Individual differences in spatial learning are correlated across tasks but not with stress response behaviour in guppies

Pamela M. Prentice[*] , Chloe Mnatzaganian, Thomas M. Houslay , Alex Thornton ,
Alastair J. Wilson

Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Cornwall, U.K.

Cognition is vital for carrying out behaviours required for survival and reproduction. Cognitive performance varies between species, but also between individuals within populations. While variation is a prerequisite for natural selection, selection does not act on traits in isolation. The extent to which cognitive traits covary with other aspects of phenotype (e.g. personality traits) may be important in shaping evolutionary dynamics. Here we adopted a multivariate approach to test spatial learning in male *Poecilia reticulata* and asked whether differences in cognitive performance are associated with (repeatable) differences in stress response behaviour. Functional links between cognitive traits and 'stress coping style' have been hypothesized. Furthermore, individual level studies of cognitive performance typically rely on multiple testing paradigms that may themselves be a stressor. There is a risk that variation in stress responsiveness is itself a cause of apparent, but artefactual variance in cognitive ability. Fish repeatedly experienced two spatial learning tasks (maze layouts) and an acute stress response test (open field trial). We found repeatable differences between individuals in performance within and across maze layouts. On average, performance improved with experience in the first maze, consistent with spatial learning, but not in the second. Individuals varied in the trajectory of mean performance with trial number in both mazes, suggesting they differ in 'learning rate'. Acute stress response behaviour was repeatable but predicted neither average time to solve the maze nor learning rate. We found no support for between-individual correlation between acute stress response and cognitive performance. However, we highlight the possibility that cumulative, chronic stress effects may nevertheless cause declines in performance 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.

Cognition 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 between species have long been studied in comparative psychology (for a review see Healy, 2019), a more recent focus in behavioural ecology has been the characterization of consistent between-individual variation within populations of nonhuman animals (Ashton et al., 2018; Boogert et al., 2018; Lucon-Xiccato & Bisazza, 2017a). This consistent between-individual variation is interesting from an evolutionary perspective, as it is a prerequisite for natural selection and genetic variation, both of which are fundamental for adaptive evolution to occur (Wilson et al., 2010). As selection does not act on traits in isolation, quantifying relationships between cognitive variation and other aspects of phenotype is therefore important. For example, functional links between cognitive performance and other aspects of behaviour (including, for example neophobia, boldness and stress responsiveness) have been widely hypothesized (Griffin et al., 2015; Medina-García et al., 2017; Quinn et al., 2012; Sweis et al., 2013). However, 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. between-individual and/or between-genotype variation; Dingemanse & Dochtermann, 2013). Nevertheless, such efforts are important if we hope to understand the

* Corresponding author.
  E-mail address: p.prentice4@exeter.ac.uk (P. M. Prentice).

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 hypothesized links between cognitive performance and a behavioural stress response (Gibelli et al., 2019; Øverli et al., 2007) in Trinidadian guppies, *Poecilia reticulata*.

Quantifying patterns of consistent between-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; Tyrone Lucon-Xiccato & Bisazza 2017; Niemelä et al., 2013 for examples). However, it is now abundantly clear that populations typically harbour high levels of consistent between-individual variation in behavioural traits more generally (Dingemanse & Réale, 2005). Consistent 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 stabilizing selection is usually predicted to erode variation (Roff, 2002), it is hypothesized 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 between cognitive domains, or alternatively between, for instance, an overall cognitive performance trait ('general intelligence'; Burkart et al., 2017; Galsworthy et al., 2005; Plomin & Spinath, 2002) and other aspects of the phenotype.

Variation in an animal's stress physiology may provide one putative source of consistent between-individual differences in both personality traits and cognitive performance between individuals (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 an adrenaline response to stressors. At the other extreme, reactive coping styles will be more behaviourally 'passive' (e.g. freezing or hiding) and show high hypothalamic-pituitary-adrenal (HPA(I)) activity leading to a cortisol response (Carere et al., 2014; Øverli et al., 2007). Various links to cognitive performance variation have been suggested. For instance, proactive styles are associated with 'bold', exploratory, risk-taking personalities that may present with more opportunities to learn initially. Conversely, greater behavioural 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 more, or less, motivated and/or are focused on sources of risk rather than environmental cues of rewards.

Although hypothesized 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). 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*, individual fish displaying less thigmotaxic behaviour (a stress-related behaviour in fish, commonly displayed as rapid swimming while 'wall hugging' the sides of the tank) 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 addressed this broad goal by testing the hypothesis that individual differences in cognitive performance and stress responsiveness are correlated in male Trinidadian guppies.

The guppy is a freshwater poeciliid fish that is widely used as a model in behavioural and evolutionary ecology. Methods for assaying consistent between-individual 'personality' variation are well established in this species generally (Burns & Rodd, 2008; White et al., 2016), while guppies have been used in cognitive studies that target learning colour discriminations (Buechel et al., 2018; Trompf & Brown, 2014), numerical discriminations (Kotrschal et al., 2013; Lucon-Xiccato & Bisazza, 2017b), reversal learning (Buechel et al., 2018), spatial learning (Lucon-Xiccato & Bisazza, 2017c) and inhibitory control (e.g. Lucon-Xiccato & Bisazza, 2016). Here, we investigated 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 maze with a different spatial configuration not just to assess variation in learning within a single spatial context, but also to ask whether, for instance, individuals performing better 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 spatial learning is thus expected to be an ecologically relevant trait (Brown et al., 2005). For our measure of stress responsiveness, we utilized 'open field trials'. Widely used across species as a paradigm for characterizing behavioural differences related to exploration, activity and 'shy–bold' type variation (Bell et al., 2009; Gosling, 2001), open field trials on this captive population of guppies have highlighted their utility for assaying the behavioural stress response (see e.g. Prentice et al., 2020). Observed behaviours expressed in these trials 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 (White and Wilson, 2019; Prentice, 2020). Furthermore, there is evidence of genetic integration between behaviour in open field trials and cortisol expression, strengthening the view that the open field trials provide an appropriate assay of the behavioural stress response (Houslay et al., 2019).

(1) We tested for evidence of learning in naïve guppies repeatedly exposed to a spatial learning task (maze) and asked (2) whether individuals differ in cognitive performance across repeated trials and, if so (3) whether performance in the first maze

predicted performance in a second spatial context (i.e. reconfigured maze). We predicted that time to complete the mazes (our proxy of cognitive performance) would, on average, improve with experience consistent with spatial 'learning', but that individuals would differ consistently in cognitive performance within each maze. We also predicted that individual performance in the first maze would be positively correlated with performance in the second, suggesting stable differences in cognitive ability, although we acknowledge proactive interference (difficulty inhibiting memory or previously learnt associations; Shettleworth, 2010) may affect performance in the second maze. Finally, (4) we tested the hypothesis that consistent individual differences in cognitive performance would be associated with consistent individual differences in stress responsiveness. Although empirical evidence suggests potential relationships in both directions between stress responsiveness and cognitive performance, with the current absence of specific models, we made no a priori predictions about the sign of the relationship here.

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 eight in separate home tanks (15 litres, 18.5×37 cm and 22 cm deep) maintained at $23-24\,°C$ on a 12:12 h 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 coloration 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.

