

MISS PAMELA M PRENTICE (Orcid ID : 0000-0002-6290-9581)

Article type : Research Papers

Genetic variance for behavioural 'predictability' of stress response

Genetic variance for 'predictability'

Pamela M. Prentice¹, Thomas M. Houslay², Julien G. A. Martin³, 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

³Department of Biology, University of Ottawa, Ottawa, K1N 6N5, Canada

Primary/ Corresponding Author: Pamela M. Prentice, Address: Stella Turk Building, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10 9FE, UK. Email: pp349@exeter.ac.uk, Mobile: 07816214113, Fax: 01326 255131

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/JEB.13601](https://doi.org/10.1111/JEB.13601)

This article is protected by copyright. All rights reserved

Abstract

Key Words: phenotypic variation; predictability; personality; behavioural stress response; double hierarchical generalized linear model; heritability; quantitative genetics; within-individual variance.

Genetic factors underpinning phenotypic variation are required if natural selection is to result in adaptive evolution. However, evolutionary and behavioural ecologists typically focus on variation among individuals in their average trait values, and seek to characterise genetic contributions to this. As a result, less attention has been paid to if and how genes could contribute towards within-individual variance, or trait “predictability”. In fact, phenotypic ‘predictability’ can vary among individuals, and emerging evidence from livestock genetics suggests this can be due to genetic factors. Here we test this empirically using repeated measures of a behavioural stress response trait in a pedigreed population of wild-type guppies. We ask (1) whether individuals differ in behavioural predictability, and (2) whether this variation is heritable and so evolvable under selection. Using statistical methodology from the field of quantitative genetics, we find support for both hypotheses and also show evidence of a genetic correlation structure between the behavioural trait mean and individual predictability. We show that investigating sources of variability in trait predictability is statistically tractable, and can yield useful biological interpretation. We conclude that, if widespread, genetic variance for ‘predictability’ will have major implications for the evolutionary causes and consequences of phenotypic variation.

Introduction

Among-individual variation in behavioural traits is widely referred to as ‘animal personality’ when individuals display behaviours that are repeatable across time and context (Bell, Hankison, & Laskowski, 2009; Gosling, 2001). Individual differences, which can sometimes account for a high proportion of the total observed behavioural variation in a population (Biro & Adriaenssens, 2013), are a prerequisite for natural selection on behaviour and there is now abundant evidence that personality traits can affect fitness. Since strong directional or stabilising selection is usually predicted to erode variation, it is widely hypothesized that personality variation within populations is maintained by fitness trade-offs (Dingemanse, Both, Drent, & Tinbergen, 2004; Godin & Davis, 1995; Réale & Festa-Bianchet, 2003; Shackleton, Jennions, & Hunt, 2005). Empirical investigations of this, and related hypotheses, have been facilitated by wide uptake of linear mixed effect models that allow partitioning of among-individual trait variation and estimation of behavioural repeatabilities (R) from data containing repeated observations of known individuals (Dingemanse & Dochtermann, 2013; Alastair J. Wilson, 2018). Where pedigree or relatedness data are also available, among-individual variance can be further decomposed to estimate behavioural heritability (Alastair J. Wilson et al., 2010). While the residual, or within-individual, component of variance is normally treated as ‘noise’ arising from plasticity (Nussey, Wilson, & Brommer, 2007) and/or measurement error, some authors have argued that it deserves more attention as a source of biological insight (Westneat, Wright, & Dingemanse, 2015). Here we follow this suggestion in a study of stress-related behaviour in wild-type guppies (*Poecilia reticulata*), and ask whether within-individual variance in behaviour should itself be viewed as a trait that can respond to selection.

A convenient, but rarely scrutinised assumption of typical statistical methods used to characterise personality is that within-individual (or residual) variation in behaviour is homogeneous across individuals. However, this need not be the case. In fact residual variation can itself differ among-individuals (Stamps, Briffa, & Biro, 2012), a phenomenon variously referred to as among-individual differences in ‘within-individual behavioural variance’, ‘intra-individual variability’, or ‘consistency’ (Biro & Adriaenssens, 2013; Müller & Schrader, 2005; Stamps et al., 2012). Here we refer to this phenomenon as ‘predictability’ following terminology used by Cleasby et al (2015) and Martin et al (2017). Predictability has been the focus of some studies in human psychology (Hoffman, 2007; MacDonald, Backman, & Nyberg, 2006) but, in recent years, has become a topic of interest in behavioural ecology (Stamps et al., 2012; Westneat, Schofield, & Wright, 2013). For example, recent empirical studies have provided evidence of variation among individuals in predictability of anti-predatory behaviours (Briffa, 2013) and parental care (Westneat et al., 2013). It is perhaps easy to envisage fitness consequences of within-individual variation (e.g., animals behaving less predictably when fleeing a predator could plausibly have a

higher escape probability), although empirical estimates of selection are scarce. Specific hypotheses for the maintenance of among-individual differences in behavioural predictability are also lacking, though trade-offs among associated life history traits could offer adaptive explanations just as they do for maintenance of variation in behavioural means (Bridger, Bonner, & Briffa, 2015; Mulder, Gienapp, & Visser, 2016; Westneat et al., 2013). ‘Speed-accuracy’ trade-offs have been proposed where observed behavioural outcomes depend on cognitive decision making (Briffa, 2013) and predictability could also be condition-dependent. In the latter case if, for instance, canalising a behavioural response is costly then low predictability may represent phenotypic instability caused by poor individual condition.

Behavioural predictability has also been postulated to have a genetic basis of variation among individuals (Martin et al., 2017). The implication is that if predictability both causes fitness variation and is heritable, it can itself be viewed as a trait that will evolve under natural selection. Though empirical tests of genetic variance for behavioural predictability are scarce more is known for non-behavioural phenotypes. In particular, quantitative genetic methods (Hill, 1984; Hill & Zhang, 2004; Rönnegård, Felleki, Fikse, Mulder, & Strandberg, 2010) have been increasingly applied to estimate genetic variation for predictability of production traits in livestock, including milk yield (Rönnegård, Felleki, Fikse, Mulder, & Strandberg, 2013), litter size (Sorensen & Waagepetersen, 2003) and body weight (Sonesson, Ødegård, & Rönnegård, 2013). While increasing the mean of such production traits is a long-standing objective of artificial selection strategies, reducing the level of variation around the means also offers increased efficiency (and profitability) in livestock production and processing. Consequently, the reality that genetic variance ‘for variance’ occurs has prompted development of strategies to select more predictable genotypes, and thus reduce variation in target traits (Hill & Mulder, 2010).

