



LETTER

Non-consumptive effects of predation: does perceived risk strengthen the genetic integration of behaviour and morphology in stickleback?

Niels J. Dingemans^{1*} 
 Iain Barber² and
 Ned A. Dochtermann³ 

Abstract

Predators can shape genetic correlations in prey by altering prey perception of risk. We manipulated perceived risk to test whether such non-consumptive effects tightened behavioural trait correlations in wild-caught stickleback from high- compared to low-risk environments due to genetic variation in plasticity. We expected tighter genetic correlations within perceived risk treatments than across them, and tighter genetic correlations in high-risk than in low-risk treatments. We identified genetic variation in plasticity, with genetic correlations between boldness, sociality, and antipredator morphology, as expected, being tighter within treatments than across them, for both of two populations. By contrast, genetic correlations did not tighten with exposure to risk. Tighter phenotypic correlations in wild stickleback may thus arise because predators induce correlational selection on environmental components of these traits, or because predators tighten residual correlations by causing environmental heterogeneity that is controlled in the laboratory. Our study places phenotypic integration firmly into an ecological context.

Keywords

Behavioural ecology, evolutionary characters, functional integration, G-matrix evolution, gene–environment interactions, modularity, morphology, personality, plasticity integration, predation.

Ecology Letters (2020) 23: 107–118

INTRODUCTION

Selection often favours trait combinations (Lande & Arnold 1983; Kingsolver *et al.* 2001; Sinervo & Svensson 2002). For example, garter snakes (*Thamnophis ordinoides*) exhibiting alternate colour patterns and anti-predator behaviours survive attempted predation best when pattern and behaviour are paired in certain combinations (Brodie 1992). Such selection can generate adaptive genetic correlations (Sinervo & Svensson 2002; Roff & Fairbairn 2012) and lead to optimal trait integration required for successfully executing particular functions (Pigliucci & Preston 2004; Klingenberg 2008, 2014). While such correlational selection characterises predator-sympatric populations (Brodie 1992, 1993), this may not be true for predator-free populations (Bell 2005; but see Herczeg *et al.* 2009). Population-specific ecological conditions may thus lead to population-specific genetic correlations (Armbruster & Schwaegerle 1996).

Population-specific trait correlations are well documented in three-spined stickleback (*Gasterosteus aculeatus*), particularly for behaviour (Bell 2005; Dingemans *et al.* 2007). In Welsh populations, behaviours associated with predation risk (aggressiveness, exploration, predator-inspection) were more tightly correlated within six populations with a history of fish predation than within six predator-naïve populations (Dingemans *et al.* 2007, 2010); similar patterns characterise North

American three-spined stickleback (Bell 2005). Predators thus shape behavioural correlations (Bell & Sih 2007; Adriaenssens & Johnsson 2013; Abbey-Lee & Dingemans 2019).

Predators can affect trait correlations in prey consumptively or non-consumptively (Lima 1998). They can selectively remove prey with particular trait combinations, a consumptive effect, or affect phenotypic correlations by inducing plastic responses in prey, a non-consumptive effect ultimately evolved from selection caused by the consumptive effect. The non-consumptive effect can result from individual (e.g. genetic) variation in plasticity (Kraft *et al.* 2006). For example, in a study with birds investigating how perceived risk affected timing of egg laying, individual plasticity underpinned changes in behavioural correlations, because explorative birds bred late under low-risk (positive correlation) but shifted to breed early under high-risk (negative correlation) (Abbey-Lee & Dingemans 2019).

In three-spined sticklebacks, anti-predator behaviours are weakly correlated in low-risk populations, but tightly correlated in high-risk populations (Dingemans *et al.* 2010). This may result from individual plasticity (Kraft *et al.* 2006; Bell & Sih 2007; Adriaenssens & Johnsson 2013) as illustrated in Fig 1a, depicting reaction norms (lines) across risk contexts for five genotypes and two behaviours (boldness and aggression) based on individual-level patterns described by Bell & Sih (2007). Here, rank order differences between genotypes change

¹Behavioural Ecology, Department of Biology, Ludwig Maximilians University of Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany

²School of Animal, Rural and Environmental Sciences, Nottingham Trent University, Brackenhurst Campus, Brackenhurst Ln, Southwell NG25 0QF, UK

³Department of Biological Sciences, North Dakota State University, 1340 Boiley Drive, Fargo, ND 58102, USA

