Unifying individual differences in personality, predictability, and plasticity: a practical guide

Running headline

Measuring individuality with multilevel models

Authors

Rose E. O'Dea[1,2*], Daniel W.A. Noble[3], Shinichi Nakagawa[1,2]

Affiliations

[1]Evolution & Ecology Research Centre, School of Biological and Environmental Sciences, University of New South Wales, Sydney, NSW, 2052, Australia.

[2]Diabetes and Metabolism Division, Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, Sydney, NSW 2010, Australia

[3]Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, Australia.

**Corresponding author:* rose.eleanor.o.dea@gmail.com

Author contributions statement

Rose O'Dea: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Visualisation; Writing – original draft; Writing – reviewing & editing.

Daniel Noble: Conceptualization; Funding acquisition; Investigation; Methodology; Resources; Writing – reviewing & editing.

Shinichi Nakagawa: Conceptualization; Funding acquisition; Investigation; Methodology; Resources; Software; Supervision; Writing – reviewing & editing.

Data availability statement

The statistical models described in this review are demonstrated in a supplementary worked example, which can be reproduced using the data, code, and model objects contained in this dedicated repository: http://doi.org/10.17605/OSF.IO/V3QAX

Keywords

coefficient of variation; DHGLM; Double Hierarchical; location-scale regression; multivariate; repeatability; rstan

Abstract

1. Organisms use labile traits to respond to different conditions over short timescales. When a population experiences the same conditions, we might expect all individuals to adjust their trait expression to the same, optimal, value, thereby minimising phenotypic variation. Instead, variation abounds. Individuals substantially differ not only from each other, but also from their former selves, with the expression of labile traits varying both predictably and unpredictably over time.

2. A powerful tool for studying the evolution of phenotypic variation in labile traits is the multilevel model. Here, we review how multilevel models are used to quantify individual differences in both means and variability, and their between-individual correlations. Individuals can differ in their average phenotypic tendencies (e.g. behavioural personalities), their intrinsic variability across time (known as 'predictability' or intra-individual variability), and their plastic response to different contexts.

3. To capture multiple facets of individuality, we provide detailed descriptions and resources for simultaneously modelling individual differences in averages, plasticity, and individual predictability. Empiricists can use these methods to quantify how traits covary across individuals and test theoretical ideas about phenotypic integration. These methods can be extended to incorporate plastic changes in predictability (termed 'stochastic malleability').

4. Overall, we showcase the unfulfilled potential of existing statistical tools to test more holistic and nuanced questions about the evolution, function, and maintenance of phenotypic variation, for any trait that is repeatedly expressed.

1 | INTRODUCTION

Life is full of variation. Phenotypic variation among taxa and species has been chronicled for centuries, but studying variation within populations, and even within individuals, is a newer venture for biologists (Westneat et al., 2015). Molecular biology has made it relatively straightforward to measure genetic differences between individuals, but we cannot simply extrapolate from genetic variation to its phenotypic consequences (Frazer et al., 2009). Much phenotypic variation is rooted in environmental variation (Stamps, 2015), either as an adaptive response to predictable environmental change, or a maladaptive consequence of environmental stress (Snell-Rood, 2013), and individuals can differ in their plastic response to environmental change (Dingemanse & Dochtermann, 2013). Still, even under controlled conditions, phenotypes vary unpredictably (Hansen et al., 2006). For labile traits that are repeatedly expressed — and can therefore be measured at multiple instances for the same individual — understanding what causes and maintains phenotypic variation both between and within individuals is a growing field.

Behavioural ecologists commonly use multilevel models to measure how behaviours vary across environments, and between individuals within populations (Allegue et al., 2017). For non-human animals, behavioural traits that consistently vary between individuals have been deemed 'personality' traits, and sometimes these individual differences are correlated in 'behavioural syndromes' (e.g. some individuals are consistently more risk-averse) (Bell, 2007; Dingemanse et al., 2010a; Dochtermann, 2010; Sih et al., 2004). While the 'consistency' of individual differences in behaviour has received much attention, in general most behavioural variation is driven not by differences between individuals, but instead by residual variation (meta-analysis of repeatability: Bell et al., 2009).

Residual variation is a composite of both biologically meaningful variation (e.g. within-individual variation), and variation that is inherent to the method of measurement. To estimate intra-individual variability from residual variance, we could quantify the precision of our measurements, and manually subtract the portion of residual variance that is attributable to measurement error. More simply, by using a measurement method that is as precise as possible, we explicitly assume that residual variance approximately represents within-individual variance. Regardless of the proportion of residual variance that we take to represent intra-individual variability, we hereafter refer to the magnitude of within-individual variance as

'predictability' (Cleasby et al., 2015).

Standard multilevel models assume that each individual has the same level of predictability (i.e. assume homogeneity of residual variances). The homogeneity assumption is violated when some individuals are more variable than others across time (Ramakers et al., 2020), and this 'heteroscedasticity' could represent interesting non-adaptive deviations from an optimal phenotype (e.g. maladaptive imprecision; Hansen et al., 2006), or adaptive variation between individuals in their level of predictability (e.g. alternative strategies; Wolf et al., 2007). The underlying mechanisms that dictate predictability are likely to be shared across different phenotypic traits. Such 'phenotypic integration' could lead to a trade-off that constrains or maintains phenotypic variance for a given trait, where individuals are more predictable than optimal for some traits, and less predictable than optimal for others (Pigliucci, 2003; Viney & Reece, 2013; Willmore et al., 2007).

Statistical methods for studying individual differences in labile (i.e. repeatedly expressed) traits will be most powerful when individual differences in averages (i.e. tendencies or personalities), plasticity, and variability (i.e. predictability) are considered together (Fig. 1). Here, we provide a guide for empiricists on methods that can be used to study factors contributing to the evolution of phenotypic variation in labile traits, while lowering the barrier to entry with a reproducible worked example. Throughout this review we describe models of behavioural traits (and therefore use terminology common in behavioural ecology), but the methods can be applied more broadly to different types of phenotypic traits, and different types of data clusters. For example, the clustering variable could be family or population origin rather than individual identity.

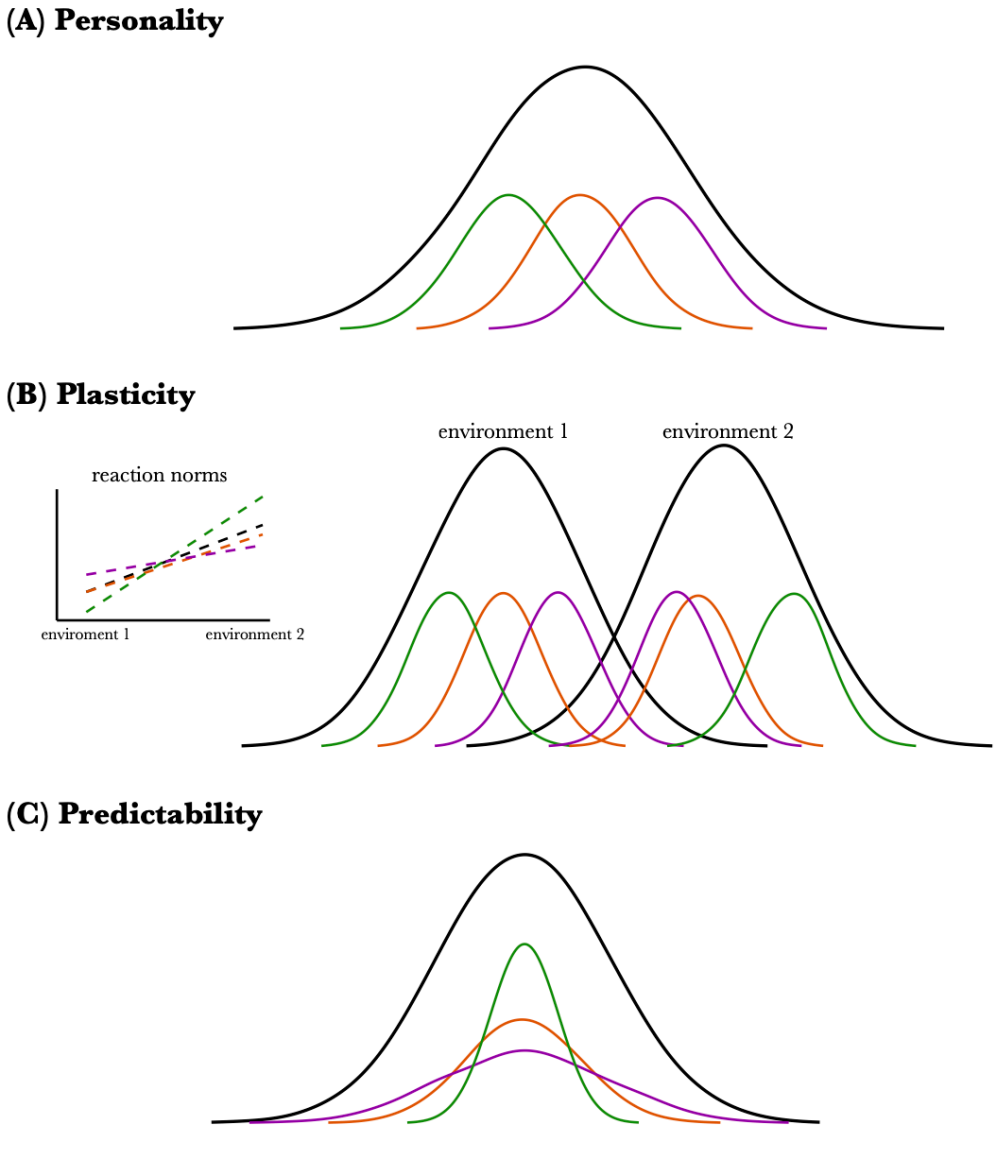


FIGURE 1

Conceptual illustration of three types of individual differences for a labile trait (in this case, behaviour). In each panel, black curves represent the normal distribution of a phenotypic trait in a population. Smaller, coloured curves represent the distribution of phenotypes expressed by an individual within that population. **(A)** 'Personality': individual differences in average behaviour. The term 'tendency' could be used for non-behavioural traits. **(B)** 'Plasticity' due to a change in the environment. In environment 2, compared with environment 1, the average phenotype of the population increases, as shown by the black distribution shifting to the right. Individuals' averages (distributions) shift to varying extents (i.e. variation in reaction norm slopes). **(C)** 'Predictability': individuals' level of variability (the breadth of individual distributions). Even individuals with the same 'personality' can show differences in predictability.

TABLE 1

Mathematical notation describing statistical models. Throughout this paper we assume that we are modelling behavioural traits in a multilevel model framework, and we are interested in the biological variables of sex, age, and individual identity. Note that when presenting square matrices, the bottom triangle elements are omitted for simplicity (as they are identical to the upper triangle).

Notation	Definition
y_{ij}	Response variable (i.e., a behavioural trait): the measured phenotypic value of trait *y* for the *j*th individual at instance *i*.
^{t1}	Superscript is used for bivariate models, to indicate model parameters for trait 1 (^{t1}) and trait 2 (^{t2}).
e_{ij}	Residual error: difference between the predicted and fitted value for the *j*th individual at instance *i*.
σ_e^2	Residual variance for single hierarchical models ('mean' model only).
$\sigma_{e_{ij}}^2$	Residual variance for double hierarchical models ('mean' and 'dispersion' models): unique value for each individual and instance.
x_{1j}	Categorical input variable for the 'sex' of individual *j* ($x_{1j} = 0$ for female, and 1 for male).
x_{2ij}	Continuous input variable for the *z*-transformed 'age' of individual *j* at instance *i* ($x_{2ij} = 0$ is the average age of the population).
β_{m0}	Population intercept for the mean model. Average value of *y* when all other input variables are set to zero (females of average age).
$\beta_{v0,exp}$	Population intercept for the dispersion (variance) model. Average value of $\ln(\sigma_{e_{ij}}^2)$when all other input variables are set to zero (females of average age). Estimated on the natural logarithm (ln) scale.
β_{m1}	Population slope for the female-male contrast for the mean model.
$\beta_{v1,exp}$	Population slope for the female-male contrast for the dispersion model). Estimated on the ln scale.
β_{m2}	Population slope. Average value of phenotypic plasticity (reaction norm) for x_{2ij} = *z*-scaled age, for the mean model.
$\beta_{v2,exp}$	Population slope. Average value of phenotypic plasticity (reaction norm) for x_{2ij} = *z*-scaled age, for the dispersion model). Estimated on the ln scale.