Evolutionary ecologists are now beginning to address the concept of predictability in relation to behavioural and life history traits with the goals of determining whether behavioural predictability consistently varies among-individuals, whether this variation has a genetic basis, and how (if at all) predictability maps to fitness. A hindrance addressing these questions stems from a lack of consensus on how best to quantify and analyse predictability. Most studies to date have taken a two-step approach by, for instance, fitting a linear model to a set of behavioural observations, then calculating an estimate of within-individual variation using model residuals for each individual, which are then used in a subsequent analysis (Biro & Adriaenssens, 2013; Highcock & Carter, 2014; Stamps et al., 2012). Though intuitive, this approach is statistically problematic for a number of reasons, not least of which is that uncertainty in the predictions of the first model is not accounted for, increasing the risk of type 1 errors and anticonservative hypothesis tests (Houslay & Wilson 2017). Fortunately, a more robust approach to model variation in behavioural predictability is provided by the double hierarchical generalized linear model (DHGLM)

developed by Lee and Nelder (2006). This model is an extension of the familiar ‘random intercept’ mixed model, however instead of only allowing random and fixed effects on the mean trait distribution, it also allows them on the residuals. In other words, it allows us to relax the assumption that residual variance is homogeneous, and ask whether it varies across levels of fixed (e.g. sex) or random (e.g. individual identity) effects (Cleasby et al. 2015; Lee & Nelder 2006).

Applied to repeated measures behavioural data, double-hierarchical models therefore allow simultaneous estimation of 1) among-individual variation in (mean) trait expression (i.e. the normal target of personality studies), 2) variation in predictability of a trait (i.e. differences in within-individual variance) (Lee & Nelder 2006) and 3) the correlation between the mean and the predictability at the individual level. Furthermore, given pedigree data, the DHGLM approach can be combined with the quantitative genetic ‘animal model’ (in a ‘double-hierarchical animal model’ DHAM), allowing among-individual variance to be further decomposed into genetic and non-genetic components. To date, only one study has used this approach to test for and estimate the genetic basis of behavioural predictability (Martin et al., 2017). Using a DHGLM, the authors of this study found evidence of among-individual variation in the predictability of docility (the reaction to being trapped and handled), as a repeatable behaviour in marmots (*Marmota flaviventris*). They also showed that individual marmots that were (on average) less docile were also less predictable. Using pedigree information, they went on to show that both (mean) behaviour and its predictability are heritable in this population, and so evolvable under selection.

Here we use a captive population of wild-derived Trinidadian guppies (*Poecilia reticulata*) to test for variation in behavioural predictability and ask whether, if present, it arises in part from genetic differences among individuals. We focus on a putatively stress-related context, specifically the way in which an individual behaves in reaction to isolation in a novel environment, such as an ‘Open Field Trial’ (OFT) arena. The OFT is a widely used paradigm for characterising personality differences related to exploration, activity, and ‘shy-bold’ type variation (Bell et al., 2009; Gosling, 2001). Previous work with *P. reticulata* has demonstrated that behaviours displayed during OFT are associated with exploration, but also risk-taking and stress response (White, Kells, & Wilson, 2016). The fact that the OFT presents a mild stressor is notable because the widely used concept of ‘stress coping style’ predicts that individuals vary along a proactive/reactive continuum of variation (Coppens et al., 2010; Koolhaas et al., 1999; Sih, Bell, & Johnson 2004), with proactive individuals tending to express more ‘fight or flight’ behaviours on average, but also forming more rigid, stereotyped routines more rapidly (Koolhaas et al., 1999). In other words, the coping style verbal model suggests variation among individuals in not only mean behaviour and behavioural predictability, but also correlation structure between these.

Previous studies of this guppy population have already shown that the behavioural responses to the OFT are repeatable, but also plastic with respect to experimentally-manipulated stressor severity (specifically perceived predation risk) (Houslay et al., 2018). We also now know from pedigree-based analysis that (average) behaviours are heritable (White, Houslay, & Wilson, 2019; White & Wilson, 2019), and that there is genetic correlation structure between OFT behaviour and cortisol expression (strengthening the view that the OFT provides an assay of behavioural stress response; (Houslay et al., 2019)). Here we aim to build on these earlier studies by, firstly confirming the repeatability and heritability of mean behaviour in an independent sample; secondly, simultaneously estimating among-individual variation in mean behaviour and predictability using a DHGLM; and thirdly, asking whether - if present - variation in predictability is itself heritable using a DHAM. Finally, we test the prediction of the stress coping style model - at both among-individual and genetic levels - that there will be (co)variance between mean behaviour and predictability, with individuals (genotypes) displaying more 'flight' type behavioural stress responses also being more predictable. In the context of a DHGLM, we are thus predicting to have a positive covariance between a 'flight' type behavioural response and its variance at the individual (genotype) level. The 'flight' type behavioural response here was a derived trait called relative area, where individuals that have a low relative area, i.e. displaying more flight type response, are expected to have a low within-individual variance (i.e. high predictability).

Methods

Husbandry

Behavioural assays were carried out on wild-type guppies from a captive population with known pedigree structure housed at the fish laboratory at the University of Exeter's Penryn campus. Data used here have not previously been published, but were collected as part of a larger study for which methods have already been extensively described elsewhere (Houslay et al., 2019; White et al., 2019, 2016; White & Wilson, 2019). In brief, all fish used were descended from wild guppies caught from the lower Aripo River, Trinidad in 2008. They were offspring of known parental crosses (as detailed in White & Wilson, 2018, for detailed breeding protocol), that had been raised in families before being tagged at maturity and then allocated to mixed family groups. Groups comprised 16-20 individuals (at 50:50 sex ratio) in 15 l tanks, with 24 tanks within each "stack" having a common sump and shared recirculating water supply. Fish were maintained at 22–24°C on a 12:12 light/dark cycle, with weekly 25% water changes on each stack, and were fed to satiation twice daily on commercial flake food and live brine shrimp (*Artemia salina*). Note these fish were part of a larger pedigree structure containing 1,518 individual fish within a genetic pedigree structure comprised of maternal full-sibships nested within 169 paternal half-sibships (as described in (Houslay et al., 2019)). Here, we pruned the full pedigree using the prunePed function in the R package MCMCglmm (Hadfield, 2010) to just include the informative individuals. Our final data set contained phenotypic data for 330 individuals from a pedigree with 2113 maternal offspring links, 1654 paternal offspring links, 218 sires and 344 dams, with a maximum depth of 4 generations.

Behavioural data collection and trait definition

Behaviour was assayed using Open Field Trials (OFT), a standardised assay of risk-related behaviours that is widely used in rodent, fish and bird studies (Boulton, Grimmer, Rosenthal, Walling, & Wilson, 2014; White et al., 2019; White & Wilson, 2019). Our assay protocol closely followed that of Boulton et al. (2014) with repeat measures on related individuals providing the data structure needed to estimate among individual and genetic variance in personality and predictability. However, here we conducted more repeats per individual with a planned maximum of 10 times. In practice some mortality occurred over the course of the data collection period (which was five weeks for each fish). Thus, in total we conducted 2970 behavioural assays on 330 individuals (a mean of 9 per fish) from 23 groups. All experimental data was collected by the same technician, and carried out in two blocks for purely logistical reasons (Batch A; n = 176, Batch B; n = 154). For each block, fish were trialled over five weeks, with data collection occurring in weeks 1, 3 and 5 at not less than 48 hour intervals (weeks 2 and 4 providing 'breaks').

