How do anthropogenic contaminants (ACs) affect behaviour? Multi-level analysis of the effects of copper on boldness in hermit crabs

White, SJ

http://hdl.handle.net/10026.1/8341

How do anthropogenic contaminants (ACs) affect behaviour? Multi-level analysis of the effects of copper on boldness in hermit crabs.

Stephen J. White[1]

Mark Briffa[2]

[1]Centre for Ecology & Conservation, University of Exeter, Cornwall Campus.

[2]Marine Biology and Ecology Research Centre, Plymouth University, Drake Circus, Plymouth

Correspondence to: Mark Briffa

                      mark.briffa@plymouth.ac.uk

Abstract

Natural animal populations are increasingly exposed to human impacts on the environment, which could have consequences for their behaviour. Among these impacts is exposure to anthropogenic contaminants (ACs). Any environmental variable that influences internal state could impact behaviour across a number of levels: at the sample mean, at the level of among individual differences in behaviour ('animal personality') and at the level of within individual variation in behaviour (intra-individual variation, 'IIV'). Here we examined the effect of exposure to seawater-borne copper on the startle response behaviour of European hermit crabs, *Pagurus bernhardus* across these levels. Copper exposure rapidly led to longer startle responses on average but did not lead to any change in repeatability indicating that individual differences were present and equally consistent in the presence and absence of copper. There was no strong evidence that copper exposure led to changes in IIV. Our data show that exposure to copper for 1 week produces sample mean level changes in the behaviour of hermit crabs. However, there is no evidence that this exposure led to changes in repeatability through feedback loops.

Keywords: feedback; startle-response; reaction-norm; pollution

Introduction

It is increasingly recognised that anthropogenic impacts on the environment, such as exposure to contaminants, can influence animal behaviour (e.g. Zala & Penn 2004; Sih et al. 2011, Tuomainen & Candolin 2011; Candolin & Wong 2013) and these effects can potentially cascade through population and community level processes (Pelletier & Garant 2013; Blumstein 2013). Behavioural change in response to anthropogenic contaminants (ACs) (or any other environmental variable) can occur on a number of levels. First, and most familiarly, there might be mean level responses that are seen when comparing across treatment groups that have been exposed to different AC conditions. Second, AC exposure may reshape behavioural variation by changing patterns of energy acquisition and allocation. In certain cases, this can lead to positive feedback between exposure to the contaminant and behaviour that may amplify among-individual behavioural differences. Equally, negative feedbacks between behaviour and AC exposure might lead to a collapse of among-individual behavioural differences. Finally, AC exposure might lead to an increase in within individual behavioural variability (see Montiglio and Royauté 2014; Royauté et al. 2015). Note also that feedback between behaviour and the environment is a potential cause of among individual variation in behaviour (Wolf et al. 2007, Dingemanse & Wolf 2010) in a more general sense (i.e. in the absence of anthropogenic impacts).

Any consistent behavioural variation among individuals is typically referred to as animal personality (Sih et al. 2004, Dall et al. 2004), which is seen in animals representing diverse taxa (Carere & Maestripieri 2013; Kralj-Fišer & Schuett 2014) and is quantified by estimating repeatability. Repeatability (R) is the ratio between among individual variation (V_{BI}) and the sum of variation among and within (V_{WI}) individuals ($V_{BI} + V_{WI}$), i.e. it is the proportion of variance explained by V_{BI}. Any changes in

3

repeatability across conditions may therefore derive in part from among individual differences in response to the change in conditions, where individuals are said to show different behavioural reaction norms (Dingemanse et al 2010). In addition, changes in the within-individual variance component could lead to changes in repeatability. This V_{WI} component is also referred to as intra-individual variation (IIV), individual consistency, predictability and stability (Stamps et al. 2012; Westneat et al. 2012; Biro & Adriaenssens 2013; Briffa 2013a; Briffa et al. 2013; Jennings et al. 2013; Highcock and Carter 2014; Bridger et al. 2015; Westneat et al. 2015). Behavioural variation in response to changing conditions could thus occur across four distinct levels: (1) at the sample mean, (2) among individuals, (3) within individuals, and in the context of detecting the effects of AC exposure, (4) the slopes of behavioural reaction norms. Previous studies have shown that in poikilothermic animals temperature variation can influence behaviour across these different levels, primarily via its effects on metabolic rate (e.g. Briffa et al 2013). Therefore, ACs, also known to cause changes in metabolism (e.g. Depledge 1984; Dissanayake et al. 2009a), might also cause changes in behaviour as we outline below.

There are many examples of studies of the behavioural consequences of contaminant exposure that have demonstrated sample mean level effects (e.g. Bell 2001; Sloman et al. 2003; Sloman 2007; Dissanayake et al. 2009b; Sopinka et al. 2010, Brodin et al. 2013; White et al. 2013). Changes at this level of behaviour might be expected because of impairment of performance or impairment of information gathering and decision making (Sloman 2007), leading to lower activity rates. Conversely, animals under AC exposure may become more active (e.g. increased foraging, reduced hiding) due to the need to meet elevated metabolic loads, required by the activation of detoxification pathways.

In addition to sample mean level responses, AC exposure also has the potential to influence the repeatability of behaviour through positive or negative feedbacks (Montiglio and Royauté 2014, Royauté et al. 2015). Positive feedbacks might occur when (a) individuals initially vary in their behaviour in ways that cause differences in exposure to the contaminant, and (b) initial exposure causes behavioural changes that subsequently lead to increasing rates of exposure. In aquatic animals, individuals with initially high ventilation rates might be exposed to more waterborne contaminants, leading to elevated metabolic demands. Since metabolic rate correlates with ventilation rate (e.g. Atlantic salmon, *Salmo salar*, Millidine et al. 2008; Dungeness crab, *Metacarcinus magister*, McMahon et al. 1979) an initial AC exposure could necessitate even greater ventilation rates and hence more exposure to the contaminant. In contrast, negative feedbacks might occur when an initial exposure causes behaviour to change in ways that reduce future exposure. So far, only a few studies have collected the longitudinal data needed to assess the potential for changes in repeatability (and its components) under AC exposure. Kolkok et al. (1998) found reduced repeatability for swimming speed in fathead minnows, *Pimephales promelas*, following exposure to waterborne heavy metals. Similarly, Royauté et al. (2015) found reduced repeatability (across a suite of behaviours) in jumping spiders, *Eris militaris*, exposed to an insecticide.

In addition to changes in the proportion of behavioural variation due to among individual differences in behaviour, ACs could lead to changes in the amount of within individual variation in behaviour. Recent studies show that IIV can differ among individuals and vary across biotic (Briffa 2013a) and abiotic conditions (Briffa et al. 2013). High levels of IIV might represent strategies for coping with risk. Metabolically

compromised individuals might be subject to greater predation risk and therefore AC exposure might also lead to facultative increases in IIV.

Copper is a widespread contaminant of coastal habitats. Like other heavy metals, copper is known to increase the metabolic load of aquatic animals leading to a range of physiological changes including elevated heart rate (Marshall et al. 2004), loss of cellular integrity and immune responses (Pipe et al. 1999; Nicholson 2003; Parry & Pipe, 2004), reduced enzyme function (Alla et al. 2006), and endocrine disruption (Lye et al. 2005). In decapods, copper exposure is associated with a decline in performance of demanding behaviour such as aggression in shore crabs, *Carcinus maenas*, (Dissanayake et al. 2009b) and hermit crabs *Pagurus bernhardus* (White et al. 2013).

P. berhardus normally show consistent among individual differences in both aggression (Mowles et al. 2012) and in the duration of a startle response (Briffa et al. 2008; Briffa and Bibost 2009; Briffa 2013a) where they withdraw into their gastropod shell when disturbed. This behaviour provides protection from attack but at the costs of reduced activity and ventilation. These two aspects of consistent among individual variation in behaviour appear to be linked within a wider behavioural syndrome. Given that aggressive behaviour is impacted by copper exposure and given the links between aggression and consistent among individual differences in startle response duration we therefore expect copper exposure to influence the duration of startle responses in hermit crabs.

