ORIGINAL ARTICLE

Weak evidence that asset protection underlies temporal or contextual consistency in boldness of a terrestrial crustacean

Darcy Beveridge[1] · David J. Mitchell[2] · Christa Beckmann[1,3] · Peter A. Biro[1]

Abstract

In recent years, many studies have investigated the potential state dependence of individual differences in behaviour, with the aim to understand the proximate and ultimate causes and consequences of animal personality. Among the potential state variables that could affect behavioural expression is size and mass, but few studies have found associations at the among-individual levels. Insufficient sampling and incorrect analysis of data are cited as impediments to detecting correlations, if they exist. Here, we conducted a study using 100 pillbugs (*Armadillidium vulgare*) and assayed their defensive behaviour 24 times each over time and across familiarity contexts, to test the asset protection hypothesis that predicts a negative correlation between boldness and mass, and with increases in mass over time. Multivariate mixed models revealed that despite mostly consistent individual behavioural differences over time (modest slope variance) and across contexts (near-parallel reaction norms), and 18-fold range in starting mass, there was no correlation between individual mean mass and boldness. However, individuals that gained more mass over time may have been more 'shy' compared to those gaining less mass, but the correlation was weak and observed variation in mass gain was small. There was also a mean level trend of increasing shyness over time that was coincident with mean level mass increases over time. Together, our study provides weak evidence for the asset protection hypothesis, whereby individuals that accumulate more resources are thought to protect them through risk averse behaviour.

Significance statement

Individual variation in 'state', such as mass or energy reserves, is thought to be a predictor of individual differences in behaviour that are consistent over time. However, few studies reveal such links, and several studies suggest insufficient sampling may explain null results in most studies. We studied 100 animals sampled 24 times each in a controlled setting to reveal stable individual differences in mean behaviour over time and across contexts; however, individual behaviour was unrelated to large differences in individual mass but weakly related to increases in mass through time whereby individuals became more shy and those growing faster were somewhat more shy. Our results provide little evidence for the asset protection hypothesis.

Communicated by T. Breithaupt

✉ Peter A. Biro
   pete.biro@deakin.edu.au

   Darcy Beveridge
   darcy.bev@hotmail.com

   David J. Mitchell
   david.mitchell@zoologi.su.se

   Christa Beckmann
   C.Beckmann@westernsydney.edu.au

1   Centre for Integrative Ecology, School of Life & Environmental Sciences, Deakin University, 75 Pigdons Road, Geelong, VIC 3216, Australia

2   Department of Zoology/Ethology, Stockholm University, Svante Arrheniusväg 18B. SE-106 91, Stockholm, Sweden

3   School of Science and Health, Western Sydney University, Parramatta, NSW 2751, Australia

Introduction

Behavioural ecologists have historically treated individuals as being highly flexible in behaviour. The last decade or two has seen a major shift in this viewpoint, with widespread acceptance now that individuals have unique personalities and thus differ in expressed behaviour at a given time and/or in a given situation, indicating they can perceive and react to internal and external stimuli quite differently. There is considerable evidence for personality in many different behaviours and across many species, evidenced by significant trait repeatability (Bell et al. 2009; Wolak et al. 2012) and heritability (Dochtermann et al. 2019), which also include energetic and neuroendocrine traits thought to covary with and affect the expression of behaviour (Nespolo and Franco 2007; White et al. 2013; Fanson and Biro 2018).

Considerable theory development and empirical studies have focussed on proximate factors that might affect the expression of behaviour—internal 'state' variables—which may help us understand personality. These can include life history traits, hormones, metabolism and acquired information which are thought to be functionally and/or genetically integrated with behaviour (discussed by Sih et al. 2015). Size and/or mass is a particularly important state variable known to affect fitness via competitive and reproductive advantages, and asset protection theory predicts that as an individual accumulates resources and size/mass its behaviour should become more risk averse (Clark 1994). An extension of that model is the prediction that individuals accumulating mass more quickly through growth or energy reserves should also reduce risk exposure (Clark 1994; Luttbeg and Sih 2010; Sih et al. 2015). Another prediction arising from that model is that if increases in state lead to reduced boldness, then over longer time horizons we would expect the erosion of individual variance due to positive feedback loops (Sih et al. 2015), converging temporal reaction norms and reduced repeatability.

Early reviews of state-behaviour relationships indicated correlations between life history productivity and individual behaviour, with traits such as activity, boldness and aggression being generally positively correlated with metabolism, growth rates, fecundity and early maturation (Biro and Stamps 2008, 2010). More recently, reviews and meta-analyses indicate that these correlations largely do hold among taxa (Healy et al. 2019), but for studies focusing on the among-individual level of variation within species, there is little evidence for state-behaviour correlations (Niemelä and Dingemanse 2018; Royauté et al. 2018), but see Biro et al. (2014), Videlier et al. (2019), and Cornwell et al. (2020). One review in particular revealed that associations between behaviour and size/mass had average correlations approaching zero and highlighted shortcomings of many of the reviewed studies: studies typically had few or no repeated measures of individuals and were poorly analysed which impeded robust inference (Niemelä and Dingemanse 2018). Although similar concerns have been explicitly raised in the past, it is clear that these problems persist. These concerns have included lack of precision and bias towards zero correlation and confounding of among- and within-individual correlations (Adolph and Hardin 2007; Wolak et al. 2012; Beckmann and Biro 2013; Brommer 2013; Biro and Stamps 2015).