*Correspondence: n.dingemans@lmu.de

between low- and high-risk for aggression (panel 2) but not for boldness (panel 1). Genetic variation in plasticity is essentially absent for boldness (panel 1) but substantial for aggression, showing crossing reaction norms (panel 2). As only some (1, 3, 5) genotypes show cross-context consistency for aggression, boldness and aggression are genetically positively correlated under both low-risk (panel 3) and high-risk (panel 4) but tighter under high-risk. Gene–environment interactions can thus change genetic correlations without requiring microevolution (Stearns *et al.* 1991; Sgro & Hoffmann 2004; Kraft *et al.* 2006; Wood & Brodie 2015). This scenario is plausible as we previously demonstrated genetic variation in plasticity for various anti-predator traits (Dingemanse *et al.* 2009b).

Biologically, non-consumptive effects on phenotypic correlations (r_P) can also result from changes in plasticity integration or heritabilities. Statistically, phenotypic correlations result from genetic (r_G) and residual correlations (r_R), whose respective influences are weighed by geometric mean heritabilities

$$r_P = r_G \sqrt{h_A^2 h_B^2} + r_R \sqrt{(1 - h_A^2)(1 - h_B^2)} \quad (1)$$

Residual correlations result from individuality in unmeasured environmental factors affecting multiple traits (i.e. plasticity integration), and are thus sometimes called “environmental correlations” (Searle 1961), despite also accumulating other sources of variation (e.g. measurement error). As predator presence generates spatiotemporal variation in risk (Lima & Bednekoff 1999), residual correlations can tighten in predator–sympatric populations, if individuals perceiving higher risk up-regulate multiple behaviours. Perceived risk may, for example, differ between edge vs. central positions in shoals formed in response to predators (Krause & Ruxton 2002; Ward *et al.* 2004). As behavioural-genetic correlations are often tighter than residual counterparts (Dochtermann 2011; see also Niemelä & Dingemanse 2018), phenotypic correlations can further tighten by increases in geometric mean heritability ($\sqrt{h_A^2 h_B^2}$). This is consistent with previous findings as perceived risk influences trait heritability in sticklebacks (Dingemanse *et al.* 2009b).

Here, we extend prior analyses of Welsh stickleback data (Dingemanse *et al.* 2009b) to test whether perceived predation risk strengthens trait correlations. We nested a perceived risk treatment within 84 full-sib families bred with standard crossing designs to establish individual plasticity as a non-consumptive mechanism affecting correlations between three behaviours (1. exploratory activity, 2. predator-inspection boldness, 3. solitariness). We applied a quantitative genetics approach to distinguish between effects of perceived risk resulting from treatment-specific i) genetic correlations, ii) residual correlations, and/or iii) geometric mean heritabilities. We viewed each behaviour (1–3) in each treatment (low risk, L, versus high risk, H), as a distinct “trait”, and estimated within- and cross-treatment genetic correlations between all six traits (1_L, 1_H, 2_L, 2_H, 3_L, 3_H; Fig 2). We then assessed relative support for five *a priori* considered hypotheses (see the legend of fig 2), implemented as structural equation models (SEMs) fitted to the genetic correlation matrix. We also applied these comparisons to morphology (body depth, body

and spine length). As anti-predator behaviour and morphology are often integrated (Niemelä & Dingemanse 2018), particularly when predators are present (Kim & Velando 2015), we further studied whether treatment affected genetic correlations between behaviour and morphology.

We replicated our study by sampling two populations. One population (Llyn Alaw) inhabited a large reservoir with a predation history exceeding 30 generations, the other (Cae Mawr) a small man-made pond with no history of piscivorous predation (Dingemanse *et al.* 2007). We compared treatment effects between these populations to assess whether effects were general versus population-specific. As these populations differed in various ecological factors (e.g. pond size, predation regime), the absence of population differences in treatment effects would indicate that conclusions were potentially more broadly generalizable (Wilson 1998), thereby answering calls for study replication in behavioural ecology (Kelly 2006; Nakagawa & Parker 2015).