Ethical Note

This work was conducted under the auspices of the Animals (Scientific Procedures Act) with approval of the University of Exeter research ethics committee, under licence from the Home Office (U.K.) (Licence Number PPL30/3256). Experimental procedures and behavioural assays were developed in accordance with the principles of the three Rs and ASAB guidelines (ASAB Ethical Committee & ABS Animal Care Committee, 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. Normal behaviour in this species is defined as fast, active swimming. Fish were proactively 'retired' into a separate tank if they were in poor body condition, not feeding or displayed non-normal behaviour, defined as laboured swimming and/or keeping to the back of the tank or tank corners. If fish did not resume feeding and normal behaviour within 48 h, they were euthanized,

($N = 9$, see details below) using a lethal overdose of MS-222, after which death was confirmed by destruction of the brain by pithing with a needle. At the end of the experiment, fish were returned to the 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 between-individual (co)variation in spatial learning performance and stress responsiveness. Spatial learning was first assessed by repeatedly trialling individuals in a maze apparatus (Fig. 1). Each individual fish was tested once per day for 11 consecutive days with reduction 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 between 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 between-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 between-individual variation as far as possible using standardized housing and husbandry conditions.

Following completion of 11 spatial learning trials using Maze A, individuals were tested for stress responsiveness three times each over a 3-week period using open field trials with a mean (range) of 4 (1–5) days between successive trials. As we wished to estimate the repeatability (or otherwise) of cognitive performance and learning across exposure to similar cognitive tasks, fish were then retested in a second reconfigured maze (Maze B) with a different spatial layout. Maze B was sufficiently different to Maze A that it could not be solved by simply reversing rules previously learnt in Maze A, but rather required learning a new route to reach the social reward. Repeat trials of Maze B were conducted (as before) once per day for 11 consecutive days. Thus, in total, the complete assay 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 open field trials over a total testing period of 43 days. Note that the sample size declined slightly across the experiment. Two mortalities occurred naturally within the testing period, both fish being found dead in their home tank of unknown cause. As a precaution against cumulative adverse effects, we also inspected fish daily and proactively 'retired' from the experiment any individuals not deemed to be feeding well or showing possible signs of stress (see Ethical Note). Thus 64 fish were used in Maze A, with $N = 60$ completing 11 trials (two mortalities and two fish retired) and the open field trials. We retired five further fish prior to trial 1 of Maze B ($N = 55$), and two more before they had completed 11 trials ($N = 53$ at trial 11).

Spatial learning trials

To facilitate more rapid data collection, a single aquarium (25×45 cm and 25 cm deep) was divided into two, with each half containing an identical version of Maze A (A1, A2). Two replicates of Maze B were similarly constructed (Fig. 1). This allowed two fish to be tested concurrently during trials. Each maze consisted of six opaque Perspex panels (8 cm), spaced 5 cm apart (Fig. 1). A visually transparent perforated panel at one end of each maze was used to separate a small holding area (12.5×10 cm and 25 cm deep) which contained two adult females selected randomly 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
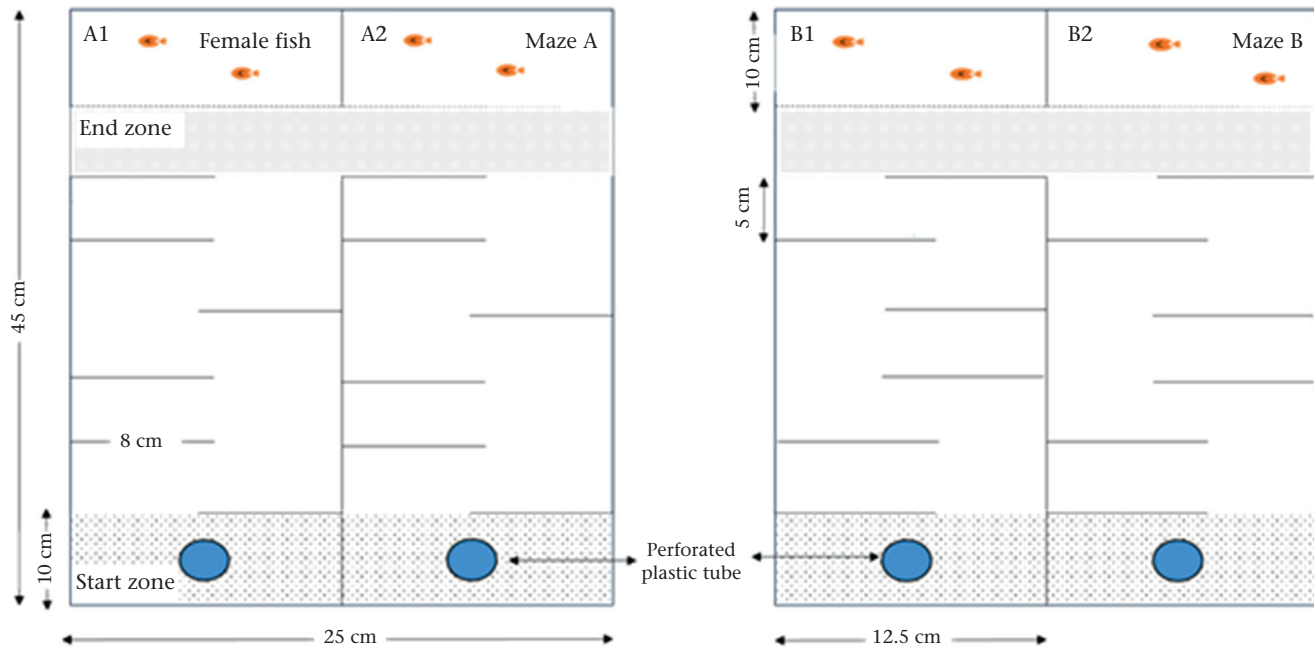
Figure 1. Aerial view of the maze designs used in the experiments (A and B); each tank was split into two identical mazes (1 and 2). Blue circles indicate perforated plastic tubes where individual fish started each trial, and time taken to reach the 'end zone' was recorded in each trial.

temperature water (ca. 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 four 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 at the start of the maze (Fig. 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 used to determine Maze time, measured as the total time taken to complete the maze after fish were released from the perforated plastic tube (with completion defined as reaching the 'end' zone; Fig. 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 their view of the females. 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 of being released from the tube were gently guided with a net behind them through the maze to the end zone and then they 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 s for Maze time.

Open field trial

Open field trials to characterize 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 from natural markings 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 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 270 s to determine track length (total distance swum (cm)) and area covered (percentage of tank area covered). These two observed behaviours, which are known to be repeatable and heritable in this population (Houslay et al., 2018; White and 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 largely cover the same small area, thus covering relatively little of the tank arena. Low values of Relative area are strongly correlated with thigmotaxis in this species, with fish swimming repeatedly across sections next to the tank walls and seeking escape from the tank arena. In contrast, high 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 restricted maximum likelihood using ASReml within R (https://www.vsni.com; Gilmour et al., 2009). By including individual identity as a random effect in these models we tested for and characterized between-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 to preserve any differences in the distributions of performance times between A and B. With traits in standard deviation units (sdu), estimates of between-individual variance (V_{ind}) can be interpreted as repeatabilities (i.e. proportion of the observed phenotypic variance explained by between-individual differences). However, we also calculated

estimates of adjusted repeatability (R), the proportion of phenotypic variance explained by consistent between-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. We deemed that acceptable based on visual inspection of the model residuals (see Appendix Fig. A1) and the fact that these mixed models are increasingly recognized as being robust to deviations from normality (Schielzeth et al., 2020).