Table continued on next page

Table 1 — *continued*

Notation	Definition
$\mathrm{ID}_{\mathrm{m0}j}$	Difference between the population intercept β_{m0} and the random intercept for individual j for the mean model.
$\mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}$	Difference between the population intercept β_{v0} and the random intercept for individual j for the dispersion model. Estimated on the ln scale.
$\mathrm{ID}_{\mathrm{m2}j}$	Difference between the population slope β_{m2} and the random slope for individual j for the mean model.
$\lvert\beta_{\mathrm{m2}} + \mathrm{ID}_{\mathrm{m2}j}\rvert$	Absolute value of the (age) slope for individual j for the mean model. Describes the magnitude of individuals' average plasticity.
$\sigma^2_{\mathrm{ID_{m0}}}$	Between-individual variance for the individual intercepts for the mean model.
$\sigma^2_{\mathrm{ID_{m2}}}$	Between-individual variance for the individual slopes for the mean model.
$\sigma^2_{\mathrm{ID_{v0,exp}}}$	Between-individual variance for the individual intercepts for the dispersion model, on the ln scale.
$\sigma^2_{\mathrm{fixed_m}}$	Variance due to fixed effects for the mean model.
$\sigma^2_{\mathrm{fixed_{v,exp}}}$	Variance due to fixed effects for the dispersion model. Estimated on the ln scale.
$\mathrm{var(a + b)}$	Variance of the sum of random variables (vectors) a and b.
$\rho(\mathrm{a,b})$	Correlation between two random variables a and b.
σ_{ab}	Covariance between two random variables a and b.

2 | INDIVIDUAL DIFFERENCES IN INTERCEPTS AND SLOPES

Individual differences in average phenotypes can be quantified with a random intercept for each individual, using a multilevel model. Other sources of variation can be modelled as fixed effects (and, if necessary, additional random effects). Throughout this paper, we will present Gaussian multilevel models containing two fixed effects: the first for sex (i.e. a fixed effect with two categories, female and male), and a second for age (i.e. a continuous fixed effect). Age is mean-centred, so that the overall intercept of the model represents the average phenotype of females at the average age of the population. Notation for all equations are explained in Table 1 (note that the same principles can be applied to non-Gaussian data too; Nakagawa &

Schielzeth, 2010).

Non-human animal behaviours are commonly deemed 'personality traits' when, after measuring the same behaviour two or more times for multiple individuals, the differences among individuals are consistent across time and contexts (i.e. non-zero between-individual variance) (Bell, 2007; Sih et al., 2004). To measure differences in personalities, our basic model can be written as:

$$y_{ij} = (\beta_{\mathrm{m0}} + \mathrm{ID}_{\mathrm{m0}j}) + \beta_{\mathrm{m1}} x_{1j} + \beta_{\mathrm{m2}} x_{2ij} + e_{ij}, \quad \text{eqn 1}$$

$$e_{ij} \sim (0, \sigma_e^2), \quad \text{eqn 2}$$

$$\mathrm{ID}_{\mathrm{m0}j} \sim (0, \sigma^2_{\mathrm{ID}_{\mathrm{m0}}}), \quad \text{eqn 3}$$

$$\sigma^2_{\mathrm{fixed_m}} = \mathrm{var}(\beta_{\mathrm{m1}} x_{1j} + \beta_{\mathrm{m2}} x_{2ij}). \quad \text{eqn 4}$$

The model described by equations 1-3 assume homoscedasticity, meaning we model differences in individuals' average behaviour, but not variability in behaviour (Fig. 1A). The spread of individual averages allows us to estimate the between-individual variance in behaviour, which is used to quantify the strength of personality traits (equations for calculating repeatability and the coefficient of individual variation are provided in Section 4, below). When fixed effects represent biological variation (rather than experimental artefacts), it is recommended to add the fixed effect variance (calculated as in equation 4) back into the total variance (de Villemereuil et al., 2018) before calculating repeatability.

When we expect a predictable relationship between a phenotypic trait and an environmental or biological context (e.g. environmental temperature, or biological age), we can model this relationship with a function called a 'reaction norm' (Gavrilets & Scheiner, 1993; Gomulkiewicz & Kirkpatrick, 1992; Stearns & Koella, 1986). In the simplest case of a linear relationship (specified by an intercept and slope), the slope (β_{m2}) describes the magnitude and direction of the population's average phenotypic plasticity. If the same individuals were measured multiple times across different contexts, we can use 'random regression' to estimate random slopes for each individual ($\beta_{\mathrm{m2}} + \mathrm{ID}_{\mathrm{m2}j}$). Individuals can vary in both intercepts (personality) and slopes (plasticity) (Fig. 1B). As a consequence, the magnitude of differences in average individual behaviour ($\sigma_{\mathrm{ID}_{\mathrm{m0}}}$) could depend upon the context at which the intercept is estimated (in this case, the value of $x_2 = 0$, which is usually set to be the 'average' environment). Whereas the model in equation 1 assumes that personality is constant, no matter the context individuals are measured in, this 'random slope' model allows for individuals to converge upon,

or diverge from, the population mean in different environments:

$$y_{ij} = (\beta_{m0} + \mathrm{ID}_{m0j}) + \beta_{m1}x_{1j} + (\beta_{m2} + \mathrm{ID}_{m2j})x_{2ij} + e_{ij}, \quad \text{eqn 5}$$

$$e_{ij} \sim (0, \sigma_e^2), \quad \text{eqn 6}$$

$$\begin{bmatrix} \mathrm{ID}_{m0j} \\ \mathrm{ID}_{m2j} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\mathrm{ID}_{m0}} & \rho(\mathrm{ID}_{m0j}, \mathrm{ID}_{m2j})\sigma_{\mathrm{ID}_{m0}}\sigma_{\mathrm{ID}_{m2}} \\ \ldots & \sigma^2_{\mathrm{ID}_{m2}} \end{bmatrix}\right). \quad \text{eqn 7}$$

Multiple individual differences are modelled together using the multivariate normal distribution (MVN), which estimates the covariance between the random intercepts and slopes across individuals. This covariance is written (in the upper triangle of equation 7) as the product of the correlation between the intercepts and slopes [$\rho(\mathrm{ID}_{m0j}, \mathrm{ID}_{m2j})$], the standard deviation for the intercepts ($\sigma_{\mathrm{ID}_{m0}}$), and standard deviation for the slopes ($\sigma_{\mathrm{ID}_{m2}}$).

2.1 | INTERCEPT-SLOPE ASSOCIATIONS

Individuals with different personalities might differ in their ability to accurately assess their environment or change their phenotype, and there are empirical observations of such 'personality-plasticity' associations (Bogacz et al., 2010; Sih & Del Giudice, 2012). For example: in a marine gastropod, boldness was negatively correlated with plasticity in response to tidal and temperature changes (Cornwell et al., 2019); in sticklebacks, exploration was positively correlated with acclimation to a novel environment (Dingemanse et al., 2012); and in house sparrows, the level of parental care was shown to be correlated with plasticity in response to brood size, nestling age, precipitation, and the provisioning effort of the breeding partner (Westneat et al., 2011). Theoretically, Dubois (2019) predicted a negative correlation between proactive personalities and adaptive plasticity, based on the assumption that proactive individuals are less capable of accurately assessing their environment, due to the higher cognitive demands of proactivity. A positive correlation, meanwhile, could represent a "rich get richer" scenario, whereby more well-resourced individuals are more proactive *and* better able to bear the costs associated with plasticity (DeWitt et al., 1998; Reznick et al., 2000). Alternatively, phenotypic plasticity can represent a maladaptive change in the phenotype (e.g. due to environmental stress), and therefore personality types that show reduced plasticity might be more resilient to environmental change (Ghalambor et al., 2007).

There are two possible types of personality-plasticity associations, the results of which are contrasted in Fig. 2. First, from the multivariate normal distribution in equation 7, we can ask

whether individual differences in intercepts are correlated with individual differences in slopes. The correlation provided by the model is the rank correlation between individual differences (i.e. the best linear unbiased predictions: BLUPs) from the average population intercept (β_{m0}) and the average population slope (β_{m2}). This correlation represents the covariance between the random intercepts and slopes ($\sigma_{\mathrm{ID_{m0}ID_{m2}}}$), divided by the product of their standard deviations:

$$\rho\left(\mathrm{ID}_{\mathrm{m0}j}, \mathrm{ID}_{\mathrm{m2}j}\right) = \frac{\sigma_{\mathrm{ID_{m0}ID_{m2}}}}{\sigma_{\mathrm{ID_{m0}}}\sigma_{\mathrm{ID_{m2}}}}. \quad \text{eqn 8}$$

Alternatively, our question might be about the magnitude of plasticity irrespective of the direction of phenotypic change. In these cases, we can measure the correlation between the magnitude of each individual's reaction norm and the difference in their intercept from the population average. Here, we take the absolute values of the sum of the population slope and the individual slope differences: $|\beta_{\mathrm{m2}} + \mathrm{ID}_{\mathrm{m2}j}|$. When fitting Bayesian multilevel models, this personality-plasticity association,

$$\rho\left(\mathrm{ID}_{\mathrm{m0}j}, \left|\beta_{\mathrm{m2}} + \mathrm{ID}_{\mathrm{m2}j}\right|\right) = \frac{\sigma_{\mathrm{ID_{m0}}|\beta_{\mathrm{m2}}+\mathrm{ID_{m2}}|}}{\sigma_{\mathrm{ID_{m0}}}\sigma_{|\beta_{\mathrm{m2}}+\mathrm{ID_{m2}}|}}, \quad \text{eqn 9}$$

can be calculated from the posterior distributions of individual differences, and the population slope. As for all calculations involving BLUPs, posterior distributions should be used when estimating equation 9 to retain uncertainty and estimate credible intervals (Hadfield et al., 2010; Postma, 2006). While bootstrapping methods could be used to estimate uncertainty from frequentist (likelihood-based) models (cf. Stoffel et al., 2017) these methods would become very difficult when predictability is incorporated into the model structure.

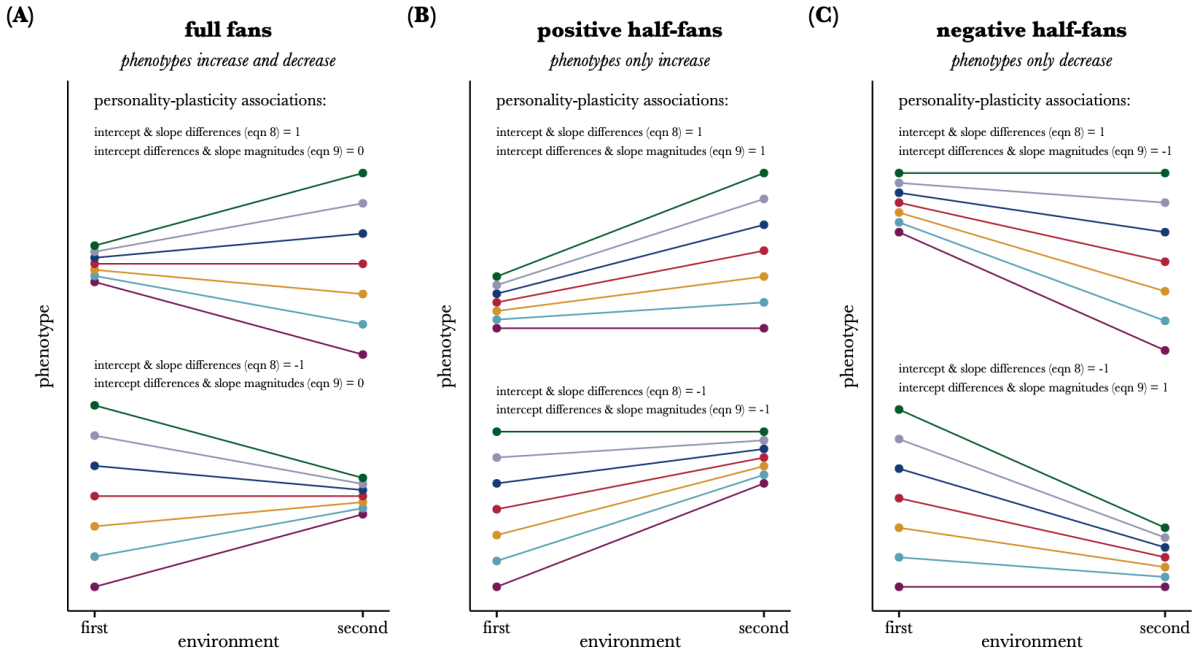


FIGURE 2

Personality-plasticity associations calculated with either slope differences, $\rho(\mathrm{ID}_{\mathrm{m0}j}, \mathrm{ID}_{\mathrm{m2}j})$, or slope magnitudes, $\rho(\mathrm{ID}_{\mathrm{m0}j}, |\beta_{\mathrm{m2}} + \mathrm{ID}_{\mathrm{m2}j}|)$, for three simplified shapes of phenotypic plasticity. Associations are shown for a population of seven individuals, where the rank order of intercepts is maintained across two environments, and phenotypes either 'fan out' (i.e. variance increases) or 'fan in' (i.e. variance decreases). Points represent each individual's phenotype in two environments. Lines represent the reaction norm depicting the direction and magnitude of phenotypic plasticity. **(A)** Full fan: individuals vary in both the magnitude and direction of their slopes, meaning that some phenotypes increase in the second environment while others decrease. The personality-plasticity association is zero for slope magnitudes, positive for slope differences that fan out, and negative for slope differences that fan in. **(B)** Positive fan: phenotypes always increase or stay the same in the second environment (i.e. individual slopes have a lower-bound at zero). Personality-plasticity associations are identical for slope differences and magnitudes, with opposite signs for rection norms that fan out or in (positive or negative correlations, respectively). **(C)** Negative fan: phenotypes always decrease or stay the same in the second environment (i.e. individual slopes have an upper-bound at zero). Personality-plasticity associations are either positive or negative, depending both on whether slope differences or magnitudes are used, and whether the reaction norms fan in or out.

