2023-07

Boldness is not associated with dynamic performance capacity in hermit crabs.

Briffa, M

https://pearl.plymouth.ac.uk/handle/10026.1/21022

1 **Boldness is not associated with dynamic performance capacity in hermit crabs.**

2

3 Winne Courtene-Jones

4 Mark Briffa

5

6 School of Biological and Marine Sciences,

7 University of Plymouth,

8 Drake Circus,

9 PL4 8AA.

10 U.K.

11

12 Author for correspondence: mark.briffa@plymouth.ac.uk

14

15

16 Boldness, the way an individual reacts to risk, is a commonly studied personality trait in
17 animals. Consistent among-individual differences in startle response durations (latency to
18 recover from a startling stimulus) are frequently assumed to reflect variation in boldness. An
19 alternative explanation is that these latencies are not directly driven by variation in responses
20 to information on risk, but by underlying differences in dynamic performance capacities. Here
21 we investigate this possibility by analysing relations between locomotory speed, a measure of
22 whole-body dynamic performance capacity in hermit crabs, and startle response duration, a
23 repeatable latency measure used as an index of boldness. Individuals differed in mean startle
24 response duration, in the consistency of their startle responses, in their reaction norms across
25 repeated observations, and mean startle responses increased with crab mass. However, there
26 were no relations between startle responses and locomotory speed. This indicates that startle
27 responses do not reflect underlying performance capacities and suggests that they provide
28 insight into differences in how individuals respond to risky situations. Since similar latencies
29 are used as measures of boldness in other animals, we suggest that potential relations between
30 apparent boldness and performance capacity should be tested.

31

32

33

34

35

INTRODUCTION

Animal personality in its narrowest sense refers to the presence of significant differences among individuals in their behavioural means derived from repeated observations, typically quantified by repeatability and its confidence intervals. Formally, repeatability, is the proportion of variance due to differences between individuals (between individual variance; V_{BA}) out of the total variance including that within individuals (V_{WA}), the latter also called unpredictability [1], predictability [2], residual [3] or intra-individual behavioural variance [4]. Within-individual variance itself can differ among individuals and thus can be considered another aspect of animal personality, along with differences between individuals in how their behaviour changes plastically over time or across gradients, where individuals show different behavioural reaction norms [5]. Demonstrating the presence of narrow sense animal personality does not in itself provide insight into its underlying proximate or evolutionary causes [6] but the terms used to describe behaviours, once demonstrated as repeatable, can carry connotations of particular causal factors [7]. For instance, of the five major axes of animal personality, as adapted from human psychology [8], 'boldness' has received particular attention. In lay-terms boldness equates to 'confidence' and biologically it has been defined as "an individual's reaction to any risky situation" [8]. Thus boldness, as generally understood, implies differences in how individuals react to information on risk, i.e. their 'risk-coping strategy'. When confronted with a potentially dangerous event animals may either flee or hide, which includes adopting a protective posture. In this latter case, bolder individuals re-emerge from hiding, and resume ongoing activity, more quickly compared with shyer individuals at the opposite end of a bold-shy continuum. An untested alternative to different risk-coping strategies is that among-individual differences in apparent boldness, instead primarily result from variation in dynamic performance capacities, defined as the ability to perform sustained and rapid movement [9]. The latter could be indirectly compatible with differences in risk-coping, if boldness covaries with performance capacities as suggested by the pace of life syndrome hypothesis (POLS) [7, 10] for example. However, risk-coping and performance capacities do not necessarily covary and performance capacities could directly drive consistent differences in apparent boldness if they constrain reaction times. Testing for a link between putative measures of boldness and dynamic performance is therefore a potentially important step in interpreting apparent boldness, particularly when measures used as indices of boldness are based on behavioural latencies that could be influenced by speed of movement.

A commonly used index of boldness is latency to recover from a startling stimulus, e.g. resumption of ongoing activities after a protective posture has been adopted [1, 11] or re-emergence from a shelter [4, 12-13]. European hermit crabs, *Pagurus bernhardus*, show both fleeing (from a visual cue [14]) and repeatable startle response durations when handled directly [12], measured as time spent tightly retracted into their empty gastropod shell, which hermit crabs use as portable shelters when threatened. This otherwise repeatable behaviour shows plasticity across gradients of risk, lasting longer and [12] and being less predictable [2] in the presence of a predator. Collectively, these results are compatible with underlying differences in risk-coping strategies, but they do not rule out the possibility that they are driven (or driven in part) by underlying differences in performance capacity, particularly as startle responses correlate with metabolic rate [15, 16]. Locomotion speed (measured as time taken to cover a set distance) has been validated an energetically significant activity [17] and a measure of dynamic performance capacity in a previous study on *P. bernhardus*, where slower moving