Here, we studied the relationship between individual mass and boldness at the among- and within-individual levels in the pillbug (*Armadillidium vulgare*). We used 100 animals of a wide range of body masses and housed them in an environment with high-quality and plentiful food; we created a situation where we had substantial among-individual variation in mass and observed within-individual variation in changes in mass (growth). We predicted that shy individuals with long latency to unroll to be larger (an among-individual correlation), and those increasing more in mass to also become shyer (a within-individual correlation)—as predicted by the asset protection hypothesis (Clark 1994). Furthermore, we predicted that individuals responding more strongly with longer latency when tested in an open and presumably more risky part of their home tank would also tend to be larger and have increased more in mass than when tested on top of their soil refuge habitat.

Methods

Study species

We used the common pillbug (*A. vulgare*) as the model animal in this study. We chose this species for its unambiguous and easily quantified anti-predator response whereby they roll themselves into a sphere, such that all appendages are contained under their protective exoskeleton. The latency to fully unroll from full 'conglobation' can thus be used as an ecologically meaningful and objective measure of boldness. This behaviour is a form of death feigning and simultaneously offers physical protection from predation and camouflages the animal among the litter, making the animal appear like a small stone rather than a potential food item. By day, they mostly live hidden under debris, and when disturbed, or uncovered by a potential predator (e.g. small mammals, birds, frogs, spiders or lizards), they roll up into a ball (Smigel and Gibbs 2008). These terrestrial crustaceans, native to areas across Europe, are an invasive species in south-eastern Australia (where this study was conducted),

as they are in much of the temperate world. These animals are isopods and feed on dead and decaying plant matter.

Sampling

We collected a total of about 200 individuals of a wide range of different sizes from the Deakin University, Waurn Ponds campus and nearby rural areas within a 30-km radius (38.1984° S, 144.2988° E). To reduce the chance of sampling related individuals, all pillbugs were collected from at least 15 different sites and separated by at least tens of metres at a given site. Animals were first held for 10 days in communal housing, containing moist compost and ad libitum food (see below). From this large sample, we next haphazardly sub-sampled 100 of these individuals with the aim of obtaining a very wide range of body sizes which was representative of the range of body sizes encountered in the field.

Husbandry

These pillbugs were then housed individually in separate clear plastic tanks (15 × 8 × 9 cm, L × W × H). Each tank received 30 g of dry compost (Bannockburn Garden Supplies), which was distributed over one half the tank (hereafter, the unexposed or familiar environment), while the other half was left empty (bare plastic floor; the exposed or unfamiliar environment). Pillbugs only regularly inhabited the unexposed half, where they buried themselves for protection. The compost was thoroughly wetted down to provide moisture, and pillbugs immediately used this substrate to burrow into and conceal themselves. Each pillbug was provided mixed frozen vegetables (carrot, peas, corn kernels and potatoes) weekly as food. The potato and carrot were in the form of small cubes (7 mm × 7 mm × 7 mm), while the pea was whole and the corn in the form of kernels. A pilot experiment indicated the pillbugs readily consumed peas and potato and to a lesser extent the corn kernels and carrots; each animal was provided one-half of a corn kernel, carrot, pea and potato for ease of eating. Visual inspections indicated consumption of the food, but excess food was present at all times; excess food was removed when it became overgrown with mould. The home tanks for each individual were distributed across laboratory bench space in a single small temperature-controlled laboratory. Several wireless weather stations were set up to record the humidity and temperature for each region of bench space location used. The humidity of the laboratory was 52% (range = ±5%) for the entire experiment, and the temperature was kept at a mean of 17.5 °C (range = ±2 °C). A moist soil environment was maintained by wetting the compost daily using a syringe containing 5 ml of deionised water. Animals were left to acclimate to their home tanks for 9 days before we commenced assays. This permitted us to standardise feeding conditions prior to commencement of assays.

Behavioural assays and timeline

Individual behaviour was assayed over the next 3 weeks, beginning 3 September 2019. Each animal was weighed at the commencement of each week (Monday), then behavioural assays occurred for the following 4 days (Tuesday to Friday; see below why we did not weigh animals more frequently). After 4 days of assays, animals were left undisturbed for 2 days at the end of each week to permit recovery from handling (Saturday and Sunday). Every individual's latency to emerge (described below) was assessed in the unexposed and exposed region of its home tank environment once per day, resulting in a total of 8 observations per week and 24 in total (12 exposed and 12 unexposed evaluations for each individual). We measured boldness 24 times per individual allowing us to quantify with precision whether individual predicted mean values were maintained over time and across exposed versus unexposed contexts. Assays were performed in the morning (between 9:00 am and 11:30 am), followed by a 3-h break for animals to recover from handling, then again in the afternoon (between 2:30 pm and 5:00 pm). Two testing environments (exposed vs. unexposed) were used in case assays in unexposed was perceived by animal as very low risk and lead to rapid habituation (this was not the case ultimately); tests alternated between morning and afternoon to avoid confounding any effect of time of day with effects due to context (exposed vs unexposed). Trial orders were randomised each day using a random number generator to avoid bias.

To estimate risk taking propensity (boldness), a predator attack was simulated on each individual, causing conglobation, and the latency to fully unroll was recorded. The experimenter (always D. B.) used a pair of tweezers to find the pillbug within the compost. Once found, a predatory "attack" was simulated by brushing past and simultaneously touching (bumping) the carapace of the animal exactly 3 times using the tweezers, which always resulted in conglobation. The individual was then picked up using the tweezers and immediately dropped back into their home tank from a height of 5 cm onto the designated environment (exposed or unexposed). This simulated an investigation and rejection of a potential prey item by a predator, such as a bird or lizard. The time recorded for the individual to unroll was considered once the organism was completely stretched out with their appendages moving in an attempt to escape. If the pillbug was still rolled up after 60 min, the investigation was stopped and a maximum latency of 3600 s was recorded for that individual; this occurred in only 5 instances out of 2400 trials, spread across 5 different individuals. After completion

of assays on each session, compost habitat was established to original condition.