Interpreting personality-plasticity associations at a given position of the intercept requires careful consideration, because multiple patterns of reaction norm slopes can produce the same correlations (as shown in Fig. 2, and noted by Stamps & Biro, 2016). When individual reaction norms 'fan out' (or in), personalities that are above (or below) the population average have a more positive slope, and individuals that are below (or above) the population average have a more negative slope. A conceptual model of 'fanning' is described by Sih et al. (2015) as resulting from within-individual feedback loops. Fanning can also occur when adaptive plasticity is condition-dependent, and only high-quality individuals can express adaptive plasticity. Individuals in poor condition (e.g. ill or injured) might express maladaptive plasticity in the opposite direction to the adaptive response. Regardless of the cause of these patterns, in a full fan scenario, the ranking of individual intercepts does not correlate with their magnitude of phenotypic plasticity (i.e. does not correlate with the absolute value of their slope). Contrasting with a full fan pattern, often we might expect all individuals in a population to respond to an environmental change with a plastic response in the *same* direction. In Fig. 2, we

call these scenarios 'positive fans' (when all phenotypes increase or stay the same) and 'negative fans' (when all phenotypes decrease or stay the same). For example, ectotherms exposed to a warmer environment will often show a plastic response in the same direction (e.g., increased activity levels).

2.2 | BIVARIATE MODEL

When two different traits are measured repeatedly for the same individuals, we can use a bivariate model to estimate the covariances (and therefore correlations) between individual differences in intercepts and slopes for these two traits (shown in equation 13, below). Between-individual correlations that span across distinct traits might reflect biologically interesting dependence, such as genetic correlations (e.g. due to linkage disequilibrium) or developmental constraints (Sih et al., 2012). Such phenotypic integration can prevent phenotypic traits from evolving independently (Fawcett et al., 2012; Pigliucci, 2003). Trait correlations could also reflect correlated selective pressures, where a change in one trait encourages an adaptive change in the other. In theory, multivariate models can estimate the dependence between many traits at once, but here, for ease of presentation, we focus on the simplest scenario of two traits ('t1' and 't2'). The bivariate model can be written as:

$$y_{ij}^{\mathrm{t1}} = \left(\beta_{\mathrm{m0}}^{\mathrm{t1}} + \mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t1}}\right) + \beta_{\mathrm{m1}}^{\mathrm{t1}} x_{1j}^{\mathrm{t1}} + \left(\beta_{\mathrm{m2}}^{\mathrm{t1}} + \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}}\right) x_{2ij}^{\mathrm{t1}} + e_{ij}^{\mathrm{t1}}, \qquad \text{eqn 10}$$

$$y_{ij}^{\mathrm{t2}} = \left(\beta_{\mathrm{m0}}^{\mathrm{t2}} + \mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t2}}\right) + \beta_{\mathrm{m1}}^{\mathrm{t2}} x_{1j}^{\mathrm{t2}} + \left(\beta_{\mathrm{m2}}^{\mathrm{t2}} + \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}}\right) x_{2ij}^{\mathrm{t2}} + e_{ij}^{\mathrm{t2}}, \qquad \text{eqn 11}$$

$$\begin{bmatrix} e^{\mathrm{t1}} \\ e^{\mathrm{t2}} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{e^{\mathrm{t1}}}^2 & \rho(e^{\mathrm{t1}}, e^{\mathrm{t2}})\sigma_{e^{\mathrm{t1}}}\sigma_{e^{\mathrm{t2}}} \\ \ldots & \sigma_{e^{\mathrm{t2}}}^2 \end{bmatrix}\right), \qquad \text{eqn 12}$$

$$\begin{bmatrix} \mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t1}} \\ \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}} \\ \mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t2}} \\ \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t1}}}^2 & \rho(\mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}})\sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t1}}}\sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t1}}} & \rho(\mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t2}})\sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t1}}}\sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t2}}} & \rho(\mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}})\sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t1}}}\sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t2}}} \\ \ldots & \sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t1}}}^2 & \rho(\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t2}})\sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t1}}}\sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t2}}} & \rho(\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}})\sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t1}}}\sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t2}}} \\ \ldots & \ldots & \sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t2}}}^2 & \rho(\mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t2}}, \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}})\sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t2}}}\sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t2}}} \\ \ldots & \ldots & \ldots & \sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t2}}}^2 \end{bmatrix}\right). \qquad \text{eqn 13}$$

Dependence between residual errors for different traits is modelled using the multivariate normal distribution (MVN) in equation 12. Similarly, in equation 13, the covariance matrix describing the relationship between individual-level differences has been expanded to include correlations both within and between traits, for both intercepts and slopes.

2.3 | SYNDROMES ACROSS TRAIT INTERCEPTS

Bivariate models can quantify the relationship between different types of personality traits (equations 10-11). When individual intercepts in behavioural traits are correlated (i.e. between-individual correlations between personality traits), those traits are said to exhibit a 'behavioural

syndrome' (Dingemanse et al., 2010a), which we can estimate as:

$$\rho\left(\mathrm{ID}^{t1}_{m0j}, \mathrm{ID}^{t2}_{m0j}\right) = \frac{\sigma_{\mathrm{ID}^{t1}_{m0}\mathrm{ID}^{t2}_{m0}}}{\sigma_{\mathrm{ID}^{t1}_{m0}}\sigma_{\mathrm{ID}^{t2}_{m0}}}. \quad \text{eqn 14}$$

While many empirical papers purport to have found these syndromes, far fewer have done so following the recommended method of decomposing total phenotypic variance into its among- and within- individual components (Dingemanse & Dochtermann, 2013; Moirón et al., 2020; Niemelä & Dingemanse, 2018). Combining both levels of the phenotypic correlation can be misleading, as selection should occur at the between-individual level, and the strength and direction of this correlation can be different from the within-individual level (i.e. violating the 'individual gambit'; Brommer, 2013).

2.4 | SYNDROMES ACROSS TRAIT SLOPES

Correlations between the reaction norm slopes for the same individuals can be measured for multiple traits, or multiple environmental manipulations. Positive correlations might be common, due to shared mechanisms in the maintenance of plasticity. Plasticity that shows phenotypic integration, or modularity, has been of enduring interest to plant scientists (Gianoli & Palacio-Lopez, 2009; Mallitt et al., 2010; Pigliucci, 2002; Schlichting, 1989). Alternatively, a negative correlation in the magnitude of plasticity could reflect trade-offs due to associated costs (DeWitt et al., 1998), while the absence of a correlation suggests the traits have been selected to be decoupled, or face independent selective pressures.

'Plasticity syndromes' are more challenging to interpret than behavioural syndromes, due to the rank order of individual differences in slopes not necessarily being correlated with the magnitude of individuals' plasticity. As with personality-plasticity associations, plasticity syndromes can be estimated in two different ways (which are compared in Fig. 3, below). Taken directly from the model, the correlation between individual slope differences,

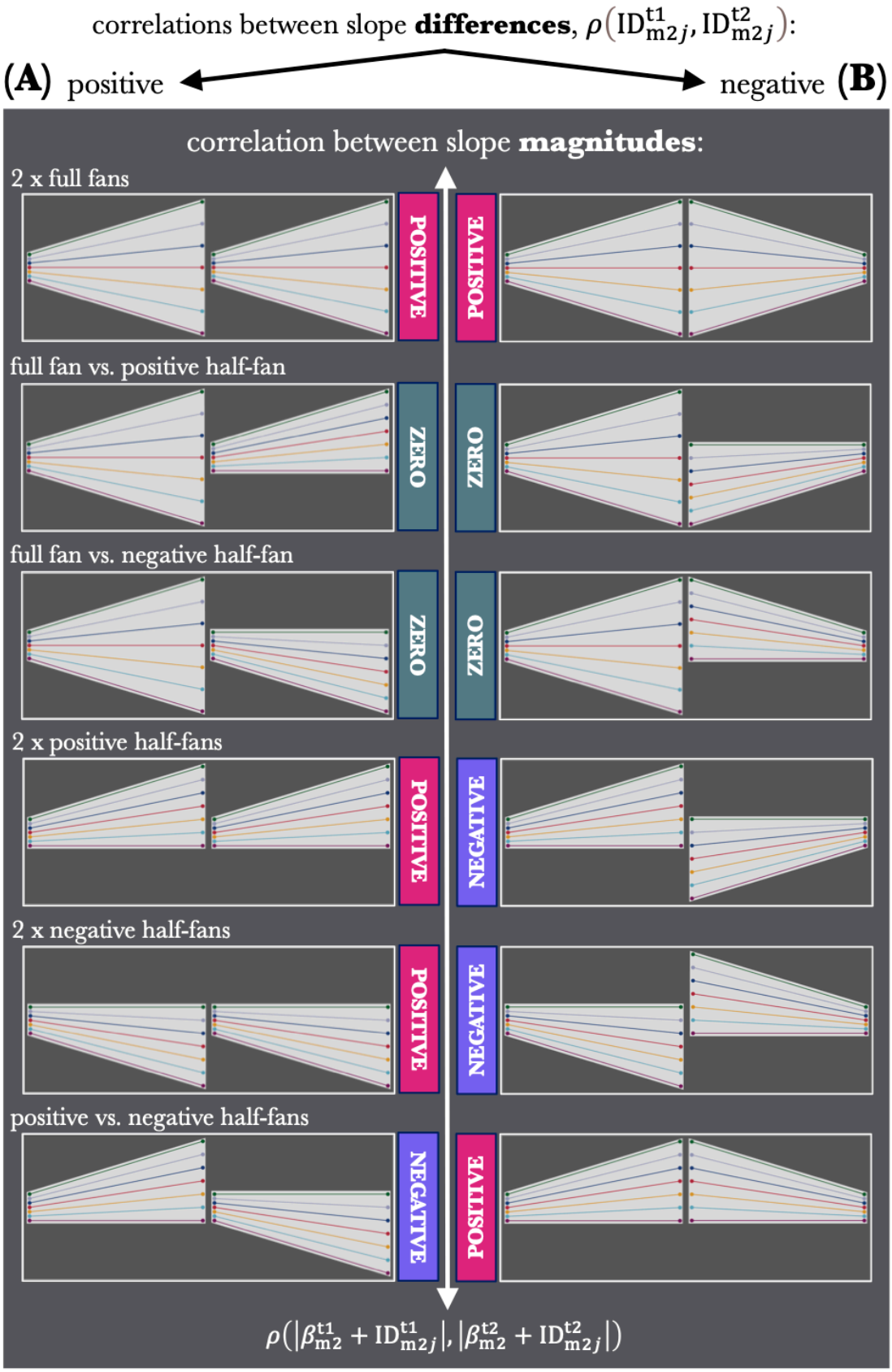
$$\rho\left(\mathrm{ID}^{t1}_{m2j}, \mathrm{ID}^{t2}_{m2j}\right) = \frac{\sigma_{\mathrm{ID}^{t1}_{m2}\mathrm{ID}^{t2}_{m2}}}{\sigma_{\mathrm{ID}^{t1}_{m2}}\sigma_{\mathrm{ID}^{t2}_{m2}}}, \quad \text{eqn 15}$$

describes whether the 'rank order' of slopes is maintained between the two traits. That is, a rank correlation from equation 15 indicates that individuals whose slopes are more positive than average in trait 1 tend to *also* be more positive than average in trait 2. This rank correlation could be useful for certain patterns of plasticity. However, when we care about slope steepness (rather than the difference from the average), we should consider the slope's magnitude. In this case, a 'plasticity syndrome' (equation 16) is calculated as the correlation between the absolute magnitude of individuals' reaction norms, such that:

$$\rho\left(\left|\beta_{\mathrm{m2}}^{\mathrm{t1}}+\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}}\right|,\left|\beta_{\mathrm{m2}}^{\mathrm{t2}}+\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}}\right|\right)=\frac{\sigma_{\left|\beta_{\mathrm{m2}}^{\mathrm{t1}}+\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}}\right|\left|\beta_{\mathrm{m2}}^{\mathrm{t2}}+\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}}\right|}}{\sigma_{\left|\beta_{\mathrm{m2}}^{\mathrm{t1}}+\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}}\right|}\sigma_{\left|\beta_{\mathrm{m2}}^{\mathrm{t2}}+\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}}\right|}}. \quad \text{eqn 16}$$

As with equation 9, correlations involving absolute values of slopes can be calculated from the posterior distributions of model estimates.

Variation in plasticity syndromes, for pairs of traits, when the rank-order of intercepts is maintained



328

FIGURE 3

Plasticity syndromes are influenced by simplified shapes of phenotypic plasticity (full fans, negative half-fans, and positive half-fans, which either 'fan in' or 'fan out'; Fig. 2). Boxes outlined in white depict reaction norms for two types of traits, where the order of individual averages (i.e. intercepts) is maintained across environments and across traits. Rows are arranged according to which fan shapes are paired together. In the centre of the figure, rotated text inside coloured boxes show variation in plasticity syndromes when measured with slope magnitudes (equation 16). Plasticity syndromes measured with slope magnitudes are always positive for a pair of full fans, always zero when a full fan is paired with a half-fan, and either positive or negative when half-fans are paired (depending on the direction of their slopes and whether reaction norms fan 'in' or 'out'). **(A)** Reaction norms in the left column all 'fan out'. When the effect of the environmental change on between-individual variance is identical for both traits (i.e. both sets of reaction norms 'fan out', or both sets of reaction norms 'fan in'), then plasticity syndromes measured with individual differences (equation 15) are always positive. **(B)** Reaction norms in the right column fan out for one trait, and fan in for the second trait. In this case, plasticity syndromes measured with individual differences (equation 15) are always negative.