individuals also performed energetically demanding 'shell rapping' (which engages the same abdominal muscles used during startle responses) more slowly during subsequent agonistic contests [18]. While the role of abdominal musculature during locomotion remains to be fully elucidated it is likely that these muscles are engaged in order to adjust shell position, so as to maintain a posture where the shell will not interfere with locomotion [17]. Here we ask whether startle response durations are associated with this measure of dynamic performance capacity. If the speed of re-emergence is directly constrained by dynamic performance capacity, or if boldness covaries with performance as predicted by POLS, we should see a positive correlation between mean startle response duration and time taken to walk a set distance (i.e. slower walking individuals hide for longer). A negative correlation (i.e. faster moving individuals hide for longer) is unexpected if latency reflects dynamic performance but could be present for other reasons, for example if investment in high movement speed is part of a wider risk-avoidance syndrome. In this case we would also expect a negative correlation between movement speed and within individual variance in startle response duration, on the assumption that less predictable startle responses mitigate risk [2]. A lack of any correlation in either direction would indicate that startle response durations vary independently of dynamic performance and may instead differ across individuals for other reasons including differences in ability to detect or process information on risk, or differences in sensitivity to such information, i.e. differences in risk-coping strategy .

METHODS

Hermit crabs were hand-collected from Hannafore Point, Cornwall UK between February-July 2013 (see ESM.1). Following transport back to the lab in Plymouth, crabs were held in groups of ~100 individuals in constantly aerated and filtered seawater at 15°C to a depth of 30cm, in a controlled 12:12 h light:dark environment, and fed *ad libitum* on small pieces of white fish. Prior to observation crabs were removed from their shells by carefully cracking the shell open using a bench vice, then sexed and weighed. Females were given new shells and returned to the sea and only males free of obvious injury or parasites were used in the experiment [19]. See ESM section 1 for further details. Male crabs were provided with a new *Littorina littorea* shell 50% of its optimal shell mass, calculated from a previous shell selection experiment, reduced shell size ensuring that continuous locomotion could be stimulated [18]. Males were then individually allocated to white plastic flat-bottomed dishes (20cm diameter), filled with a 5cm depth of aerated 15˚C seawater. Startle responses were evoked by manually lifting a crab out of the seawater, causing it to retract into its shell [20], and replacing the shell, aperture upwards on the base of the dish. The duration of the response was timed using a stopwatch, until the crab re-emerged to the point where its second pair of walking legs contacted the substrate [12, 20]. In 52 males 8 startle responses were recorded per crab, over a period of 4 days, alternating between 16- and 5-hour intervals between successive observations. Following this, each crab was placed into a clear plastic circular raceway (3.25m in outer circumference, 6cm wide raceway) filled to a depth of 5cm with aerated 15˚C seawater. Each crab was stimulated to walk by a series of light taps on its shell using a wooden probe. Once in motion the crab was followed by the probe at its own speed at a distance of one body length. If a crab stopped moving it immediately received an additional light tap, which caused walking to resume throughout the trial, and did not cause crabs to withdraw into their shell. We obtained two measures of locomotory performance: The time taken to cover 13m (4 laps), a measure of

126   overall performance, and the duration of the fastest lap of the four, a measure of maximum
127   exertion [18]. Due to some data exclusions (see ESM section 2) a final sample size of 407
128   startle responses across 51 crabs was obtained.

129   *Statistical methods*

130   To determine whether variance in startle response durations differed among individuals we
131   conducted an initial Levene's test. This revealed the presence of significant among individual
132   differences in V_{WI} (see below), so prior to further analysis we Log_{10} $(1+x)$ transformed the data,
133   which yielded homogeneity of variance across individuals, and improved the normality of
134   residuals in subsequent models. There were no correlations between either measure of dynamic
135   performance capacity and mass (see ESM section 2), so we then ran a linear mixed effects
136   model to determine the effects of crab mass, time taken to complete 4 laps, and observation
137   number, on the duration of startle responses. Time taken to complete the fastest lap was used
138   as a predictor in a further model. Random effects included in the initial model were individual
139   specific intercepts and slopes across observations 1-8, with an assumed correlation between
140   them. The model was first fitted using ML estimation so that it could be compared to alternative
141   models were (a) random intercepts and slopes were uncorrelated and (b) only random intercepts
142   were included, using ΔAICc values, where a more complex model was favoured over a simpler
143   one if its AICc value was at least three points lower. Once the appropiate random effects
144   structure was established we re-ran the model using REML estimation to test the fixed effects,
145   using the Satterthwaite method to calculate degrees of freedom. Finally, we calculated (LMM
146   based) repeatability of startle response duration. See ESM section 3 for details of the code and
147   R packages used.