METHODS

Study populations

Both populations were from man-made lakes on the island of Anglesey (North Wales, U.K.). Llyn Alaw (N50°20'23" W04°26'20") is a 3.09 km² reservoir formed by the impoundment of River Alaw in 1966; Cae Mawr (N50°17'06" W04°23'31") is a 214 m² pond constructed ca. 1980 with no natural input or outflow. The two were 5.5 km apart and experienced different predation regimes (Dingemanse *et al.* 2007). Predatory fish (perch, *Perca fluviatilis*, and rainbow trout *Oncorhynchus mykiss*) were introduced into Llyn Alaw immediately after impoundment; stocking programmes maintained predator populations since. Predatory fish were neither introduced into Cae Mawr nor observed during ecological surveys.

Experimental protocol

We provide here a summary of the experimental protocol detailed elsewhere (Dingemanse *et al.* 2009b). Adult stickleback were captured in May 2006 and housed at Aberystwyth University, U.K. We applied split-clutch *in vitro* fertilisation techniques and standard egg husbandry protocols (Barber & Arnott 2000). We conducted partial North Carolina II breeding designs and full crosses (Lynch & Walsh 1998), involving a total of 22 males and 15 females from Llyn Alaw, and 30 males and 28 females from Cae Mawr (see Table S3 in Dingemanse *et al.* 2009b). Following egg fertilisation (day 0), each portion of the split-clutch was incubated in isolation. At hatching (day 8), fry in each split-clutch were divided equally between two 30 × 20 × 20 cm³ tanks (hereafter, ‘holding tanks’; 14.72 ± 0.466 (mean ± SE) fry per holding tank). One of two holding tanks was randomly assigned to a low-risk, the other to a high-risk treatment. We thereby generated four groups (i.e. two treatment groups for each population). Fry were fed a progressive diet of LiquifryTM (days 6–14; twice a day), and newly hatched *Artemia* (from day 10 onwards; *ad libitum*). For further details, see Dingemanse *et al.* (2009b).

