org/10.1016/j.neubiorev.2006.10.006>

Plomin, R., & Spinath, F. M. (2002). Genetics and general cognitive ability (g). In *Trends in Cognitive Sciences*, 6 (4), 169–176. [https://doi.org/10.1016/S1364-6613\(00\)01853-2](https://doi.org/10.1016/S1364-6613(00)01853-2)

Prentice, P. M., Houslay, T. M., Martin, J. G. A., & Wilson, A. J. (2020). Genetic variance for behavioural ‘predictability’ of stress response. *Journal of Evolutionary Biology*, jeb.13601. <https://doi.org/10.1111/jeb.13601>

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), 20150184.

- 764 <https://doi.org/10.1098/rstb.2015.0184>
- 765 Raoult, V., Trompf, L., Williamson, J. E., & Brown, C. (2017). Stress profile influences learning
766 approach in a marine fish. *PeerJ*, 2017(6). <https://doi.org/10.7717/peerj.3445>
- 767 Réale, D., Garant, D., Humphries, M. M., Bergeron, P., Careau, V., & Montiglio, P. O. (2010).
768 Personality and the emergence of the pace-of-life syndrome concept at the population level.
769 *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*,
770 365(1560), 4051–4063. <https://doi.org/10.1098/rstb.2010.0208>
- 771 Réale, Denis, Reader, S. M., Sol, D., McDougall, P. T., & Dingemanse, N. J. (2007). Integrating
772 animal temperament within ecology and evolution. *Biological Reviews*, 82(2), 291–318.
- 773 Roff, D. A. (2002). *Life history evolution*. Sinauer Associates.
- 774 Rowe, C., & Healy, S. D. (2014). Measuring variation in cognition. In *Behavioral Ecology* 25(6),
775 1287–1292. <https://doi.org/10.1093/beheco/aru090>
- 776 Salak-Johnson, J. L., & Mcglone, J. J. (2007). Making sense of apparently conflicting data: Stress and
777 immunity in swine and cattle 1. *J. Anim. Sci*, 85, 81–88. <https://doi.org/10.2527/jas.2006-538>
- 778 Segerstrom, S. C., & Miller, G. E. (2004). Psychological stress and the human immune system: A
779 meta-analytic study of 30 years of inquiry. *Psychological Bulletin*, 130(4), 601–630.
780 <https://doi.org/10.1037/0033-2909.130.4.601>
- 781 Shaw, R. C., MacKinlay, R. D., Clayton, N. S., & Burns, K. C. (2019). Memory Performance
782 Influences Male Reproductive Success in a Wild Bird. *Current Biology*, 29(9), 1498-1502.e3.
783 <https://doi.org/10.1016/j.cub.2019.03.027>
- 784 Shettleworth, S. J. (2010). Getting Around: Spatial Cognition. In *Cognition, Evolution, and Behavior*
785 (second, pp. 261–312). Oxford University Press.
- 786 Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: an ecological and evolutionary

overview. *Trends in Ecology & Evolution*, 19(7), 372–378.

Sih, A., Bell, A. M., Johnson, J. C., & Ziemba, R. E. (2004). Behavioral Syndromes: An Integrative Overview. *The Quarterly Review of Biology*, 79(3), 241–277. <https://doi.org/10.1086/422893>

Sih, A., & Del Giudice, M. (2012). Linking behavioural syndromes and cognition: a behavioural ecology perspective. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367(1603), 2762–2772. <https://doi.org/10.1098/rstb.2012.0216>

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.e3. <https://doi.org/10.1016/j.cub.2019.01.006>

Sterne, J. A. C., & Smith, G. D. (2001). Sifting the evidence—what’s wrong with significance tests? *Physical Therapy*, 81(8), 1464–1469. <https://doi.org/10.1093/ptj/81.8.1464>

Stram, D., Biometrics, J. L.-, & 1994, undefined. (n.d.). Variance components testing in the longitudinal mixed effects model. *JSTOR*. Retrieved March 3, 2020, from <https://www.jstor.org/stable/2533455>

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

Thornton, A., & Wilson, A. J. (2015). In search of the Darwinian Holy Trinity in cognitive evolution: A comment on Croston et al. In *Behavioral Ecology*, 26(6), 1460–1461. <https://doi.org/10.1093/beheco/arv119>

van Horik, J. O., Langley, E. J. G., Whiteside, M. A., & Madden, J. R. (2017). Differential participation in cognitive tests is driven by personality, sex, body condition and experience. In *Behavioural Processes* 134, 22–30. <https://doi.org/10.1016/j.beproc.2016.07.001>

- 810 van Horik, J. O., & Madden, J. R. (2016). A problem with problem solving: Motivational traits, but
811 not cognition, predict success on novel operant foraging tasks. *Animal Behaviour*, 114, 189–198.
812 <https://doi.org/10.1016/j.anbehav.2016.02.006>
- 813 Warren, E. W., & Callaghan, S. (1976). The response of male guppies (*Poecilia reticulata*, Peters) to
814 repeated exposure to an open field. *Behavioral Biology*, 18(4), 499–513.
815 [https://doi.org/10.1016/S0091-6773\(76\)92524-4](https://doi.org/10.1016/S0091-6773(76)92524-4)
- 816 White, S. J., Houslay, T. M., & Wilson, A. J. (2019). Evolutionary genetics of personality in the
817 Trinidadian guppy II: sexual dimorphism and genotype-by-sex interactions. *Heredity*, 122(1),
818 15–28. <https://doi.org/10.1038/s41437-018-0083-0>
- 819 White, S. J., Kells, T., & Wilson, A. J. (2016). Metabolism, personality and pace of life in the
820 Trinidadian guppy, *Poecilia reticulata*. *Behaviour*, 153(13–14), 1517–1543.
821 <https://doi.org/10.1163/1568539X-00003375>
- 822 White, S. J., & Wilson, A. J. (2019). Evolutionary genetics of personality in the Trinidadian guppy I:
823 maternal and additive genetic effects across ontogeny. *Heredity*, 122(1), 1–14.
824 <https://doi.org/10.1038/s41437-018-0082-1>
- 825 Wolf, M., Van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs favour the
826 evolution of animal personalities. *Nature*, 447(7144), 581–584.
- 827 Wong, R. Y., French, J., & Russ, J. B. (2019). Differences in stress reactivity between zebrafish with
828 alternative stress coping styles. *Royal Society Open Science*, 6(5), 181797.
829 <https://doi.org/10.1098/rsos.181797>
- 830 Wong, S. C., Dykstra, M., Campbell, J. M., & Earley, R. L. (2008). Measuring water-borne cortisol in
831 convict cichlids (*Amatitlania nigrofasciata*): Is the procedure a stressor? *Behaviour*, 145(10),
832 1283–1305. <https://doi.org/10.1163/156853908785765863>
- 833 Zidar, J., Balogh, A., Favati, A., Jensen, P., Leimar, O., Sorato, E., & Løvlie, H. (2018). The

834 relationship between learning speed and personality is age- and task-dependent in red
 835 junglefowl. *Behavioral Ecology and Sociobiology*, 72(10), 168. <https://doi.org/10.1007/s00265->
 836 018-2579-2
 837