Each OFT comprised a fish being netted from its home tank and placed into an 'arena' comprised of a 30 × 20 cm fish tank filled to 5 cm water depth and lit from below with a light box. (Three identical arena 'set-ups' A, B and C were used concurrently during data collection to facilitate high throughput phenotyping, with fish allocated haphazardly among them). Following a 30 s acclimation period, individuals were tracked for 180 s from a Sunkwang C160 video camera fixed above each tank and the tracking software Viewer II (<http://www.biobserve.com>). Each fish was then returned to its home tank. Behavioural experiments were conducted under license from the Home Office (UK) and under the auspices of the Animal (Scientific Procedures) 1986 Act, and with local ethical approval from the University of Exeter.

A number of specific variables assayed by OFT have been used to assay 'risk-prone/risk-averse', or 'shy-bold' type personality variation in fishes including guppies (Sih, Bell, & Johnson, 2004; White & Wilson, 2019). Here we extracted two variables from the video - total track length swum (cm) and the area covered (percent of tank area explored, %). While both are expected a priori to be repeatable and heritable (Houslay et al., 2018; White et al., 2019; White & Wilson, 2019), previous work has failed to detect a strong positive (among-individual) correlation. This is notable since, if fish move randomly in the OFT arena, we expect area covered to increase as a monotonic function of track length. The lack of expected correlation actually arises from variation in how fish respond behaviourally to the stressor stimulus of the OFT. This variation is revealed by calculating the derived trait of relative area –defined as the difference between observed area covered and the predicted area covered given a 'random swim' of the track length actually observed (Houslay et al., 2019). To do this we (i) simulated 'random swims' in the arena across the full range of observed track lengths; (ii) estimated the 'null' relationship between simulated area and simulated track length using a fourth order polynomial regression (which captured 97.85% of the variation); and (iii) used the regression equation to predict area covered given a 'random swim' corresponding to each observed track length. Code and a full description of the simulation approach is provided in Houslay et al. (2019).

Biologically, high values of relative area arise from efficient exploration of the arena by a (putatively) less stressed individual (Fig 1 a). In contrast, low values of relative area arise from trials in which fish swim rapidly (yielding a high track length) but also display thigmotaxis (i.e. staying close to the tank wall resulting in a low area covered) (Fig 1 a). This scenario is commonly observed and is biologically interpretable as a 'flight'-type stress response (i.e. the fish is seeking escape from the arena).

Statistical analyses

First, we sought to confirm our expectations from previous work on this population that there would be among-individual variation for mean behaviour (relative area). We did this using a simple univariate linear mixed-effects model fit by REML (restricted maximum likelihood) using ASReml-R 3.0 (Butler 2009; Gilmour et al. 2002) within R version 3.4.1 (R Core Team, 2017) in which relative area was modelled with random effects of individual identity and social housing group (Model 1). In addition to the mean, we included fixed effect factors of arena set-up and fish sex, as well as within-group trialling order (as a continuous variable to account for any cumulative disturbance effect of removing fish sequentially from the home tank). Conditional F-statistics were used to determine the significance of fixed effects although we note they were simply included to control statistically for sources of variance not directly relevant to our present goals. Random effects were tested using likelihood ratio tests (LRT), assuming twice the difference in log-likelihood between full and reduced models is distributed as a 50:50 mix of χ^2_0 and χ^2_1 as recommended by Visscher (2006). We make the standard assumptions that random effects and residuals are normally distributed with means of zero and variances to be estimated. Importantly in the current context we also make the standard (but rarely stated) assumption that ‘residual’ variance is homogeneous across individuals (and fixed effect classes). We also calculated an estimate of the adjusted repeatability (conditional on fixed effects) as the intraclass correlation $R = V_I/V_P$, where V_I is the among-individual variance and V_P is the total phenotypic variance. V_P is therefore calculated as $V_I + V_{GR} + V_R$, where V_{GR} is the among-group variance (which accounts for environmental and social sources of variation among groups within home tanks) and V_R is the residual (within-individual) variance. The adjusted repeatability R is thus the proportion of phenotypic variance explained by among-individual differences in behavioural mean, after controlling for fixed effects (Nakagawa & Schielzeth, 2010).

We then extended this model by including the individual genetic merit for (mean) behaviour as an extra random effect (Model 2). This becomes the standard repeated measures animal model of quantitative genetics (with additional fixed and random effects as described above), and allowed us to utilise the pedigree data to partition V_I into additive genetic (V_A) and non-genetic, permanent environment (V_{PE}) components. We tested the significance of V_A by LRT (as described above) and estimated the narrow sense heritability h^2 (where $h^2 = V_A/V_P$ and V_P is the sum of the variance components and thus conditional on fixed effects).

To estimate among-individual variation of predictability of the behaviour (relative area), we used a double hierarchical generalized linear effect model (DHGLM) (Cleasby et al., 2015; Lee & Nelder, 2006) of relative area (Model 3). The DHGLM allows for the simultaneous analysis of a mean level model and a

dispersion level model each including fixed and random effects. We estimated not only the among-individual variation in residual variance (i.e. variation in predictability (V_{I_v})) but also the correlation between the mean behaviour and its predictability at the individual level (Cleasby et al., 2015). To simplify slightly, we included as fixed effects in the mean model only those variables that were statistically significant in Model 1, while for the dispersion part of the model, we included a fixed effect of sex (i.e. males and females are permitted to differ in average predictability). We included group and individual identity as random effects in both the mean and the dispersion part of the model. We also modelled the covariance (at group and individual levels) between the random means and the predictabilities of relative area.

Finally we extended Model 3 to include random genetic effects on both the mean and the predictability of the behaviour in a double hierarchical animal model (DHAM) (Rönnegård et al., 2010; SanCristobal-Gaudy, Elsen, Bodin, & Chevalet, 2009). The DHAM thus allows us to partition among-individual (co)variance into genetic and non-genetic (permanent environment) components using the pedigree. This DHAM (Model 4) has the same fixed effect structure as Model 3 for both mean and dispersion parts of the model. For the random effects, we included a permanent environment, an additive genetic and a group effect on both the mean and the dispersion models. Thus the double hierarchical models (Model 3 and 4) relate to each other in the same way as the ‘normal’ mixed models with random effects on the mean behaviour only (Models 1 and 2). In both model 3 and 4, the residual variance is dependent on the fixed and random effects included in the dispersion part of the model. However, it is possible to estimate an average residual variance for DHGLMs, \overline{V}_R . Assuming fixed effects in the dispersion part of the model are centred, we can estimate the (average) residual variance in model 4 as follow:

$$\overline{V}_R = \exp\left(\eta + \frac{V_{PE_v}}{2} + \frac{V_{A_v}}{2} + \frac{V_{GR_v}}{2}\right)$$

where η is the intercept of the dispersion model and V_{PE_v} , V_{A_v} and V_{GR_v} are the variance components associated with the permanent environment, genetic and group random effects in the dispersion part of the model respectively.