Assessment of state (mass)

Individual pillbugs were weighed (± 0.0001 g) at commencement of the study and then once a week over the 3-week period; resulting in 3 masses for each individual. We did not weigh animals daily due to time constraints on days when we conducted behavioural assays and to avoid excessive handling of the animals as they were already being handled twice daily on those days. We weighed animals at midday, before adding any water to their cage on that day. The procedure had the animals out of their cage habitat for about 2 h before being returned to their home tanks. For the purpose of statistical analysis, the mass determined at the start of each week was assumed to be constant for the subsequent 4 days of behavioural sampling. There was 18-fold variation in mass among individuals, ranging from 0.0104 to 0.1808 g (mean = 0.07 g).

Mass increased over time in almost every individual, even in the absence of moulting, indicating increases in stored reserves. This was confirmed by the fact that at least 42% were confirmed to have moulted during the course of behavioural sampling. Individuals found to be moulting or recently moulted were not assayed on that day to avoid injury to the pillbug, leading to missing observations in the data set.

At the end of the experiment, sex was determined as indicated in Fig. 1. Females have a distinct clear line of exoskeleton down the middle of their abdomen, whereas males exhibit a penis-like structure at the top of the abdomen extending down (Wright 1997). Pillbugs were then humanely killed by first placing all animals into micro test tubes (1 ml) with a drop of water to prevent drying out; tubes were then refrigerated at 5 °C for 4 h, then transferred to a
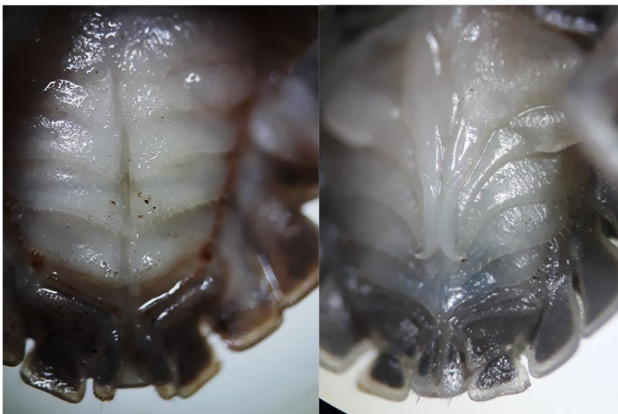

Fig. 1 A peek under the pillbug bonnet: photographs of two representative individuals showing genital features. Left is a female and right a male. Photos credit to PAB

(−40 °C) freezer for 24 h. Animals were killed as they are an invasive pest species and to facilitate sex determination.

Of the 100 animals used, 53 were female and 47 male. A total of 17 animals died during the course of the experiment. Many of these were found dead and desiccated in the morning, located in the open portion of the tank, suggesting it had attempted to climb out of the tank and fell over on its back and was unable to right itself on the smooth surface; this was observed on several occasions where we were able to intervene. Mortality may also have been due to injury when performing latency assays if done just after moulting as it was not always obvious moulting had occurred recently; if obvious, we did not test the animal that day, leading to several missing observations in the data set. We tested for an effect of death during the experiment on latency to emerge values prior to death, but found none (see results). In total, we used data containing 24 observations of latency to unroll from 100 individuals (total $n = 2400$, less those that died or moulted).

Statistical analysis

Our approach was to first fit univariate models on latency to emerge and mass and use those results to inform model structure for the much more complex multivariate (MV) mixed models. In particular, the aim was to determine which contextual random slope effects were significant, and thus supported for inclusion into MV models, in an effort to reduce model complexity which can become problematic for model fitting and especially for interpretation. All fixed effects that were included in MV models were also included in the univariate models, and random intercept effects were included in all models.

Exploratory univariate models were implemented using lme4 (Bates et al. 2015) in the R environment. Day and context were both centred on the mean so that individual predicted values (from the random intercept effect) represented the predicted mean for each individual at the mid-point of the 18-day assay timeline, and average of the two contexts. With respect to latency to emerge, we included temperature, day and context as random slope effects (and all covariances) and a number of fixed effects described below. The inclusion of temperature as a random slope did not have model support (LR test, $\chi_4^2 = 1.42$, $P = 0.84$); thus, its associated covariances were removed from the model (it was also not supported as a sole random effect in the model: LR test $\chi_2^2 = 0.42$, $P = 0.81$). Context (LR test $\chi_3^2 = 12.47$, $P = 0.006$) and day (LR test $\chi_3^2 = 79.2$, $P < 0.0001$) were significant random slope terms and thus retained for the MV models. With respect to the mass data, day was the only random slope effect considered and was substantial and highly significant (LR test, $\chi_2^2 = 1500$, $P < 0.0001$): this result confirmed individual variation in short-term growth and thus was retained

for inclusion into the MV model. We ran these univariate models with and without individuals that died during the experiment, and this made no substantial difference to the random effects variances.