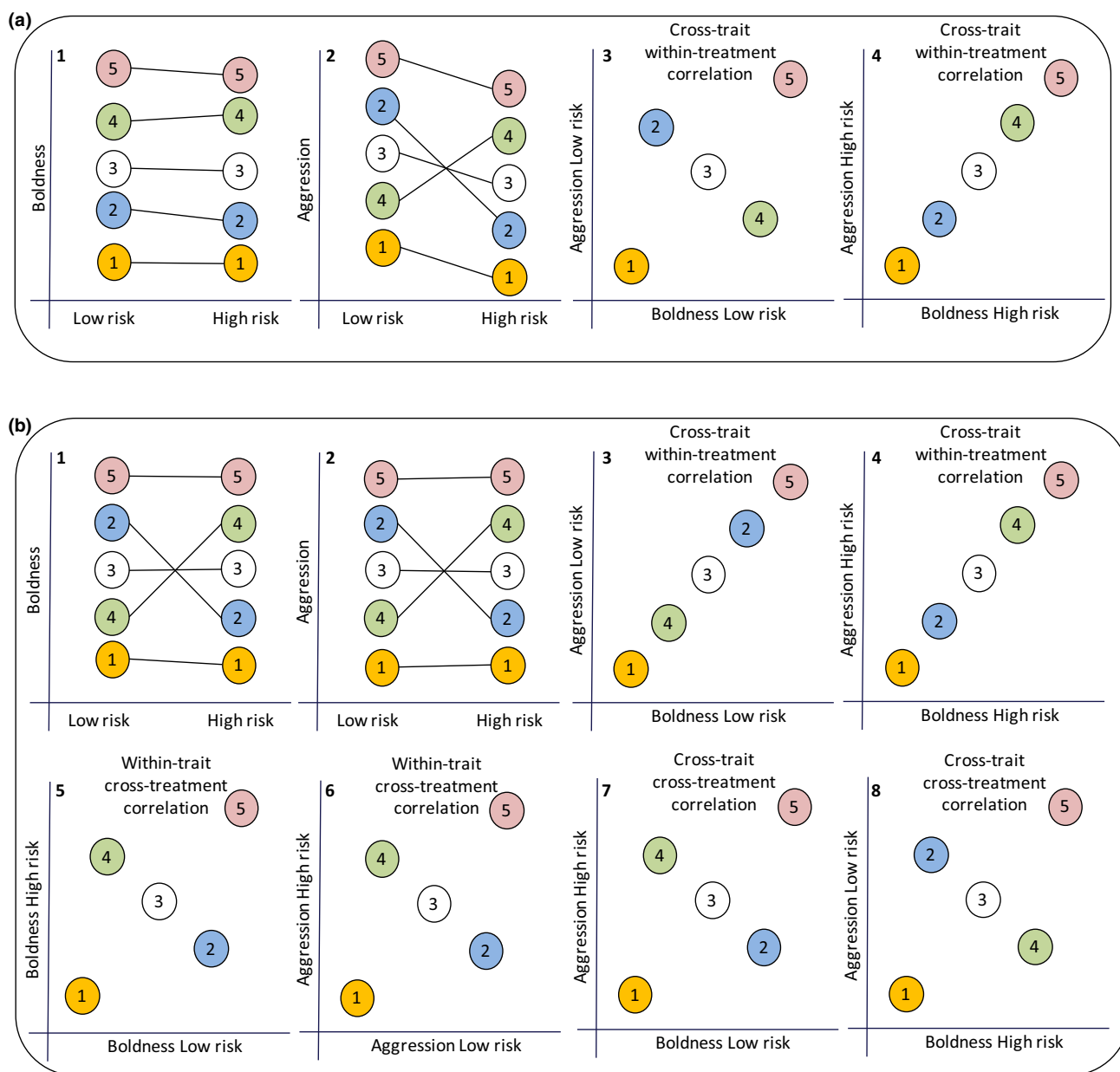


Figure 1 (a) Conceptual illustration of the hypothesised role of genetic variation in phenotypic plasticity as a mechanism explaining differences between low-risk and high-risk environments in the strength of genetic correlations. We sketch here a simple scenario where genetic correlations vary with perceived risk because of genetic variation in plasticity in only one trait. The example is inspired by patterns of individual variation in aggressiveness and boldness described for stickleback by Bell & Sih (2007). We show here reaction norms for five genotypes (numbers) as a function of a manipulated two-level environmental gradient (low vs. high perceived risk) for boldness (panel 1) and aggressiveness (panel 2). Genetic variation in phenotypic plasticity is largely absent boldness but is substantial for aggressiveness, with only the latter showing crossing reaction norm slopes. The example also illustrates Bell & Sih's (2007) finding that aggressiveness but not boldness decreases with perceived risk within the average individual. As some genotypes (1, 3, 5) show cross-context consistency for both traits, the genetic correlation between aggression and boldness is positive within both the low-risk (panel 3) and the high-risk (panel 4) treatment. Genetic correlations are nevertheless tighter under high-risk (panel 4) because some genotypes (2, 4) exhibit genetic variation in plasticity in aggressiveness. (b) Conceptual illustration of the multivariate pattern of genetic variation in phenotypic plasticity supported by our study. Here, genetic variation in phenotypic plasticity exists for both aggression and boldness as both show crossing reaction norm slopes (boldness: panel 1; aggressiveness: panel 2). Within-treatment genetic correlations (panels 3 and 4) are tighter than cross-treatment genetic correlations (within-traits: panels 5-6; cross-traits: panels 7-8) but within-treatment genetic correlations do not differ between treatment groups (compare panel 3 and 4). The tighter within-versus cross-treatment genetic correlations give rise to support for structural equation models hypothesising two treatment-specific latent variables; nonzero cross-treatment genetic correlations lends further support to a model hypothesising a correlation between these two latent variables (as in Model 1E; Fig 2)

Holding tanks assigned to the high-risk treatment were given chemical, visual and behavioural stimulation to mimic conditions experienced when living sympatrically with

predatory fish. The treatment consisted of repeated visual exposures to a live perch (a 30-min 'live predator trial' undertaken on alternate days between day 29 and day 49),