Both DHGLM and DHAM were fitted in a Bayesian framework using Stan (Carpenter et al., 2017) within R version 3.4.1 (R Core Team, 2017) via the package RStan version 2.18.0 (Stan Development Team, 2018). In order to optimize model specification in Stan, the residual variance was modelled on the log-normal scale and the covariance matrices of random effects were estimated as standard deviations and correlation matrices. We used uninformative (or weak) priors on all parameters. For fixed effect priors we

used a normal distribution with mean of zero and a variance of 100. We used a half-cauchy distribution (cauchy(0,5)) for standard deviations and for the correlation matrices, we used a LKJ correlation distribution, parametrized in terms of its Cholesky factor (allowing for a uniform distribution between -1 and 1 for the correlation). Model 3 was fitted using 5 Markov chains each including 6,000 iterations, 2,000 burn-in iterations and a thinning interval of 10. Model 4 was fitted using 5 chains each with 43,000 iterations, 3,000 burn-in iterations and a thinning interval of 100. Convergence was first assessed by visually inspecting the trace plots, which were also used to identify an appropriate number of burn-in iterations. We then checked that the Monte Carlo error was less than 1-5% of the posterior standard deviation, that the Brooks-Gelman-Rubin (BGR) diagnostic converged to 1 ± 0.2 and that the autocorrelation was below 0.05 for all parameters (Kass, Gilks, Richardson, & Spiegelhalter, 1997). The mode and 95% Highest Posterior Density Intervals (HPDI) were used to summarise the posterior distributions of the model parameters. For all calculated parameters (e.g., R , h^2), the parameters are calculated at each iteration and we reported their posterior mode and HPDI. Consequently, their estimates might differ slightly from the calculation done directly on the posterior mode of their components.

Results

Model 1 revealed significant among-individual variation in relative area, (*repeatability*(with *SE*), $R = 0.288$ (0.024), $\chi^2_{0,1} = 517.44$, $P < 0.001$) (Table 1). Under this model the social group effect was also significant ($\chi^2_{0,1} = 10.63$, $P = 0.001$ though it only explained 3.7% (1.9%) of the total variance. Comparison between model 1 and 2 provided strong evidence for significant additive genetic variance in mean behaviour ($\chi^2_{0,1} = 10.88$, $P < 0.001$), with the latter yielding an estimated heritability for relative area of $h^2 = 0.110$ (0.052), conditional on fixed effects (Table 1). These results are consistent with previously reported estimates using independent data from the same population ($h^2 = 0.080$ (0.003)) (Houslay et al., 2019), but note that both estimates are conditional on fixed effects that differ slightly. Fixed effects estimated from the current (and subsequent) models are presented in the Supplementary Material (Table S1 and S2) for completeness, although are not directly relevant to our hypotheses in this study.

The mean part of model 3 yields very similar point estimates of individual and group level variances to model 1, although the posterior of the latter was not very clearly distinct from zero. The repeatability of (mean) behaviour under Model 3 is the same as that obtained in the standard repeat measures mixed model (Model 1) with $R = 0.288$ (95% CrI, 0.248-0.348). More notably Model 3 provided evidence of among-individual variance in predictability, and also of a strong negative correlation between the individual mean and predictability of behaviour (Table 1). While frequentist-type P values are not applicable given the Bayesian inference, the 95% credible interval of the individual level variance in the dispersion part of the mode ($V_{I,\nu}$) is clearly distinct from zero. Similarly, the credible interval of the individual mean-predictability correlation is narrow and does not span zero (from which we can conclude statistical 'significance').

These findings are mirrored at the genetic level. Thus Model 4 yields very similar estimates for the heritability of (mean) behaviour. However, the DHAM also shows that both the variation in individual predictability and the mean-predictability correlation estimated in Model 3 have a genetic basis (Table 1). More precisely, both the genetic and permanent environment correlations between the mean behaviour and the variation (i.e. predictability) are strongly negative with 95% CrI that do not overlap zero (Table 1, Fig 2). Thus, individuals - and genotypes - that are more explorative (express high mean relative area) are also more predictable (i.e. less variable) in their behavioural response to the OFT. Point estimates of the corresponding group level correlations are similarly strongly negative, though we reiterate that the amount of variance in mean behaviour explained by group is low. Fixed effect estimates from all models are not discussed here but are reported in full in the Supplementary Information (Supp Information, Tables S1 and S2).

Table 1: Estimated variance components and derived parameters for mean and dispersion parts of the four models of relative area. Subscripts denote residual (R), group (GR), individual (I), permanent environment (PE) and additive genetic (A) components of variance and the corresponding mean-predictability correlations (r). We use the second subscripts (v) to denote variance in the dispersion part of the model (applicable to models 3 and 4 only). Also shown are the familiar measures of repeatability (R) and heritability (h^2) of mean behaviour (estimated at an average residual variance in models 3 and 4). Values in parentheses indicate approximate standard errors for models 1 and 2, and 95% credible intervals for models 3 and 4.

Parameter	Model			
	1 (repeatability model)	2 (animal model)	3 (DHGLM)	4 (DHAM)
V_R	0.640 (0.018)	0.640 (0.018)	0.643 (0.587, 0.716)	0.617 (0.563, 0.74)
Mean				
V_{GR}	0.035 (0.019)	0.034 (0.018)	0.020 (0.004, 0.059)	0.021 (<0.001, 0.056)
V_I	0.273 (0.028)	-	0.234 (0.191, 0.295)	-
V_{PE}	-	0.173 (0.044)	-	0.138 (<0.001, 0.226)
V_A	-	0.105 (0.052)	-	0.052 (0.004, 0.302)
R	0.288 (0.024)	-	0.269 (0.221, 0.309)	-
pe^2	-	0.182 (0.047)	-	0.165 (<0.001, 0.242)
h^2	-	0.110 (0.052)	-	0.092 (0.009, 0.306)
Predictability (dispersion)				
V_{GRv}	-	-	<0.001 (<0.001, 0.038)	<0.001 (<0.001, 0.038)
V_{Iv}	-	-	0.328 (0.244, 0.421)	-
V_{PEv}	-	-	-	0.157 (<0.001, 0.264)
V_{Av}	-	-	-	0.146 (0.034, 0.420)
Mean-predictability correlation				
r_{GR}	-	-	-0.482 (-0.922, 0.486)	-0.603 (-0.951, 0.473)
r_I	-	-	-0.955 (-0.988, -0.858)	-
r_{PE}	-	-	-	-0.956 (-0.998, -0.199)
r_A	-	-	-	-0.921 (-0.987, -0.623)

Discussion

Here, we show evidence of among-individual variation in stress-related behaviour in the guppy *P. reticulata* and show that variation arises partly through heritable differences among fish. The present data thus provide confirmation of earlier results showing genetic variation for individual mean behaviours expressed during open field trials (Houslay et al., 2019, 2018; White et al., 2019; White & Wilson, 2019). However, while previous analyses were limited to individual means, we now also show that (i) fish differ in behavioural predictability of relative area; (ii) variation in predictability is underpinned by additive genetic effects, and (iii) individual mean behaviour is genetically correlated with predictability. Our results show that the assumption of homogeneous residuals, which is typical to linear (mixed) models applied in personality research (Brommer, 2013; Dingemans & Dochtermann, 2013), is violated. Fortunately, this will not generally bias measurement of among-individual or additive genetic variance in mean behaviours. However, it does highlight how standard analytical approaches will necessarily miss interesting and important components of variation among-individuals. In what follows we first discuss our findings in relation to the behavioural stress response in guppies. We then broaden our focus with the aim of highlighting several consequences of (genetic) variance in predictability. We argue that this phenomenon has interesting implications for the evolution of phenotypes under selection that are more general than the current behavioural context.