148

149   **RESULTS**

150   Individuals differed in within-individual variance (Levene's test: $F_{50,356} = 3.2$, $P < 0.0001$) but
151   startle responses were still repeatable ($R = 0.61$ [95% CI = 0.473, 0.699], $P_{LRT} <00001$). The
152   model including correlated random intercepts and slopes provided the better fit for the data
153   compared with the model containing uncorrelated random effects (ΔAICc = 3.7) or random
154   intercepts only (ΔAICc = 16.9), indicating significant variation among individuals in how their
155   startle responses changed across observations (**Figure 1**). Startle response duration did not vary
156   across observations ($F_{1,50.04} = 1.3$, $P = 0.26$) or correlate with time to complete four laps ($F_{1,48}$
157   $= 2.85$, $P = 0.1$; **Figure 2a**) but the duration increased on average with crab mass ($F_{1,48} = 5.26$,
158   $P = 0.026$; **Figure 2b**). Results using time taken to complete the fastest lap were qualitatively
159   identical and are reported in ESM section 2, which also outlines an alternative analytical
160   approach.

161

162   **DISCUSSION**

163   Startle response duration is frequently used as an index of boldness (e.g. [1, 4, 11-14]), and for
164   probing motivational states (e.g. [21, 22]). Here there was no association in either direction
165   between locomotion speed and mean startle responses. Thus, it is unlikely that locomotive
166   performance contributes to repeatable variation in startle response duration in hermit crabs,
167   either as a constraint on dynamic performance capacity, or as an alternative strategy to hiding.

In a previous study [23] we found a link between startle response duration and a potential casual factor for dynamic performance capacity, haemocyanin concentration, the respiratory pigment that determines oxygen carrying capacity. In this case, the correlation was in the opposite direction to the expectation under constraints on performance limiting the speed of re-emergence because startle response duration increased rather than decreased with haemocyanin concentration. Thus, it appeared that individuals in good physiological condition behaved in a more risk averse way rather than being more risk prone, perhaps because greater oxygen storage allowed them to tolerate a longer period of respiratory limitation while withdrawn into the shell [23], or because (as suggested above) investment in haemocyanin is part of a wider risk-mitigation syndrome.

The current null result for dynamic performance provides further evidence to support the view that startle response duration in hermit crabs provides a direct index of boldness, i.e. variation in risk-coping, albeit one where oxygen carrying capacity may contribute to an upper limit of hiding times [23]. While there was no relation with dynamic performance, other patterns in the current analysis are consistent with previous studies of boldness in hermit crabs. Boldness was repeatable [12, 20, 23], it increased with crab mass [24] and individuals differed in variance around their means [2, 4, 25]. Individuals also differed in how their startle responses changed over repeated observation [4]. Visual inspection of individual slopes indicates that some crabs sensitised (i.e. startle response durations increased with observation number), some habituated (i.e. startle response durations decreased with observation number) but for most crabs there was limited overall change in either direction.

Similar patterns are seen in other animal personality studies focussed on boldness and using startle responses or analogous latency measures (e.g. see [4]). We suggest that testing for links between such measures and dynamic performance would clarify the extent to which these patterns represent variation in boldness *per se*. The presence of a correlation where latency decreases as performance capacity increases (note that in the current experiment this would equate to a positive correlation between startle response duration and time taken to complete the locomotory task) could be due to an indirect link between the repeatable behaviour and risk-coping. Alternatively, such a correlation could be present because the repeatable behaviour under test primarily relates to dynamic performance capacity rather than risk-coping. Thus, additional data would then be needed to determine the underpinnings of repeatable latency behaviour. One potential approach would be to collect repeated measures of dynamic performance per individual, ideally time-matched with the collection of repeated startle response data. This would enable between- and within- individual covariation to be distinguished [25] across different conditions of risk exposure (e.g. a predator cue absent and present) and energetic state (which will impact on dynamic performance). Then the relative contribution of each to the means and variances of startle response duration could be assessed. Furthermore, an interaction between risk-level and performance capacity could be tested for. For instance, in systems where dynamic performance capacities do impose constraints on recovery time, such constraints might be greater under low-risk situations where latencies are expected to be relatively short in species that show behavioural plasticity over gradients of risk. In contrast, under high-risk situations with elevated hiding times, we would not expect performance capacity to constrain latency to recover. In the current study though, we found no evidence that startle responses covary with dynamic performance capacity indicating that latency of re-emergence is independent of this measure. This allows us to be less "agnostic"

[6] in our interpretation of repeatable startle response durations and lends more confidence to the assumption that they represent consistent differences in risk-coping strategy, or 'boldness' in terms of how this phrase is generally understood.