If copper exposure leads to elevated metabolic rate in hermit crabs (as in other decapod crustaceans) hiding times might be reduced in order to perform the elevated respiration rates (and perhaps food acquisition) required to service these increased metabolic demands. This in turn could lead to greater exposure to the AC, and such positive feedback could enhance among individual differences in behaviour as

described above. On the other hand, we might also expect changes in IIV with increasing AC exposure, and if exposure leads to increased IIV this would have tend to reduce the repeatability of hiding times. Therefore, to fully understand the effects of ACs on repeatable behaviour it is important to investigate changes in both components, V_{BI} and V_{WI}.

Although the effects of contaminants on animal personality have been discussed in recent studies (e.g. Brodin et al. 2013) the longitudinal data required to assess changes in repeatability and residual variance have rarely been incorporated into studies investigating the effects of AC exposure on behaviour (see Royaute 2015 for a recent example). Our aim in this study is to use longitudinal data on startle response duration in hermit crabs to test the hypotheses that (a) AC exposure will lead to changes in startle response durations at the level of the sample mean, (b) individuals will vary in their responses to copper exposure (i.e. there is an 'individual x environment interaction effect', Dingemanse et al. 2010), (c) copper exposure influences among individual variance in startle responses (V_{BI}) and (d) that copper exposure influences IIV (V_{WI}) in startle responses.

Materials and Methods

Collection and maintenance of crabs

We collected crabs from Hannafore Point, Cornwall, UK during June 2012 and August 2014. In June 2012 we also recorded individual startle responses in the field, in order to provide a baseline measure of startle response duration in the animal's natural setting. We lifted each crab out of the rock pool by hand, inverted it and held it in this position for five seconds, which causes the crab to withdraw into its gastropod shell. We then replaced it on the substrate (sand in the base of a rock pool in the case of field

observations, a plastic dish in the case of lab observations) in this inverted position and timed (in seconds) the duration of the startle response from the point at which the crab was replaced to the point where it re-emerged from the shell and first contacted the substrate with its walking legs (Briffa et al. 2008). There was no mean level difference in field collected startle response durations between crabs allocated to the two treatment groups in the laboratory experiment described below ($t_{38} = 0.63$, NS). We then transferred each crab to an individually labelled container containing seawater, for transport back to the laboratory in Plymouth. Here, we isolated each crab in a 17cm diameter plastic dish containing constantly aerated seawater at 15°C.

Laboratory experiment

We used an experimental design adapted from previous studies (de la Haye et al. 2011) investigating the effects of seawater parameters on behaviour that allows us to test for the effects of contaminant exposure on individual behaviour, while controlling for time. We conducted the experiment across two seven-day periods, A and B, with two treatments with n = 32 crabs in each group (2012; n = 20 individuals per treatment group, 2014; n= 12 individuals per treatment group), following exclusion of any crabs that had missing appendages or obvious parasites. In the first group, 'NN', crabs were held in normal seawater, without added copper, for both periods. In the second group, 'NC', crabs were held in normal seawater for period A and then seawater with added copper (see below for details) for period B. Thus we were able to determine whether exposure to copper led to a change in startle response behaviour while controlling for time held in the laboratory and for the effect of repeated exposure to the startling stimulus. An alternative approach would have been to use a crossover design (e.g. Briffa et al. 2013), where all individuals receive both treatments but in opposite

treatment orders. This approach, however, would have been poorly suited to the current experiment because copper is likely to accumulate in biological tissues and thus be retained in the body post-exposure.

In the first two days of each period we left the crabs undisturbed to allow acclimation to the experimental conditions. During the following five days we induced a startle as described above on each day, and recorded the duration of the response. See Figure 1 for a schematic description of the experimental design and sampling pattern. Due to logistical constraints we terminated observations after a maximum startle response duration of 600s, resulting in censored data for 77 out of 640 (640 = 10 observations per crab for 32 crabs in each treatment group) total observations. We randomised the order of observations among crabs across days. At the end of the experiment, the crabs were removed from their shells by cracking the shells in a bench vice, the crabs were sexed, weighed and examined for obvious parasites and damage. We also retained the shell fragments and weighed these after drying. The number of females:males in each treatment group was NN = 17:15 and NC = 11:21 (χ^2_1 = 1.6, P = 0.27). For each crab, we calculated their shell weight as a proportion of the crab's preferred shell weight, using regression lines that relate crab weight to optimal shell weight obtained from a previous shell selection experiment (Briffa & Elwood 2007). There was no difference between treatment groups in mean crab weight (t_{62} = 0.68, P = 0.5) or in proportion of optimal shell weight (t_{62} = 0.57, P = 0.57).

Seawater parameters and copper manipulation
As described previously (White et al. 2013) routine monitoring of copper pollution data does not exist for Hannafore Point, but it is considered to be a high quality site where

copper levels are unlikely to increase above background levels and annual median values are expected to be lower than $5\mu g\ l^{-1}$, the accepted natural background level of copper in seawater (Langston et al. 2003). As in previous studies (e.g. Fernández & Beiras 2001; Krång & Ekerholm 2006) we produced our copper exposed treatment by adding high purity Copper(II)Chloride ($Cu(II)Cl_2$), (Sigma No. 203149, 99.999% trace metal basis) to the seawater of 33ppt salinity. We chose an exposure level of $2mg\ l^{-1}$ of $Cu(II)Cl_2$ following White *et al* (2013) and pilot observations that indicated that this was a sub-lethal exposure for *P. bernhardus*, which did not lead to any obvious changes in observed locomotory behaviour. For $2mg\ l^{-1}$ of $Cu(II)Cl_2$, the maximum nominal concentration (the theoretical amount in solution assuming zero adherence to glassware) of copper ions in solution is $0.944\ mg\ l^{-1}$, which is towards the lower end of the nominal concentration range that has been used in previous studies (e.g. $10mg\ l^{-1}$ in Depledge 1984, $2mg\ l^{-1}$ in Cross et al. 2001, $0.5mg\ l^{-1}$ in Krång & Ekerholm 2006) on the effects of copper exposure on marine invertebrates. This nominal concentration is, however, towards the upper end of concentrations to which hermit crabs might be exposed in anthropogenically impacted coastal waters, with values at copper polluted locations reported as ranging from $61\mu g\ l^{-1}$ to $1mg\ l^{-1}$ (see Fernández & Beiras 2001; Krång & Ekerholm 2006) and values of over $500\ \mu g\ l^{-1}$ have been reported in the UK (DETR 1998). *P. bernhardus* is widely distributed in coastal habitats, including rocky and sandy shores, as well estuaries, throughout northwestern Europe, and hermit crabs in general are distributed globally. Therefore the copper exposures used in this study reflect levels that *P. berhhardus* and other hermit crabs could experience in the field. Moreover, further analyses reported in White et al. (2013) indicated that the actual concentrations of the exposure treatments were likely to be substantially lower than the nominal concentration, with actual concentrations of $0.0017mg\ l^{-1} \pm$ S.E. $= 0.0005mg$

l^{-1} (i.e. trace levels) for samples without added Copper(II)Chloride and 0.25mg ± S.E. = 0.0007mg l^{-1} for samples with a nominal concentration of 0.944 mg l^{-1}.