First, we found variation in (mean) risk-related behaviour in this population of guppies at both the individual and genetic level. This is consistent with our previous work on the same population (Houslay et al., 2019), other species of wild-type poeciliid (Boulton et al. 2018), and the growing empirical evidence of heritable ‘personality’ variation across taxa (Dochtermann et al. 2014). Second, and of greater novelty, is the finding that behavioural predictability differs among individuals. Furthermore, our analyses demonstrate correlations between mean and predictability such that individuals expressing low relative area (i.e. more flight-type behavioural responses) are also less predictable (i.e. more variation in response to the OFT). The presence of correlation structure between behavioural mean and predictability is consistent with findings at the phenotypic (among individual) level from several other recent studies (Mitchell, Fanson, Beckmann, & Biro, 2016; Stamps et al., 2012). For example, a negative phenotypic correlation between mean activity rates and within-individual variation was previously estimated in guppies, where individuals that were more active, were also more predictable (Mitchell et al., 2016). Here, by using pedigree analysis we are also able to show that these individual-level patterns are underpinned by correlated genetic effects on behavioural means and predictabilities. To our knowledge only one previous study has attempted to measure a genetic correlation between mean behaviour and predictability (Martin et al., 2017). This study of docility in marmots estimated a negative correlation between mean behaviour and predictability, though the genetic correlation was not statistically significant.

Although the stress coping style (SCS) model does propose a relationship between average behavioural response to a stressor stimulus, and the predictability of behaviour, our results do not fully align with its specific predictions. This is because the structural pattern of observed variation in relative area found is not consistent with the proactive-reactive model of SCS (Coppens et al., 2010; Koolhaas et al., 1999; Sih, Bell, & Johnson, 2004) in which proactive individuals are expected to express more ‘fight or flight’ stress responses on average, but are also expected to be ‘bolder’ and/or more exploratory than reactive types. In fact, variation in relative area is orthogonal to this expectation, because it discriminates between a (putatively stressed) ‘flight’ response to the OFT (low relative area) and a (putatively less stressed) exploratory response (high relative area). In other words, relative area is probably better interpreted as measuring the magnitude, rather than ‘style’ of the behavioural stress response. Thus, while SCS predicts that high (mean) ‘flight’ behaviour will be linked to high predictability within the proactive coping style, we find it is linked to low predictability instead and likely reflects a high magnitude of stress responsiveness. While this means the stress coping model does not provide a good description of guppy responses to the OFT (Houslay et al., 2019), we nevertheless argue that it provides a useful heuristic framework precisely because it emphasises the need to evaluate integration among stress-response components in a multivariate empirical framework. Here we show links between mean behaviour and predictability, but there is also evidence of genetic integration between (mean) behaviour and glucocorticoid (GC) physiology (flight type behaviours being associated with higher GC levels; Houslay et al. 2019). It therefore seems likely that predictability will also be genetically correlated with GC responses and their rates of habituation to repeated or chronic stressor exposure (Houslay et al., 2019) though this remains to be confirmed.

Before considering the evolutionary implications of this genetic covariance structure further, it is perhaps worth noting that mean-variance (or predictability) relationships may sometimes be inevitable given trait definitions and distributions (Tatliyer, Cervantes, Formoso-Rafferty, & Gutiérrez, 2019). Here we derived the trait of relative area as a biologically relevant measure of behavioural stress response using observed data on the actual area covered and the distance swum. Specifically, relative area is defined as the difference between observed area covered and the predicted area covered given a ‘random swim’ as long as the observed track length. It is inevitable that the possible range of observed area covered is restricted for trials of low track lengths (i.e. a fish cannot cover 100% of the area with a very short track length) and, thus it is possible that (genetic) variance in predictability of relative area arises in part from (genetic) variance in mean track length. Although such dependencies might in principle also drive the (genetic) correlation between mean and predictability of relative area, this would lead in the present case to a positive correlation (Tatliyer et al., 2019), not a negative one as estimated here. However, to explore this further we fitted a post hoc DHAM model of the observed area covered (Supp Information, Table S3). We found the

same pattern as reported above with respect to relative area; negative correlations were present between the mean behaviour and predictability at the individual and genetic levels. Thus, while it is necessarily true that our quantitative results for relative area depend on track length, we do not think there is any sense in which our qualitative conclusions are driven by mathematical artefacts of trait definition. For completeness, we also ran a further DHAM for track length (another biologically relevant trait often used to investigate personality traits such as boldness or activity) (Burns & Rodd, 2008; White & Wilson, 2019) (Supp Information, Table S4). We found among-individual differences in the trait mean and within-individual variance, which was in part due to additive genetic effects. This further suggests that the presence of genetic variance in predictability is not a particularly trait-limited phenomenon.

Our results add to the small but emerging set of studies evidencing among-individual and genetic variance for predictability (or intra-individual variation). If widespread, this could have major implications beyond the present focus on stress response and coping strategies. Variance among individuals means that behavioural predictability could be a direct target of selection, and if this does occur, the fact that it is heritable means it could evolve under selection. Furthermore, genetic correlation between the individual mean and the variation around it will allow correlated evolution of predictability in response to selection on ‘personality’ (individual average behaviour) and vice versa. Clearly our study tells us nothing about the fitness consequences of behavioural predictability in wild guppies. Nonetheless, low predictability can sometimes be selectively advantageous for prey species under specific predation threats (Briffa, 2013; Chang, Teo, Norma-Rashid, & Li, 2017). For instance, in the jumping spider (*Cosmophasis umbratical*) low predictability is advantageous when faced with aggressive predators (Chang et al., 2017). There is also some evidence for predictability-fitness associations in the pill bug (*Armadillidium vulgare*), where individuals become less predictable in risk-taking behaviour in unfamiliar, rather than familiar environments (Horváth et al., 2019) (but see Richardson et al. 2018 for a counter-example). Predator-mediated direct selection on predictability thus seems at least a plausible hypothesis in guppies (though indirect selection arising from causal effects of the genetically correlated mean behaviour could be more important).