REFERENCES

1. Maskrey DK, Sneddon LU, Arnold KE, Wolfenden DCC, Thomson JS. 2021. Temperature-driven changes in behavioural unpredictability and personality in the beadlet sea anemone, *Actinia equina. Anim. Behav.* **181**, 13-27.
2. Briffa M. 2013. Plastic proteans: reduced predictability in the face of predation risk in hermit crabs. *Biol. Letts.* **9**, 20130592-20130592.
3. Westneat DF, Wright J, Dingemanse NJ. 2014. The biology hidden inside residual within-individual phenotypic variation. *Biol. Rev.* **90**, 729-743.
4. Stamps JA, Briffa M, Biro PA. 2012 Unpredictable animals: individual differences in intraindividual variability (IIV). *Anim. Behav.* **83**, 1325-1334.
5. Dingemanse NJ, Kazem AJN, Réale D, Wright J. 2010. Behavioural reaction norms: animal personality meets individual plasticity. *TREE* **25**, 81-89.
6. Bell AM. 2017. There is no special sauce: a comment on Beekman and Jordan. *Behav. Ecol.* **28**, 626-627.
7. Careau V, Garland T. 2012. Performance, Personality, and Energetics: Correlation, Causation, and Mechanism. *Physiol. Biochem. Zool.* **85**, 543-571.
8. Réale D, Reader SM, Sol D, McDougall PT, Dingemanse NJ. 2007. Integrating animal temperament within ecology and evolution. *Biol. Rev.* **82**, 291-318.
9. Jerry F. Husak, Duncan J. Irschick, Stephen D. McCormick, Ignacio T. Moore. 2009. Hormonal regulation of whole-animal performance: Implications for selection. *Integr. Comp. Biol.* **49**, 349-353.
10. Réale D, Garant D, Humphries MM, Bergeron P, Careau V, Montiglio P-O. 2010. Personality and the emergence of the pace-of-life syndrome concept at the population level. *Phil. Trans. R. Soc. B* **365,** 4051-4063.
11. Briffa M, Greenaway J. 2011. High *In Situ* Repeatability of Behaviour Indicates Animal Personality in the Beadlet Anemone *Actinia equina* (Cnidaria). *PLoS ONE* **6**, e21963-e21963.
12. Briffa M, Rundle SD, Fryer A. 2008. Comparing the strength of behavioural plasticity and consistency across situations: animal personalities in the hermit crab *Pagurus bernhardus. Proc. R. Soc. B* **275**, 1305-11.
13. Dalesman S, Inchley C. (2008). Interaction between olfactory and visual cues affects flight initiation and distance by the hermit crab, *Pagurus bernhardus. Behaviour* **145**, 1479-1492.
14. Brown C, Jones F, Braithwaite V. 2005. *In situ* examination of boldness–shyness traits in the tropical poeciliid, *Brachyraphis episcopi. Anim. Behav.* **70**, 1003-1009.