Statistical methods

In this experimental design an effect of copper exposure on startle response duration would be indicated by an interaction effect between period (A or B) and treatment group (no copper in period A followed by no copper in period B [NN] or no copper in period A followed by copper in period B [NC]). Therefore, we analysed the effect of treatment group, period and their interaction on mean startle response duration. Note that this interaction effect would indicate a significant effect both at the sample mean level and at the level of variance around the sample mean, which are both modelled in the analysis described below. Since multiple observations were collected from each individual we allowed random intercepts for individuals. As well as accounting for the non-independence of data collected from the same individual, random intercept effects may be interpreted biologically. If individuals differ (significantly) in their intercepts this indicates that there is significant among individual variation in the mean values of startle response durations. We also allowed random slopes across periods, since changes in response across the two periods might vary across individuals. Random slope effects may also be interpreted biologically. A significant random slope across periods would indicate significant variation among individuals in how they respond to the change in period, i.e. there would be significant variation in behavioural reaction norms. In this experiment the change in period equates to a change in conditions for the NC treatment group, and to constant conditions for the NN treatment group. We also included crab weight, deviation from preferred shell weight and observation number (1-10) as covariates. An initial analysis suggested no difference in mean response durations or in residual variance of startle response durations between the data sets

collected in each year, so year was not included as an effect in our subsequent analyses. In order to assess the influence of these fixed and random effects on mean startle response durations and the variance around mean startle response durations, we used a doubly hierarchical generalised linear model (DHGLM), fit via Markov chain Monte Carlo (MCMC) sampling (see Supplement 1 for a detailed explanation). The two levels of modelling are termed the mean and standard deviation (henceforth 'SD') models respectively. Data were Log_{10} transformed to improve the normality of residuals. This analysis was conducted using JAGS (Plummer 2003), which we controlled from within the R statistical computing environment (R core team 2014) using the package RJAGS (3.13) (Plummer 2014).

Analyses presented here exclude the 77 data points that exceeded our maximum observation time, since right censored data are difficult to normalise through transformation. However, analyses with these data points included gave qualitatively identical results to those reported below (see Table S1a, b).

To provide direct estimates of R, V_{BI} and V_{WI} we calculated these values specific to each block of data. We constructed a hierarchical general linear model (HGLM) using the MCMCglmm R package (Hadfield 2010). We then calculated posterior modes and 95% CIs for the differences in these values (i.e ΔR , ΔV_{BI}, ΔV_{WI}) across periods and treatment groups (Royauté et al. 2015). If the 95% CIs did not overlap zero, differences were deemed to be significant. See Supplement 1 for further details and table S.3 for a summary of the HGLM.

Ethical note

During collection crabs were placed into a large bucket containing ample seawater collected form the study site, at approximately 15°C, and this water was changed prior

to transport back to the laboratory, a trip that takes ca 45 minutes by car. Although we do not monitor seawater temperature or quality during this short transport period it is unlikely that any significant changes to seawater parameters would take place during this short trip. As in numerous previous studies, no crabs were observed to suffer injury or mortality during transport. In the lab crabs were maintained under a 12:12hr dark: light regime and were fed *ad libitum* on ca 0.5g portions of Atlantic pollock (*Pollachius* sp.). No crabs were injured during the process of removing them from shells by use of the bench vice. Due to the fact that some crabs had been exposed to copper, which could potentially accumulate in their tissues, we did not immediately return the crabs to the sea following the experiment. Rather, they were held in constantly filtered seawater in the laboratory for a period of 3 months, after which we judged that the risk of introducing significant quantities of copper to the sea was very low.

Results

Mean level effects of copper exposure

Full details of parameter estimates from the mean model and their 95% credible intervals are given in Table 1a, but here (for brevity) we give *P*-values only for fixed effects. There was no effect of crab weight ($P = 0.74$), crab sex ($P = 0.61$) or observation number ($P = 0.42$) on startle response duration but as deviation from preferred shell weight increased startle responses became longer in duration ($P < 0.02$). There was no effect of period ($P = 0.09$) but crabs in the NN treatment group had shorter startle responses overall than those in the NC treatment group ($P < 0.02$). A significant interaction effect ($P = 0.0005$) (Figure 2) between period and treatment group, showed that this effect of treatment group was due to an increase in startle responses during period B in crabs in the NC treatment group, whereas those in the NN treatment group

13

showed a slight decline in startle response duration from period A to period B. For both the random intercept and random slope effects, the 95% CIs did not cross zero, indicating significant among individual differences in startle response duration over the ten observations (Table 1) and that individuals varied in their responses (reaction norms) to the change in period (Table 1a). Note that for crabs in the NN treatment group the conditions were stable across periods A and B, whereas for crabs in the NC treatment group the transition from period A to period B represents a change in conditions. Since the random slope effect might have been caused by differences between the two treatment groups in the effect of period (rather than among individual differences in response to the change from period A to B, within each treatment group) we further investigated the effect of period (and the additional covariates described above) separately for each treatment group. In both cases the 95% CIs for the random intercept effect and for the random slope effect did not cross zero (see Table S2 for a full report of each model). This confirms the presence of among individual differences in behaviour across the 10 observations. Moreover, this indicates that although individuals varied in their responses to the change from period A to period B, this is unlikely to reflect among individual differences in response to copper exposure, since the effect was present in both treatment groups, NN and NC.

Variance (SD) level effects of copper exposure

Full details of mean parameter estimates and their 95% credible intervals are given in Table 1b, but here (for brevity) we give *P*-values only for fixed effects. The average amount of residual variance in behaviour expressed by the crabs did not differ between males and females ($P = 0.28$) or across periods ($P = 0.5$). However, the amount of residual variance did differ between the two treatment groups ($P < 0.02$), lighter crabs

14

were less consistent than heavier crabs ($P < 0.05$) and consistency in behaviour decreased with deviation from preferred shell weight ($P < 0.006$). There was no significant interaction effect between treatment group and period ($P = 0.99$), but note that this effect was estimated with low precision (see discussion below). For the random intercept effect the 95% CIs did not cross zero indicating the presence of among individual variation in within individual variance (Table 1b).

Temporal dynamics of startle response duration in the absence and presence of copper
It is clear from the above that for crabs exposed to copper during Period B (those in the NC treatment group) startle response durations were longer on average. However, crabs were observed five times in period B and the analysis above does not tell us anything about how the effect develops during this period. One possibility is that there is an immediate effect, apparent as soon as the crabs are exposed to copper; alternatively, the effect might develop more gradually as exposure accrues across the five observations. Therefore, to investigate the temporal dynamics of the observed change in behaviour, we performed an additional analysis, on data from period B only. We again used a DHGLM, as described above except for the following changes: In the mean model, period and its interaction with treatment group were omitted as fixed factors and instead we specified an interaction effect between treatment group and observation number. We again allowed random intercepts for each individual and, to account for the possibility that individuals might vary in their responses to repeated observation, we also allowed random slopes with respect to observation number. In the SD model, observation number and random slope effects were again absent.

The mean model revealed significant effects for treatment group ($P < 0.02$) and observation ($P < 0.02$) but also a significant interaction effect between treatment group

and observation, which suggests that startle responses declined over the five observations for crabs in the NN treatment group but increased for those in the NC treatment group ($P < 0.02$) (Table 2a; Figure 3). There was a significant random intercept indicating among individual variation in behaviour and a random slope indicating variation among crabs in their responses to repeated observations. There was no evidence that any fixed parameter influenced variance in behaviour during this period, but a random intercept effect indicates that among individual variation in IIV was still present (see Table 2b).

Comparison of repeatabilities across treatment combinations

For each block of data in the experiment startle response durations were significantly repeatable, as the 95% CIs did not overlap zero (Table 3). However, there were no significant differences in repeatability among blocks of the experiment as the 95% CIs for ΔR crossed zero in each case (Table 3). Similarly there was no evidence that among individual variance, V_{BI}, differed between blocks of the experiment (Table S4a). In the case of within individual variance, V_{WI}, there was some evidence for greater within individual variance in crabs that were exposed to copper during period B compared to crabs that were not exposed to copper during period B ($\Delta V_{WI} = 0.18$, 95% CI = 0.01, 0.36; Table S4b). However, we also note that crabs in the NN treatment group had rather low values for V_{WI} during both periods of the experiment and that for crabs in the NC treatment group there was no change in V_{WI} across periods. Furthermore, the lower CI for the delta value was very close to zero.