2.5 | SUMMARY OF INDIVIDUAL DIFFERENCES IN INTERCEPTS AND SLOPES

Individuals can have different personalities, as measured by differences in average behaviour (intercepts), and differences in plasticity, as measured by differences in their average change in behaviour across a covariate (slopes). The covariate can be an external variable such as temperature, or an internal variable such as developmental age or circulating hormone concentrations. Between-individual variation in plasticity has implications for estimates of personality (and related correlations), because the magnitude of between-individual variation can depend upon the point at which the 'intercept' in behaviour is measured. From individual differences in two BLUPs (personality and plasticity) we can measure three types of biologically relevant correlations: first, personality-plasticity associations are a correlation between reaction norm intercepts and slopes (either individual differences or slope magnitudes); second, behavioural syndromes are a correlation between individual intercepts for more than one trait; third, plasticity syndromes are a correlation between the differences or magnitudes of individual slopes for more than one trait, or the same trait measured across more than one covariate.

When interpreting rank correlations involving slopes, which have both a direction and magnitude, researchers should plot each individual's reaction norm to consider the 'shape' of phenotypic plasticity. For some research questions, the magnitude of plasticity could be more relevant than the direction of change away from the population average; for example, under thermal stress, are some individuals consistently better at maintaining homeostasis in physiological traits? In these circumstances, researchers can perform additional calculations to capture the absolute value of individual slopes, rather than individual differences from the average slope. Performing vector calculations on posterior distributions (from a Bayesian model) ensures that uncertainty in model estimates is carried forward.

3 | INDIVIDUAL DIFFERENCES IN PREDICTABILITY

The effect animals have on their surroundings depends not only on their average behaviour, but also on how their behaviour fluctuates through time. Individual differences can be consistent yet small, and these might not have a material impact on fitness (and therefore might not respond to selection). Despite the variability of individuals' behaviour being biologically important, it is currently rare for behavioural studies to distinguish between individuals who are very consistent through time, and those whose behaviour fluctuates enormously (an early example is seen in Westneat et al., 2013). Individual differences in within-individual variance (i.e. heteroscedasticity) can be modelled with a Double Hierarchical Generalized Linear Model (DHGLM; Cleasby et al., 2015). The 'double' in DHGLM refers to a random effect being included in both the mean model, and the residual variance model. In the social and medical sciences, DHGLMs are also known as location-scale regression models (with 'location' indicating the mean, and 'scale' indicating the variance; e.g. Lin et al., 2018; Rast et al., 2012). The residual variance model, hereafter referred to as the 'dispersion' model, is usually estimated on the natural logarithm scale. Fitting a random intercept for individual identity at both levels of the model allows individuals to vary in both their average behaviour and residual variance (equations 17-18, below). We can therefore consider a third individual difference alongside personality and plasticity: predictability.

3.1 | MODELLING INDIVIDUAL DISTRIBUTIONS

For labile traits, by allowing individuals to vary in both personality (Fig. 1A) and predictability (Fig. 1C), we effectively estimate a different distribution for each individual. Extending the univariate model shown in equations 5-8, we can write the double hierarchical model as:

$$y_{ij} = (\beta_{\mathrm{m0}} + \mathrm{ID}_{\mathrm{m0}j}) + \beta_{\mathrm{m1}}x_{1j} + (\beta_{\mathrm{m2}} + \mathrm{ID}_{\mathrm{m2}j})x_{2ij} + e_{ij}, \quad \text{eqn 17}$$

$$\ln\left(\sigma^2_{e_{ij}}\right) = (\beta_{\mathrm{v0,exp}} + \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}) + \beta_{\mathrm{v1,exp}}x_{1j} + \beta_{\mathrm{v2,exp}}x_{2ij}, \quad \text{eqn 18}$$

$$e_{ij} \sim N(0, \sigma^2_{e_{ij}}) \quad \text{eqn 19}$$

$$\begin{bmatrix} \mathrm{ID}_{\mathrm{m0}j,} \\ \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}} \\ \mathrm{ID}_{\mathrm{m2}j} \end{bmatrix} \sim MVN\left(\begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix} \sigma^2_{\mathrm{ID_{m0}}} & \rho(\mathrm{ID}_{\mathrm{m0}j}, \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}})\sigma_{\mathrm{ID_{m0}}}\sigma_{\mathrm{ID_{v0,exp}}} & \rho(\mathrm{ID}_{\mathrm{m0}j}, \mathrm{ID}_{\mathrm{m2}j})\sigma_{\mathrm{ID_{m0}}}\sigma_{\mathrm{ID_{m2}}} \\ \dots & \sigma^2_{\mathrm{ID_{v0,exp}}} & \rho(\mathrm{ID}_{\mathrm{v0}j}, \mathrm{ID}_{\mathrm{m2}j})\sigma_{\mathrm{ID_{v0}},\mathrm{exp}}\sigma_{\mathrm{ID_{m2}}} \\ \dots & \dots & \sigma^2_{\mathrm{ID_{m2}}} \end{bmatrix}\right). \quad \text{eqn 20}$$

Estimating individual variances requires many repeated measurements at the individual level, which is relatively uncommon in animal personality studies (sample size recommendations depend on the number of individuals and the magnitude of heteroscedasticity, which is explored in Cleasby et al., 2015). Note that equations 17-20 vary from equations 19-24 in Cleasby et al. (2015), as the dispersion model is based on residual variances, rather than residual standard deviations (which has some benefits for summarising the magnitude of individual differences; see Section 4.3, below).

3.2 | ASSOCIATIONS BETWEEN THREE INDIVIDUAL DIFFERENCES

From the correlation between individual intercepts in both the mean and dispersion models, we can estimate whether some personality types are more prone to being unpredictable than others. From the multivariate distribution in equation 20, we have:

$$\rho(\mathrm{ID}_{\mathrm{m0}j}, \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}) = \frac{\sigma_{\mathrm{ID_{m0}ID_{v0,exp}}}}{\sigma_{\mathrm{ID_{m0}}}\sigma_{\mathrm{ID_{v0,exp}}}}. \quad \text{eqn 21}$$

Interpreting equation 21 is somewhat unintuitive; remember that an individual having more residual variance is likely to be less predictable (i.e. because they have more within-individual variance). Therefore, a positive correlation between mean and dispersion intercepts represents a negative correlation between personality and predictability. When presenting results, we prefer to multiply correlations involving dispersion intercepts by minus 1, to make their interpretation intuitive (e.g. a positive correlation signifies a bolder individual is more predictable, with a smaller residual variance), such that:

$$\rho(\mathrm{ID}_{\mathrm{m0}j}, -1 \times \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}) = -\frac{\sigma_{\mathrm{ID_{m0}ID_{v0,exp}}}}{\sigma_{\mathrm{ID_{m0}}}\sigma_{\mathrm{ID_{v0,exp}}}}. \quad \text{eqn 22}$$

Our supplementary example presents this sign-reversed correlation for personality-predictability associations. Although little theory exits on the personality-predictability association, we might expect risker personality types to also be less predictable; being more variable can be a risky strategy. Alternatively, riskier individuals could be closer to a hypothetical 'ceiling', whereby a fluctuation beyond that point would be fatal to the individual. Riskier individuals might therefore show greater precision around their mean phenotype, to avoid crossing some point of no return (a similar idea around stability of more 'extreme' personalities is discussed in Stamps & Groothuis, 2010).

Broadly, plasticity is the expression of different phenotypes by the same genotype in a different environment (Stamps, 2015). The environment will always be slightly different each time an individual expresses a labile trait because of variation in endogenous variables (internal and developmental), and uncontrolled fluctuations in the external environment (Flatt, 2005; Hansen et al., 2006). Therefore, predictability is a special type of 'stochastic plasticity', because there are stochastic changes in internal and external environments that prevent us from knowing exactly which phenotype will be expressed at any point in time. From the slope in the mean model and the intercept in the dispersion model, we can estimate whether individual differences in predictable and stochastic plasticity are correlated. There is theoretical interest in whether different types of plasticity (or 'flexibility' or 'responsiveness') are related to each other (e.g. through shared mechanisms), but to date this type of question has received little empirical attention (Stamps & Biro, 2016). Less predictable individuals express a larger range of behaviours, which could imply a greater scope for showing an average plastic response to an environmental change. Predictability and plasticity could therefore be positively correlated. The rank correlation between individual differences from mean slopes, and dispersion intercepts,

$$\rho\left(\mathrm{ID}_{\mathrm{m2}j}, \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}\right) = \frac{\sigma_{\mathrm{ID}_{\mathrm{m2}}\mathrm{ID}_{\mathrm{v0,exp}}}}{\sigma_{\mathrm{ID}_{\mathrm{m2}}}\sigma_{\mathrm{ID}_{\mathrm{v0,exp}}}}, \qquad \text{eqn 23}$$

measures whether individuals that are further away from the average level of plasticity are more or less predictable than average. The correlation between the magnitudes of mean slopes and dispersion intercepts,

$$\rho\left(\left|\beta_{\mathrm{m2}} + \mathrm{ID}_{\mathrm{m2}j}\right|, -1 \times \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}\right) = -\frac{\sigma_{|\beta_{\mathrm{m2}}+\mathrm{ID}_{\mathrm{m2}}|\mathrm{ID}_{\mathrm{v0,exp}}}}{\sigma_{|\beta_{\mathrm{m2}}+\mathrm{ID}_{\mathrm{m2}}|}\sigma_{\mathrm{ID}_{\mathrm{v0,exp}}}}, \qquad \text{eqn 24}$$

estimates whether individuals who are more plastic (in either direction) are more or less predictable. Multiplying by minus 1 makes this correlation interpretable is a 'plasticity-

predictability association'.

3.3 | SYNDROMES ACROSS INDIVIDUAL DIFFERENCES IN PREDICTABILITY

Up to this point, we have discussed fives types of correlations between individual differences: behavioural syndromes (Fig. 4A); plasticity syndromes (Fig. 4B); personality-plasticity associations (Fig. 4D); personality-predictability associations (Fig. 4E); and plasticity-predictability associations (Fig. 4F). Given sufficient data, one bivariate with three individual differences can estimate all of these correlations, plus a sixth: predictability syndromes (Fig. 4C). The bivariate model can be written as:

$$y_{ij}^{t1} = \left(\beta_{m0}^{t1} + ID_{m0j}^{t1}\right) + \beta_{m1}^{t1} x_{1j}^{t1} + \left(\beta_{m2}^{t1} + ID_{m2j}^{t1}\right) x_{2ij}^{t1} + e_{ij}^{t1}, \quad \text{eqn 25}$$

$$y_{ij}^{t2} = \left(\beta_{\mathrm{m0}}^{t2} + \mathrm{ID}_{\mathrm{m0}j}^{t2}\right) + \beta_{\mathrm{m1}}^{t2} x_{1j}^{t2} + \left(\beta_{\mathrm{m2}}^{t2} + \mathrm{ID}_{\mathrm{m2}j}^{t2}\right) x_{2ij}^{t2} + e_{ij}^{t2}, \quad \text{eqn 26}$$

$$\ln\left(\sigma_{e_{ij}^{\mathrm{t1}}}^2\right) = \left(\beta_{\mathrm{v0}}^{t1} + \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}^{t1}\right) + \beta_{\mathrm{v1}}^{t1} x_{1j}^{t1} + \beta_{\mathrm{v2}}^{t1} x_{2ij}^{t1}, \quad \text{eqn 27}$$

$$\ln\left(\sigma_{e_{ij}^{\mathrm{t2}}}^2\right) = \left(\beta_{\mathrm{v0}}^{t2} + \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}^{t2}\right) + \beta_{\mathrm{v1}}^{t2} x_{1j}^{t2} + \beta_{\mathrm{v2}}^{t2} x_{2ij}^{t2}, \quad \text{eqn 28}$$

$$\begin{bmatrix} e_{ij}^{t1} \\ e_{ij}^{t2} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{e_{ij}^{\mathrm{t1}}}^2 & \rho\left(e_{ij}^{t1}, e_{ij}^{t2}\right)\sigma_{e_{ij}^{\mathrm{t1}}}\sigma_{e_{ij}^{\mathrm{t2}}} \\ \ldots & \sigma_{e_{ij}^{\mathrm{t2}}}^2 \end{bmatrix}\right), \quad \text{eqn 29}$$