Next, we fit two different but complimentary bivariate models which quantified the among- and within-individual covariances between latency and mass, implemented using the R package *brms* (Bürkner 2017). First, we fit a bivariate model containing random intercept and slope effects on both latency and mass with respect to time and random slopes of latency with respect to context. We assessed all among-individual covariances of these effects (i.e. a 5×5 unstructured variance–covariance matrix; see 'Reaction Norm Model' which is found in Supplement 1 for annotated model structure and code). We were particularly interested in the covariance between individual predicted mean values of each trait (the intercept-intercept covariance), as predicted by the asset protection model. We were also interested in the covariance between random slope effects for day on each trait (the within-individual covariance, specifically the covariance among slopes between traits), because it evaluates whether individuals that increase in mass over time also increase in latency (i.e. become shyer). Because we did not weigh individuals on each sampling occasion for reasons outlined above, we did not fit a residual covariance.

We note here that there are no assumptions or statistical problems created by the disparity in sample size between traits, and it is common practice to relate a labile trait to even a single point estimate using MV mixed models (Houslay and Wilson 2017). Furthermore, mass is estimated with precision and is not a labile trait (in comparison to behaviour), with trait repeatability of $R = 0.97$ in this study (see 'Results'), requiring far less sampling to estimate parameters with precision, as proven by simulation (Adolph and Hardin 2007; Wolak et al. 2012) and evidenced by small SEs on estimates (see Fig. 4). On a practical note, for each individual, there are 21 out of 24 rows without data. The measures of mass do not need to be 'aligned' with individual observations of behaviour, but only with any row that has the right combination of individual and week (= burst) identifiers. This is because the covariance is only estimated at the among-individual and among-burst level (i.e. there is no residual covariance). To implement this, we tell the model to use a subset of the observations (using 'subset' command) for the trait mass, which removes all the rows with missing data from being read into the model (see Supplementary information for code).

Our second and complimentary bivariate analysis adopted a 'character state' approach. Here, each week of samples (i.e. a 'burst' of samples) is viewed as a separate but potentially correlated 'trait' (Roff 1992). This allowed us to evaluate the among-individual correlations of predicted mean values across weeks to assess temporal consistency and to assess correlations between individual predicted mean behaviour in a given week to its corresponding mass for that week. Treating behaviour, mass and their correlation in this way does not make any assumptions about how behaviour may change across weeks (no linearity assumption) and permits us to evaluate whether and how behaviour is related to state at the beginning, middle and end of the experiment. A downside to this extra analysis is that the model becomes very complex and parameter heavy, and so we relegate most those results to the Supplement, highlighting only a few key correlations of interest. A random intercept was fit for each burst and context combination for latency (i.e. burst_a/unexposed, burst_a/exposed and so on), and a random intercept for burst for mass measurements, yielding a 9×9 covariance matrix. For mass, the residual standard deviation was fixed to 0.05, as there were no repeated measures within a burst (the Gaussian model requires a tiny inconsequential residual error in order to perform). These random intercepts are assumed to follow a multivariate normal distribution, with an unstructured variance–covariance matrix (i.e. evaluating all correlations of the 9×9 matrix). Consistent with the random slopes model, we also fit the same fixed effects. See 'Character-state Model' which is found in Supplement 1 for annotated model structure and code.

Latency values were $\log_{10}$-transformed to achieve normality within and among individuals and stabilise variance. Mass residuals were nearly symmetric on the raw data, but square root transformation improved normality of residuals and intercepts; normality is an assumption for the residuals as is familiar to most readers, but it is also an assumption for the predicted intercepts and predicted slopes for individuals (Zuur et al. 2009). Latency and mass data were additionally 'z-transformed' to yield a mean zero and variance of one to aid in convergence and prior specification. We retained a fixed effect for animals that were alive for the whole experiment versus those that died at some point to test for any differences in their latency or mass. All predictors were mean centred, observation day (1–18), time of day (am, pm), context, temperature, humidity, sex and death (animal died during experiment or not); thus, the random intercept variance represents among-individual differences in predicted mean values in the average context and mid-point in time during the longitudinal sampling. Fixed effects and correlation parameters with credible intervals not overlapping zero were considered 'significant' (see also Supplementary information for prior specifications and full model output).

Results

Latency to emerge

At the level of the average individual, latency to emerge increased across days suggesting sensitization, with

animals becoming more shy over time (Est = 0.019, 95% CRI = 0.009–0.030; red line in Fig. 2a). On average, latency increased from unexposed to exposed contexts as expected, with animals being more shy when tested on the bare plastic floor in the 'exposed' context (Est = 0.156, 95% CRI = 0.079–0.233; red line in Fig. 2b), and latencies also increased in the afternoon (Est = 0.12, 95% CRI = 0.043–0.196). All other predictors had no significant effect (Suppl. 1, Reaction Norm Analysis unless indicated otherwise).

Individuals substantially differed from one another in average latency to emerge at the mid-point of the experiment (intercept sd = 0.677, 95% CRI = 0.582–0.788) and also differed somewhat in terms of their temporal patterns of activity across days (Fig. 2a, slope sd = 0.039, 95% CRI = 0.030–0.049; Table 1, see also Suppl. 1). There was some indication of a relationship between individual intercepts (mean centred time) and slopes over time that would indicate a fanning out of reaction norms, but this was modest and uncertain (corr = 0.222, 95% CRI = -0.019–0.451; Fig. 2a). The character state model indicated moderate to high correlations between the individual predicted values in week 1 of sampling compared to weeks 2 and 3 ('burst' a, b, and c in the output) within a given context, which ranged from 0.569 to 0.864; see Suppl. 1 Character State

Analysis); additionally, among-individual variation in each week and context did not vary substantially or significantly (unexposed context, Est = 0.67 week 1, 0.68 week 2, 0.68 week 3; exposed context, Est = 0.60 week 1, 0.65 week 2, 0.68 week 3; all CIs overlapping heavily with each other, See Supplement).