258   15.   Velasque M, Briffa M. 2016. The opposite effects of routine metabolic rate and
259   metabolic rate during startle responses on variation in the predictability of behaviour in
260   hermit crabs. *Behaviour* **153**, 13-14.
261   16.   Velasque M, Denton JA, Briffa M. 2023. Under the influence of light: How light
262   pollution disrupts personality and metabolism in hermit crabs. *Environ. Pollut.* **316**,
263   120594.
264   17.   Herreid CF, Full RJ. 1986. Energetics of hermit crabs during locomotion: the cost of
265   carrying a shell. *J. Exp. Biol.* **120**: 297–308.
266   18.   Mowles SL, Cotton PA, Briffa M. 2010. Whole-organism performance capacity
267   predicts resource-holding potential in the hermit crab Pagurus bernhardus. *Anim.*
268   *Behav.* **80**, 277-282.
269   19.   Elwood RW, Neil SJ. 1992. *Assessments and decisions: A study of information*
270   *gathering by hermit crabs*. London: Chapman & Hall.
271   20.   Briffa M, Bibost A-L. 2009. Effects of shell size on behavioural consistency and
272   flexibility in hermit crabs. *Can. J. Zool.* **87**,. 597-603.
273   21.   Briffa M, Elwood RW. 2001. Motivational change during shell fights in the hermit crab
274   *Pagurus bernhardus. Anim. Behav.* **62**, 505-510.
275   22.   Moorhouse JE, Fosbrooke IH, Ludlow AR. 1987. Stopping a walking locust with
276   sound: an analysis of variation in behavioural threshold. Exp. Biol. **46**, 193-201.
277   23.   Bridger D, Bonner SJ, Briffa M. 2015. Individual quality and personality: bolder males
278   are less fecund in the hermit crab *Pagurus bernhardus. Proc. R. Soc B* **282**, 20142492.
279   24.   Briffa M, Archer R. 2022. Size specific boldness associated with differences in resource
280   requirements and habitat use: a cross-sectional study in hermit crabs. *Curr. Zool.* p.
281   zoac049.
282   25.   Niemelä PT, Dingemanse NJ. 2018. On the usage of single measurements in
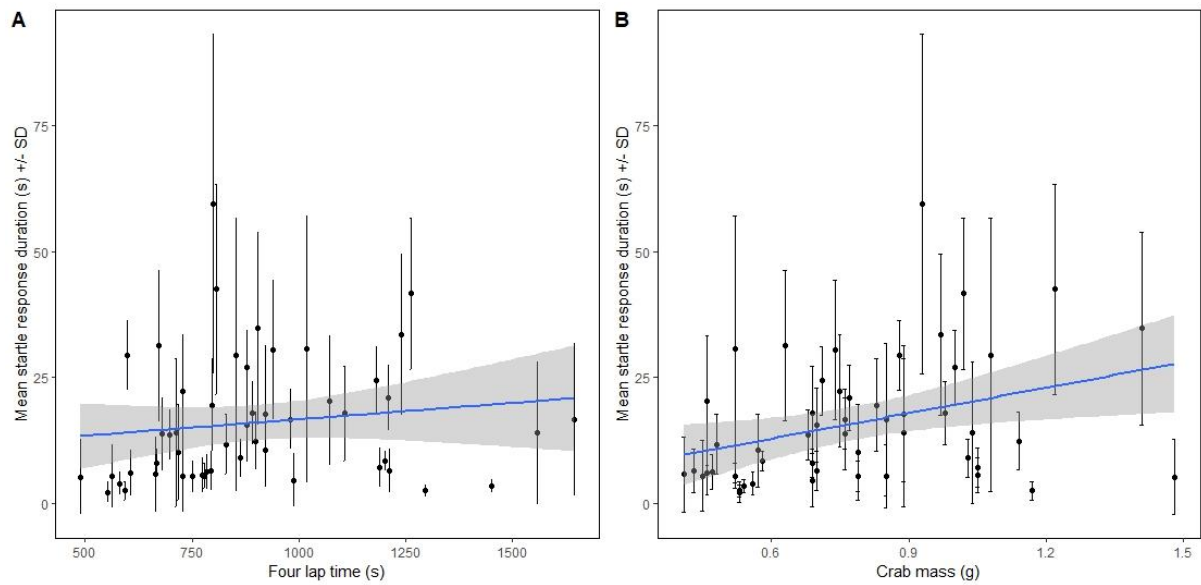283   behavioural ecology research on individual differences. *Anim. Behav.* **145**, 99-105.
284

285

286
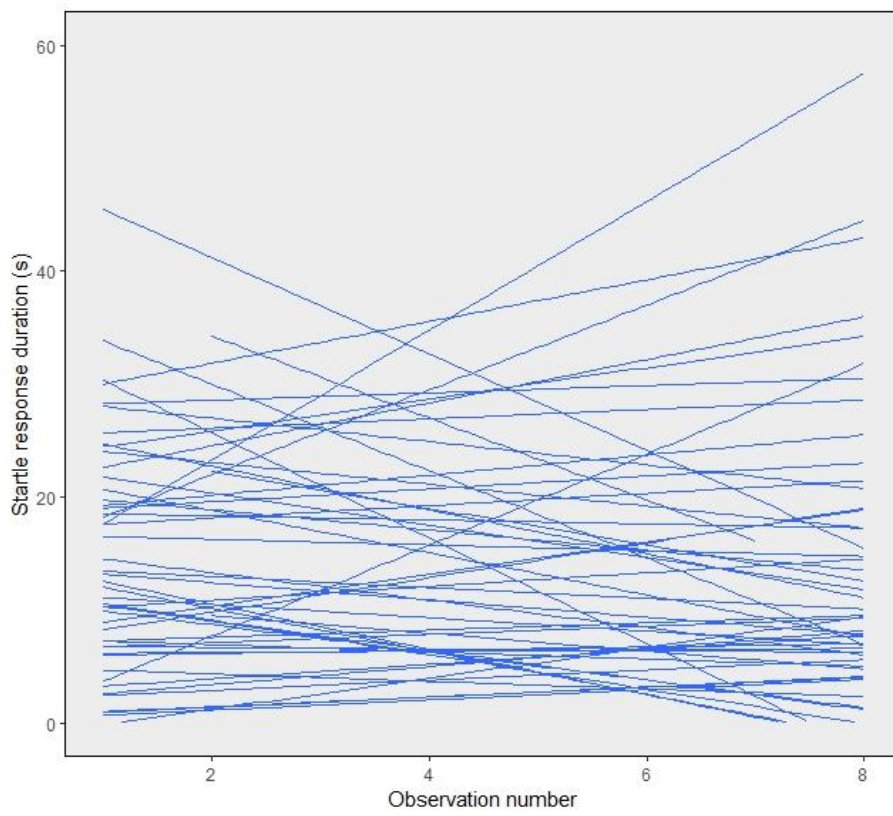
Figure 1: Fitted lines representing individual norms of reaction in startle response duration across repeated observations.