Discussion

16

Here we investigated the effects of AC exposure on behaviour at multiple levels of analysis. In common with the majority of studies looking at the links between ACs and behaviour we first looked at the effect of contaminant exposure at the sample mean level. *A priori*, we expected that copper exposure might lead to shorter startle responses (since exposed crabs might require higher ventilation) but here we found the opposite; hermit crabs exposed to waterborne copper showed, on average, an increase in startle response durations compared to those that were not exposed. Similar to previous studies on hermit crabs, we found no mean level effects of crab weight or observation number (e.g. Briffa et al 2008, 2013) and here we found no effect of sex.

We also looked at the effect of copper on among individual differences in startle response durations. For crabs in the NC treatment group, most individuals showed an increase in startle response duration following copper exposure but this did not occur across all individuals and there were significant differences in reaction norms. However these differences in individual responses did not result in any detectable differences in repeatability estimates, as assessed by calculating delta R values and their 95% CIs. Furthermore, using a similar approach, of assessing delta V_{BI}, there was no evidence for an effect of copper exposure on among individual variation in behaviour. The SD portion of our main analysis provided no evidence of a significant treatment group x period interaction effect, indicating that there was no effect of copper exposure on IIV. However, the confidence intervals around the estimate of this effect were very large, such that we were unable to estimate it with a high degree of precision within the DHGLM. This lack of precision indicates that our sample size may not have been adequate for modelling interaction effects on residual variance (see van de Pol 2012). Indeed, in our experimental design there is only a single estimate of V_{WI} for each individual during each period of the experiment. We also assessed changes in IIV by

assessing changes in the posterior mode of V_{WI} (calculated from a simpler HGLM model). Here there was some relatively weak evidence that V_{WI} differed between treatment groups in period B. However, it seems unlikely that this is actually due to copper exposure. Rather it appears that crabs in the NN treatment group had relatively low baseline levels of within individual variation in behaviour compared to those in the NC treatment group. Given that crabs were allocated to the two groups at random, the cause of this difference is somewhat puzzling. It is clear, however, that startle response durations are significantly repeatable both before and after exposure to this particular AC.

Although mean startle responses continued to diverge during the five observations in period B of the experiment, this difference between treatments had already started to emerge during the first observation of this period. This is not surprising given that the crabs had already been exposed to the different conditions for 48 h prior to the resumption of observations during period B. Therefore it seems unlikely that clearer changes in repeatability would have been captured by a longer observation window. In the following section we discuss some possible explanations for longer startle responses in copper exposed crabs.

First, longer startle responses might have arisen because copper exposure leads to reduced performance (White et al. 2013). There are two ways that reduced performance could lead to longer duration of startle responses. First they might be less capable of re-emerging from their shells rapidly. However, the act of actually re-emerging from the shell contributes little to the total duration of the startle response and re-emerging is unlikely to be an energetically demanding behaviour. A second possibility is that individuals with impaired performance might be more vulnerable to attack by predators and hence they behave more cautiously. However, in a previous

18

study we found that individuals under elevated risk conditions also show greater IIV (Briffa 2013a), an effect that we found no strong evidence for here. An alternative explanation is that longer startle responses are due to impaired decision-making. Anthropogenic contaminants are known to degrade the ability to acquire and process information (Sloman 2007; Briffa et al. 2012). Indeed, previous studies have shown that startle response duration is sensitive to water-borne chemical cues (Briffa et al. 2008; Briffa 2013a) and contaminants such as copper might prevent their detection. Moreover, even in the absence of specific chemical cues (as was the case in this study) ACs may disrupt information processing and thus the decision to re-emerge could be delayed in copper exposed crabs.

Irrespective of the specific mechanisms underlying the changes in behaviour detected here, longer hiding times would reduce the time available for other activities such as foraging and the acquisition of new shells, which are required as the crab grows. Indeed previous studies (Briffa 2013b, Mowles et al 2012) have shown that low boldness is associated with reduced shell acquisition. Therefore copper exposure could reduce the rates of shell acquisition within hermit crab populations, leading to reduced growth and fecundity. The potential for behavioural changes in response to an AC to cascade through to population and community level effects is difficult to predict. ACs might impact a range of mechanisms that underpin the observed behaviour and in natural environments animals are often exposed to a 'cocktail' of ACs, with the possibility that each component in the mix has different effects on behaviour (e.g. Kortenkamp 2007, Sopinka et al. 2010). Nevertheless, we suggest that using longitudinal data to understand the range of impacts across different levels of analysis for a single contaminant represents a first step in predicting the wider consequences of AC exposure.

Although the mean level effects of copper exposure had emerged after seven days, longer term studies, allowing for the possibility of contaminant accumulation (Montiglio & Royauté 2014), and acclimation (e.g. see Spicer & Weber 1992) are also warranted. Indeed, in common with other studies investigating the links between behaviour and ACs, we did not investigate the tissue burden of contaminant exposure (e.g. Dissanayake *et al*. 2009b; Sopinka et al. 2010), a factor that might vary among individuals and contribute to their different responses to copper exposure. To the best of our knowledge, however, this study represents the most detailed analysis to date of the effects of AC exposure across different levels of behavioural variation. The results at the sample mean level are clear, individuals exposed to copper show increased startle response durations. Our analysis at the levels typically investigated in animal personality studies raises a number of further questions for investigation. We suggest that a useful approach would be to couple the physiological analyses typical of eco-toxicological studies with the longitudinal data on behaviour that defines animal personality research. As well as providing insights into the likely ecological effects of human induced rapid environmental change (Sih 2013) manipulative studies on ACs might also provide insights into the maintenance of behavioural variation among individuals, by providing a way of testing the idea that positive feedback between behaviour and internal state can maintain such variation.

Acknowledgements

References

Alla AA, Mouneyrac C, Durou C, Moukrim A, Pellerin J (2006) Tolerance and biomarkers as useful tools for assessing environmental quality in the Oued Souss estuary (Bay of Agadir, Morocco). Comparative Biochemistry Physiology C 143: 23-29.

Bell A (2001) Effects of an endocrine disrupter on courtship and aggressive behaviour of male three-spined stickleback, *Gasterosteus aculeatus*. Animal Behaviour, 62: 775–780.

Biro PA, Adriaenssens B (2013) Predictability as a personality trait: consistent differences in intraindividual behavioral variation. American Naturalist, 182: 621-629.

Blumstein DT (2013) Social behaviour. In, Candolin, U. & Wong B.B. (Eds.) Behavioural responses to a changing world, mechanisms and consequences. Oxford: Oxford University Press.

Bridger D, Bonner SJ, Briffa M (2015) Individual quality and personality: bolder males are less fecund in the hermit crab *Pagurus bernhardus*. Proceedings of the Royal Society B 282:20142492. doi: 10.1098/rspb.2014.2492

Briffa M (2013a) Plastic proteans: reduced predictability in the face of predation risk in hermit crabs. Biology Letters 9:20130592. doi: 10.1098/rsbl.2013.0592

Briffa M (2013b) The influence of personality on a group-level process: Shy hermit crabs make longer vacancy chains. Ethology 119:1014-1023.

Briffa M, Bibost A-L (2009) Effects of shell size on behavioural consistency and flexibility in hermit crabs. Canadian Journal of Zoology 87:597–603. doi: 10.1139/Z09-047

Briffa M, Bridger D, Biro PA (2013) How does temperature affect behaviour? Multilevel analysis of plasticity, personality and predictability in hermit crabs. Animal Behaviour 86:47–54. doi: 10.1016/j.anbehav.2013.04.009

Briffa M, Elwood RW (2007) Monoamines and decision making during contests in the hermit crab *Pagurus bernhardus*. Animal Behaviour 73: 605-612.