A plot of the raw data for four individuals displaying a range of mean latencies, including two with the lowest and highest mean latencies, is shown to illustrate the very large range in mean latencies observed (ca. 100-fold comparing individual 1 and 46) and also the within-individual and day to day variation observed in the raw data (Fig. 3).

Although there was significant variance among individuals in their responses to changing test context (sd = 0.226, 95% CRI = 0.093–0.336), this effect was modest, and individual reaction norms were essentially parallel, indicating maintenance of rank order differences across contexts (Fig. 2b). There was no relationship between individual intercepts and slopes with respect to context (corr = −0.177, 95% CRI = −0.56 to 0.22; see Table 1). The character state model also confirmed consistency of individual means, whereby the among-individual correlations from one context to the other were high within each week (week 1 corr = 0.720, week 2 corr = 0.816, week 3 corr = 0.859; see Supplement 1). Trait repeatability of latency, estimated at

Fig. 2 Predicted latency to emerge for 100 individual pillbugs measured **a** repeatedly over time and **b** across unexposed and exposed (bare half of tank) contexts. Each black line represents the model predicted individual trendline (reaction norm); the red lines indicate the mean level temporal and contextual trendline, respectively
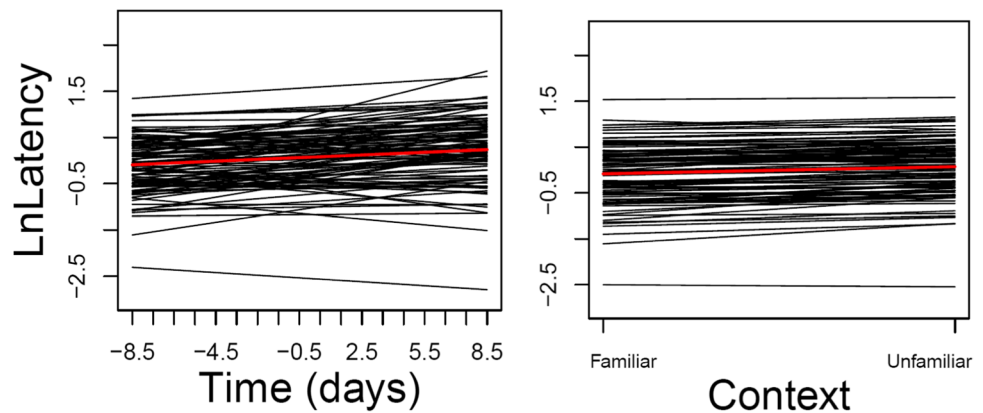
Table 1 Matrix of random effect SDs and covariances among traits from a bivariate mixed effects model containing random intercept and slope effects on each trait. Random effect SDs are shown on diagonal for each trait and effect, and all possible covariances among traits and effects on off-diagonal, expressed as correlations. Credible intervals for each parameter estimate is given in main text, and also in Supplement 1 where full model code and output is found

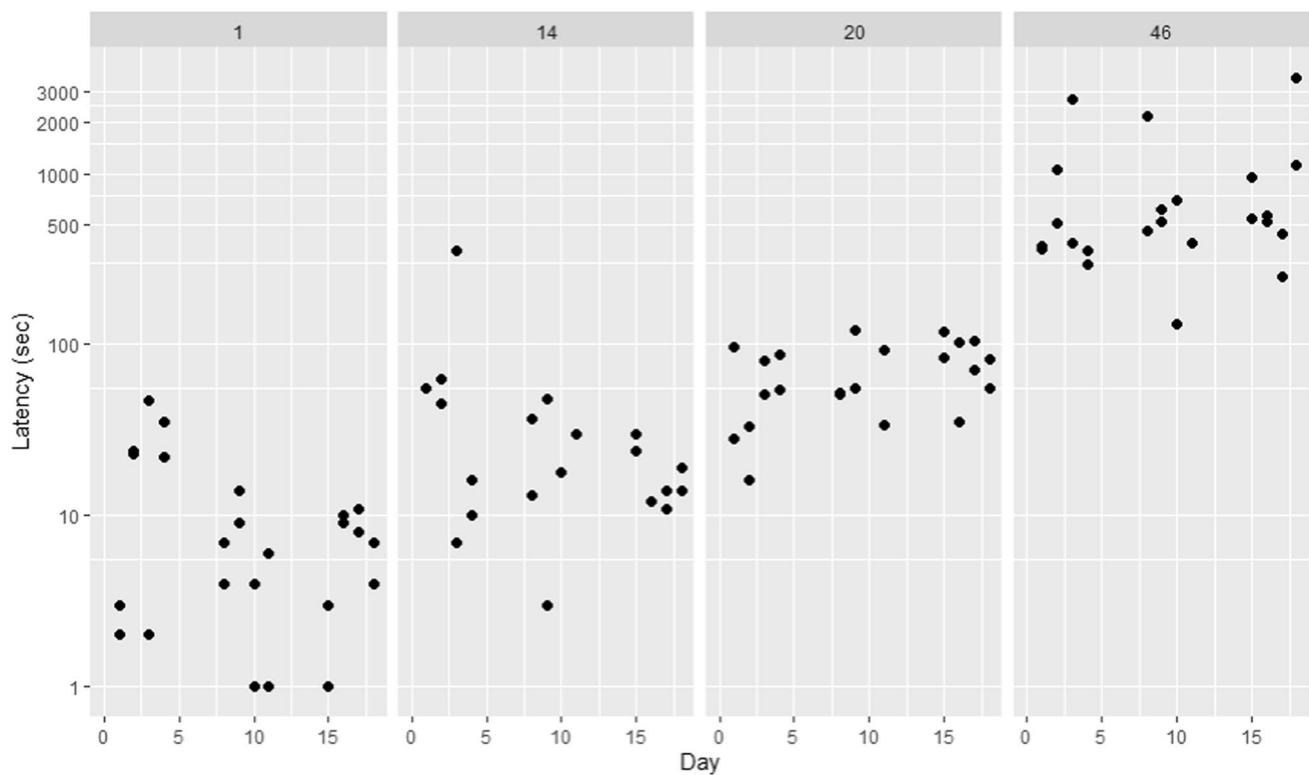| Trait		Latency			Mass	
		Intercept	Slope—time	Slope—context	Intercept	Slope—time
Latency						
	Intercept	0.677				
	Slope—time	0.222	0.039			
	Slope—context	−0.093	−0.177	0.226		
Mass	Intercept	−0.086	0.084	0.168	0.999	
	Slope—time	0.322	0.121	−0.164	−0.265	0.012