Univariate analyses of maze performance and spatial learning

We used Maze time as our observed measure of performance. Here we describe in full the univariate analysis of data collected in Maze A (subsequently Maze time$_A$). Identical procedures were then applied to data from Maze B. First, we visualized the distribution of Maze time$_A$ across trial repeats using box plots and also plotted the proportion of mazes completed as a function of trial repeat to see if a pattern of increasing average performance (i.e. decreasing Maze time and/or increasing proportion of successful completion) was immediately apparent. Next a series of three nested models with identical fixed effects but differing random effect structure were fitted to the centred and scaled Maze time$_A$ data. All models included a fixed effect of trial number (the cumulative number of trials experienced by an individual, treated as a continuous variable), allowing us to test for improvement in the mean (indicative of learning). Additional fixed effects were included as statistical controls for potential sources of variance not relevant to hypotheses being tested here. These included time of day (in minutes after 0900), maze replicate (as a factor denoting position 1 or 2 in the maze tank) and order caught from the home tank. The latter was to account for any cumulative disturbance effect of removing fish sequentially from the home tank and/or build-up of chemical cues in the maze between water changes.

The first univariate model of Maze time contained no random effects, while the second contained a random intercept of individual identity. LRT comparison of these models was conducted to test the hypothesis that individuals differ in their average performance (Maze time$_A$) across the 11 repeated trials, and we estimated the (adjusted) repeatability of performance under the second model. For the LRT we assumed twice the difference in model log-likelihoods is distributed as a 50:50 mix of χ^2_1 and χ^2_0 following Stram and Lee (1994). The third model was a first-order random regression (i.e. a random slope and intercept model) in which each individual's deviation from the fixed effect mean Maze time can change as a linear function of trial number (1–11). Variation in random slopes means that there is between-individual variation around the mean Maze time$_A$–trial number relationship. Thus, LRT comparison of the second and third models provides a test for between-individual variation in learning rate. The LRT comparison was conducted assuming the test statistic is distributed as χ^2_2, since the third model has two extra parameters (a slope variance and a slope–intercept covariance). Note that between-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, between-individual variance in learning (slope) means that between-individual variance Maze time$_A$ changes with trial number (Appendix Fig. A2). Thus, to understand the biological effect size of estimated variance in slopes, we used the third model to predict

between-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 also Supplementary Material Table S1 for a didactic explanation of the linear algebra behind this). We note that between-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 this 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 fitted 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 0900 in which each trial took place) as well as a random effect of individual identity. Adjusted repeatability (R) of Relative area was calculated and the significance of between-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 χ^2_1 and χ^2_0 as above).

Multivariate modelling of Maze A, Maze B and open field trials 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 5×5 between-individual covariance matrix (**ID**) between these effects. Since each observation of a fish provided data on a single trait only, residual covariances between traits were fixed to zero. After fitting the model, we compared it to a simplified fit in which all between-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 $r_{xy} = COV_{xy}/(V_x V_y)^{0.5}$). This allowed us to scrutinize 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.

This model also yields an estimate of the individual level correlation between the cognitive performance measures (final Maze time performance, learning) across both 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 (Barbosa & Morrissey, 2021). 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 tended 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 decreased across trials (Fig. 2), and this pattern is qualitatively consistent with expectations if (average) performance
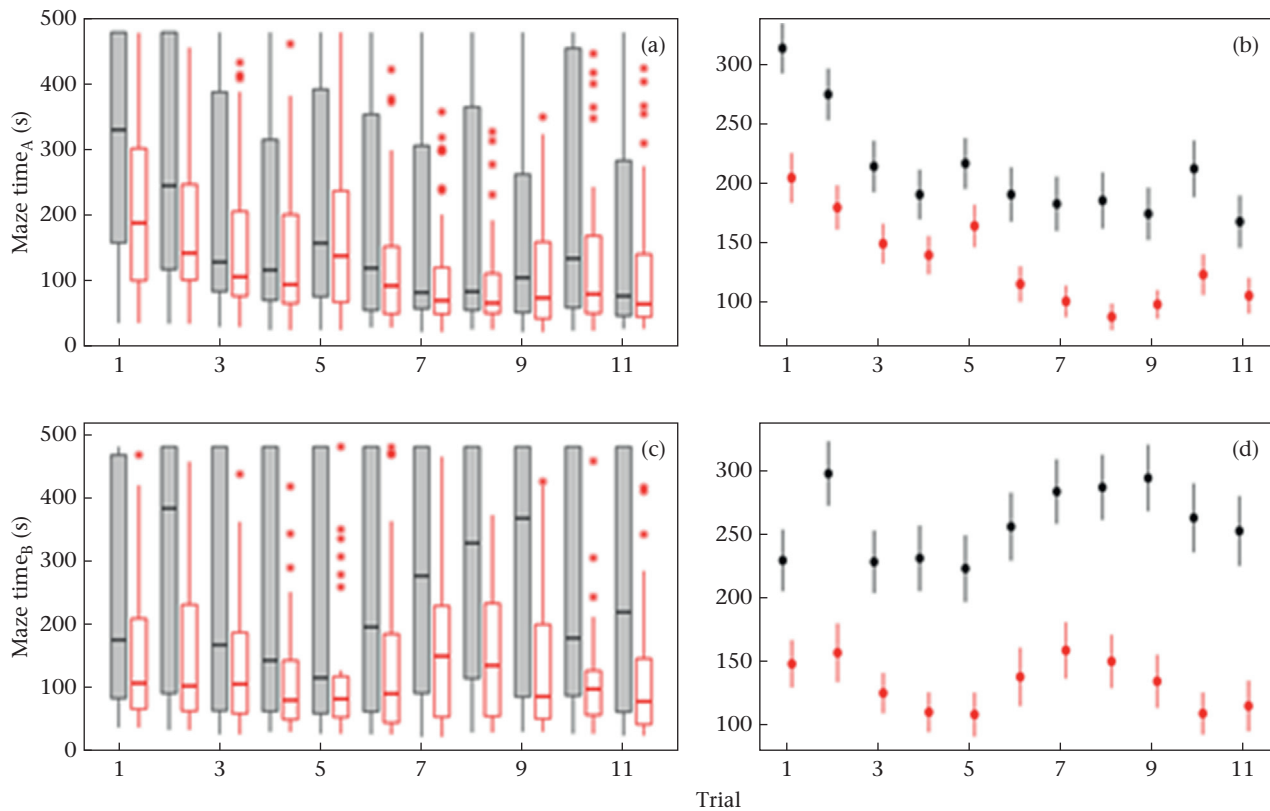
Figure 2. Plots of raw data of Maze time across both maze designs. (a, c) Data distributions for time to complete Maze A and Maze B, respectively, across the 11 trials. Grey boxes: data of all individuals; red boxes: those individuals that successfully completed the task within 480 s. Horizontal lines within boxes 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. (b, d) Mean and SE for time to complete Maze A and B, respectively. Colours represent the same groups; black: all individuals; red: those individuals that successfully completed the maze in the allocated time.

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 = −0.058 (0.014) sdu, $F_{1,59.8} = 17.890$, $P < 0.001$). This effect size equates to an estimated decrease of 102 s in average Maze time over the 11 trials. Other fixed effects (order caught and maze position) were nonsignificant (see Appendix Table A1). LRTs confirmed between-individual variation in Maze time$_A$ (comparison of null and random intercept models: $\chi^2_{0,1} = 184.713$ $P < 0.001$). Under the random intercept model, repeatability of Maze time$_A$ conditional on fixed effects was estimated as $R_A = 0.379$ (0.052).