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*. Proceedings of the Royal Society B 275:1305–11. doi: 10.1098/rspb.2008.0025

Brodin T, Fick J, Jonsson M, Klaminder J (2013) Dilute concentrations of psychiatric drug alter behaviour of fish from natural populations. Science 339: 814-815.

Candolin U, Wong BB (2013) Behavioural responses to a changing world, mechanisms and consequences. Oxford: Oxford University Press.

Carere C, Maestripieri D (2013) Animal Personalities: Behavior, Physiology and Evolution. University of Chicago Press. Chicago, USA.

Cross MA, Irwin SWB, Fitzpatrick SM (2001) Effects of heavy metal pollution on swimming and longevity in cercariae of *Cryptocotyle lingua* (Digenea: Heterophyidae). Parasitology 123: 499-507.

Dall SRX, Houston AI, McNamara JM (2004) The behavioural ecology of personality: consistent individual differences from an adaptive perspective. Ecology Letters 7: 734-739. doi: 10.1111/j.1461-0248.2004.00618.x

de la Haye K, Spicer JI, Widdicombe S, Briffa M (2011) Reduced seawater pH disrupts information gathering, resource assessment and decision making in the hermit crab *Pagurus bernhardus*. Animal Behaviour 82: 495-501.

Depledge MH (1984) Disruption of circulatory and respiratory activity in shore crabs (Carcinus maenas (l)) exposed to heavy-metal pollution. Comparative Biochemistry and Physiology C 78: 445-459.

DETR (1998) Digest of Environmental Statistics No. 20 1998. Department of the Environment, Transport and the Regions. HMSO.

Dingemanse NJ, Kazem AJN, Réale D, Wright J (2010) Behavioural reaction norms: animal personality meets individual plasticity. Trends in Ecology & Evolution 25:81–89. doi: 10.1016/j.tree.2009.07.013

Dingemanse NJ, Wolf M (2010) Recent models for adaptive personality differences: a review. Philosophical Transactions of the Royal Society B: Biological Sciences 365: 3947-3958.

Dissanayake A, Galloway TS, Jones MB (2009a) Responses of shore crabs to physiological challenges following exposure to selected environmental challenges and contaminants. Journal of Experimental Marine Biology & Ecology 375:57–63. doi: 10.1016/j.jembe.2009.05.007

Dissanayake A, Galloway TS, Jones MB (2009b) Physiological condition and intraspecific agonistic behaviour in *Carcinus maenas* (Crustacea: Decapoda). Journal of Experimental Marine Biology & Ecology 375:57–63. doi: 10.1016/j.jembe.2009.05.007

Fernández N, Beiras R (2001) Combined toxicity of dissolved mercury with copper, lead and cadmium on embryogenesis and early larval growth of the *Paracentrotus lividus* sea-urchin. Ecotoxicology 10:263-271.

Hadfield JD (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software, 33: 1-22.

Highcock L, Carter AJ (2014) Intraindividual variability of boldness is repeatable across contexts in a wild lizard. PLoS One 9:e95179. doi: 10.1371/journal.pone.0095179

Jennings DJ, Hayden TJ, Gammell MP (2013) Personality and predictability in fallow deer fighting behaviour: the relationship with mating success. Animal Behaviour 86:1041–1047. doi: 10.1016/j.anbehav.2013.09.009

Kortenkamp A (2007) Ten years of mixing cocktails: a review of combination effects of endocrine-disrupting chemicals. Environmental Health Perspectives: 115(Suppl.) S98-S105.

Kralj-Fišer S, Schuett W (2014) Studying personality variation in invertebrates: why bother? Animal Behaviour 91: 42-52.

Krång A-S, Ekerholm M (2006) Copper reduced mating behaviour in male shore crabs (*Carcinus maenas* (L.)). Aquatic Toxicology 80: 60-69.

Langston WJ, Chesman BS, Burt GR, Hawkins SJ, Readman J, Worsfold P (2003) Characterisation of the South West European Marine Sites. Plymouth Sound and estuaries cSAC, SPA. Marine Biological Association of the United Kingdom occasional publication No. 9.

Lye CM, Bentley MG, Clare AS, Sefton EM (2005) Endocrine disruption in the shore crab *Carcinus maenas* - a biomarker for benthic marine invertebrates? Marine Ecology Progress Series 288: 221-32.

Marshall DJ, Peter R, Chown SL (2004) Regulated bradycardia in the pulmonate limpet *Siphonaria* (Gastropoda : Mollusca) during pollutant exposure: implication for biomarker studies. Comparative Biochemistry Physiology A 139: 309-16.

McMahon BR, McDonald DG, Wood CM (1979) Ventilation, oxygen uptake and haemolymph oxygen transport, following enforced exhausting activity in the dungeness crab *Cancer magister*. Journal of Experimental Biology 80: 171-285.

Millidine KJ, Metcalfe NB, Armstrong JD (2008) The use of ventilation frequency as an accurate indicator of metabolic rate in juvenile Atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Sciences, 65: 2081-2087.

Montiglio P-O, Royauté R (2014) Contaminants as a neglected source of behavioural variation. Animal Behaviour 88:29–35. doi: 10.1016/j.anbehav.2013.11.018

Mowles SL, Cotton PA, Briffa M (2012) Consistent crustaceans: the identification of stable behavioural syndromes in hermit crabs. Behavioral Ecology & Sociobiology 66:1087–1094. doi: 10.1007/s00265-012-1359-7

Nicholson S (2003) Lysosomal membrane stability, phagocytosis and tolerance to emersion in the mussel *Perna viridis* (Bivalvia : Mytilidae) following exposure to acute, sublethal, copper. Chemosphere 52: 1147-51.

Parry HE, Pipe RK (2004) Interactive effects of temperature and copper on immunocompetence and disease susceptibility in mussels (*Mytilus edulis*). Aquatic Toxicology 69: 311-325.

Pelletier F, Garant D (2013) Population consequences of individual variation in behaviour. In, Candolin, U & Wong BB (Eds.) Behavioural responses to a changing world, mechanisms and consequences. Oxford: Oxford University Press.

Pipe RK, Coles JA, Carissan FMM, Ramanathan K (1999) Copper induced immunomodulation in the marine mussel, *Mytilus edulis*. Aquatic Toxicology 46: 43-54.

Plummer M (2003) JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. In Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003). March, pp. 20–22. (doi:10.1.1.13.3406)

Plummer M (2014) rjags: Bayesian graphical models using MCMC. pp. http://CRAN.R–project.org/package=rjags.

van de Pol A (2012) Quantifying individual variation in reaction norms: how study design affects the accuracy, precision and power of random regression models. Methods in Ecology and Evolution 3: 268-280.

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Royauté R, Buddle CM, Vincent C (2015) Under the influence: sublethal exposure to an insecticide affects personality expression in a jumping spider. Functional Ecology 29: 962-970.

Sih A (2013) Understanding variation in behavioural responses to human-induced rapid environmental change: a conceptual overview. Animal Behaviour 85: 1077-1088.

Sih A, Bell A, Johnson JC (2004) Behavioural syndromes: an ecological and evolutionary overview. Trends in Ecology & Evolution 19: 372 – 377. doi:10.1016/j.tree.2004.04.009

Sih A, Ferarri MCO, Harris DJ (2011) Evolution and behavioural responses to human-induced rapid environmental change. Evolutionary Applications 4: 367-387.

Sloman KA, Baker DW, Ho CG, McDonald DG, Wood CM (2003) The effects of trace metal exposure on agonistic encounters in juvenile rainbow trout, *Oncorhynchus mykiss*. Aquatic Toxicology 63:187-96.

Sloman K (2007) Effects of trace metals on salmonid fish: The role of social hierarchies. Applied Animal Behaviour Science 104:326–345.

Sopinka NM, Marentette JR, Balshine S (2010) Impact of contaminant exposure on resource contests in an invasive fish. Behavioral Ecology & Sociobiology 64: 1947–1958.