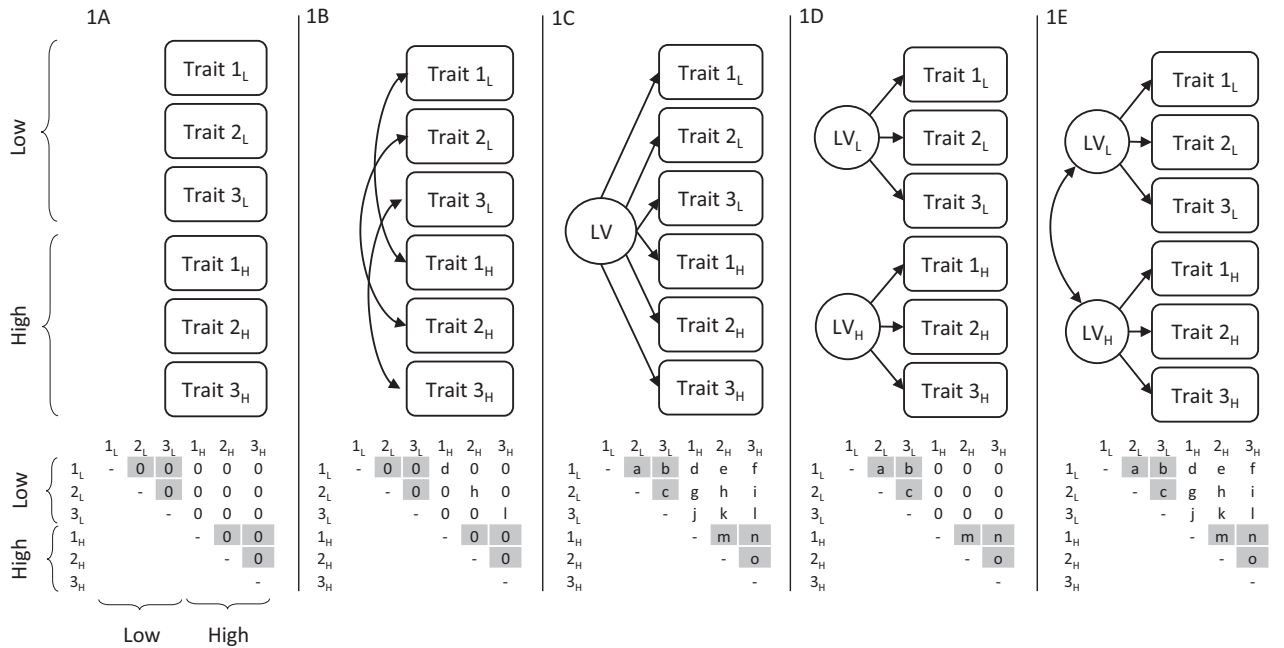


Figure 2 Patterns of hypothesised correlation structures with associated structural equation models (SEMs). Viewing each behaviour (1-3) in each treatment (low risk, L, versus high risk, H) as a distinct ‘trait’, we estimated within- and cross-treatment genetic correlations between all six traits (1_L , 1_H , 2_L , 2_H , 3_L , 3_H). We then assessed relative support for five *a priori* considered hypotheses, implemented as SEMs fitted to the genetic correlation matrix. These SEMs differed in which elements (correlations) in the matrix were assumed zero (“0”) versus non-zero (letters a-o). Grey blocks in the correlation matrix indicate within-treatment correlations, all other blocks cross-treatment correlations. Model 1A posits no genetic correlations (Coleman & Wilson 1998), hypothesising a correlation matrix containing only zeros. Model 1B posits (positive) genetic correlations between the *same* behaviour expressed in different treatments (as previously documented for this dataset; Dingemanse et al. 2009b), indicated by letters *d*, *h*, *l* in its hypothesised correlation matrix, but no correlations between the three *different* behaviours (similar to Model 1A). Model 1C posits an overarching behavioural syndrome (Sih et al. 2004), where all traits are correlated (letters *a*-*o*). Model 1D posits genetic correlations between the behaviours within low-risk (letters *a*-*c*) and high-risk treatments (letters *m*-*o*) but no cross-treatment genetic correlations, and therefore support for two treatment-specific latent variables (LV_L and LV_H). Model 1D is implausible, owing to previous evidence for positive cross-treatment genetic correlations within the same behaviour (see above). Model 1E addresses this issue by allowing cross-treatment genetic correlations between the latent variables (double-headed arrow connecting LV_L and LV_H), which we hypothesised to be positive based on previous analyses (Dingemanse et al. 2009b). Models 1C-1E are all compatible with our hypothesis that perceived risk can tighten correlations, provided that genetic correlations are weaker under low-risk (letters *a*-*c*) versus high-risk (letters *m*-*o*), and provided that path coefficients connecting latent variables to observed traits (single-headed arrows) are also weaker under low- vs. high-risk

reinforced by ‘chasing’ by a model perch (a 2-min ‘chasing trial’ every four days between day 30-46), while simultaneously introducing a chemical stimulus (water that had held live perch). The chasing trial ensured that ‘high-risk’ treatment stickleback perceived the perch as threatening. Each trial was executed at a random time of day (08h00-18h00). ‘Low-risk’ holding tanks were given the same treatment, with the exception that (i) instead of a live perch, a similar-sized stone was introduced; (ii) instead of perch water, water was added from a source that had not held perch; and (iii) instead of moving the perch model through the water, a metal rod was used with no perch model attached. All trials were executed at similar times of the day, thereby ensuring that the timing of encounters with predators was equally unpredictable. For detailed methodology see Dingemanse *et al.* (2009b).