$$\begin{bmatrix} \mathrm{ID}_{\mathrm{m0}j}^{t1} \\ \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}^{t1} \\ \mathrm{ID}_{\mathrm{m2}j}^{t1} \\ \mathrm{ID}_{\mathrm{m0}j}^{t2} \\ \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}^{t2} \\ \mathrm{ID}_{\mathrm{m2}j}^{t2} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t1}}}^2 & \rho\left(\mathbf{ID}_{\mathbf{m0}j}^{\mathbf{t1}}, \mathbf{ID}_{\mathbf{v0}j,\mathbf{exp}}^{\mathbf{t1}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m0}}}^{\mathbf{t1}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{v0,exp}}}^{\mathbf{t1}} & \boldsymbol{\rho}\left(\mathbf{ID}_{\mathbf{m0}j}^{\mathbf{t1}}, \mathbf{ID}_{\mathbf{m2}j}^{\mathbf{t1}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m0}}}^{\mathbf{t1}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m2}}}^{\mathbf{t1}} & \boldsymbol{\rho}\left(\mathbf{ID}_{\mathbf{m0}j}^{\mathbf{t1}}, \mathbf{ID}_{\mathbf{m0}j}^{\mathbf{t2}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m0}}}^{\mathbf{t1}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m0}}}^{\mathbf{2}} & \rho\left(\mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}^{\mathrm{t2}}\right)\sigma_{\mathrm{ID}_{\mathrm{m0}}}^{\mathrm{t1}}\sigma_{\mathrm{ID}_{\mathrm{v0,exp}}}^2 & \rho\left(\mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}}\right)\sigma_{\mathrm{ID}_{\mathrm{m0}}}^{\mathrm{t1}}\sigma_{\mathrm{ID}_{\mathrm{m2}}}^2 \\ \ldots & \sigma_{\mathrm{ID}_{\mathrm{v0,exp}}^{\mathrm{t1}}}^2 & \boldsymbol{\rho}\left(\mathbf{ID}_{\mathbf{v0}j,\mathbf{exp}}^{\mathbf{t1}}, \mathbf{ID}_{\mathbf{m2}j}^{\mathbf{t1}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{v0,exp}}}^{\mathbf{t1}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m2}}}^{\mathbf{t1}} & \rho\left(\mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t2}}\right)\sigma_{\mathrm{ID}_{\mathrm{v0,exp}}}^{\mathrm{t1}}\sigma_{\mathrm{ID}_{\mathrm{m0}}}^2 & \boldsymbol{\rho}\left(\mathbf{ID}_{\mathbf{v0}j,\mathbf{exp}}^{\mathbf{t1}}, \mathbf{ID}_{\mathbf{v0}j,\mathbf{exp}}^{\mathbf{t2}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{v0,exp}}}^{\mathbf{t1}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{v0,exp}}}^{\mathbf{2}} & \rho\left(\mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t2}}\right)\sigma_{\mathrm{ID}_{\mathrm{v0,exp}}}^{\mathrm{t1}}\sigma_{\mathrm{ID}_{\mathrm{m2}}}^2 \\ \ldots & \ldots & \sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t1}}}^2 & \rho\left(\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{m0}j}^{\mathrm{t2}}\right)\sigma_{\mathrm{ID}_{\mathrm{m2}}}^{\mathrm{t1}}\sigma_{\mathrm{ID}_{\mathrm{m0}}}^2 & \rho\left(\mathrm{ID}_{\mathrm{m2}j}^{\mathrm{t1}}, \mathrm{ID}_{\mathrm{v0}j,\mathrm{exp}}^{\mathrm{t2}}\right)\sigma_{\mathrm{ID}_{\mathrm{m2}}}^{\mathrm{t1}}\sigma_{\mathrm{ID}_{\mathrm{v0,exp}}}^2 & \boldsymbol{\rho}\left(\mathbf{ID}_{\mathbf{m2}j}^{\mathbf{t1}}, \mathbf{ID}_{\mathbf{m2}j}^{\mathbf{t2}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m2}}}^{\mathbf{t1}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m2}}}^{\mathbf{2}} \\ \ldots & \ldots & \ldots & \sigma_{\mathrm{ID}_{\mathrm{m0}}^{\mathrm{t2}}}^2 & \boldsymbol{\rho}\left(\mathbf{ID}_{\mathbf{m0}j}^{\mathbf{t2}}, \mathbf{ID}_{\mathbf{v0}j,\mathbf{exp}}^{\mathbf{t2}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m0}}}^{\mathbf{2}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{v0,exp}}}^{\mathbf{2}} & \boldsymbol{\rho}\left(\mathbf{ID}_{\mathbf{m2}j}^{\mathbf{t2}}, \mathbf{ID}_{\mathbf{v0}j,\mathbf{exp}}^{\mathbf{t2}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m2}}}^{\mathbf{2}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{v0,exp}}}^{\mathbf{2}} \\ \ldots & \ldots & \ldots & \ldots & \sigma_{\mathrm{ID}_{\mathrm{v0,exp}}^{\mathrm{t2}}}^2 & \boldsymbol{\rho}\left(\mathbf{ID}_{\mathbf{m2}j}^{\mathbf{t2}}, \mathbf{ID}_{\mathbf{v0}j,\mathbf{exp}}^{\mathbf{t2}}\right)\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{m2}}}^{\mathbf{2}}\boldsymbol{\sigma}_{\mathbf{ID}_{\mathbf{v0,exp}}}^{\mathbf{2}} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \sigma_{\mathrm{ID}_{\mathrm{m2}}^{\mathrm{t2}}}^2 \end{bmatrix}\right). \quad \text{eqn 30}$$

The variance-covariance matrix in equation 30 emphasises, in bold, the off-diagonal elements that comprise the six types of correlations we are interested in (shown in Fig. 4).

Predictability syndromes describe whether individuals who are less (or more) predictable in one trait are also less (or more) predictable in another trait (i.e., a positive, or negative, correlation between within-individual variances), such that:

$$\rho\left(\mathrm{ID}^{\mathrm{t1}}_{\mathrm{v0}j,\mathrm{exp}}, \mathrm{ID}^{\mathrm{t2}}_{\mathrm{v0}j,\mathrm{exp}}\right) = \frac{\sigma_{\mathrm{ID}^{\mathrm{t1}}_{\mathrm{v0,exp}}\mathrm{ID}^{\mathrm{t2}}_{\mathrm{v0,exp}}}}{\sigma_{\mathrm{ID}^{\mathrm{t1}}_{\mathrm{v0,exp}}}\sigma_{\mathrm{ID}^{\mathrm{t2}}_{\mathrm{v0,exp}}}}. \quad \text{eqn 31}$$

(Following the notations described in Table 1, the numerator $\sigma_{\mathrm{ID}^{\mathrm{t1}}_{\mathrm{v0,exp}}\mathrm{ID}^{\mathrm{t2}}_{\mathrm{v0,exp}}}$ is the covariance between $\mathrm{ID}^{\mathrm{t1}}_{\mathrm{v0}j,\mathrm{exp}}$ and $\mathrm{ID}^{\mathrm{t2}}_{\mathrm{v0}j,\mathrm{exp}}$, while the denominator $\sigma_{\mathrm{ID}^{\mathrm{t1}}_{\mathrm{v0,exp}}}\sigma_{\mathrm{ID}^{\mathrm{t2}}_{\mathrm{v0,exp}}}$ is the product of their standard deviations). As with behavioural syndromes, the presence of a 'predictability syndrome' implies some modularity, or phenotypic integration (which can represent correlated selective pressures, or genetic correlations; Pigliucci, 2003). The absence of a predictability syndrome implies that different types of traits might be selected to have different levels of predictability.

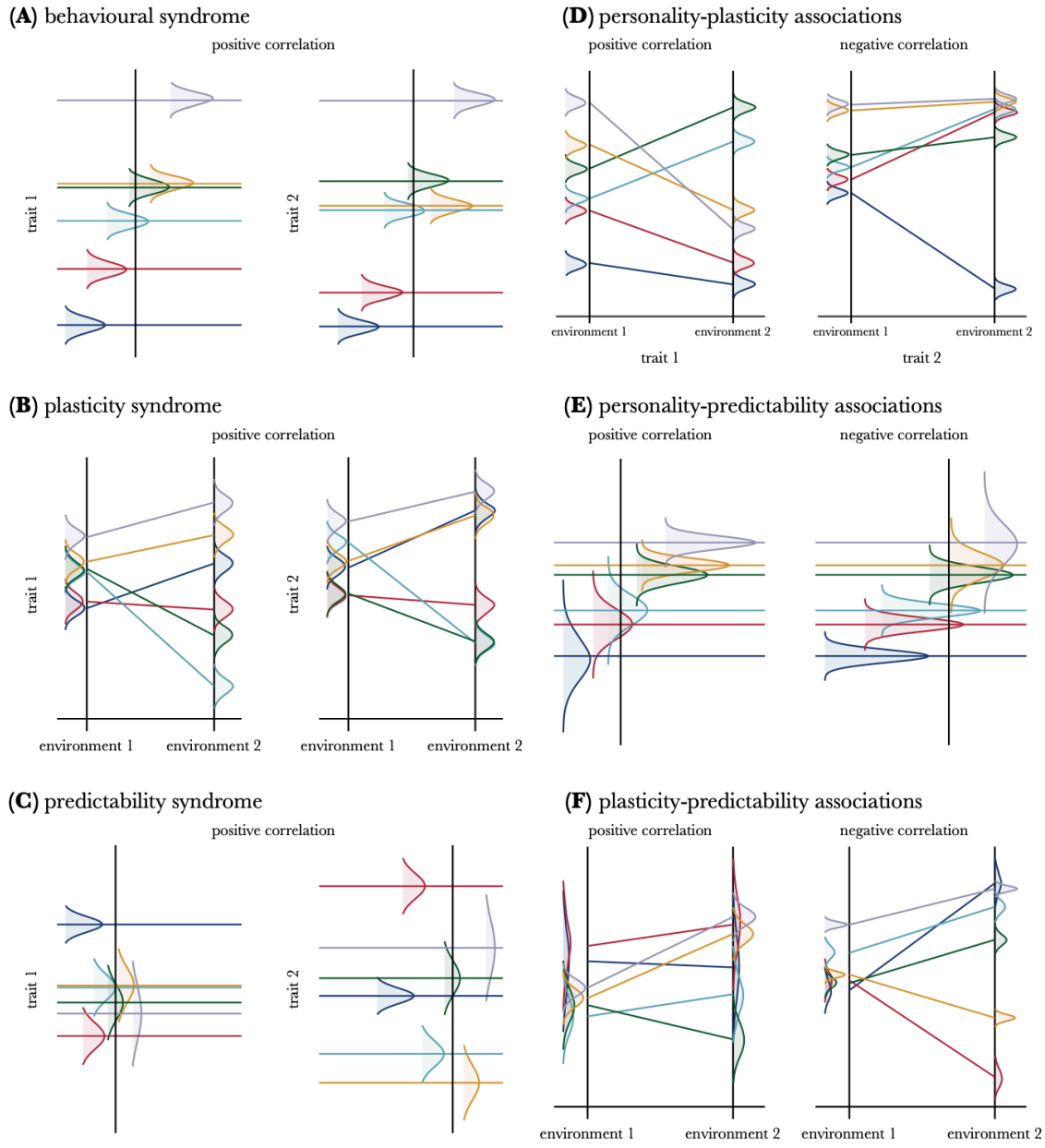


FIGURE 4

Conceptual illustration of six types of correlations, from three types of individual differences (personality, individual differences in plasticity, and individual differences in predictability). Each coloured line and distribution represents a different individual from the same population. The left column shows positive between-trait correlations ('syndromes'), where individual differences are correlated with each other for multiple traits. The right column shows within-trait correlations between pairs of individual differences. (**A**) Behavioural syndrome: individual differences in average behaviour (measured by random intercepts) are positively correlated between two traits, meaning that the 'rank order' of intercepts is maintained (equation 14). (**B**) Plasticity syndrome: the

magnitudes of random slopes are positively correlated (equation 16). (**C**) Predictability syndrome: individuals that are less predictable in one trait (shown by a wider distribution) are less predictable in the second trait (equation 31). (**D**) Personality-plasticity association: individuals with a higher ranking in average behaviour (intercept) have larger *absolute* slopes (equation 9). (**E**) Personality-predictability association: individuals' intercepts are correlated with the (reversed) magnitude of within-individual variance (equation 22). (**F**) Plasticity-predictability syndrome: the magnitude of individual slopes correlates with the ranking of (reversed) individual variances (equation 24).

3.4 | INTRODUCING STOCHASTIC MALLEABILITY

As a future extension to the methods reviewed here, it is possible (given sufficient data) to include a random slope in the dispersion model (i.e. to add $\mathrm{ID}_{\mathrm{v2}j,\mathrm{exp}}$ into equation 18), to estimate individual differences in 'stochastic malleability' (i.e. plasticity in predictability, or simply 'malleability'). While it would require many repeated measurements across different contexts (data simulations are required to estimate the minimum sample size requirements), a fourth type of individual difference, in malleability, could answer three additional questions about phenotypic integration (Fig. 5, below): (1) is the level of malleability integrated across traits (i.e. *malleability syndromes*), or can individuals be malleable in one trait and show fixed predictability in another? (2): do individuals with more flexibility in average phenotypes show more flexibility in variability (i.e. *plasticity-malleability associations*)? (3) are some personality types more or less likely to change their level of predictability in response to an environmental change (i.e. *personality-malleability associations*)? Stochastic malleability could be an important aspect of learning or adapting to novel conditions: naïve individuals (i.e. individuals who are young, or in an unfamiliar environment) might increase variability, to 'sample' a wider array of options. As individuals gain more experience, they might hone in upon the optimal phenotype, and therefore become more predictable (McNamara et al., 2006). An interesting avenue of future research, therefore, could be to incorporate individual differences in malleability into studies of learning or invasion biology (c.f. Chapple et al., 2012; Griffin et al., 2015).