Spicer JI, Weber RE (1992) Respiratory impairment by water-borne copper and zinc in the edible crab *Cancer pagurus* (L.) (Crustacea: Decapoda) during hypoxic exposure. Marine Biology 112: 429-435.

Stamps JA, Briffa M, Biro PA (2012) Unpredictable animals: individual differences in intraindividual variability (IIV). Animal Behaviour 83:1325–1334. doi: 10.1016/j.anbehav.2012.02.017

Tuomainen U, Candolin U (2011) Behavioural responses to human-induced environmental change. Biological Reviews 86: 640-657.

Westneat DF, Schofield M, Wright J (2012) Parental behavior exhibits among-individual variance, plasticity, and heterogeneous residual variance. Behavioural Ecology 24:598–604. doi: 10.1093/beheco/ars207

Westneat DF, Wright J, Dingemanse NJ (2015) The biology hidden inside residual within-individual phenotypic variation. Biological Reviews, in press.

White SJ, Pipe RK, Fisher A, Briffa M (2013) Asymmetric effects of contaminant exposure during asymmetric contests in the hermit crab *Pagurus bernhardus*. Animal Behaviour 86: 773-781.

Wolf M, Van Doorn GS, Leimar O, Weissing FJ (2007) Life-history trade-offs favour the evolution of animal personalities. Nature 447: 581-584.

Zala SM, Penn DJ (2004) Abnormal behaviours induced by chemical pollution: a review of the evidence and new challenges. Animal Behaviour 68: 649–664.

Table 1: Posterior summary statistics with standard deviation, upper and lower 95% credible intervals and pseudo-*P* values for Model 1 used to assess the effects of copper across both periods of the experiment. Group refers to the treatment group (NN or NC). Mean model (a) and SD model (b).

Variable	Estimate	SD	2.5% CI	97.5% CI	*P*
(a) Mean model					
Intercept	0.210	0.178	-0.13	0.56	0.21
Group	-0.552	0.210	-0.95	-0.12	<0.02
Period	-0.276	0.157	-0.57	0.04	0.09
Sex	0.116	0.216	-0.30	0.55	0.61
Weight	-0.045	0.131	-0.29	0.11	0.74
Observation	0.107	0.134	-0.16	0.36	0.42
% PSW	0.123	0.047	0.03	0.22	<0.02
Group x Period	0.631	0.182	0.27	0.98	0.0005
Random intercept	0.790	0.084	0.64	0.96	
Random slope[1]	0.576	0.084	0.42	0.75	
(b) SD model					
Intercept	-0.426	0.083	-0.59	-0.26	<0.0001
Group	-0.234	0.092	-0.41	-0.05	<0.02
Period	-0.059	0.082	-0.22	0.10	0.50
Sex	0.099	0.091	-0.08	0.28	0.28
Weight	-0.126	0.056	-0.23	-0.01	<0.05
%PSW	0.158	0.053	0.05	0.26	<0.006
Group x Period	0.112	100.9	-195.1	198.2	0.99
Random intercept	0.187	0.066	0.04	0.31	

[1]In this model the random slope effect is specified across the two periods of the experiment.

Table 2: Posterior summary statistics with standard deviation, upper and lower 95% credible intervals and pseudo-*P* values for Model 2 used to assess the effects of copper across observations within period B of the experiment. Group refers to the treatment group (NN or NC). Mean model (a) and SD model (b).

Variable	Estimate	SD	2.5% CI	97.5% CI	*P*
(a) Mean model					
Intercept	0.242	0.209	-0.17	0.66	0.25
Group	-0.657	0.266	-1.17	-0.13	<0.02
Sex	0.192	0.290	-0.38	0.77	0.47
Weight	-0.095	0.241	-0.57	0.40	0.69
%PSW	0.216	0.243	-0.28	0.69	0.37
Observation	0.233	0.096	0.05	0.42	<0.02
Group x Observation	-0.320	0.131	-0.58	-0.07	<0.02
Random intercept	0.734	0.111	0.55	0.98	
Random slope[1]	0.241	0.091	0.04	0.41	
(b) SD model					
Intercept	-0.33	0.11	-0.55	-0.11	0.006
Group	-0.17	0.16	-0.48	0.14	0.28
Sex	-0.13	0.17	-0.45	0.21	0.42
Weight	0.04	0.15	-0.26	0.33	0.76
%PSW	0.05	0.14	-0.22	0.32	0.69
Random intercept	0.13	0.09	0.004	0.34	

[1]In this model the random slope effect is specified across observations 6 to 10 within period B of the experiment.

Table 3: MCMC repeatability estimates for each block of data, with ΔR for the difference between periods of the experiment (ΔR = B-A) and for the difference between treatment groups (ΔR = NC-NN). Upper and lower 95% CIs for R and ΔR values are given in square brackets.

	A	B	ΔR (B-A)
NN	0.54 [0.39, 0.71]	0.70 [0.56, 0.81]	-0.14 [-0.06, 0.34]
NC	0.59 [0.43, 0.72]	0.59 [0.46, 0.75]	-0.004 [-0.21, 0.22]
ΔR (NC-NN)	0.03 [-0.16, 0.26]	-0.02 [-0.28, 0.10]	

NN treatment	Seawater							Seawater					
NC treatment	Seawater							Seawater + Copper					
Period	A							B					
Observation			1	2	3	4	5			6	7	8	9	10
Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Fig. 1. Schematic representation of the experimental design, showing how observations and sea water conditions were distributed across the 14 days of experiment for the two treatment groups (n = 32 *Pagurus bernhardus* per treatment group). Grey periods indicate days when crabs were undisturbed.

SUPPLEMENTARY MATERIAL

Details of DHGLM for the effects of treatment group and period on mean and variances around startle response durations

A common approach for analysis of mean level responses would be to use a linear mixed effects model, where fixed factors represent experimental treatment effects (and the effects of covariates) and random intercepts allow for variation across individuals over repeated observations. Here, however, we took an alternative approach of using a doubly hierarchical generalised linear model (DHGLM), fit using Bayesian methods via Markov chain Monte Carlo (MCMC) sampling. This hierarchical method allows the simultaneous modelling of mean and variance level effects, allowing us to ask an additional question; whether V_{WI} (i.e. IIV or residual variance) of startle responses also varies (on average) across treatments (Cleasby et al. 2015). If there is a change in V_{WI} as a response to exposure to this AC we would expect a significant 'treatment group x period' interaction whereby residual variance differs between periods A and B for the NC treatment group but remains stable for the NN treatment group. In doubly hierarchical models we may also include the random effects described above, allowing us to ask whether mean level responses, behavioural reaction norms and IIV vary among individuals. This method also provides the ability to cope with non-heterogeneous residual errors, allowing fixed effects to be assessed more robustly in comparison to LMM. The two levels of modelling (mean level standard deviation level responses) are termed the mean and standard deviation (henceforth 'SD') models respectively. Data were Log_{10} transformed to improve the normality of residuals. For the mean model we included the fixed and random effects described above. For the SD model it was not possible to include fixed effects for observation number or a random slope effect (since we only obtained one set of repeated measures within each period per individual, allowing a single estimate of residual variance per individual during each period) but for each crab we did allow a random intercept. Sampling from the posterior distributions of the model parameters was conducted using the freely available software JAGS (Plummer 2003), which we controlled from within the R statistical computing environment (R core team 2014) using the package RJAGS (3.13) (Plummer 2014). Following the usual MCMC setup, the parameters in each model are updated conditional on the remaining parameters to generate random draws from their posterior distribution. The standard deviations of the random effects and error terms in both the mean and SD models were assigned weakly informative scaled half-t prior distributions with 3 degrees of freedom (Gelman et al. 2008) while the fixed effects parameters were assigned non-informative normal prior distributions. Three chains were run in parallel so that convergence could be assessed and each chain was run with an adaptive phase ('burn in') of 5000 iterations and a sampling phase of 20000 iterations. Convergence across chains was assessed using the Gelman-Rubin diagnostic, which was <1.1 for each model parameter, indicating that the adaptive phase was adequate. We made inferences about the parameters in each model based on their posterior means and 95% credible intervals. We based the primary assessment of the significance of each predictor on whether or not the 95% credible intervals for the corresponding parameter covered zero. In the case of fixed effects we are also able to express these metrics in a more familiar way, by generating values analogous to classical P-values (Bridger et al. 2015). These pseudo P-values (referred to as 'P' in the main text) are obtained by

calculating the tail probability for each fixed parameter. They express, as a value between 0 and 1, the probability over the set of all equal tailed credible intervals that cover zero. This calculation, however, was not possible for random effects as they are constrained to be positive. For these effects we infer significance by examination of 95% credible intervals as described above.