LRT comparison of the random intercept and first-order random regression models showed the latter to be a significantly better fit to the data ($\chi^2_2 = 26.990$, $P < 0.001$). This comparison provides evidence for between-individual variance in the rate of change of Maze time$_A$ across repeated trials. We interpret this (with caveats discussed below) as variation in the rate of spatial learning. Between-individual variance in intercepts (int) and slope (slp) were estimated as $V_{\mathrm{ind_{int}}} = 0.402$ (0.104) and $V_{\mathrm{ind_{slp}}} = 0.006$ (0.032), respectively, while the between-individual intercept—slope correlation was estimated as $r_{\mathrm{ind_{int},ind_{slp}}} = -0.375$ (0.168). Biological interpretation of these parameters is not completely straightforward. Given the scaling of trial number in the random effect structure of the model (see Supplementary Material Table S1) $V_{\mathrm{ind_{int}}}$ is interpretable as between-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 means that between-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 $V_{\mathrm{ind_{A1}}} = 0.402$ (0.104), and $V_{\mathrm{ind_{A11}}} = 0.677$ (0.157) at first and last trial in maze A, respectively, suggesting more between-individual variation in performance at the end of trials than at the beginning. The negative intercept—slope correlation ($r_{\mathrm{ind_{A,int,A,slp}}} = -0.375$ (0.168)) means that individuals with higher intercepts (high Maze time$_A$ at trial 1) tended to have lower (i.e. more negative) slopes indicative of faster learning. The corresponding predictions of repeatability at first and last observed trial are $R_{A1} = 0.429$ (0.070) and $R_{A11} = 0.560$ (0.063). These patterns are represented visually in Fig. 3a, 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).

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. The mixed model analysis confirms the lack of improvement in the mean Maze time$_B$, with a (nonsignificant) positive estimate of the trial repeat number effect (from random slope and intercept model; coefficient = 0.019 (0.014) , $F_{1, 53.8} = 2.054$, $P = 0.175$). Effects of order caught and maze position were not significant (Appendix Table A1). LRTs between the univariate random intercept model and the null model with no random effect show the presence of significant between-individual variation for Maze time$_B$ ($\chi^2_{0,1} = 200.048$, $P < 0.001$), with a corresponding repeatability estimate of $R_B = 0.423$ (0.056). The random slope model was a significantly better fit again ($\chi^2_2 = 15.926$, $P = 0.001$) providing evidence of between-individual variation in the performance—trial number relationship. Between-
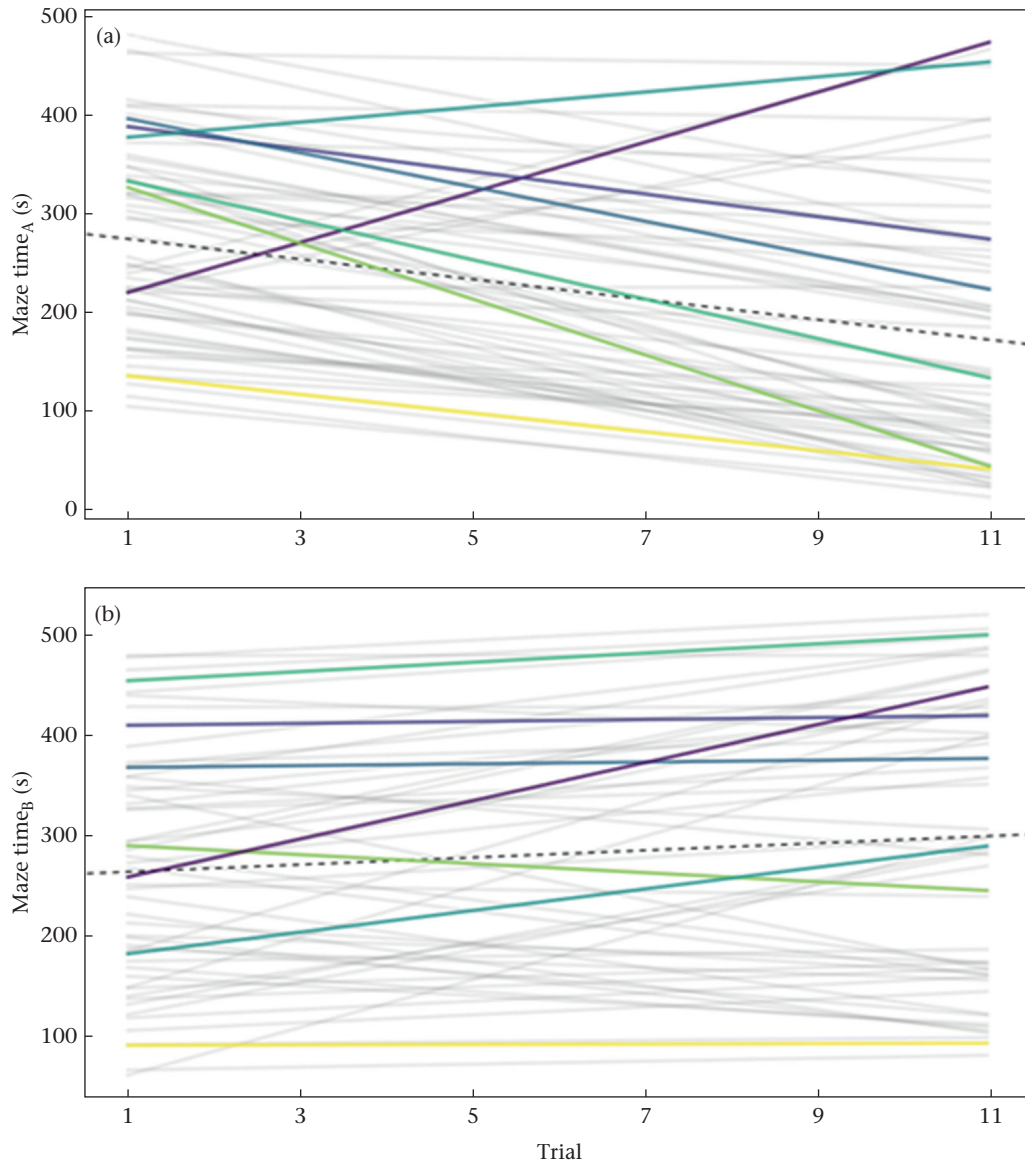
Figure 3. Spatial learning traits across Maze A and Maze B as a function of trial number, (a) Maze time$_A$ and (b) Maze time$_B$. Grey lines represent individual predicted reaction norms (BLUPs) from univariate random intercept and slope models for each trait. Coloured lines are used to illustrate reaction norms for a small random set of arbitrarily chosen individuals tested in both mazes (colours represent the same individuals across panels). The black dashed line represents the trend in fixed effect mean Maze time across repeat trials.

individual variance in intercepts (int) and slope (slp) were estimated as 0.472 (0.125) and 0.005 (0.034), respectively. The intercept—slope correlation was negative as in Maze A ($= -0.332$ (0.190)). These estimates mean predicted values of $V_{\mathrm{ind}_{B1}} = 0.472$ (0.125) and $V_{\mathrm{ind}_{B11}} = 0.659$ (0.162) which correspond to repeatabilities of $R_{B1} = 0.471$ (0.072) and $R_{B11} = 0.554$ (0.067). Although there is no (significant) effect of trial number on mean Maze time$_B$ the presence of between-individual variance in slope suggests that some individuals improved (consistent with learning) while for others performance tended to get worse across trials in Maze B (Fig. 3b).

Open Field Trial Behaviour

We found evidence of significant between-individual variation in Relative area (repeatability(with SE) : $R = 0.465$ (0.089), $\chi^2_{0,1} = 20.421, P < 0.001$). This replicates previous findings in the same population (Prentice et al., 2020) although the current estimate of

repeatability is somewhat higher, likely due to differences in study design (e.g. the current study used a shorter interobservation interval and was limited to males only). Fixed effects from the open field trials behaviour models are presented in Appendix Table A2 for completeness, although they are not directly relevant to our hypotheses in this study.