Fig. 3 Raw data plots illustrating the longitudinal sampling of behaviour. Shown are four individual pillbugs with mean latencies that span some of the shortest (bold) and longest (shy) latencies observed. Note the two samples taken each day, 4 days per week, across a total of 3 weeks (where each 4 days per 'week' is referred to as a burst of sampling). Data are plotted on a $\log_{10}$ scale, which is the transformation used in our analyses

centred time and context using the reaction norm model was $R = 0.48$ (95% CRI = 0.41–0.57).

Mass

Mass increased on average across days (Est = 0.016, 0.011–0.021), and males were smaller than females (Est = − 0.460, 95% CRI = − 0.856 to − 0.061; Suppl. 1). On the raw scale, there was 18-fold variation in mass among individuals, ranging from 0.0104 to 0.1808 g (mean = 0.07 g). Therefore, due to this deliberate sampling, the majority of mass variance was among individuals (sd = 0.999, 95% CRI = 0.863–1.163), and modest variation in growth (individual slope variation) was small but significant (sd = 0.012, 95% CRI = 0.002–0.020; Table 1). Importantly, there was no correlation between individual intercepts (predicted mean mass) and slopes (short-term growth) indicating that large individuals were not necessarily growing slower due to ageing or senescence (corr = − 0.265, 95% CRI = − 0.653 to 0.165; see Table 1, also Suppl. 1). Not surprisingly, trait repeatability of mass estimated at centred time was $R = 0.97$ (95% CRI = 0.96–0.98). Thus, with substantial variation in both behaviour, state and modest

within-individual changes in state, covariances between them are possible which we report next.

Covariance between latency and mass

We found no evidence for an among-individual correlation between average latency and average mass (intercept-intercept correlation; corr = − 0.086, 95% CRI = − 0.281 and 0.116, Fig. 4a); there was also no evidence that individual changes in latency over time were correlated with individual changes in mass over time (slope-slope covariance; corr = 0.121, 95% CRI = − 0.340 to 0.564; Table 1). Other potential correlations between individual intercepts and slopes with respect to mass, and aspects of contextual responsiveness and temporal trends in boldness, were all not significant (Table 1, see also Suppl. 1); the absence of correlations among these individual attributes was not surprising given only slight variation in temporal plasticity and near-zero variation in contextual plasticity (see above, also Suppl. 1). This result was mirrored in the character state model, with week- and context-specific predicted values for individuals showing no correlations whatsoever with their mass at that time (all CRIs centred about zero, Suppl. 1). However, there was an indication of a possible correlation

Fig. 4 Among-individual relationships between latency to unroll and **a** mean mass and **b** short-term growth in mass. Dots represent model predicted mean values and error bars the SE of those estimates
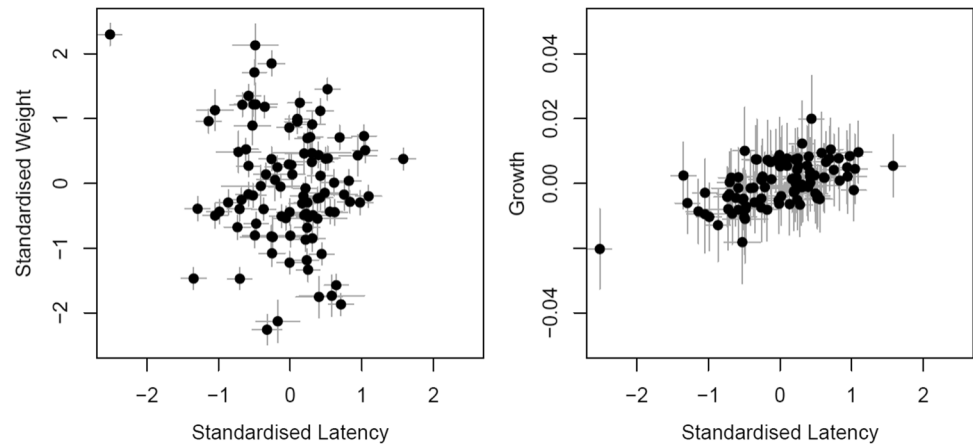
between mean behaviour and changes in mass, whereby individuals with longer latencies to emerge on average may have had relatively greater increases in mass over time (intercept-slope correlation; $r = 0.322$, SE $= 0.203$, 95% CRI $= -0.095$ to 0.704, Fig. 4b, Suppl. 1).