We also note that, while advantages of low predictability do occur in a behavioural context (Briffa, 2013), it seems likely that high within-individual variation may more often be costly. For instance, given a single (constant) phenotypic optimum, an individual predictably expressing this value across multiple selective events will have higher fitness than a less predictable individual with the same mean phenotype. Where canalisation of some continuously distributed trait around the optimum is itself ‘expensive’, individuals of lower ‘quality’ or ‘condition’ may also be less predictable in trait expression and incur costs as a result (Westneat et al., 2015). In other words, low predictability can be a symptom of inability to buffer trait expression against environmental effects. A complementary perspective at the genetic level is gained

by recognising that heritable differences in predictability can equally be viewed as ‘genetic heterogeneity of environmental variance’ (Mulder, Bijma, & Hill, 2007). Thus, this phenomenon is a manifestation of genotype x environment (GxE) interaction, in which the genotype-phenotype map is sensitive to one or more environmental parameters (Nussey et al., 2007). Although GxE are normally investigated across gradients of some environmental parameter defined a priori, our results show that some guppy genotypes are more phenotypically plastic (i.e. less predictable) than others in respect to unknown (and uncontrolled) environmental variables. This shows that application of DHAM could be a useful strategy for characterising the potential importance of GxE in scenarios where the most relevant or appropriate descriptor of environmental variation is itself unclear (e.g. wild populations experiencing complex multivariate changes in environmental state). It is notable, for instance, that (linear) reaction norm models applied to wild vertebrates in naturally variable environments have generally detected limited support for GxE (e.g. Hayward et al. 2018), while evidence from experimental studies that manipulate environment conditions is compelling (Des Marais, Hernandez, & Juenger, 2013; Ingleby, Hunt, & Hosken, 2010; Pigliucci, Whitton, & Schlichting, 1995). One explanation for this might be the (univariate) environmental descriptors used in the former, and/or the assumption that reaction norms are linear, have been inadequate or inappropriate choices. Typically ‘extrinsic’ variables (e.g. measures of climate) have been used, though some studies used environment specific trait means (Ramakers, Culina, Visser, & Gienapp, 2018) or other measures of average ‘performance’ (e.g., annual mortality; Wilson et al., 2006) as proxies for overall environmental quality. This approach is common in plant studies (following Finlay & Wilkinson, 1963) and may well have wider utility in evolutionary ecology. Regardless, demonstrating the presence of genetic variance for ‘predictability’ in behaviours or other traits could be a useful starting point for more targeted investigation of which specific environmental factors genotypes are responding to, and of what functional form those responses take.

In conclusion, here we build on previous studies highlighting genetic variation in mean behavioural stress response traits, to show that variance is also present in predictability of behaviour. This variation among individuals is itself underpinned by additive genetic effects, meaning behavioural predictability can be viewed as trait with adaptive potential under selection. Furthermore, this is one of the first studies to estimate a genetic correlation between mean behaviour and predictability (i.e. within-individual variation), and so highlights the expectation that these aspects of phenotype will coevolve under selection. We recommend wider application of double hierarchical models, including the DHAM used here, to investigate the presence and causes of among-individual heterogeneity in environmental sensitivity of phenotypes generally (including but not limited to behaviours). By doing this we will gain a more complete picture of how variation is structured within and across hierarchical levels, and consequently a deeper understanding of the evolutionary ecology of labile traits in general.

Acknowledgements

We are grateful to Jip Ramakers and one anonymous reviewer for constructive comments on this MS. We would also like to thank A. Grimmer, W. Lammers, A. Campbell and R. Woodward for technical and animal husbandry support. This work was supported by the Biotechnology and Biological Sciences Research Council (grant BB/L022656/1) and by a BBSRC studentship to PMP.

Data accessibility

Data available from the Dryad Digital Repository: doi:10.5061/dryad.r4xgxd28b

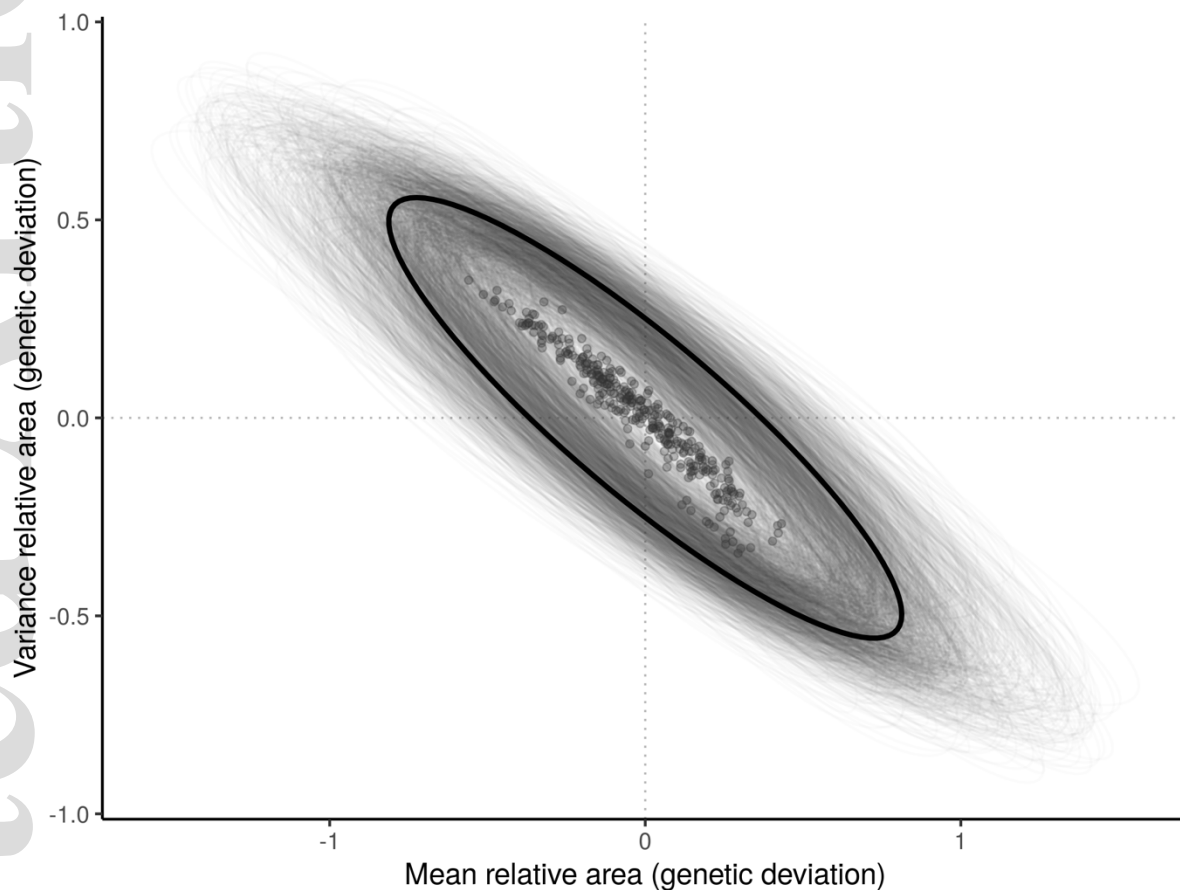
Figure 1. (a) Illustration of contrasting Open Field Trials (OFT), showing a) tracks swum by two individual fish (1, 2) as blue lines, and b) the resulting trait data. Here both fish swim a very similar track length, but individual 1 also covers a high percentage of the tank (Area covered = 65.7%) and displays an exploratory phenotype. By comparison Individual 2 covers much less area and is exhibiting a characteristic stress response of fast swimming along the tank walls. This results in very different values of relative area (RA) (where $RA = AC - ACTL$, AC = area covered, = track length, and $ACTL$ is the expected area covered in a random swim of observed TL , predicted by a fourth order polynomial regression fitted to simulated data; see Houslay et al 2019 for further details).

a)

b)

Trait Individual	Track Length (cm)	Area Covered (%)	Relative Area
1	661.1	65.7	15.0
2	633.2	12.4	-37.7

Figure 2. Estimated genetic matrix of relative area mean and trait predictability. The black line captures 95% of variance in the genetic correlation between mean and predictability in relative area. The grey ellipses are bootstrapped replicates from model 4, showing uncertainty around the estimated matrix. Individual points are best linear unbiased predictions (BLUPs) of genetic values from individuals in our data set.