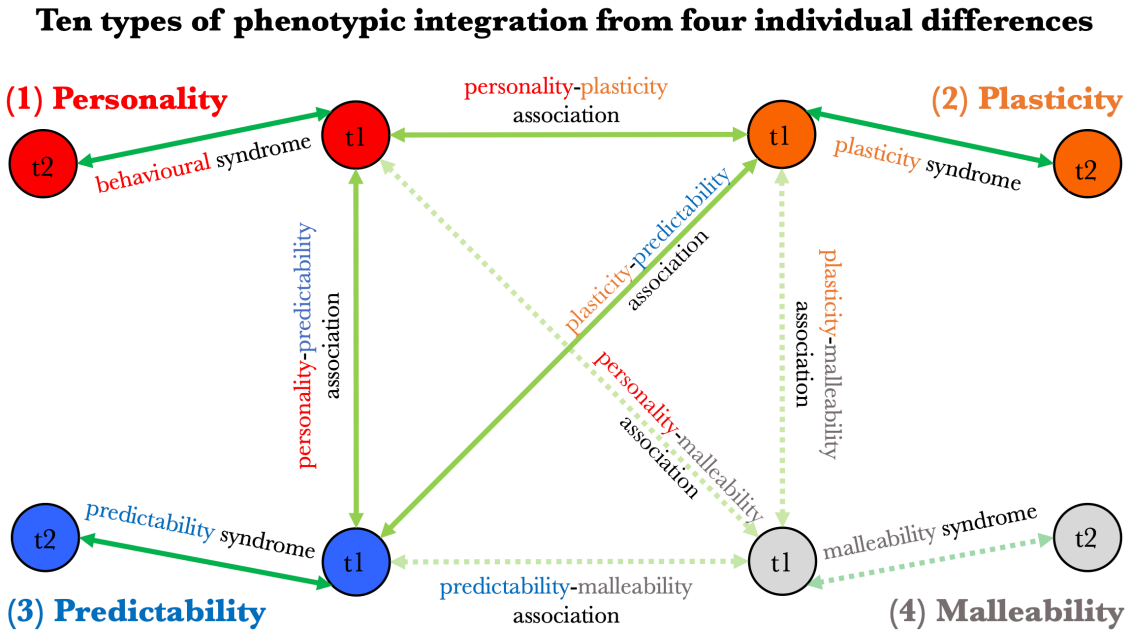


FIGURE 5

Ten types of between-individual correlations can be modelled in a bivariate DHGLM (t1 = trait 1, and t2 = trait 2; as in Table 1), containing four individual differences: (1) personality (random intercept in mean models); (2) plasticity (random slope in mean models); (3) predictability (random intercept in dispersion models); and (4) malleability (random slope in dispersion models). Solids lines indicate correlations that were modelled in the supplementary worked example; our dataset was not suitable to model the correlations shown by dashed lines.

3.5 | SUMMARY OF PREDICTABILITY

With two individual differences — a random intercept and slope in the mean model — we can look at three correlations: two types of syndromes (between traits; Fig. 4A and Fig. 4B) and one intercept-slope association (within trait; Fig. 4D). Modelling predictability adds a third individual difference — a random intercept in the dispersion model. Using a bivariate (multivariate) model, we can simultaneously model these three individual differences in two (or more) types of traits (equations 25-30), and estimate three additional correlations: (1) a predictability syndrome (between traits; Fig. 4C); (2) an association (within traits) between personality and predictability (Fig. 4E); and (3) an association between plasticity and predictability (Fig. 4F). Given sufficient sample sizes, this model can be extended to quantify how much individuals differ in their change in predictability

in different contexts (i.e. 'stochastic malleability'; Fig. 5). Given that most of the variation in behaviour is contained within residual variation — the lowest level of the phenotypic hierarchy — we cannot meaningfully explain important biological variation without considering the variability of individuals.

4 | SUMMARY STATISTICS FOR META-ANALYSIS

The preceding sections described how multilevel models can be used to quantify individual differences in averages, plasticity, and predictability, but how can we compare our results to those from other studies? For between-study comparisons and synthesis (including meta-analyses), the magnitude of individual differences in personality and predictability can be quantified with two different summary statistics: repeatability (Rp), which is variance-standardised, and the coefficient of individual variation (CV_{ID}), which is mean-standardised (and only suitable for ratio-scale measurements; Houle et al., 2011).

For ratio-scale data, both repeatability and the coefficient of individual variation are phenotypic analogues for statistics relating to evolutionary potential (the utility of which are, themselves, debatable; Hansen et al., 2011). Repeatability roughly sets the upper limit on narrow-sense heritability (but see: Dohm, 2002), whereas the coefficient of individual variation is analogous to the coefficient of additive genetic variance, CV_A (Dochtermann & Royauté, 2019; Holtmann et al., 2017; Houle, 1992). Notably, by definition, a repeatability estimate from the dispersion model will always be smaller than its counterpart from the mean model, whereas estimates of the coefficient of individual variation for means and variances are comparable to each other.

Below we describe the calculations required to obtain Rp and CV_{ID} from DHGLM model described by equations 17-20. Supplementary *R* code is available to calculate Rp and CV_{ID} for all models described above and, with some minor modifications, the formulas are broadly applicable for other model specifications too.

4.1 | REPEATABILITY AND THE COEFFICIENT OF INDIVIDUAL VARIATION

Repeatability for the mean model ($\mathrm{Rp_m}$) and dispersion model ($\mathrm{Rp_v}$) are given by:

$$\mathrm{Rp_m} = \frac{\sigma^2_{\mathrm{ID_m}}}{\sigma^2_p}, \quad \text{eqn 32}$$

$$\mathrm{Rp_v} = \frac{\sigma^2_{\mathrm{ID_v}}}{\sigma^2_{\sigma^2_p}}, \quad \text{eqn 33}$$

where σ^2_p is the total phenotypic variance, $\sigma^2_{\sigma^2_p}$ is the total variance in phenotypic variance, and $\sigma^2_{\mathrm{ID_m}}$ and $\sigma^2_{\mathrm{ID_v}}$ are the variance components for between-individual differences in the mean and dispersion models, respectively (Nakagawa & Schielzeth, 2010).

Coefficients of individual variation (similar to CV for additive genetic variance; Mulder et al., 2007; Sae-Lim et al., 2015) for the mean model ($\mathrm{CV_{IDm}}$) and dispersion model ($\mathrm{CV_{IDv}}$) are given by:

$$\mathrm{CV_{IDm}} = \frac{\sigma_{\mathrm{ID}_{\mu_p}}}{\mu_p}, \quad \text{eqn 34}$$

$$\mathrm{CV_{IDv}} = \frac{\sigma_{\mathrm{ID_v}}}{\overline{\sigma}^2_w}. \quad \text{eqn 35}$$

where μ_p is the average individual phenotype, $\overline{\sigma}^2_w$ is the average within-individual variance (the '*w*' represents 'within', and the bar represents the average), and $\sigma_{\mathrm{ID}_{\mu_p}}$ and $\sigma_{\mathrm{ID_v}}$ are the standard deviations for between-individual differences in the mean and dispersion models, respectively. If no transformations have been applied to the response variable, *y*, then $\sigma_{\mathrm{ID}_{\mu_p}} = \sqrt{\sigma^2_{\mathrm{ID_m}}}$ (i.e., the square-root of the numerator for repeatability of the mean, equation 32), and the population mean is calculated for an even sex ratio at the average age of the population ($\mu_p = \frac{2\beta_{\mathrm{m0}} + \beta_{\mathrm{m1}}}{2}$).

4.2 | OBTAINING EACH PARAMETER

Converting parameters from the dispersion model

When calculating Rp and $\mathrm{CV_{ID}}$ from DHGLM models it is essential that all parameters from the dispersion model are first converted back from the natural logarithm (ln) scale

onto the same scale as the mean model. In general, if we have a mean and variance that are estimated on the ln scale, $\mu_{y,\text{exp}}$ and $\sigma^2_{y,\text{exp}}$, then we can convert them back to the normal (observed) scale as follows:

$$\mu_y = \exp\left(\mu_{y,\text{exp}} + \frac{\sigma^2_{y,\text{exp}}}{2}\right), \quad \text{eqn 36}$$

$$\sigma^2_y = \left(\exp\left(\sigma^2_{y,\text{exp}}\right) - 1\right)\exp\left(2\mu_{y,\text{exp}} + \sigma^2_{y,\text{exp}}\right), \quad \text{eqn 37}$$

where μ_y and σ^2_y are the mean and variance on the observed scale. Note that simply taking the exponent of the mean on the ln scale, $\exp(\mu_{y,\text{exp}})$, gives the median estimate on the observed scale, rather than the mean.

Within-individual variance

Usually, the within-individual variance $\overline{\sigma}^2_w$ is assumed to be equal to the average residual variance, $\overline{\sigma}^2_e$. However, there could be a scenario where we calculate $\overline{\sigma}^2_w < \overline{\sigma}^2_e$ by removing an artificial source of variance from the dispersion model (e.g. estimated measurement error). For now, let us assume all the variance in y is biologically meaningful (i.e. we assume $\sigma^2_p = \sigma^2_y$) (de Villemereuil et al., 2018). We therefore take the total variance from the dispersion model as $\sigma^2_{\text{v,exp}} = \sigma^2_{\text{ID}_{\text{v0,exp}}} + \sigma^2_{\text{fixed}_{\text{v,exp}}}$.

On the ln-normal scale, the mean residual variance is the 'population intercept' from the dispersion model, $\beta_{\text{pv0,exp}} = \frac{2\beta_{\text{v0,exp}} + \beta_{\text{v1,exp}}}{2}$, assuming an equal sex ratio with individuals at an average age, $x_{2ij} = 0$ (where $\beta_{\text{v0,exp}}$ is the female intercept, and $\beta_{\text{v1,exp}}$ is the female-male contrast; Table 1). By substituting the ln-normal mean and variance into the mean conversion formula for a ln-normal distribution (i.e., μ_y in equation 36), we obtain $\overline{\sigma}^2_w$ as:

$$\overline{\sigma}^2_w = \exp\left(\beta_{\text{pv0,exp}} + \frac{\sigma^2_{\text{ID}_{\text{v0,exp}}} + \sigma^2_{\text{fixed}_{\text{v,exp}}}}{2}\right). \quad \text{eqn 38}$$

Different model structures will require modifications of the above (and below) equations, for example, when $\sigma^2_y \neq \sigma^2_p$ and/or $\overline{\sigma}^2_e \neq \overline{\sigma}^2_w$.

Between-individual variance and total phenotypic variance

The variance components from the mean model (including variance due to fixed effects) can be summed to obtain $\sigma^2_{\text{ID}_\text{m}}$ and σ^2_p (Allegue et al., 2017). In our case (equations 17-20), modelling individual differences in intercepts (ID_{m0}) and slopes (ID_{m2}) across age (x_2), the

variances are written as:

$$\sigma^2_{\mathrm{ID_m}} = \sigma^2_{\mathrm{ID_{m0}}} + \sigma^2_{\mathrm{ID_{m2}}}\sigma^2_{x_2} + \mu_{x_2}\sigma^2_{\mathrm{ID_{m2}}}, \quad \text{eqn 39}$$

$$\sigma^2_p = \sigma^2_{\mathrm{ID_m}} + \sigma^2_{\mathrm{fixed_m}} + \overline{\sigma}^2_w, \quad \text{eqn 40}$$

$$x_{2ij} \sim \mathrm{D}\left(\mu_{x_2}, \sigma^2_{x_2}\right). \quad \text{eqn 41}$$

The predictor variable x_2 has a mean of μ_{x_2} and a variance of $\sigma^2_{x_2}$, with an arbitrary distribution, D (because no assumptions are made about the distribution of predictors). From equation 39, we can see that when individual differences in personality and plasticity are modelled at the same time, the magnitude of individual differences will depend upon the 'environment' or 'context' at which intercepts are estimated. Typically, continuous predictor variables are mean-centred, so that intercepts are estimated at the average value for that trait ($\mu_{x_2} = 0$). When the predictor is also *z*-transformed ($\sigma^2_{x_2} = 1$), the between-individual variance is simply $\sigma^2_{\mathrm{ID_m}} = \sigma^2_{\mathrm{ID_{m0}}} + \sigma^2_{\mathrm{ID_{m2}}}$ (this is the case in our worked example; Supplementary Information).