At ages 44, 46, 48, 50 and 52 days, one or two randomly selected individuals were captured from each holding tank. Subsequently, each individual was behaviourally assayed (fully detailed in Dingemanse *et al.* 2009b). Briefly, following capture, a focal subject was released alone in a similar-sized novel tank and its exploratory activity filmed for 300 s (“novel environment test”; conducted between 12h30-13h15). Subsequently, a

transparent barrier was placed, and the subject left to acclimate overnight. The following morning, an unrelated, age- and population-matched stimulus conspecific was introduced into the empty compartment to quantify the subject’s tendency to shoal for 300 s (“sociability test”; between 8h30-10h00), and then removed. Two hours later, a live perch was introduced behind the barrier, and the subject’s behaviour recorded for 10 minutes (“predator inspection test”; between 10h00-12h00). All tests were filmed from above using digital video cameras (Model GR-D340EK, JVC, Yokohama, Japan) and behavioural measures taken from the recordings (listed in Table S4 in Dingemanse *et al.* 2009b). Subjects were subsequently sacrificed with an overdose of Benzocaine anaesthetic (10 mgL^{-1}) and photographed laterally to acquire three morphological measurements (detailed in Dingemanse *et al.* 2009a): standard length (horizontal distance from the anterior tip of the upper lip to the caudal border of the hypural plate), body depth (depth of the body at the base of the first dorsal spine), and length of the first dorsal spine (horizontal distance from the base to the tip of the first dorsal spine). Sample sizes were 577 (Cae Mawr), 464 (Llyn Alaw) individuals, but rare events (e.g. camcorder malfunctions) slightly reduced sample sizes for some assays.

Statistical analyses

Behavioural summaries

We scored multiple behaviours within each test type, which we summarised using Principal Component Analyses (PCA) to acquire a single integrative measure per test type. This reduced the number of ‘traits’ in our analyses to one per test. Common PCA confirmed that test-specific phenotypic variance–covariance input matrices did not differ between populations/treatments; we therefore ran a single PCA for all data per test type (see Table S4 in Dingemanse *et al.* 2009b). PCA summarised response variables into principal components (PCs), two (sociability test) or one (all other tests) of which were significant (Dingemanse *et al.* 2009b). Briefly, the single PC (Comp. A1; variance explained: 82%) extracted for the novel environment test had high positive loadings for movement variables and was labelled “exploratory activity”. The first PC (Comp. D1; 40%) extracted for the sociability test measured lack of sociability (labelled “solitariness”), with high positive loadings for both latency to reach the stimulus conspecific and mean distance between the focal and stimulus during the test: social fish approached the conspecific quickly and stayed close (without showing aggression), whereas solitary fish remained near the refuge. The second PC (Comp. D2; 39%) aggregated variables measuring activity and was not taken forward as it did not measure the target behaviour. The single PC (Comp. E1; 59%) extracted for the predator-inspection test described shyness, with high negative loadings for variables describing movement and number of inspections, and high positive loadings for latency to begin moving and minimum approach distance: inspecting fish approached the predator (shortly following exposure), turned, returned to cover and initiated new inspections, whereas non-inspecting fish remained still and did not approach. For further analyses, PC E1 was multiplied by -1 to reflect “boldness” instead.