References

- Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: a meta-analysis. *Animal Behaviour*. <https://doi.org/10.1016/j.anbehav.2008.12.022>
- Biro, P. A., & Adriaenssens, B. (2013). Predictability as a Personality Trait: Consistent Differences in Intraindividual Behavioral Variation. *The American Naturalist*, 182(5), 621–629. <https://doi.org/10.1086/673213>
- 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>
- Boulton, K., Walling, C. A., Grimmer, A. J., Rosenthal, G. G., & Wilson, A. J. (2018). Phenotypic and genetic integration of personality and growth under competition in the sheepshead swordtail, *Xiphophorus birchmanni*. *Evolution*, 72(1), 187–201. <https://doi.org/10.1111/evo.13398>
- Bridger, D., Bonner, S. J., & Briffa, M. (2015). Individual quality and personality: bolder males are less fecund in the hermit crab *Pagurus bernhardus*. *Proc Biol Sci*, 282(1803), 20142492. <https://doi.org/10.1098/rspb.2014.2492>
- Briffa, M. (2013). Plastic proteans: Reduced predictability in the face of predation risk in hermit crabs. *Biology Letters*, 9(5), 20130592. <https://doi.org/10.1098/rsbl.2013.0592>
- Brommer, J. E. (2013, June 26). On between-individual and residual (co)variances in the study of animal personality: Are you willing to take the “individual gambit”? *Behavioral Ecology and Sociobiology*, 67(6), 1027–1032. <https://doi.org/10.1007/s00265-013-1527-4>
- 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>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan : A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1). <https://doi.org/10.18637/jss.v076.i01>
- Chang, C. C., Teo, H. Y., Norma-Rashid, Y., & Li, D. (2017). Predator personality and prey behavioural predictability jointly determine foraging performance. *Scientific Reports*, 7(1), 40734. <https://doi.org/10.1038/srep40734>
- Cleasby, I. R., Nakagawa, S., & Schielzeth, H. (2015). Quantifying the predictability of behaviour: Statistical approaches for the study of between-individual variation in the within-individual variance. *Methods in Ecology and Evolution*, 6(1), 27–37. <https://doi.org/10.1111/2041-210X.12281>
- 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 of London. Series B, Biological Sciences*, 365(1560), 4021–4028. <https://doi.org/10.1098/rstb.2010.0217>

Des Marais, D. L., Hernandez, K. M., & Juenger, T. E. (2013). Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), 5–29. <https://doi.org/10.1146/annurev-ecolsys-110512-135806>

Dingemanse, N. J., Both, C., Drent, P. J., & Tinbergen, J. M. (2004). Fitness consequences of avian personalities in a fluctuating environment. *Proceedings. Biological Sciences / The Royal Society*, 271(1541), 847–852. <https://doi.org/10.1098/rspb.2004.2680>

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>

Dochtermann, N. A., Schwab, T., & Sih, A. (2014). The contribution of additive genetic variation to personality variation: Heritability of personality. *Proceedings of the Royal Society B: Biological Sciences*, 282(1798). <https://doi.org/10.1098/rspb.2014.2201>

Finlay, K.W. & Wilkinson, G.N. (1963) The Analysis of Adaptation in a Plant-Breeding Programme. *Australian Journal of Agricultural Research*, 14, 742-754. <https://doi.org/10.1071/AR9630742>

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

Godin, J.-G. J., & Davis, S. A. (1995). Who Dares, Benefits: Predator Approach Behaviour in the Guppy (*Poecilia reticulata*) Deters Predator Pursuit. *Proceedings of the Royal Society B: Biological Sciences*, 259(1355), 193–200. <https://doi.org/10.1098/rspb.1995.0028>

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>

Hadfield, J. (2010). MCMC Methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*. 33, 1–22.

Hayward, A. D., Pemberton, J. M., Berenos, C., Wilson, A. J., Pilkington, J. G., & Kruuk, L. E. B. (2018). Evidence for selection-by-environment but not genotype-by-environment interactions for fitness-related traits in a wild mammal population. *Genetics*, 208(1), 349–364. <https://doi.org/10.1534/genetics.117.300498>

Highcock, L., & Carter, A. J. (2014). Intraindividual variability of boldness is repeatable across contexts in a wild lizard. *PLoS ONE*, 9(4). <https://doi.org/10.1371/journal.pone.0095179>

Hill, W. G. (1984). On selection among groups with heterogeneous variance. *Animal Production*, 39(03), 473–477. <https://doi.org/10.1017/s0003356100032220>

Hill, W. G., & Mulder, H. A. (2010). Genetic analysis of environmental variation. *Genetics Research*,

92(5–6), 381–395. <https://doi.org/10.1017/S0016672310000546>

Hill, W. G., & Zhang, X. S. (2004). Effects on phenotypic variability of directional selection arising through genetic differences in residual variability. *Genetical Research*, 83(2), 121–132.

<https://doi.org/10.1017/S0016672304006640>

Hoffman, L. (2007). Multilevel Models for Examining Individual Differences in Within-Person Variation and Covariation Over Time. *Multivariate Behavioral Research*, 42(4), 609–629.

<https://doi.org/10.1080/00273170701710072>

Horváth, G., Garamszegi, L. Z., Bereczki, J., Urszán, T. J., Balázs, G., & Herczeg, G. (2019). Roll with the fear: environment and state dependence of pill bug (*Armadillidium vulgare*) personalities. *Science of Nature*, 106(3–4). <https://doi.org/10.1007/s00114-019-1602-4>

Houslay, T. M., Earley, R. L., White, S. J., Lammers, W., Grimmer, A. J., Travers, L. M., ... Wilson, A. J. (2019). Genetic integration of the stress response. *BioRxiv*, 770586. <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/ax023>

Ingleby, F. C., Hunt, J., & Hosken, D. J. (2010). The role of genotype-by-environment interactions in sexual selection. *Journal of Evolutionary Biology*, 23(10), 2031–2045. <https://doi.org/10.1111/j.1420-9101.2010.02080.x>

Kass, R. E., Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1997). Markov Chain Monte Carlo in Practice. *Journal of the American Statistical Association*, 92(440), 1645.

<https://doi.org/10.2307/2965438>

Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H., ... Blokhuis, H. J. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience & Biobehavioral Reviews*, 23(7), 925–935.

Lee, Y., & Nelder, J. A. (2006). Double hierarchical generalized linear models - Discussion. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 55(2), 139–185. <https://doi.org/10.1111/j.1467-9876.2006.00538.x>

MacDonald, S. W. S., Backman, L., & Nyberg, L. (2006). Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences*, 29(8), 474–480.