Discussion

Individual behaviour and state could be linked through a variety of biological processes. One important aspect of state is body mass, and it is thought that mass and behaviour may be linked through the asset protection principle with a main prediction being that larger/heavier individuals should be more risk averse with the aim to protect accumulated assets (Clark 1994). Here, we tested for evidence of state dependence in the behaviour of the pillbug and found that (a) individual mean boldness (latency to unroll) was unrelated to their mean mass, but (b) individuals that gained more mass during the course of the study were shy (longer latency to unroll), and (c) a mean level trend of increasing shyness over time was consistent with mean level mass increases over time. Together, our results provide evidence for the asset protection hypothesis which states that individuals with more accumulated resources in the form of mass or energy should aim to protect them and be risk averse.

The lack of any relationship between individual mean boldness and mean mass in this study ($r = -0.086$) was observed despite large sample sizes and high statistical power—allowing us to reject correlations larger than $r = -0.28$ (the lower bound of the credible interval; Suppl. 1). In addition to our large sample size ($n = 100$ animals, and 24 repeated measures per animal), we observed substantial variation in latency to emerge and in mass, and so correlation between these traits was certainly possible but was not found. However, we did find that individuals that increased more in mass during the study may also have tended to be shy (among individual correlation, but highly uncertain),

and that on average increasing latency over time was coincident with increases in mass over time on average—both of which are consistent with predictions for asset protection.

Reasons for why we observed correlations between behaviour and within-individual changes in state, but not overall average state, are unclear. One possibility is a mismatch in the timing of how state and behaviour are linked, with different measures of state confusing matters. For instance, mean mass may not be the most pertinent measure of current state if it mostly reflects past events influencing growth over extended periods of time in the field prior to collection, and if so we should not expect correlations between this measure of state and current behaviour. Another potential reason for why we detected weak within- but not among-individual correlations between behaviour and state may be due to the fact that sampling pillbugs of a wide range of sizes from the field has introduced not only size effects, but also unknown age effects. For instance, our samples likely contain animals of similar size but different age, and so our sample may contain individuals that are large in size but were very slow growing (and therefore older), or contain senescent individuals. However, it seems we can exclude this possibility because we did not observe any correlation between individual mean mass and their short-term growth rate (given by the intercept-slope covariance). By contrast, short-term increases in mass observed during the study do reflect current improvements in state that is measured concurrently with behaviour, and those improving more are also shy; this was also reflected in the mean level increases in both shyness and mass over time.

Previous studies which have found links between behaviour and mass included species where size and predation risk are related (Dewitt et al. 1999) and where large size is associated with boldness and dominance (Colléter and Brown 2011). Yet, other studies find no association between size and behaviour (Harris et al. 2010; Royauté et al. 2015; Underhill et al. 2021). Furthermore, a meta-analysis demonstrates more broadly that correlations

between mass/size and behaviour at the among individual level show no general trend for positive or negative correlations, and the correlations on average are near zero (Niemelä and Dingemanse 2018). As already discussed in the 'Introduction', small sample sizes and few or no repeated measures may tend to bias studies towards a nil result and account for lack of correlations generally in the literature.

In addition to our results on state-behaviour correlations, our study also provides evidence for consistency of individual differences over time and across contexts. It is noteworthy because in this study we were able to rigorously partition temporal and contextual plasticity both experimentally and statistically, and show that individual predicted means are somewhat consistent over time (temporal slope variance low, week to week correlations high) and across contexts (near-parallel reaction norms). This is subtly, but importantly, different from estimation of repeatability (R) of behaviour, because R provides a measure of the relative within-individual correlation of individual scores, not the consistency of predicted means over time or rank order stability, though they are related (Biro and Stamps 2015). The variation in temporal trend-lines among individuals (Fig. 2) can be expressed as a correlation between predicted mean values estimated during each week, as we did using the complimentary character state model; this generated among-individual correlations between successive 3 weeks of sampling (three 'bursts' of sampling) that ranged between 0.6 and 0.9, with the higher correlations occurring between weeks 2 and 3. Cross-context correlations between predicted values within a given week ranged from $r = 0.72$ in week 1, $r = 0.82$ in week 2 and $r = 0.86$ in week 3. Our results stand in contrast to many studies showing evidence of acclimation at the mean level and substantial among-individual differences in acclimation in short-term studies such as ours, and also in contrast with many studies showing mean responses to changing contexts and among individual differences in responses to changing contexts (Westneat et al. 2011; Mathot et al. 2012; Forsman 2015; Saltz et al. 2017; Stamps et al. 2018).

In conclusion, our study has provided evidence for consistent individual differences in behaviour over time and across contexts, but weak evidence of state-dependent individual behaviour supporting asset protection whereby improvements in state were associated with greater shyness (increased latency to unroll). Future studies might improve upon our study by examining state-dependent behaviour using individuals reared on similar diets to test whether innate differences in state affect behaviour, and/or manipulations of individual state where behaviour is measured before and after the manipulation to examine changes in associations between state and behaviour at the mean-, among- and within-individual levels.

Declarations

Ethics Not required.

Conflict of interest The authors declare no competing interests.