Multivariate Model

The full multivariate model of Maze time$_A$, Maze time$_B$ and Relative area provides evidence of some significant covariance structure between traits at the individual level (comparison of the full model to one in which all between-individual between-trait covariances are fixed to zero; $\chi^2_8 = 48.844, P < 0.001$). Examination of the estimated covariances and correlations (Table 1) suggests this result is largely driven by a strong positive correlation between the individual intercepts for Maze time$_A$ and Maze time$_B$

Table 1
Between individual variance–covariance–correlation matrix of random effects from the final trivariate model of Maze time$_A$, Relative area and Maze time$_B$

	Random effect	(Biological meaning)	Maze time$_A$		Relative area ID	Maze time$_B$	
			ID	ID*Trial		ID	ID*Trial
Maze time$_A$	ID	(Trial 1 performance)	**0.402 (0.104)**[a]	−0.019 (0.012)	0.096 (0.079)	*0.305 (0.089)*	−0.005 (0.010)
	ID*Trial	(Change in maze time over trials)	*−0.374 (0.168)*[c]	**0.007 (0.002)**[b]	−0.002 (0.011)	0.005 (0.011)	0.001 (0.001)
Relative area	ID	(Stress responsiveness)	0.226 (0.178)	−0.034 (0.202)	**0.452 (0.118)**	0.035 (0.086)	0.012 (0.011)
Maze time$_B$	ID	(Trial 1 performance)	*0.704 (0.127)*	0.094 (0.203)	0.075 (0.186)	**0.467 (0.122)**	−0.016 (0.012)
	ID*Trial	(Change in maze time over trials)	−0.103 (0.226)	0.186 (0.246)	0.255 (0.221)	−0.335 (0.189)	**0.005 (0.002)**

Variances are shown on the diagonal (bold font), with covariances above and correlations below. Standard errors are shown in parentheses. Nominally significant values, assuming 95% CI of ± 1.96 SE, are shown in italic font. Also included are biological meanings of the individual level random effects. An example interpretation is given in the superscript footnotes below.

  [a] The between-individual variance in performance at trial 1 in Maze A is presented in the upper left cell of this submatrix and is estimated as 0.402. Since data are analysed in standard deviation units this means 40.2% of the variation in Maze time$_A$ at trial 1 is due to differences between individuals.

  [b] The ID*Trial variance (lower right cell) is also nominally significant, supporting the presence of between-individual variation in the rate of change in Maze time$_A$ with repeated experience (i.e. variation in learning).

  [c] The estimated correlation between individuals between Maze time$_A$ at trial 1 and the rate of learning is -0.374. Since learning decreases Maze time$_A$, the significant negative correlation means individuals that initially perform poorly tend to have higher rates of improvement (as seen in Fig. 3a).

$(r_{\text{ind}_{A.int,B.int}} = 0.704\ (0.127))$. Biologically, this means that individual performance at first trial is strongly positively correlated across mazes. At final trial (i.e. 11), the individual correlation of performance across mazes is estimated at $r_{\text{ind}_{A11.B11}} = 0.629\ (0.119)$. Thus, our results show strong positive correlations of individual performance as measured by maze time across trials and mazes. This is not just true at first and last trial, but also for intermediate trial numbers within and across mazes (see Supplementary Material Table S1 for estimates and an explanation of their derivation from the random regression model output). However, taking a reaction norm interpretation of results, we do not find strong support for repeatable variation in learning rate across mazes. The individual correlation of reaction norm slopes is positive $(r_{\text{ind}_{A.slp,B.slp}} = 0.186\ (0.246))$, but relatively weak and not statistically significant (assuming approximate 95% confidence intervals of ±1.96 SE). Nor do we find statistical support for between-individual correlation between maze performance intercepts or slopes (for either maze) and Relative area (Table 1).

DISCUSSION

Here, we have shown evidence of between-individual differences in performance, measured as time to complete a maze, in guppies exposed to a spatial learning test paradigm. Individual performance was 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 (Maze A) we found evidence of improvement in mean performance consistent with learning (on average). We also find between-individual variation in this rate of improvement, and so, putatively, their rate of learning. However, the same fish exposed to Maze B showed (on average) no increase in performance across successive trials. We found between-individual correlation structure between performances (i.e. time in the maze) but not learning (i.e. rate of improvement) across the two spatial learning tasks. We did not 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 improved across repeated trials. This improvement suggests that spatial learning was 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 saw evidence of consistent, repeatable differences between in performance between individuals in Maze A. This is shown in our reaction norm models as significant between-individual variance in intercepts, which can be understood as performance at first trial. However, using between-individual variation in intercepts and slope to predict the corresponding variance at, and correlation between, all trials (see Supplementary Material Table S1 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 were faster than average at completing Maze A in their first trial tended 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 et al., 2017 for an overview). We note that a contributing factor is likely to be the short interobservation period (here 24 h) typical of cognitive studies, since behavioural repeatabilities generally decline 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 between-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). This methodology avoids statistically problematic 'two-step' analyses, whereby individual random effect predictions are extracted from fitted mixed models and used in subsequent analyses (e.g. to test correlations or regressions) without carrying forward the uncertainty in the estimates (Houslay & Wilson, 2017). Although the strategy we used here is now widely used in studies of behavioural plasticity, it has not yet been widely adopted by researchers focusing specifically on animal cognition (but see e.g. Langley et al., 2020). In addition to finding variance in slopes (learning), we estimated a negative between-individual intercept-slope correlation using the Maze A data: individuals with higher intercepts (i.e. Maze time at first trial) tended to have lower (more negative) slopes. While it is therefore the case that those fish performing poorly initially exhibited higher rates of learning, it is also true, as noted above, that individual performance (Maze time) was positively correlated across trials 1–11. These two results are

compatible because differences in learning (slope) are not sufficiently pronounced that initially poor performing (but fast learning) fish 'overtake' initially better performing (but slow learning) individuals by trial 11. We cannot comment on what fitness consequences, if any, the variation detected here would have in wild fish. Nevertheless, this finding does highlight a danger with any common presumptions that cognitive abilities may be 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), if for no other reason than they start engaging with the task faster. Here we do not have data to test this possibility, but it is plausible that some individuals are, at least initially, more reluctant to swim through the openings in the maze. Similarly, there could be between-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 increased slightly, although not significantly, across trials. Thus, there is between-individual variation in intercept (Maze time at trial 1) and in slope. Given that there is no (significant) change in mean performance, but there is significant variation in slopes, we conclude that some individuals improved (learned) in Maze B while others became 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 whether the average rate of learning across repeated trials differs between maze A and B it is reasonable to conclude it does (given no overlap of approximate 95% confidence intervals calculated as the slope ±1.96 SE). Several possibilities may explain the finding of spatial learning on average in maze 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 spatial learning tasks (Fong et al., 2019; Kotrschal et al., 2015; Lucon-Xiccato & Bisazza, 2017b), we consider this unlikely. Second, it may be that the layout of Maze B was 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, for example due to proactive interference (difficulty inhibiting memory; Shettleworth, 2010). In this case the second maze may require more trials to detect improvement. There is some evidence for such effects in guppies. For instance, Lucon-Xiccato and Bisazza (2014) found that on average guppies took 14.6 trials to learn a reversed colour cue

association, and Fong et al. (2019) found that on average 15.3 trials were required for guppies to learn a reversed maze layout. A third explanation could be that the perceived value of the reward declined across tasks. As the social reward (visual access to a female) remained the same in both mazes, males may have learned that they cannot gain physical access to the females. This may have reduced the perceived value of the reward stimulus, and thus reduced reinforcement and motivation to complete the task over repeat trials (Berridge, 2018).