Retrieved from <https://www.sciencedirect.com/science/article/pii/S0166223606001251>

Martin, J. G. A., Pirotta, E., Petelle, M. B., & Blumstein, D. T. (2017). Genetic basis of between-individual and within-individual variance of docility. *Journal of Evolutionary Biology*, 30(4), 796–805.

<https://doi.org/10.1111/jeb.13048>

- Mitchell, D. J., Fanson, B. G., Beckmann, C., & Biro, P. A. (2016). Towards powerful experimental and statistical approaches to study intraindividual variability in labile traits. *Royal Society Open Science*, 3(10), 160352. <https://doi.org/10.1098/rsos.160352>
- Mulder, H. A., Bijma, P., & Hill, W. G. (2007). Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. *Genetics*, 175(4), 1895–1910. <https://doi.org/10.1534/genetics.106.063743>
- Mulder, H. A., Gienapp, P., & Visser, M. E. (2016). Genetic variation in variability: Phenotypic variability of fledging weight and its evolution in a songbird population. *Evolution*, 70(9), 2004–2016. <https://doi.org/10.1111/evo.13008>
- Müller, R., & Schrader, L. (2005). Behavioural consistency during social separation and personality in dairy cows. *Behaviour*, 142(9–10), 1289–1306.
- 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>
- 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>
- Pigliucci, M., Whitton, J., & Schlichting, C. D. (1995). Reaction norms of *Arabidopsis*. I. Plasticity of characters and correlations across water, nutrient and light gradients. *Journal of Evolutionary Biology*, 8(4), 421–438. <https://doi.org/10.1046/j.1420-9101.1995.8040421.x>
- R Development Core Team. (2008). *Control*, 55, 275–286. Retrieved from <https://stat.ethz.ch/pipermail/r-help/2008-May/161481.html>
- Ramakers, J., Culina, A., Visser, A., & Gienapp, P. (2018). Environmental coupling of heritability and selection is rare and of minor evolutionary significance in wild populations. *Nature Ecology and Evolution*, 2(7), 1093–1103. <https://doi.org/10.1038/s41559-018-0577-4>
- Réale, D., & Festa-Bianchet, M. (2003). Predator-induced natural selection on temperament in bighorn ewes. *Animal Behaviour*, 65(3), 463–470. <https://doi.org/10.1006/anbe.2003.2100>
- Richardson, G., Dickinson, P., Burman, O. H. P., & Pike, T. W. (2018). Unpredictable movement as an anti-predator strategy. *Proceedings of the Royal Society B: Biological Sciences*, 285(1885), 20181112. <https://doi.org/10.1098/rspb.2018.1112>
- Rönnegård, L., Felleki, M., Fikse, F., Mulder, H. A., & Strandberg, E. (2010). Genetic heterogeneity of residual variance - Estimation of variance components using double hierarchical generalized linear models. *Genetics Selection Evolution*, 42(1), 8. <https://doi.org/10.1186/1297-9686-42-8>
- Rönnegård, L., Felleki, M., Fikse, F., Mulder, H. A., & Strandberg, E. (2013). Variance component and breeding value estimation for genetic heterogeneity of residual variance in Swedish Holstein dairy

- cattle. *Journal of Dairy Science*, 96(4), 2627–2636. <https://doi.org/10.3168/jds.2012-6198>
- SanCristobal-Gaudy, M., Elsen, J.-M., Bodin, L., & Chevalet, C. (2009). Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. *Genetics Selection Evolution*, 30(5), 423. <https://doi.org/10.1186/1297-9686-30-5-423>
- Shackleton, M. A., Jennions, M. D., & Hunt, J. (2005). Fighting success and attractiveness as predictors of male mating success in the black field cricket, *Teleogryllus commodus*: The effectiveness of no-choice tests. *Behavioral Ecology and Sociobiology*, 58(1), 1–8. <https://doi.org/10.1007/s00265-004-0907-1>
- Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: an ecological and evolutionary overview. *Trends in Ecology & Evolution*, 19(7), 372–378. <http://www.jstor.org/stable/10.1086/432265>
- 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>
- Sonesson, A. K., Ødegård, J., & Rønnegård, L. (2013). Genetic heterogeneity of within-family variance of body weight in Atlantic salmon (*Salmo salar*). *Genetics Selection Evolution*, 45(1). <https://doi.org/10.1186/1297-9686-45-41>
- Sorensen, D., & Waagepetersen, R. (2003). Normal linear models with genetically structured residual variance heterogeneity: a case study. *Genetical Research*, 82(3), 207–222. <https://doi.org/10.1017/s0016672303006426>
- Stamps, J. A., Briffa, M., & Biro, P. A. (2012). Unpredictable animals: Individual differences in intraindividual variability (IIV). *Animal Behaviour*, 83(6), 1325–1334. <https://doi.org/10.1016/j.anbehav.2012.02.017>
- Stan Development Team. (2018). RStan: the R interface to Stan. R package version 2.17.3. [https://doi.org/10.3168/jds.S0022-0302\(63\)89186-9](https://doi.org/10.3168/jds.S0022-0302(63)89186-9)
- Tatliyer, A., Cervantes, I., Formoso-Rafferty, N., & Gutiérrez, J. P. (2019). The Statistical Scale Effect as a Source of Positive Genetic Correlation Between Mean and Variability: A Simulation Study. *G3: Genes|Genomes|Genetics*, 9(9), 3001–3008. <https://doi.org/10.1534/g3.119.400497>
- Visscher, P. M. (2006). A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin Research and Human Genetics*, 9(4), 490–495. <https://doi.org/10.1375/183242706778024928>
- Westneat, D. F., Schofield, M., & Wright, J. (2013). Parental behavior exhibits among-individual variance, plasticity, and heterogeneous residual variance. *Behavioral Ecology*. <https://doi.org/10.1093/beheco/ars207>
- Westneat, D. F., Wright, J., & Dingemanse, N. J. (2015). The biology hidden inside residual within-individual phenotypic variation. *Biological Reviews*, 90(3), 729–743.

<https://doi.org/10.1111/brv.12131>

White, S. J., Houslay, T. M., & Wilson, A. J. (2019). Evolutionary genetics of personality in the Trinidadian guppy II: sexual dimorphism and genotype-by-sex interactions. *Heredity*, 122(1), 15–28.

<https://doi.org/10.1038/s41437-018-0083-0>

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. (2018). How should we interpret estimates of individual repeatability? *Evolution Letters*, 2(1), 4–8. <https://doi.org/10.1002/evl3.40>

Wilson, A. J., Pemberton, J. M., Pilkington, J. G., Coltman, D. W., Mifsud, D. V., ... Kruuk, L. E. B. (2006). Environmental coupling of selection and heritability limits evolution. *PLoS Biology*, 4(7), 1270–1275. <http://doi.org/10.1371/journal.pbio.0040216>

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., ... Nussey, D. H. (2010). An ecologists guide to the animal model. *J Anim Ecol*, 79(1), 13–26. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2656.2009.01639.x/full>