Details of HGLM for the estimation and comparison of variance components across treatment groups and periods

Although the lack of a treatment group x period interaction effect in the SD model reported above indicates that copper exposure did not influence within individual variation in behaviour (V_{WI}), this does not directly test the hypothesis that copper exposure influences repeatability, since R = V_{BI}/(V_{BI}+V_{WI}). Therefore we also assessed repeatability for each block of data. Following Royauté et al. (2015) we first constructed a HGLM (using the MCMCglmm R package, Hadfield 2010) where random intercepts were estimated for each group, without including a random slope effect. As in the mean part of the DHGLM described above, we included fixed effects for treatment, period and treatment group x period. Since %PSW also influenced mean level responses we included this as a covariate, but we did not carry forward the additional covariates that had no effect on startle response duration in the DHGLM. For a summary of the HGLM see Table S3. To assess convergence of each HGLM model we calculated the autocorrelation factors (ACF) for each effect and each specific variance component. We then extracted the posterior modes for V_{BI} and V_{WI} and used these to calculate adjusted (on %PSW) repeatabilities (and 95% CIs) for each combination of treatment group and period. We then calculated the difference in repeatability and its credible intervals (ΔR +/- 95% CI) between treatment groups within each period, and between periods within each treatment group. Similar calculations were also performed for the V_{BI} and V_{WI} variance components to give ΔV_{BI} and ΔV_{WI} respectively. As well as providing effect sizes for any changes in repeatability or variance components, we may assess the significance of such changes by determining overlap of zero by 95% CIs.

Table S1. Summary of a DHGLM analysing the effects of copper based on the full dataset, including right censored data. Group refers to treatment group.

Variable	Estimate	SD	2.5% CI	97.5% CI	P
(a) Mean model					
Intercept	0.206	0.156	-0.08	0.52	0.24
Group	-0.529	0.207	-0.94	-0.14	<0.02
Period	-0.236	0.149	-0.52	0.07	0.12
Sex	0.038	0.206	-0.35	0.46	0.89
Weight	-0.048	0.117	-0.28	0.18	0.72
Occasion	0.103	0.131	-0.15	0.36	0.48
% PSW	0.140	0.045	0.05	0.23	<0.003
Group x Period	0.608	0.177	0.27	0.96	<0.001
Random intercept	0.754	0.080	0.61	0.93	
Random slope	0.584	0.086	0.43	0.77	
(b) SD model					
Intercept	-0.458	0.092	-0.64	-0.28	<0.0001
Group	-0.143	0.107	-0.35	0.07	0.18
Period	-0.051	0.081	-0.21	0.11	0.52
Sex	0.015	0.110	-0.20	0.23	0.9
Weight	-0.140	0.065	-0.27	-0.01	<0.05
%PSW	0.150	0.065	0.02	0.28	<0.03
Group x Period	0.027	100.4	-195.5	198.2	0.99
Random intercept	0.297	0.0004	0.19	0.42	

Table S2. DHGLM models fitted separately for each treatment group, NN and NC.

Variable	Estimate	SD	2.5% CI	97.5% CI	*P*
NN treatment group					
(a) Mean model					
Intercept	-0.009	0.261	-0.530	0.539	0.99
Period	0.275	0.189	-0.094	0.660	0.16
Sex	-0.273	0.361	-1.005	0.430	0.43
Weight	0.293	0.201	-0.106	0.691	0.12
Occasion	-0.002	0.221	-0.440	0.433	0.99
% PSW	0.067	0.078	-0.086	0.219	0.41
Random intercept	0.826	0.135	0.604	1.130	
Random slope	0.596	0.129	0.363	0.871	
(b) SD model					
Intercept	-0.597	0.106	-0.800	-0.385	<0.0001
Period	0.017	0.106	-0.192	0.224	0.86
Sex	0.195	0.123	-0.053	0.435	0.11
Weight	-0.083	0.070	-0.217	0.060	0.223
%PSW	0.091	0.077	-0.062	0.238	0.23
Random intercept	0.116	0.076	0.006	0.285	
NC treatment group					
(a) Mean model					
Intercept	-0.03611	0.1926	-0.40525	0.3345	0.77
Period	-0.11566	0.1749	-0.46119	0.2289	0.54
Sex	0.35642	0.3105	-0.26404	0.9736	0.20
Weight	-0.15931	0.1793	-0.50102	0.2164	0.33
Occasion	0.04141	0.1738	-0.30905	0.3711	0.78
% PSW	0.19238	0.0721	0.05139	0.3336	<0.01
Random intercept	0.7267	0.1194	0.5274	0.9928	
Random slope	0.4922	0.1424	0.2140	0.7798	
(b) SD model					
Intercept	-0.495	0.103	-0.699	-0.297	<0.0001
Period	0.0235	0.107	-0.183	0.239	0.84
Sex	-0.025	0.152	-0.325	0.280	0.88
Weight	-0.169	0.095	-0.342	0.029	0.09
%PSW	0.219	0.088	0.046	0.397	<0.02
Random intercept	0.274	0.109	0.056	0.494	

Table S3. HGLM with random effects specified for each treatment group.

Variable	Estimate	2.5% CI	97.5% CI	P_{MCMC}
(a) Mean model				
Intercept	-0.539554	-1.245809	0.098568	0.110
Group	0.123241	-0.254276	0.535784	0.542
Period	0.498477	0.016266	0.940636	0.036
% PSW	0.215052	-0.110811	0.563071	0.198
Group x Period	-0.646687	-1.246905	-0.006083	0.038
Random intercepts				
NN.A	0.4520	0.2146	0.7124	
NN.B	0.7378	0.3849	1.1548	
NC.A	0.7296	0.4071	1.1781	
NC.B	0.7464	0.3761	1.1925	
Residual variance				
NN.A	0.3666	0.2791	0.4606	
NN.B	0.3321	0.2397	0.4196	
NC.A	0.4896	0.3728	0.6220	
NC.B	0.4976	0.3746	0.6303	

Table S4. Posterior modes for (a) among and (b) within individual variation in startle response duration with ΔV for the difference between periods of the experiment (B-A) and for the difference between treatment groups (NC-NN). Upper and lower 95% CIs for V and ΔV values are given in square brackets and significant delta values are shown in bold.

(a) Among individual variation, V_{BI}	A	B	ΔV_{BI} (B-A)
NN	0.41 [.021, 0.71]	0.73 [0.38, 1.15]	0.38 [-0.18, 0.81]
NC	0.64 [0.41, 1.18]	0.57 [0.37, 1.19]	-0.03 [-0.56, 0.64]
ΔV_{BI} (NC-NN)	0.17 [-0.19, 0.81]	-0.05 [-0.56, 0.66]	

(b) Within individual variation, V_{WI}	A	B	ΔV_{WI} (B-A)
NN	0.36 [0.28, 0.46]	0.33 [0.24, 0.42]	-0.05 [-0.17, 1.00]
NC	0.48 [0.36, 0.62]	0.47 [0.37, 0.63]	-0.03 [-0.19, 0.17]
ΔV_{WI} (NC-NN)	0.10[-0.04, 0.30]	**0.18 [0.01, 0.36]**	

References

Bridger D, Bonner SJ, Briffa M (2015) Individual quality and personality: bolder

    males are less fecund in the hermit crab *Pagurus bernhardus*. Proceedings of the

    Royal Society B 282:20142492. doi: 10.1098/rspb.2014.2492.