A further possible explanation could be that the potential to detect learning in Maze B is limited by the effects of 'trial fatigue' or changes in physiological or informational 'state' that impact cognitive performance and/or motivation. For example, chronic stress effects may arise from repeated capture and handling experienced in the experimental design (Huntingford et al., 2006; Warren & Callaghan, 1976; Wong et al., 2008). Such an effect could manifest in our Maze B data as an increase in the proportion of maze trials not successfully completed in <480 s (assuming trial fatigue increases the probability of not completing), coupled with a decline in average time across trial repeats for those trials that were successfully finished (assuming better performing fish are learning). However, visual inspection of the data revealed no obvious trend in proportion of trials completed in <480 s, and post hoc analysis of Maze B trials with censored records excluded provided no evidence of improvement either (results not shown). Instead, it is possible that successful individuals in Maze A may have learned generalizable rules (e.g. avoiding corners) that then helped them to navigate through Maze B. Indeed, the average time taken for successful individuals to complete Maze B was similar to the time taken to complete the last trial in Maze A, suggesting that some individuals were already performing near ceiling levels in the second task. Empiricists obviously seek to minimize the possibility of chronic stress or learned experience confounding conclusions from cognitive studies but it can be difficult to validate the implicit assumption that individuals' stress or knowledge remains (equally) unaffected over experimental periods requiring repeated observations. Here we suggest this is a plausible hypothesis but not one we can currently test.

While the influence, or otherwise, of chronic stress on our results necessarily remains speculative, our experiment does confirm between-individual variation in acute stress behaviour as measured by Relative area in the open field trials. This replicates earlier results using independent data sets of fish from the same captive population (Houslay et al., 2019; Prentice, 2020; White, 2016). Acute stressor exposure can affect cognitive performance in spatial learning tasks in both mammals and fish (Gaikwad et al., 2011; Wong et al., 2019). At the individual level, there is also evidence that acute stress responses can predict outcomes under longer term chronic and/or repeated stressor exposure (Salak-Johnson & McGlone, 2007; Segerstrom & Miller, 2004; Øverli et al., 2007). However, we found no evidence of strong relationships between acute stress behaviour and performance or learning in either maze. Thus, we found no support for the prediction, made under the stress coping style model, that (acute) stress responsiveness will covary with cognitive performance (Coppens et al., 2010; Griffin et al., 2015; Sih & Del Giudice, 2012).

In summary, here we have evidence of consistent differences between individuals in spatial task performance in the guppy. Individual performance was repeatable across trials within and between two different spatial tasks (i.e. maze layouts). This between-individual variation in performance is consistent with underlying differences in cognitive factors but differences in 'personality' (e.g. neophobia, exploratory tendency) may also contribute. We also

found 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' note that performance declined for some individuals, especially in the second maze where there was no improvement in average time across 11 trials. We have shown 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 that cumulative, chronic stress effects may have contributed 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 characterization of cognitive variation in animal studies.

Author Contributions

Pamela M. Prentice: Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing Original Draft, Review & Editing. **Chloe Mnatzaganian**: Conceptualization, Methodology, Investigation. **Thomas M. Houslay**: Visualization, Writing – Review & Editing. **Alex Thornton**: Supervision, Conceptualization, Methodology, Writing – Review & Editing. **Alastair J. Wilson**: Supervision, Conceptualization, Methodology, Formal Analysis, Resources, Writing – Review & Editing.

Acknowledgments

Supplementary Material

Supplementary material associated with this article is available, in the online version, at https://doi.org/10.1016/j.anbehav.2022.04.009.

References

ASAB Ethical Committee & ABS Animal Care Committee. (2020). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour, 159*, I–XI. https://doi.org/10.1016/j.anbehav.2019.11.002

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

Barbosa, M., & Morrissey, M. B. (2021). The distinction between repeatability and correlation in studies of animal behaviour. *Animal Behaviour, 75*, 201–217. https://doi.org/10.1016/j.anbehav.2021.03.008

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*, 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

Berridge, K. C. (2018). Evolving concepts of emotion and motivation. *Frontiers in Psychology, 9*, 1647. https://doi.org/10.3389/fpsyg.2018.01647

Bókony, V., Lendvai, Á. Z., Vágási, C. I., Pătraş, 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. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*, 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 sheepshead swordtail, *Xiphophorus birchmanni*. *Behavioral Ecology and Sociobiology, 68*(5), 791–803. https://doi.org/10.1007/s00265-014-1692-0

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

Bridger, D., Bonner, S. J., & Briffa, M. (2015). Individual quality and personality: Bolder males are less fecund in the hermit crab *Pagurus bernhardus*. *Proceedings of the Royal Society B: Biological Sciences, 282*(1803), 20142492. https://doi.org/10.1098/rspb.2014.2492

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

Brust, V., & Guenther, A. (2017). Stability of the guinea pig's personality – cognition – linkage over time. *Behavioural Processes, 134*, 4–11. https://doi.org/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), 20172031. https://doi.org/10.1098/rspb.2017.2031

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

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

Carazo, P., Noble, D. W. A., 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. https://doi.org/10.1098/rspb.2013.3275

Carere, C., Maestripieri, D., Caramaschi, D., Carere, C., Sgoifo, A., & Koolhaas, J. M. (2014). Neuroendocrine and autonomic correlates of animal personalities. In C. Carere, & D. Maestripieri (Eds.), *Animal personalities* (pp. 353–377). University of Chicago Press. https://doi.org/10.7208/chicago/9780226922065.003.0013.

Cauchoix, M., Hermer, E., Chaine, A. S., & Morand-Ferron, J. (2017). Cognition in the field: Comparison of reversal learning performance in captive and wild passerines. *Scientific Reports, 7*(1), 12945. https://doi.org/10.1038/s41598-017-13179-5

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. https://doi.org/10.1016/j.anbehav.2010.11.025

Coppens, C. M., de Boer, S. F., & Koolhaas, J. M. (2010). Coping styles and behavioural flexibility: Towards underlying mechanisms. *Philosophical Transactions of the Royal Society Series B, Biological Sciences, 365*(1560), 4021–4028. https://doi.org/10.1098/rstb.2010.0217

Croft, D. P., Albanese, B., Arrowsmith, B. J., Botham, M., Webster, M., & Krause, J. (2003). Sex-biased movement in the guppy (*Poecilia reticulata*). *Oecologia, 137*(1), 62–68. https://doi.org/10.1007/s00442-003-1268-6

Dingemanse, N. J., Both, C., Drent, P. J., & Tinbergen, J. M. (2004). Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society Series B: Biological Sciences, 271*(1541), 847–852.

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

Dingemanse, N. J., & Réale, D. (2005). Natural selection and animal personality. *Behaviour, 142*(9/10), 1159–1184. https://doi.org/10.1163/156853905774539445

Fong, S., Buechel, S. D., Boussard, A., Kotrschal, A., & Kolm, N. (2019). Plastic changes in brain morphology in relation to learning and environmental enrichment in the guppy (*Poecilia reticulata*). *Journal of Experimental Biology, 222*(10), Article jeb200402. https://doi.org/10.1242/jeb.200402

Fraser, H., Parker, T., Nakagawa, S., Barnett, A., & Fidler, F. (2018). Questionable research practices in ecology and evolution. *PLoS One, 13*(7), Article e0200303. https://doi.org/10.1371/journal.pone.0200303

Gaikwad, S., Stewart, A., Hart, P., Wong, K., Piet, V., Cachat, J., & Kalueff, A. V. (2011). Acute stress disrupts performance of zebrafish in the cued and spatial memory tests: The utility of fish models to study stress-memory interplay. *Behavioural Processes, 87*(2), 224–230. https://doi.org/10.1016/j.beproc.2011.04.004

Galsworthy, M. J., Paya-Cano, J. L., Liu, L., Monleón, S., Gregoryan, G., Fernandes, C., Schalkwyk, L. C., & Plomin, R. (2005). Assessing reliability, heritability and

general cognitive ability in a battery of cognitive tasks for laboratory mice. *Behavior Genetics, 35*(5), 675–692. https://doi.org/10.1007/s10519-005-3423-9

Gibelli, J., Aubin-Horth, N., & Dubois, F. (2019). Individual differences in anxiety are related to differences in learning performance and cognitive style. *Animal Behaviour, 157*, 121–128. https://doi.org/10.1016/j.anbehav.2019.09.008

Gilmour, A. R., Gogel, B. J., Cullis, B. R., Thompson, R., & Butler, D. (2009). *ASReml user guide release 3.0*. VSN International.

Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin, 127*(1), 45–86. https://doi.org/10.1037/0033-2909.127.1.45

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

Guillette, L. M., Hahn, A. H., Hoeschele, M., Przyslupski, A. M., & Sturdy, C. B. (2015). Individual differences in learning speed, performance accuracy and exploratory behaviour in black-capped chickadees. *Animal Cognition, 18*(1), 165–178. https://doi.org/10.1007/s10071-014-0787-3

Healy, S. D. (2019). The face of animal cognition. *Integrative Zoology, 14*, 132–144. https://doi.org/10.1111/1749-4877.12361

van Horik, J. O., & Madden, J. R. (2016). A problem with problem solving: Motivational traits, but not cognition, predict success on novel operant foraging tasks. *Animal Behaviour, 114*, 189–198. https://doi.org/10.1016/j.anbehav.2016.02.006

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Medina-García, A., Jawor, J. M., & Wright, T. F. (2017). Cognition, personality, and stress in budgerigars, *Melopsittacus undulatus. Behavioral Ecology, 28*(6), 1504–1516. https://doi.org/10.1093/beheco/arx116

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

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), Article 0038987. 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*, 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 behavioural 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. *Neuroscience and Biobehavioral Reviews, 31*, 396–412. https://doi.org/10.1016/j.neubiorev.2006.10.006

Plomin, R., & Spinath, F. M. (2002). Genetics and general cognitive ability (g). *Trends in Cognitive Sciences, 6*, 169–176. 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, 33*(5), 642–652. https://doi.org/10.1111/jeb.13601

Quinn, J. L., Cole, E. F., Bates, J., Payne, R. W., & Cresswell, W. (2012). Personality predicts individual responsiveness to the risks of starvation and predation. *Proceedings of the Royal Society B: Biological Sciences, 279*(1735), 1919–1926. https://doi.org/10.1098/rspb.2011.2227

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). https://doi.org/10.1098/rstb.2015.0184

Raoult, V., Trompf, L., Williamson, J. E., & Brown, C. (2017). Stress profile influences learning approach in a marine fish. *PeerJ, 2017*(6). https://doi.org/10.7717/peerj.3445

Réale, D., Garant, D., Humphries, M. M., Bergeron, P., Careau, V., & Montiglio, P. O. (2010). Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philosophical Transactions of the Royal Society Series B, Biological Sciences, 365*(1560), 4051–4063. https://doi.org/10.1098/rstb.2010.0208

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.

Roff, D. A. (2002). *Life history evolution*. Sinauer Associates.

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

Salak-Johnson, J. L., & McGlone, J. J. (2007). Making sense of apparently conflicting data: Stress and immunity in swine and cattle. *Journal of Animal Science, 85*, 81–88. https://doi.org/10.2527/jas.2006-538

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution, 11*(9), 1141–1152. https://doi.org/10.1111/2041-210X.13434

Segerstrom, S. C., & Miller, G. E. (2004). Psychological stress and the human immune system: A meta-analytic study of 30 years of inquiry. *Psychological Bulletin, 130*(4), 601–630. https://doi.org/10.1037/0033-2909.130.4.601

Shettleworth, S. J. (2010). *Cognition, evolution, and behavior* (2nd ed.). Oxford University Press https://books.google.com/books/about/Cognition_Evolution_and_Behavior.html?id=-Qs1qGys0AwC.

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., & 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

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. O., & Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics, 50*(4), 1171. https://doi.org/10.2307/2533455

Sweis, B. M., Veverka, K. K., Dhillon, E. S., Urban, J. H., & Lucas, L. R. (2013). Individual differences in the effects of chronic stress on memory: Behavioral and neuro-chemical correlates of resiliency. *Neuroscience, 246*, 142–159. https://doi.org/10.1016/j.neuroscience.2013.04.052

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. *Behavioral Ecology, 26*, 1460–1461. https://doi.org/10.1093/beheco/arv119

Trompf, L., & Brown, C. (2014). Personality affects learning and trade-offs between private and social information in guppies, *Poecilia reticulata. Animal Behaviour, 88*, 99–106.

Warren, E. W., & Callaghan, S. (1976). The response of male guppies (Poecilia reticutata, Peters) to repeated exposure to an open field. *Behavioral Biology, 18*(4), 499–513. https://doi.org/10.1016/S0091-6773(76)92524-4

White, S. J., Kells, T., & Wilson, A. J. (2016). Metabolism, personality and pace of life in the Trinidadian guppy, *Poecilia reticulata*. *Behaviour, 153*(13−14), 1517−1543. https://doi.org/10.1163/1568539X-00003375

White, S. J., & Wilson, A. J. (2019). Evolutionary genetics of personality in the Trinidadian guppy I: Maternal and additive genetic effects across ontogeny. *Heredity, 122*(1), 1−14. https://doi.org/10.1038/s41437-018-0082-1

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., & Nussey, D. H. (2010). An ecologists guide to the animal model. *Journal of Animal Ecology, 79*, 13−26. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2656.2009.01639.x/full.

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.

Wong, S. C., Dykstra, M., Campbell, J. M., & Earley, R. L. (2008). Measuring waterborne cortisol in convict cichlids (*Amatitlania nigrofasciata*): Is the procedure a stressor? *Behaviour, 145*(10), 1283−1305. https://doi.org/10.1163/156853908785765863

Wong, R. Y., French, J., & Russ, J. B. (2019). Differences in stress reactivity between zebrafish with alternative stress coping styles. *Royal Society Open Science, 6*(5), 181797. https://doi.org/10.1098/rsos.181797

Zidar, J., Balogh, A., Favati, A., Jensen, P., Leimar, O., Sorato, E., & Løvlie, H. (2018). The relationship between learning speed and personality is age- and task-dependent in red junglefowl. *Behavioral Ecology and Sociobiology, 72*(10), 168. https://doi.org/10.1007/s00265-018-2579-2

Appendix

Table A1
Fixed effect estimates from the full random intercept and random slope model for Maze time in both maze A and B

Maze	Fixed effect	Effect size (SE)	df	F	P
A	Intercept	0.355 (0.119)	1, 115.1	1.137	<0.001
	Trial	−0.058 (0.014)	1, 59.8	17.890	<0.001
	Maze position (top)	0.122 (0.062)	1, 599.1	3.950	0.047
	Order	−0.029 (0.015)	1, 631.5	4.153	0.057
B	Intercept	0.031 (0.135)	1, 102.8	0.001	0.563
	Trial	0.019 (0.014)	1, 53.8	2.054	0.175
	Maze position (top)	0.071 (0.066)	1, 535.4	1.163	0.281
	Order	−0.378 (0.135)	1, 559.7	4.949	0.034

Standard errors are shown in parentheses.