References

Adolph S, Hardin J (2007) Estimating phenotypic correlations: correcting for bias due to intraindividual variability. Funct Ecol 21:178–184

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear mixed-effects models using lme4 67:48

Beckmann C, Biro PA (2013) On the validity of a single (boldness) assay in personality research. Ethol 119:937–947

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783

Biro PA, Adriaenssens B, Sampson P (2014) Individual and sex-specific differences in intrinsic growth rate covary with consistent individual differences in behaviour. J Anim Ecol 83:1186–1195

Biro PA, Stamps JA (2008) Are animal personality traits linked to life-history productivity? Trends Ecol Evol 23:361–368

Biro PA, Stamps JA (2010) Do consistent individual differences in metabolic rate promote consistent individual differences in behavior? Trends Ecol Evol 25:653–659

Biro PA, Stamps JA (2015) Using repeatability to study physiological and behavioural traits: ignore time-related change at your peril. Anim Behav 105:223–230

Brommer JE (2013) On between-individual and residual (co)variances in the study of animal personality: are you willing to take the "individual gambit"? Behav Ecol Sociobiol 67:1027–1032

Bürkner P-C (2017) brms: an R package for Bayesian multilevel models using Stan. J S Soft 80:1–28

Clark CW (1994) Antipredator behavior and the asset-protection principle. Behav Ecol 5:159–170

Colléter M, Brown C (2011) Personality traits predict hierarchy rank in male rainbowfish social groups. Anim Behav 81:1231–1237

Cornwell TO, McCarthy ID, Biro PA (2020) Integration of physiology, behaviour and life history traits: personality and pace of life in a marine gastropod. Anim Behav 163:155–162

Dewitt TJ, Sih A, Hucko JA (1999) Trait compensation and cospecialization in a freshwater snail: size, shape and antipredator behaviour. Anim Behav 58:397–407

Dochtermann NA, Schwab T, Anderson Berdal M, Dalos J, Royauté R (2019) The Heritability of Behavior: A Meta-analysis. J Hered 110:403–410

Fanson KV, Biro PA (2018) Meta-analytic insights into factors influencing the repeatability of hormone levels in agricultural, ecological, and medical fields. Am J Phys-Reg Int Comp Phys 316:R101–R109

Forsman A (2015) Rethinking phenotypic plasticity and its consequences for individuals, populations and species. Heredity 115:276–284

Harris S, Ramnarine IW, Smith HG, Pettersson LB (2010) Picking personalities apart: estimating the influence of predation, sex and body size on boldness in the guppy Poecilia reticulata. Oikos 119:1711–1718

Healy K, Ezard THG, Jones OR, Salguero-Gómez R, Buckley YM (2019) Animal life history is shaped by the pace of life and the distribution of age-specific mortality and reproduction. Nat Ecol Evol 3:1217–1224

Houslay TM, Wilson AJ (2017) Avoiding the misuse of BLUP in behavioural ecology. Behav Ecol 28:948–952

Luttbeg B, Sih A (2010) Risk, resources and state-dependent adaptive behavioural syndromes. Phil Trans Roy Soc B 365:3977–3990

Mathot KJ, Wright J, Kempenaers B, Dingemanse NJ (2012) Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. Oikos 121:1009–1020

Nespolo RF, Franco M (2007) Whole-animal metabolic rate is a repeatable trait: a meta-analysis. J Exp Biol 210:2000–2005

Niemelä PT, Dingemanse NJ (2018) Meta-analysis reveals weak associations between intrinsic state and personality. Proc Roy Soc B 285:20172823

Roff DA (1992) The evolution of life histories. Chapman and Hall, New York

Royauté R, Berdal MA, Garrison CR, Dochtermann NA (2018) Paceless life? A meta-analysis of the pace-of-life syndrome hypothesis. Behav Ecol Sociobiol 72:64

Royauté R, Greenlee K, Baldwin M, Dochtermann NA (2015) Behaviour, metabolism and size: phenotypic modularity or integration in Acheta domesticus? Anim Behav 110:163–169

Saltz JB, Lymer S, Gabrielian J, Nuzhdin SV (2017) Genetic Correlations among Developmental and Contextual Behavioral Plasticity in Drosophila melanogaster. Am Nat 190:61–72

Sih A, Mathot KJ, Moirón M, Montiglio P-O, Wolf M, Dingemanse NJ (2015) Animal personality and state–behaviour feedbacks: a review and guide for empiricists. Trends Ecol Evol 30:50–60

Smigel JT, Gibbs AG (2008) Conglobation in the pill bug, Armadillidium vulgare, as a water conservation mechanism. J Insect Sci 8:1–9

Stamps JA, Biro PA, Mitchell DJ, Saltz JB (2018) Bayesian updating during development predicts genotypic differences in plasticity. Evolution 72:2167–2180

Underhill V, Pandelis GG, Papuga J, Sabol AC, Rife A, Rubi T, Hoffman SMG, Dantzer B (2021) Personality and behavioral syndromes in two Peromyscus species: presence, lack of state dependence, and lack of association with home range size. Behav Ecol Sociobiol 75:9–16

Videlier M, Rundle HD, Careau V (2019) Sex-specific among-individual covariation in locomotor activity and resting metabolic rate in Drosophila melanogaster. Am Nat 194:E164–E176

Westneat DF, Hatch MI, Wetzel DP, Ensminger AL (2011) Individual variation in parental care reaction norms: integration of personality and plasticity. Am Nat 178:652–667

White CR, Schimpf NG, Cassey P (2013) The repeatability of metabolic rate declines with time. J Exp Biol 216:1763–1765

Wolak ME, Fairbairn DJ, Paulsen YR (2012) Guidelines for estimating repeatability. Meth Ecol Evol 3:129–137

Wright DJ (1997) PILLBUGS. In: Nels H. Troelstrup J (ed). South Dakota Department of Game, Fish and Parks, Division of Wildlife, Pierre, SD., Northern State University

Zuur A, Ieno EN, Walker N, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.