289

290

Figure 2: Individual means (black circles) with standard deviations (black vertical lines) of startle response durations plotted against (A) time taken to walk 13m and (B) crab mass. Blue lines represent OLS linear regressions of mean startle response durations against (A) time taken to complete four laps and (B) crab mass, and shaded areas show the standard error of the regression.

ELECTRONIC SUPPLEMENTARY MATERIAL

298 **Boldness is not associated with dynamic performance capacity in hermit crabs.**

299

300 Winne Courtene-Jones

301 Mark Briffa

302

303

304 **ESM. 1 STRANGE statement**

305 Hermit crabs were collected from all areas of rockpools, which were searched across open
306 water areas, under algae, in cervices and around the periphery of each pool, to reduce the
307 chance of collection biases for certain behavioural types, which may select different
308 microhabitats and be differently trappable, due to unquantified variation in social experience,
309 age, genetic make-up or preferred environment [1]. Although it is not possible to age hermit
310 crabs data included in the analysis contained only contained crabs of a single size class defined
311 by the preferred species of occupied gastropod shell, *Littorina littorea* [2]. As crabs were wild-
312 caught rather than lab-reared we did not consider potential sampling biases due to differences
313 in rearing history.

314

315 **ESM. 2 Data exclusions and supplementary analyses**

316 Initial inspection of the data indicated that one crab was of a different size class to the rest of
317 the sample (mean mass with large individual removed = 0.79g, SD = 0.25g) and because shell
318 preferences and startle responses vary across size class [2] this single large individual (mass =
319 2.66g) was excluded from the analyses. Additionally, one crab produced a single startle
320 response of unusually long duration (392s) compared to the average (mean with long response
321 removed = 16.05s, SD = 16.28s) and Grubb's test confirmed that this was an outlying value (G
322 = 15.17, P < 0.0001). To check whether this outlying value exerted undue leverage on any
323 effects in the LMM, we analysed the data with this data point both included and excluded. The
324 results of both analyses were qualitatively identical (and quantitatively similar) but we chose
325 to report the analysis with the outlying data point excluded in the main Results section (see
326 Table 1), to aid visualisation of the results (in Figure 1). The version of the analysis including
327 the outlier is reported below in Table ESM 1, along with versions of the analysis using time
328 taken to complete the fastest lap, rather than time taken to complete four laps. As mentioned in
329 the main text, prior to these analyses we checked for a correlation between crab mass and the
330 two measures of locomotory performance. To do this we used a pair of linear models where
331 crab mass was the predictor and locomotion was the response. There was no effect of crab mass
332 on the time taken to complete the fastest lap ($F_{1,49} = 0.07$, P = 0.79) or on the time taken to
333 complete four laps ($F_{1,49} = 0.06$, P = 0.81).

334   In previous studies we analysed similar data using a different analysis approach, doubly
335 hierarchical general linear models (DHGLM), implemented using Bayesian techniques. The
336 advantage of these is that the effects of predictors can be simultaneously modelled on the mean

and variance of the response variable, and random effects can be specified for both levels of analysis, in this way streamlining the analysis. Here we were able to implement such a model that achieved convergence but (on the basis of the Raftery-Lewis diagnostic) it was not possible to generate 95% confidence intervals of the posterior means with sufficient accuracy, and therefore the results of this analysis would need to be interpreted cautiously. Nevertheless, the results of this version of the analysis are congruent with those obtained via LMM, reported in the main text. Briefly, there was no effect of dynamic performance or observation number on mean startle response duration but there was a positive association between crab mass and startle response duration. There were no effects of dynamic performance or mass on the variance of startle response durations. It was not possible to include random slopes in mean model (due to convergence issues) but 95% credible intervals of the random intercept in the mean model did not cross zero, congruent with the finding of significant repeatability. Furthermore, 95% credible intervals of the random intercept in the dispersal model did not cross zero, analogous to the significant Levene's test showing heterogenous variance among individuals, reported in the main text. For comparison of results yielded by the two approaches we include below the results of a DHGLM, where time taken to complete 4 laps was used as a predictor, and the single outlying startle response duration was excluded from the data (Table ESM 2). As with the LMMs, further analyses including the outlying data point and using fastest lap time as a predictor produced qualitatively identical results.