Cleasby IR, 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: 27-37.

Gelman A, Jakulin A, Pittau MG, Su, Y-S (2008) A weakly informative default prior

    distribution for logistic and other regression models. Annals of Applied Statistics

    2, 1360–1383.

Hadfield JD (2010) MCMC Methods for Multi-Response Generalized Linear Mixed

    Models: The MCMCglmm R Package. Journal of Statistical Software, 33: 1-22.

Plummer M (2003) JAGS: A Program for Analysis of Bayesian Graphical Models

    Using Gibbs Sampling. In Proceedings of the 3rd International Workshop on

    Distributed Statistical Computing (DSC 2003). March, pp. 20–22.

Plummer M (2014) rjags: Bayesian graphical models using MCMC. pp.

    http://CRAN.R–project.org/package=rjags.

R Core Team (2014) R: A language and environment for statistical computing. R

    Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-

    project.org/.

Royauté R, Buddle CM, Vincent C (2015) Under the influence: sublethal exposure to

an insecticide affects personality expression in a jumping spider. Functional
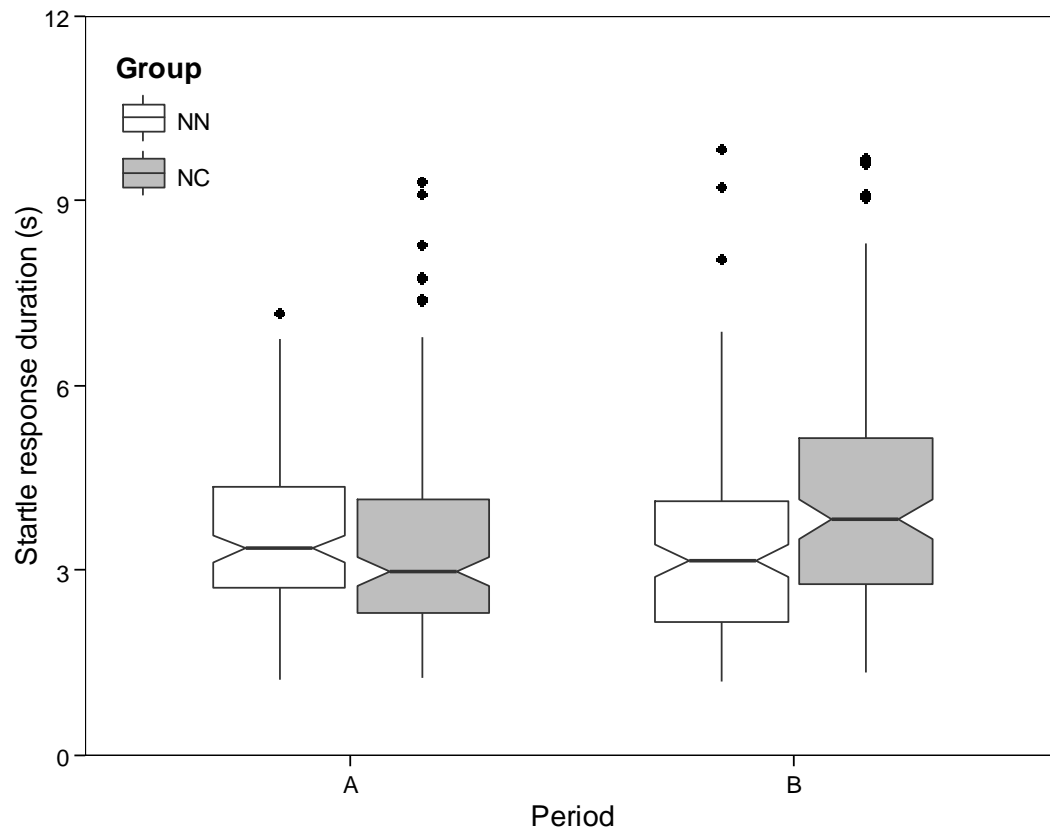
Ecology 29: 962-970.

Fig. 2. The significant interaction effect between treatment group and period on the duration of startle responses. Thick horizontal bars show the median, boxes show the interquartile range (IQR) from first to third quartiles, whiskers show the nominal range of the data (maximum and minimum values that are within 1.5 x IQR) and data falling outside the nominal range are shown as dots. Notches indicate the 95% CIs of the median.
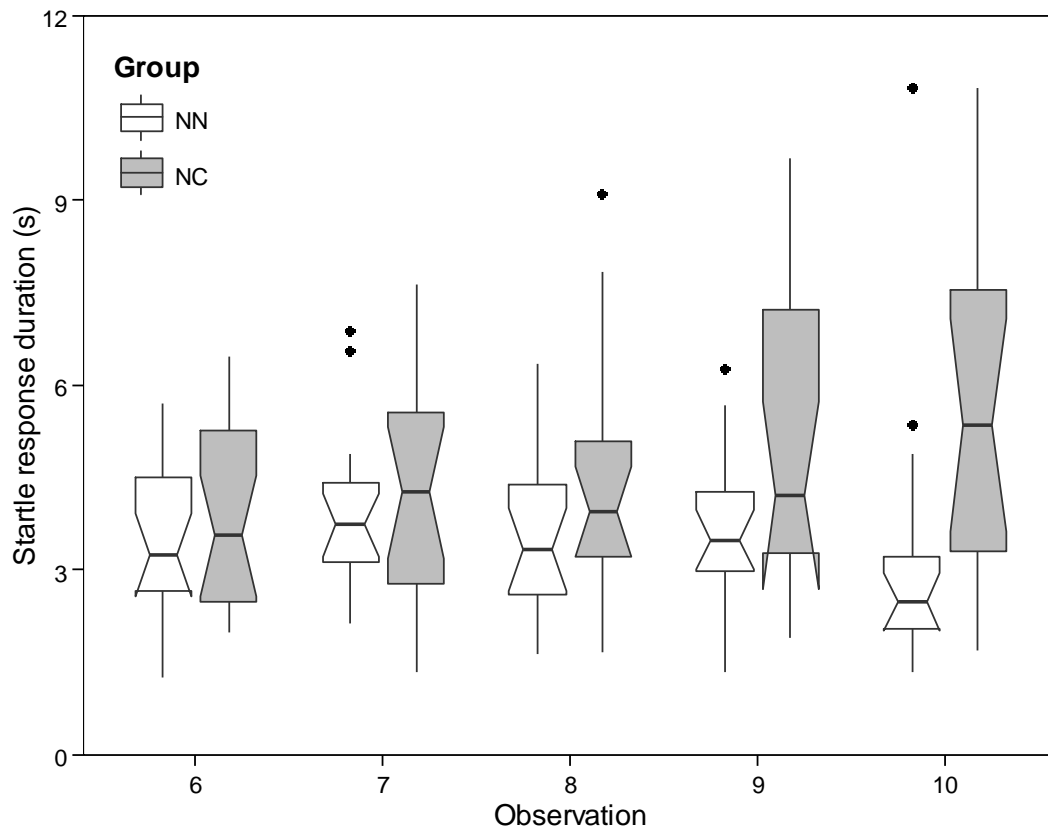
Fig. 3. The significant interaction effect between treatment group and observation
number during period B of the experiment (observations 6 to 10). Thick horizontal
bars show the median, boxes show the interquartile range (IQR) from first to third
quartiles, whiskers show the nominal range of the data (maximum and minimum
values that are within 1.5 x IQR) and data falling outside the nominal range are shown
as dots. Notches indicate the 95% CIs of the median. Note that for the NC group the
lower 95% CI exceeds the first quartile at observation 9.