Variance in total phenotypic variance

Variance of the total phenotypic variance, $\sigma^2_{\sigma^2_p}$, is hard to conceptualise, but we can estimate it from the total variance in residual variance on the ln-normal scale, $\sigma^2_{\mathrm{ID_{v0,exp}}} + \sigma^2_{\mathrm{fixed_{v,exp}}}$, and the average residual variance, $\beta_{\mathrm{pv0,exp}}$. To obtain the variance of within-individual variances on the observed scale, $\sigma^2_{\sigma^2_w}$, we substitute these values into the conversion of variance from a ln-normal distribution (equation 37), such that:

$$\sigma^2_{\sigma^2_w} = \left(\exp(\sigma^2_{\mathrm{ID_{v0,exp}}} + \sigma^2_{\mathrm{fixed_{v,exp}}}) - 1\right)\exp\left(2\beta_{\mathrm{pv0,exp}} + \sigma^2_{\mathrm{ID_{v0,exp}}} + \sigma^2_{\mathrm{fixed_{v,exp}}}\right), \quad \text{eqn 42}$$

The formula for $\sigma^2_{\sigma^2_p}$ is then provided by Mulder et al. (2007) as:

$$\sigma^2_{\sigma^2_p} = 2\sigma^4_p + 3\sigma^2_{\sigma^2_w}, \quad \text{eqn 43}$$

where the value for σ^2_p is shown in equation 40.

Between-individual variance for the within-individual variance

In our case, the between-individual variance for the within-individual variances is $\sigma^2_{\mathrm{ID_v}} = \sigma^2_{\mathrm{ID_{v0}}}$, so we need to convert $\sigma^2_{\mathrm{ID_{v0,exp}}}$ (from the ln-normal scale) to $\sigma^2_{\mathrm{ID_{v0}}}$. Our first thought might be to apply the same transformation to $\sigma^2_{\mathrm{ID_{v0,exp}}}$ as we did for $\sigma^2_{\mathrm{ID_{v0,exp}}} +$

$\sigma^2_{\text{fixed}_{\text{v,exp}}}$ (i.e. equation 37). However, because the ln-transformation is non-linear, we cannot simply disentangle $\sigma^2_{\text{ID}_{\text{v0,exp}}}$ from $\sigma^2_{\text{fixed}_{\text{v,exp}}}$. The solution, provided by Mulder et al. (2007), is to assume that the proportionality of variance components is preserved across different scales (see also Sae-Lim et al., 2015) so that:

$$\sigma^2_{\text{ID}_{\text{v0}}} = \sigma^2_{\sigma^2_w}\left(\frac{\sigma^2_{\text{ID}_{\text{v0,exp}}}}{\sigma^2_{\text{ID}_{\text{v0,exp}}}+\sigma^2_{\text{fixed}_{\text{v,exp}}}}\right), \quad \text{eqn 44}$$

where $\sigma^2_{\sigma^2_w} = \sigma^2_{\text{ID}_{\text{v0}}} + \sigma^2_{\text{fixed}_{\text{v}}}$. Thus, we are assuming the ratio of variance components on the ln-normal scale is the same as the ratio of variance components on the observed scale: $\frac{\sigma^2_{\text{ID}_{\text{v0,exp}}}}{\sigma^2_{\text{ID}_{\text{v0,exp}}}+\sigma^2_{\text{fixed}_{\text{v,exp}}}} = \frac{\sigma^2_{\text{ID}_{\text{v0}}}}{\sigma^2_{\text{ID}_{\text{v0}}}+\sigma^2_{\text{fixed}_{\text{v}}}}$ (we refer to this assumption as 'the preservation of proportionality').

4.3 | COMPARING ESTIMATES BETWEEN STUDIES

When standardising variance estimates it is important to consider the scale of measurement, mean-variance relationship, and whether or not the data were transformed prior to analysis. An accessible summary of the limitations of coefficients of variation is provided by Pélabon et al. (2020).

Any between-study comparison of the magnitude of individual differences would ideally start with a re-examination and analysis of the original data (which are increasingly made publicly available by authors in ecology and evolution). Standardising the way Rp and CV_{ID} are calculated is important because between-study variance in estimates can be increased by variation in statistical methods and chosen formulas (e.g., was fixed effect variance included or excluded from the total phenotypic variance?). Calculating Rp and CV_{ID} from scratch also allows sampling variance to be estimated for meta-analytic models.

In addition to being influenced by analysis decisions, Rp and CV_{ID} can vary due to different experimental and sampling designs (Wilson, 2018). For instance, a statistical difference between individuals could reflect the effects of measuring individuals in different conditions (e.g., due to being sampled at different times), rather than true between-individual differences (e.g. 'pseudo-repeatability'; Dingemanse & Dochtermann, 2013). Likewise, a short sampling interval between repeated measurements is likely to

inflate estimates of individual differences, due to temporal autocorrelation. It is also important to consider the impact that sampling intervals have on individual's behavioural responses (e.g. habituation) and, within studies, standardise these intervals across individuals.

For comparisons of CV_{ID}, two additional points are important to consider. First, were data transformed prior to analysis? If so, estimated parameters need to be brought back to the observed scale (this applies both to comparisons across studies, and comparisons within studies for different phenotypic traits). The supplementary worked exampled describes how to reverse linear transformations (e.g., z-scaling) and non-linear transformations (e.g., log- or square-root transformations, which are commonly done to improve the normality of residuals). Second, when comparing estimates of $\mathbf{CV_{IDv}}$ to another study, did that study also use residual variances as the response variable for the dispersion model, or did it use residual standard deviations, as in Cleasby et al. (2015)? In the latter case, we can convert $\mathbf{CV_{IDv}}$ to $\mathbf{CV_{IDsd}}$, and *vice versa*, using basic properties of logarithms and variance (full details of these conversions are provided in the Supplementary Information).

5 | CONCLUSIONS AND FUTURE DIRECTIONS

Incorporating predictability into studies of personality and plasticity creates an opportunity to test more nuanced questions about how phenotypic variation is maintained, or constrained. For some traits, it might be adaptive to be unpredictable, such as in predator-prey interactions (Briffa, 2013). For other traits, selection might act to minimise maladaptive imprecision around an optimal mean (Hansen et al., 2006). The supplementary worked example shows phenotypic integration of predictability across multiple behavioural traits, and some integration of predictability with personality and plasticity. Phenotypic integration could hint at genetic integration too; other studies have found additive genetic variance in predictability (Martin et al., 2017; Prentice et al., 2020). Given that different traits might have different optimal levels of unpredictability, phenotypic integration of predictability could constrain variation in one trait (resulting in

724 lower than optimal variability) and maintain variation in another (resulting in greater than
725 optimal variability). Associations with personality and plasticity mean that variation in
726 predictability — the lowest level of the phenotypic hierarchy — could have cascading
727 effects upwards (Westneat et al., 2015). Empirical estimates of the strength of these
728 associations can inform theoretical models on the simultaneous evolution of means and
729 variances.

730 *Beyond behaviour*

731 We focussed this paper on animal behaviour (the field we are most familiar with), but the
732 models are broadly adaptable. Individuals can show differences in predictability for any
733 trait that is repeatedly expressed. For example, medical researchers might want to
734 quantify the variability of patient's drug responses (Nettles et al., 2006), and selective
735 breeders of plants might want to reduce individual variability in seed or fruit mass
736 (Herrera, 2017). The review by Herrera (2017) discusses the overlooked importance of
737 variability within the structures of an individual plant, including for plant-animal
738 interactions. Given the large sample sizes required to estimate multiple individual
739 differences, the most tractable tests of the synchronous evolution of means and variances
740 could come from non-animal systems. Clonal species can also be used to estimate
741 individual differences in predictability of non-labile traits.
742

743 *Conclusions*

744 While many studies quantify consistent individual differences in repeatedly expressed
745 traits, such as behaviour, much of the mystery of phenotypic variation is obscured within
746 residual variation. Individuals impact the world not only through their 'average'
747 phenotype, but also through their extremes. Given that evolution can act on both averages
748 and variances, to understand the evolution of labile traits, we need to measure both the
749 magnitude and consistency of individual differences, as well as their phenotypic
750 integration. Limitations of the concepts and tools presented here include the high sample
751 sizes required to accurately estimate variance components and co-variances, and concerns
752 about inflated rates of false-positive findings when estimating many parameters. Future
753 simulation work is required to help empiricists design adequate sampling methods to
754 chronicle the integration of multiple levels of phenotypic variation in diverse systems. In
755 doing so we can improve our understanding of the factors promoting and constraining
756 variability, as well as the evolution, and ecological consequences, of individuality.

Acknowledgements

This work was conducted with permission from the Garvan Institute Animal Ethics Committee (approval number 15/15) with funding support from UNSW. Daniel Hesselson provided access to the zebrafish population featured in the supplementary worked example, which were looked after by staff within the Garvan Institute of Medical Research. Data were collected with help from Malgorzata Lagisz, Melissa Fangmeier, Takuji Usui, Marissa Baptista, Alex Lincoln-Dodgson, Hamza Anwer, Harry Thomas, and Alexander Aloy. Joel Pick helped write the code for bivariate predictability models. RO was supported by an Australian Government Research Training Program scholarship, DN was supported by an ARC DECRA fellowship (DE150101774), SN was supported by an ARC Discovery grant (DP180100818), and all authors declare no conflicts of interest. We would like to thank Bob Wong and Alison Bell for their comments on an earlier version of this manuscript.

References

Allegue, H., Araya-Ajoy, Y. G., Dingemanse, N. J., Dochtermann, N. A., Garamszegi, L. Z., Nakagawa, S., et al. (2017). Statistical Quantification of Individual Differences (SQuID): An educational and statistical tool for understanding multilevel phenotypic data in linear mixed models. *Methods in Ecology and Evolution*, *8*(2), 257–267. http://doi.org/10.1111/2041-210X.12659

Bell, A. M. (2007). Future directions in behavioural syndromes research. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1611), 755–761. http://doi.org/10.1098/rspb.2006.0199

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

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed–accuracy tradeoff. *Trends in Neurosciences, 33*(1), 10–16. http://doi.org/10.1016/j.tins.2009.09.002

Briffa, M. (2013). Plastic proteans: Reduced predictability in the face of predation risk in hermit crabs. *Biology Letters*, *9*(5), 20130592–4.

http://doi.org/10.1098/rsbl.2013.0592

Brommer, J. E. (2013). 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. http://doi.org/10.1007/s00265-013-1527-4

Chapple, D. G., Simmonds, S. M., & Wong, B. B. M. (2012). Can behavioral and personality traits influence the success of unintentional species introductions? *Trends in Ecology & Evolution*, *27*(1), 57–64. http://doi.org/10.1016/j.tree.2011.09.010

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. http://doi.org/10.1111/2041-210X.12281

Cornwell, T. O., McCarthy, I. D., Snyder, C. R. A., & Biro, P. A. (2019). The influence of environmental gradients on individual behaviour: Individual plasticity is consistent across risk and temperature gradients. *Journal of Animal Ecology*, *88*(4), 511–520. http://doi.org/10.1111/1365-2656.12935

Dahlbom, S. J., Lagman, D., Lundstedt-Enkel, K., Sundström, L. F., & Winberg, S. (2011). Boldness predicts social status in zebrafish (*Danio rerio*). *PLoS One*, *6*(8), e23565. http://doi.org/10.1371/journal.pone.0023565

de Villemereuil, P., Morrissey, M. B., Nakagawa, S., & Schielzeth, H. (2018). Fixed-effect variance and the estimation of repeatabilities and heritabilities: Issues and solutions. *Journal of Evolutionary Biology*, *31*(4), 621–632. http://doi.org/10.1111/jeb.13232

Dereje, S., Sawyer, S., Oxendine, S. E., Zhou, L., Kezios, Z. D., Wong, R. Y., et al. (2012). Comparing behavioral responses across multiple assays of stress and anxiety in zebrafish (*Danio rerio*). *Behaviour*, *149*(10-12), 1205–1240. http://doi.org/10.1163/1568539X-00003018

DeWitt, T. J., Sih, A., & Wilson, D. S. (1998). Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution*, *13*(2), 77–81. http://doi.org/10.1016/S0169-5347(97)01274-3

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

Dingemanse, N. J., Barber, I., Wright, J., & Brommer, J. E. (2012). Quantitative

822 genetics of behavioural reaction norms: Genetic correlations between personality
823 and behavioural plasticity vary across stickleback populations. *Journal of Evolutionary*
824 *Biology*, *25*(3), 485–496. http://doi.org/10.1111/j.1420-9101.2011.02439.x
825 Dingemanse, N. J., Dochtermann, N., & Wright, J. (2010a). A method for exploring the
826 structure of behavioural syndromes to allow formal comparison within and between
827 data sets. *Animal Behaviour*, *79*(2), 439–450.
828 http://doi.org/10.1016/j.anbehav.2009.11.024
829 Dochtermann, N. A. (2010). Behavioral syndromes: Carryover effects, false discovery
830 rates, and a priori hypotheses. *Behavioral Ecology*, *21*(3), 437–439.
831 http://doi.org/10.1093/beheco/arq021
832 Dochtermann, N. A., & Royauté, R. (2019). The mean matters: Going beyond
833 repeatability to interpret behavioural variation. *Animal Behaviour*, *153*, 147–150.
834 http://doi.org/10.1016/j.anbehav.2019.05.012
835 Dohm, M. R. (2002). Repeatability estimates do not always set an upper limit to
836 heritability. *Functional Ecology*, *16*(2), 273–280. http://doi.org/10.1046/j.1365-
837 2435.2002.00621.x
838 Dubois, F. (2019). Why are some personalities less plastic? *Proceedings of the Royal Society B:*
839 *Biological Sciences*, *286*(1908), 20191323–7. http://doi.org/10.1098/rspb.2019.1323
840 Fawcett, T. W., Hamblin, S., & Giraldeau, L. A. (2012). Exposing the behavioral
841 gambit: The evolution of learning and decision rules. *Behavioral Ecology*, *24*(1), 2–11.
842 http://doi.org/10.1093/beheco/ars085
843 Flatt, T. (2005). The evolutionary genetics of canalization. *The Quarterly Review of Biology*,
844 *80*(3), 287–316. http://doi.org/10.1086/432265
845 Frazer, K. A., Murray, S. S., Schork, N. J., & Topol, E. J. (2009). Human genetic
846 variation and its contribution to complex traits. Nature Reviews Genetics, 10(4),
847 241–251. http://doi.org/10.1038/nrg2554
848 Gavrilets, S., & Scheiner, S. M. (1993). The genetics of phenotypic plasticity. V.
849 Evolution of reaction norm shape. *Journal of Evolutionary Biology*, *6*(1), 31–48.
850 http://doi.org/10.1046/j.1420-9101.1993.6010031.x
851 Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive
852 versus non-adaptive phenotypic plasticity and the potential for contemporary
853 adaptation in new environments. *Functional Ecology*, *21*(3), 394–407.
854 http://doi.org/10.1111/j.1365-2435.2007.01283.x
855 Gianoli, E., & Palacio-Lopez, K. (2009). Phenotypic integration may constrain

phenotypic plasticity in plants. *Oikos*, *118*(12), 1924–1928. http://doi.org/10.1111/j.1600-0706.2009.17884.x