Estimation of quantitative genetic parameters

We applied multi-response animal models (Kruuk 2004; Wilson *et al.* 2010) implementing the known pedigree stemming from the partial North Carolina II design and full crosses (Lynch & Walsh 1998). This allowed the partitioning of phenotypic (co)variances into additive genetic and residual (co)variances. The number of response variables differed between models as detailed below. The six traits were treated as distinct response variables for each treatment (e.g. ‘exploratory activity under low-risk’ versus ‘exploratory activity under high-risk’), and all 66 within- and cross-treatment genetic covariances estimated. Data of each population was analysed in two parts. We first generated **G**-matrices for testing a set of *a priori* considered alternative hypotheses, detailed in Fig 2 (Models 1A–1E), for behaviour and morphology separately. We fitted four multivariate models with six response variables each (i.e. 3 traits \times 2 environments; see Fig 2), one for each combination of population and trait type (behaviour/morphology). Combined behaviour–morphology **G**-matrices were then used to test hypotheses 2A–2C (detailed in the legend of Fig 3), which concerned the integration of behaviour and morphology, and were generated *a posteriori* based on the best-fitting model arising from separate analyses of behaviour and morphology. We fitted two multivariate models

with twelve response variables each (i.e. 6 traits \times 2 environments), one for each population. These multi-response models also estimated **R**-matrices, estimating all 30 within-treatment residual covariances (15 for the low-risk and 15 for the high-risk treatment). The 36 cross-treatment residual covariances were non-estimable because treatment did not vary within individuals (Dingemanse & Dochtermann 2013). All models estimated trait-specific effects of age fitted as a fixed effect covariate; common environment (i.e. growth tank ID) and maternal effects (i.e. dam ID) were previously found to be non-significant (Dingemanse *et al.* 2009b) and not modelled to avoid over-parametrised models. **G**- and **R**-matrices derived from our two 12-response variable models matched those from our simpler four 6-response variable models (Results not shown). **G**- and **R**-matrices are provided in Table S1; our Results solely focus on relative fit of alternative hypotheses (implemented as SEMs, detailed below) explaining multivariate structures.

We used a Bayesian framework to model phenotypic traits assuming Gaussian errors with Markov chain Monte Carlo sampling of 1.04×10^6 iterations, with a 2.4×10^5 iteration burn-in, thinning intervals of 800. Despite the complexity of the models, these chain lengths led to good mixing even of the twelve trait models, though mixing was better within treatment than between (average effective sample size out of 1000: **G**_{between}: 740.2, **G**_{within}: 896.2, **R**_{within}: 903.4). We used a minimally informative inverse-Wishart ($V = 1$, $\nu = 0.002$) following Hadfield (2010). We used the MCMCglmm package (Hadfield 2010) in *R* (v3.3.0, R Core Team 2016), estimating additive genetic variances, heritabilities, additive genetic covariances (correlations), and residual variances and covariances (correlations).

Animal models estimated population- and treatment-specific narrow-sense heritabilities, and trait-specific cross-context genetic correlations. These were quantitatively similar to those based on previous REML-based analyses of these data (Dingemanse *et al.* 2009b), though behavioural heritability estimates were somewhat higher, and trait-specific cross-context genetic correlations somewhat lower (see Table S1). Briefly, (i) behavioural heritabilities were lower than morphological heritabilities, (ii) heritabilities were generally higher for Cae Mawr vs. Llyn Alaw, and (iii) exposure to predators during ontogeny did not consistently decrease/increase the expression of additive genetic variance (Table S1 and Dingemanse *et al.* 2009b). Our Results focus exclusively on estimates of parameters not previously published.

Structural Equation Modelling (SEM)

We used SEM-comparisons to assess behavioural and/or morphological trait integration (Dochtermann & Jenkins 2007; Dingemanse *et al.* 2010; Roff & Fairbairn 2011; Araya-Ajoy & Dingemanse 2014; Moiron *et al.* 2019). Because we aimed to evaluate specific hypotheses of trait correlation, we compared SEMs (Dochtermann & Jenkins 2007) rather than describing treatment-specific differences in major axes of variation (Aguirre *et al.* 2014). We compared *a priori* models (Figs 2 and 3) using the estimated additive genetic correlation matrices as input. SEMs were fitted using Lavaan (Rossee 2012) in *R* (v3.3.0, R Core Team 2016) and models with different structures compared using Akaike’s Information