Table ESM 1. Supplementary analyses of fixed effects using LMM. (a) LMM using total time to complete 4 laps of the race way as a predictor and including the outlying startle response datapoint in the analysis. (b) LMM using time to complete the quickest single lap of the race way as a predictor and excluding the outlying startle response datapoint from the analysis. (c) LMM using time to complete the quickest single lap of the race way as a predictor and including the outlying startle response datapoint from the analysis. All three supplementary analyses retained the random effects structure of the model reported in the main text.

	F	df	*P*
Model (a)			
Four lap time	3.40	1,48	0.07
Crab mass	6.04	1,48	0.02
Observation	1.83	1,50	0.18
Model (b)			
Fastest lap time	1.84	1,48	0.18
Crab mass	5.11	1,48	0.03
Observation	1.30	1,50	0.23
Model (c)			
Fastest lap time	2.17	1,48	0.15
Crab mass	5.82	1,48	0.02
Observation	1.83	1,50	0.18

Table ESM 2: Posterior summary statistics for each component of a supplementary DHGLM analysis.

	Posterior mean	SD	Lower 95% CI	Upper 95% CI
Mean model				
Fixed effects				
Intercept	1.19	0.51	0.18	2.19
Four lap time	0.06	0.04	-0.02	0.14
Crab mass	0.92	0.42	0.08	1.75
Observation	0.00	0.01	-0.03	0.02
Random effect				
Intercept (Crab ID)	0.70	0.08	0.68	0.88
Dispersion model				
Fixed effects				
Intercept	-1.08	0.50	-2.04	-0.09
Four lap time	-0.04	0.04	-0.12	0.03
Crab mass	0.40	0.42	-0.44	1.19
Random effect				
Intercept (Crab ID)	0.53	0.12	0.31	0.78

ESM 3. Details of LMM analysis

Software used

All analyses were conducted using R 4.2 [3] and using the following packages:

outliers [4]: Grubb's to check for outliers.

lmerTest [5]: Provides df and P for fixed effects in models fitted using lme4.

lme4 [6]: Fitting mixed effects models.

AICcmodavg [7]: Calculating AICc values.

rptR [8]: Estimating LMM based repeatability.

car [9]: Levene's Test.