Gomulkiewicz, R., & Kirkpatrick, M. (1992). Quantitative genetics and the evolution of reaction norms. *Evolution*, *46*(2), 390–411. http://doi.org/10.1111/j.1558-5646.1992.tb02047.x

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

Hadfield, J. D., Wilson, A. J., Garant, D., Sheldon, B. C., & Kruuk, L. E. B. (2010). The misuse of BLUP in ecology and evolution. *The American Naturalist*, *175*(1), 116–125. http://doi.org/10.1086/648604

Hansen, T. F., & Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology*, *21*(5), 1201–1219. http://doi.org/10.1111/j.1420-9101.2008.01573.x

Hansen, T. F., Carter, A. J. R., & Pélabon, C. (2006). On adaptive accuracy and precision in natural populations. *The American Naturalist*, *168*(2), 168–181. http://doi.org/10.1086/505768

Hansen, T. F., Pélabon, C., & Houle, D. (2011). Heritability is not Evolvability, *38*(3), 258–277. http://doi.org/10.1007/s11692-011-9127-6

Herrera, C. M. (2017). The ecology of subindividual variability in plants: Patterns, processes, and prospects. *Web Ecology*, *17*(2), 51–64. http://doi.org/10.5194/we-17-51-2017

Holtmann, B., Lagisz, M., & Nakagawa, S. (2017). Metabolic rates, and not hormone levels, are a likely mediator of between-individual differences in behaviour: A meta-analysis. *Functional Ecology*, *31*(3), 685–696. http://doi.org/10.1111/1365-2435.12779

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, *130*, 195–204.

Kern, E. M. A., Robinson, D., Gass, E., Godwin, J., & Langerhans, R. B. (2016). Correlated evolution of personality, morphology and performance. *Animal Behaviour*, *117*, 79–86. http://doi.org/10.1016/j.anbehav.2016.04.007

Lin, X., Mermelstein, R. J., & Hedeker, D. (2018). A 3-level Bayesian mixed effects location scale model with an application to ecological momentary assessment data. *Statistics in Medicine*, *37*(13), 2108–2119. http://doi.org/10.1002/sim.7627

Mallitt, K. L., Bonser, S. P., & Hunt, J. (2010). The plasticity of phenotypic integration in response to light and water availability in the pepper grass, *Lepidium bonariense*. *Evolutionary Ecology*, *24*(6), 1321–1337. http://doi.org/10.1007/s10682-010-9373-6

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. http://doi.org/10.1111/jeb.13048

McNamara, J. M., Green, R. F., & Olsson, O. (2006). Bayes' theorem and its applications in animal behaviour. *Oikos*, *112*(2), 243–251. http://doi.org/10.1111/j.0030-1299.2006.14228.x

Mitchell, D. J., & Biro, P. A. (2017). Is behavioural plasticity consistent across different environmental gradients and through time? *Proceedings of the Royal Society B: Biological Sciences*, *284*(1860), 20170893. http://doi.org/10.1098/rspb.2017.0893

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), 160352–10. http://doi.org/10.1098/rsos.160352

Moirón, M., Laskowski, K. L., & Niemelä, P. T. (2020). Individual differences in behaviour explain variation in survival: a meta-analysis. *Ecology Letters*, *23*(2), 399–408. http://doi.org/10.1111/ele.13438

Moretz, J. A., Martins, E. P., & Robison, B. D. (2007). Behavioral syndromes and the evolution of correlated behavior in zebrafish. *Behavioral Ecology*, *18*(3), 556–562. http://doi.org/10.1093/beheco/arm011

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. http://doi.org/10.1534/genetics.106.063743

Mustafa, A., Roman, E., & Winberg, S. (2019). Boldness in male and female zebrafish (*Danio rerio*) is dependent on strain and test. *Frontiers in Behavioral Neuroscience*, *13*, 20150509. http://doi.org/10.3389/fnbeh.2019.00248

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

Nettles, R. E., Kieffer, T. L., Parsons, T., Johnson, J., Cofrancesco, J., Gallant, J. E., et al. (2006). Marked intraindividual variability in antiretroviral concentrations may limit the utility of therapeutic drug monitoring. *Clinical Infectious Diseases*, *42*(8),

924 1189–1196. http://doi.org/10.1086/501458
925 Niemelä, P. T., & Dingemanse, N. J. (2018). On the usage of single measurements in
926 behavioural ecology research on individual differences. *Animal Behaviour*, *145*, 99–
927 105. http://doi.org/10.1016/j.anbehav.2018.09.012
928 O'Dea, R.E., Nakagawa, S. (2020). Worked Example: Personality, Plasticity, and
929 Predictability. *Open Science Framework*. http://doi.org/10.17605/OSF.IO/V3QAX
930 Pélabon, C., Hilde, C. H., Einum, S., & Gamelon, M. (2020). On the use of the
931 coefficient of variation to quantify and compare trait variation. *Evolution Letters*, 4(3),
932 180–188. http://doi.org/10.1002/evl3.171
933 Pigliucci, M. (2002). Touchy and bushy: Phenotypic plasticity and integration in
934 response to wind stimulation in *Arabidopsis thaliana*. *International Journal of Plant*
935 *Sciences*, *163*(3), 399–408. http://doi.org/10.1086/339158
936 Pigliucci, M. (2003). Phenotypic integration: Studying the ecology and evolution of
937 complex phenotypes. *Ecology Letters*, *6*(3), 265–272. http://doi.org/10.1046/j.1461-
938 0248.2003.00428.x
939 Postma, E. (2006). Implications of the difference between true and predicted breeding
940 values for the study of natural selection and micro-evolution. *Journal of Evolutionary*
941 *Biology*, *19*(2), 309–320. http://doi.org/10.1111/j.1420-9101.2005.01007.x
942 Prentice, P. M., Houslay, T. M., Martin, J. G. A., & Wilson, A. J. (2020). Genetic
943 variance for behavioural "predictability" of stress response. *Journal of Evolutionary*
944 *Biology*, *39*(03), 473. http://doi.org/10.1111/jeb.13601
945 R Core Team. (2020). *R*: A language and environment for statistical computing.
946 Vienna, Austria: R Foundation for Statistical Computing. Retrieved from
947 http://www.R-project.org.
948 Ramakers, J. J. C., Visser, M. E., & Gienapp, P. (2020). Quantifying individual
949 variation in reaction norms: Mind the residual. *Journal of Evolutionary Biology*, *33*, 352-
950 366. http://doi.org/10.1111/jeb.13571
951 Rast, P., Hofer, S. M., & Sparks, C. (2012). Modeling individual differences in within-
952 person variation of negative and positive affect in a mixed effects location scale
953 model using BUGS/JAGS. *Multivariate Behavioral Research*, *47*(2), 177–200.
954 http://doi.org/10.1080/00273171.2012.658328
955 Reznick, D., Nunney, L., & Tessier, A. (2000). Big houses, big cars, superfleas and the
956 costs of reproduction. *Trends in Ecology & Evolution*, *15*(10), 421–425.
957 http://doi.org/10.1016/S0169-5347(00)01941-8

958 Sae-Lim, P., Kause, A., Janhunen, M., Vehviläinen, H., Koskinen, H., Gjerde, B., et al.
959 (2015). Genetic (co)variance of rainbow trout (*Oncorhynchus mykiss*) body weight and
960 its uniformity across production environments. *Genetics Selection Evolution*, *47*(1), 303.
961 http://doi.org/10.1186/s12711-015-0122-8
962 Schlichting, C. D. (1989). Phenotypic integration and environmental change. *BioScience*,
963 *39*(7), 460–464. http://doi.org/10.2307/1311138
964 Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: An ecological and
965 evolutionary overview. *Trends in Ecology & Evolution*, *19*(7), 372–378.
966 http://doi.org/10.1016/j.tree.2004.04.009
967 Sih, A., Cote, J., Evans, M., Fogarty, S., & Pruitt, J. (2012). Ecological implications of
968 behavioural syndromes. *Ecology Letters*, *15*(3), 278–289.
969 http://doi.org/10.1111/j.1461-0248.2011.01731.x
970 Sih, A., & Del Giudice, M. (2012). Linking behavioural syndromes and cognition: a
971 behavioural ecology perspective. *Philosophical Transactions of the Royal Society B:*
972 *Biological Sciences, 367*(1603), 2762–2772. http://doi.org/10.1098/rstb.2012.0216
973 Sih, A., Mathot, K. J., Moirón, M., Montiglio, P. O., Wolf, M., & Dingemanse, N. J.
974 (2015). Animal personality and state–behaviour feedbacks: a review and guide for
975 empiricists. *Trends in Ecology & Evolution*, 30(1), 50–60.
976 http://doi.org/10.1016/j.tree.2014.11.004
977 Snell-Rood, E. C. (2013). An overview of the evolutionary causes and consequences of
978 behavioural plasticity. *Animal Behaviour*, *85*(5), 1004–1011.
979 http://doi.org/10.1016/j.anbehav.2012.12.031
980 Stamps, J. A. (2015). Individual differences in behavioural plasticities. *Biological Reviews*,
981 *91*(2), 534–567. http://doi.org/10.1111/brv.12186
982 Stamps, J. A., & Biro, P. A. (2016). Personality and individual differences in plasticity.
983 *Current Opinion in Behavioral Sciences*, *12*, 18–23.
984 http://doi.org/10.1016/j.cobeha.2016.08.008
985 Stamps, J. A., & Groothuis, T. G. G. (2010). Developmental perspectives on personality:
986 implications for ecological and evolutionary studies of individual differences.
987 *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*,
988 *365*(1560), 4029–4041. http://doi.org/10.1098/rstb.2010.0218
989 Stearns, S. C., & Koella, J. C. (1986). The evolution of phenotypic plasticity in life-
990 history traits: Predictions of reaction norms for age and size at maturity. *Evolution*,
991 *40*(5), 893–913. http://doi.org/10.1111/j.1558-5646.1986.tb00560.x

992 Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2017). *rptR*: Repeatability estimation and
993 variance decomposition by generalized linear mixed-effects models. *Methods in*
994 *Ecology and Evolution*, *8*(11), 1639–1644. http://doi.org/10.1111/2041-210X.12797
995 Viney, M., & Reece, S. E. (2013). Adaptive noise. *Proceedings of the Royal Society B:*
996 *Biological Sciences*, *280*(20131104). http://doi.org/10.1098/rspb.2013.1104
997 Westneat, D. F., Hatch, M. I., Wetzel, D. P., & Ensminger, A. L. (2011). Individual
998 variation in parental care reaction norms: Integration of personality and plasticity.
999 *The American Naturalist*, *178*(5), 652–667. http://doi.org/10.1086/662173
1000 Westneat, D. F., Schofield, M., & Wright, J. (2013). Parental behavior exhibits among-
1001 individual variance, plasticity, and heterogeneous residual variance. *Behavioral*
1002 *Ecology*, *24*(3), 598–604. http://doi.org/10.1093/beheco/ars207
1003 Westneat, D. F., Wright, J., & Dingemanse, N. J. (2015). The biology hidden inside
1004 residual within-individual phenotypic variation. *Biological Reviews*, *90*(3), 729–743.
1005 http://doi.org/10.1111/brv.12131
1006 Willmore, K. E., Young, N. M., & Richtsmeier, J. T. (2007). Phenotypic variability: Its
1007 components, measurement and underlying developmental processes. *Evolutionary*
1008 *Biology*, *34*(3-4), 99–120. http://doi.org/10.1007/s11692-007-9008-1
1009 Wilson, A. J. (2018). How should we interpret estimates of individual repeatability?
1010 *Evolution Letters*, 2(1), 4–8. http://doi.org/10.1002/evl3.40
1011 Wolf, M., van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs
1012 favour the evolution of animal personalities. *Nature*, *447*(7144), 581–584.
1013 http://doi.org/10.1038/nature05835