Table A2
Fixed effect estimates from the full random intercept and random slope model for Relative area from the open field trials

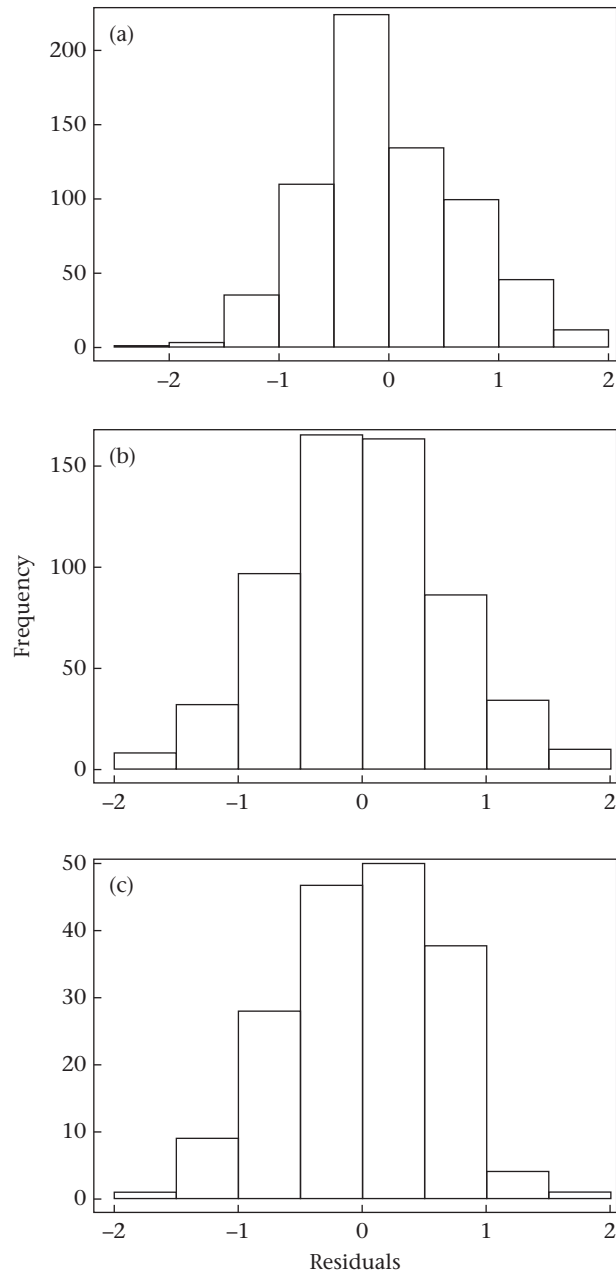| Model	Fixed effect	Effect size (SE)	df	F	P
Open field trial	Intercept	0.229 (0.451)	1, 159	0.003	0.613
	Trial	−0.262 (0.066)	1, 116.8	15.710	<0.001
	Time	0.000 (0.000)	1, 150.3	0.003	0.953

Standard errors are shown in parentheses.

Figure A1. Histogram of model residuals from univariate mixed models for each trait. Model residual distributions are used for visual assessment of conformity of Gaussian error structure. Residuals are shown from the full random regression models for (a) Maze time$_A$ and (b) Maze time$_B$, and from the random intercepts only model for (c) Relative area.
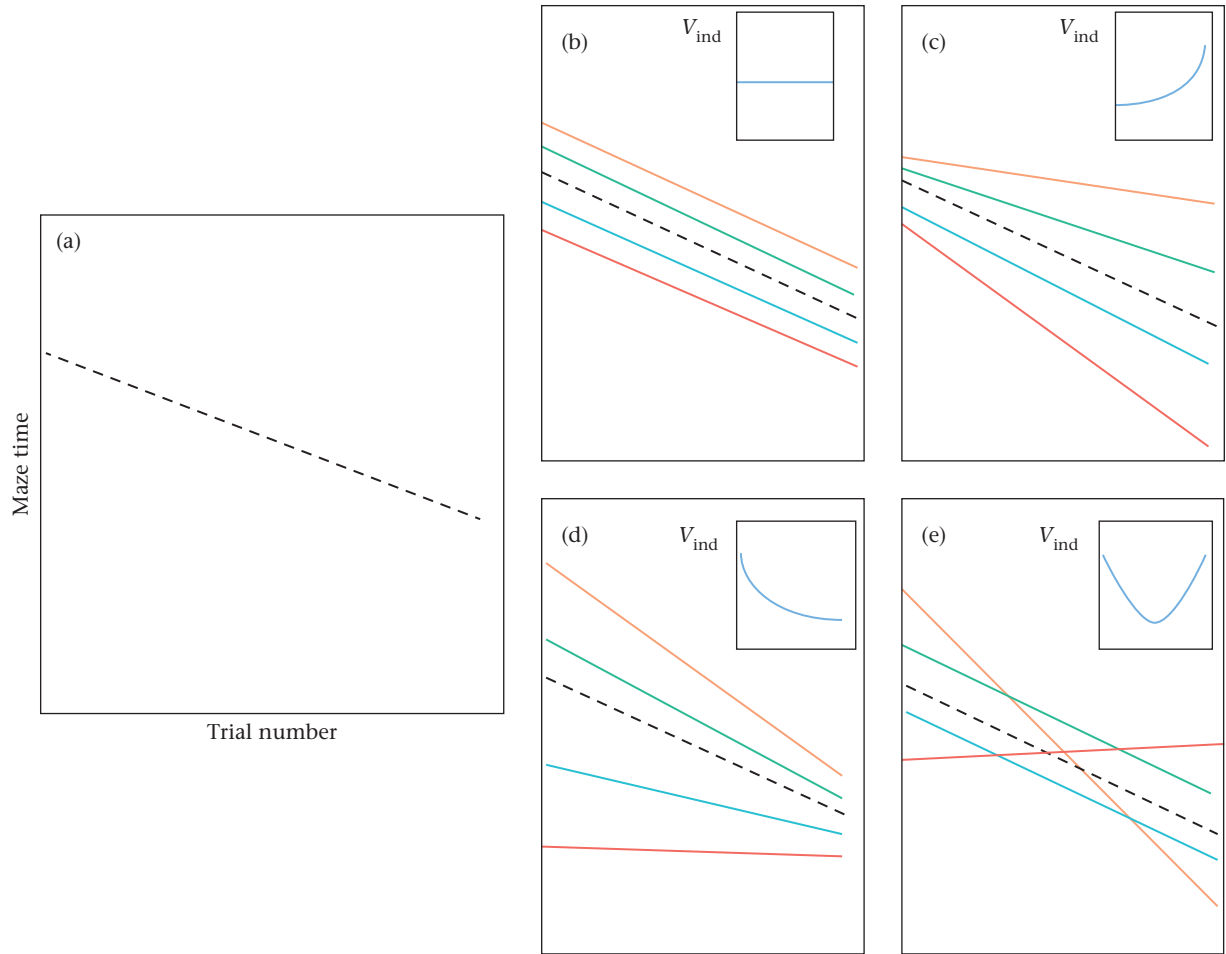
Figure A2. Characterizing individual variation in learning performance. (a) Average (black dashed line) decrease in maze time with trial number from 1 to 11 consistent with learning. Inset panels in (b) − (e) show how individual trajectories may vary around this because of differences in (b) reaction norm intercepts or (c-e) slopes. Where slopes vary (c-e), a corollary of this is that the between-individual variance (V_{ind}) in Maze time will change across trials. This could potentially increase (c) or decrease (d) monotonically, or there could be an intermediate trial number at which variance is minimized (e) or maximized (not shown). Where reaction norms tend to cross a lot within the range of trial numbers explored (e), this will result in low (and potentially negative) between-individual correlations between early and late trials.