Code

```
data <-read.csv("DPSR_LM data.csv")
#re-scaling dynamic performance predictor to optimise convergence
four.lap.sc <-data$Four.lap.s/100
fast.lap.sc  <-data$Fastest.lap.s/100
```

```
385

386    newdata<- data.frame(data$ID,data$DURATION, data$Wt, data$Occasion,
387    data$Four.lap.s, four.lap.sc, fast.lap.sc)

388    #(Retains unscaled versions for use in figures)

389

390    #Outlier test and version of data frame excluding it

391    grubbs.test(newdata$data.DURATION)

392

393    newdata.2 <-  newdata[!newdata$data.DURATION > 390,]

394    names (newdata.2)

395

396    #Can 4 lap speed and mass both be included as predictors? (Do they
397    correlate)

398    newdata.3 <-  newdata[!newdata.2$data.Occasion > 1.1,]

399    plot(newdata.3$four.lap.sc,newdata.3$data.Wt)

400    lm1<-lm(newdata.3$four.lap.sc~newdata.3$data.Wt)

401    lm2<-lm(newdata.3$fast.lap.sc~newdata.3$data.Wt)

402    anova(lm1)

403    anova(lm2)

404    # no correlation, they may both be included.

405

406    #Initial random slopes models, Log10(X+1) and raw

407    m1<-lmer(log(1+data.DURATION) ~ four.lap.sc + data.Wt +
408    data.Occasion + (data.Occasion|data.ID), data = newdata.2)

409    m2<-lmer(data.DURATION ~ four.lap.sc + data.Wt + data.Occasion +
410    (data.Occasion|data.ID), data = newdata.2)

411

412

413    #diagnostics log version m1

414    plot(m1)

415    leveneTest(log(1+data.DURATION)~data.ID, data = newdata.2)

416    par(mfrow=c(1,3))

417    eij=residuals(m1)

418    qqnorm(residuals(m1))
```

```
419   hist(eij,main="Histogram of residuals")

420   plot(density(eij),main="Density plot of
421   residuals",ylab="Density",xlab="Residuals")

422   par(mfrow=c(1,1))

423   #Diagnostics raw version m2

424

425   plot(m2)

426   leveneTest(data.DURATION~data.ID, data = newdata.2)

427

428   par(mfrow=c(1,3))

429   eij=residuals(m2)

430   qqnorm(residuals(m2))

431   hist(eij,main="Histogram of residuals")

432   plot(density(eij),main="Density plot of
433   residuals",ylab="Density",xlab="Residuals")

434   par(mfrow=c(1,1))

435

436   ## choose Log10 model

437   #Random slopes v random intercepts

438

439   m1a<-lmer(log(1+data.DURATION) ~ four.lap.sc + data.Wt +
440   data.Occasion + (data.Occasion|data.ID), data = newdata.2, REML = F)

441   m1b<-lmer(log(1+data.DURATION) ~ four.lap.sc + data.Wt +
442   data.Occasion + (1|data.ID) + (0+data.Occasion|data.ID), data =
443   newdata.2, REML = F)

444   m1c<-lmer(log(1+data.DURATION) ~ four.lap.sc + data.Wt +
445   data.Occasion + (1|data.ID), data = newdata.2, REML = F)

446

447   random_slopes_and_intercept_corrleated      <- AICc(m1a)

448   random_slopes_and_intercept_uncorrleated    <- AICc(m1b)

449   random_intercept_only                       <- AICc(m1c)

450

451   AICc <-
452   c(random_slopes_and_intercept_corrleated,random_slopes_and_intercept
453   _uncorrleated,random_intercept_only)
```

```
454  AICc

455  delta.AICc.a.c = random_slopes_and_intercept_corrleated [1] -
456  random_intercept_only [1]

457  delta.AICc.a.b = random_slopes_and_intercept_corrleated [1] -
458  random_slopes_and_intercept_uncorrleated [1]

459  delta.AICc.a.c

460  delta.AICc.a.b

461

462  ## choose random slopes + intercepts

463  anova(m1)

464

465  #Supplementary analyses

466  mSa<-lmer(log(1+data.DURATION) ~ four.lap.sc + data.Wt +
467  data.Occasion + (data.Occasion|data.ID), data = newdata)

468  anova(mSa)

469  mSb<-lmer(log(1+data.DURATION) ~ fast.lap.sc + data.Wt +
470  data.Occasion + (data.Occasion|data.ID), data = newdata.2)

471  anova(mSb)

472  mSc<-lmer(log(1+data.DURATION) ~ fast.lap.sc + data.Wt +
473  data.Occasion + (data.Occasion|data.ID), data = newdata)

474  anova(mSc)

475

476  #Repeatability

477  rpt(log(1+data.DURATION) ~ (1 | data.ID), grname = "data.ID", data =
478  newdata.2, datatype = "Gaussian",

479      nboot = 1000, npermut = 1000)

480  #supplementary version with outlier included

481  rpt(data.DURATION ~ (1 | data.ID), grname = "data.ID", data =
482  newdata.2, datatype = "Gaussian",

483      nboot = 1000, npermut = 1000)

484

485

486

487

488
```

ESM References

1. Hills A, Webster MM. 2022. Sampling biases and reproducibility: experimental design decisions affect behavioural responses in hermit crabs. *Anim. Behav.* **194**, 101-110.

2. Briffa M, Archer R. 2022. Size specific boldness associated with differences in resource requirements and habitat use: a cross-sectional study in hermit crabs. *Curr. Zool.* p. zoac049. https://doi.org/10.1093/cz/zoac049

3. R Core Team . 2023. R: A Language and Environment for Statistical Computing}. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

4. Komsta L. 2022. outliers: Tests for Outliers. R package version 0.15. https://CRAN.R-project.org/package=outlier.

5. Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* **82**, 1-26. doi:10.18637/jss.v082.i13

6. Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67**, 1-48. doi:10.18637/jss.v067.i01.

7. Mazerolle MJ. 2020 AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). R package version 2.3-1. https://cran.r-project.org/package=AICcmodavg.

8. Stoffel MA, Nakagawa S, Schielzeth H. 2017. rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods Ecol. Evol. 8, 1639-1644.

9. Fox J, Weisberg S. 2019. *An {R} Companion to Applied Regression*, Third Edition. Thousand Oaks CA: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/