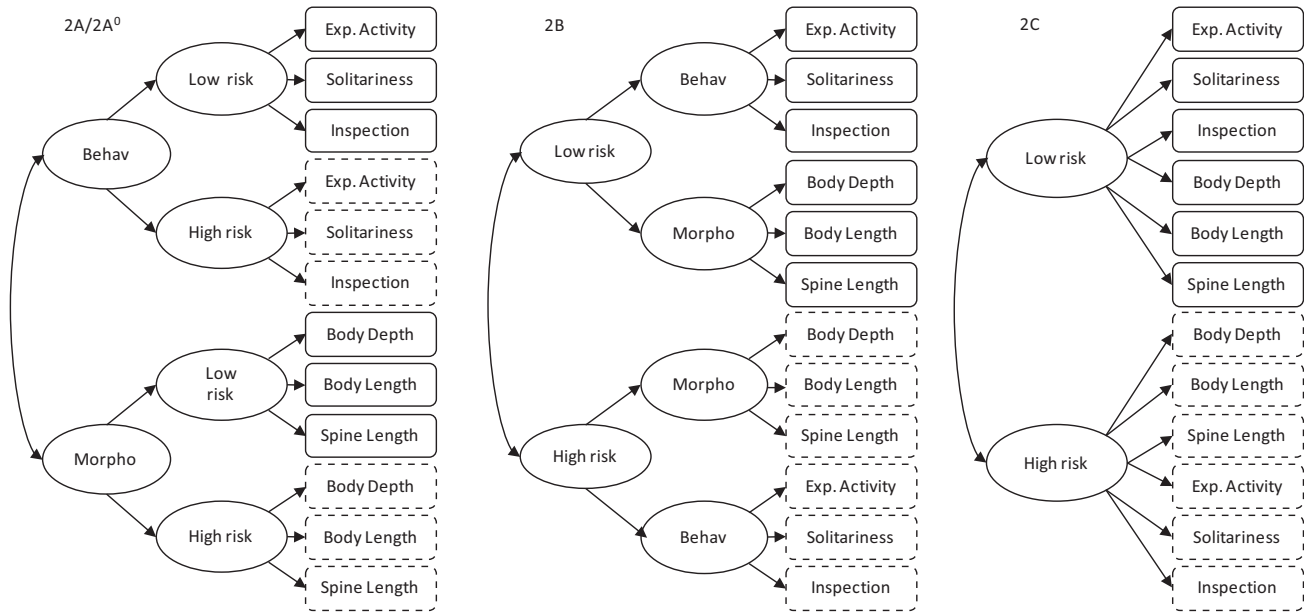


Figure 3 Visualisations of models (hypotheses) explaining correlations among three behavioural and three morphological traits. Models described here were constructed *a posteriori* to be compatible with tighter within- versus cross-treatment genetic correlations posited by Model 1E (Fig 2), hypothesising a structure with two genetically correlated treatment-specific latent variables supported in separate analyses of behaviour versus morphology (Fig 4). Solid boxes are for traits measured in the low-risk and dashed boxes for the same traits measured in the high-risk treatment. **Model 2A** proposes tighter within-versus cross-treatment genetic correlations (as in Model 1E) for the behavioural versus morphological traits separately, combined with an overall treatment-independent genetic correlation between behaviour and morphology (indicated by a double-headed arrow between latent variables “Behav” and “Morpho”). A genetic correlation between behaviour and morphology of zero would equate to removing this double-headed arrow (**Model 2A⁰**); **Model 2B** instead proposes that all six traits (whether behaviour or morphology) are genetically correlated (double-headed arrow) but more tightly so within- versus cross-treatments. At the same time, behaviour-morphology correlations within treatments are weaker compared to within-treatment correlations among the three behaviours or among the three morphological traits. **Model 2C** alleviates this latter assumption, but is otherwise identical to Model 2B

Criterion (AIC) scores. For the 1000 posterior estimates of the correlation matrix, we calculated each SEM’s AIC relative to the SEM that best fit that posterior correlation matrix ($\Delta\text{AIC} = 0$), thereby taking forward uncertainty in correlation estimates (Mutzel *et al.* 2013; Araya-Ajoy & Dingemanse 2014). We then calculated each SEM’s model likelihood as $\exp(-0.5 \times \Delta\text{AIC})$ (Burnham & Anderson 1998), and the proportion of times (out of 1000 posteriors) that the focal SEM had a ΔAIC value equal to zero (Table 1). Model likelihoods therefore represent support for a particular SEM and the proportion an SEM best fit the data an estimate of certainty around that support. For one posterior estimate, models could not be fit and reported proportions based on 999 posterior estimates (Table 1).

We verified that SEM-based interpretations were consistent with a visual inspection of **G**-matrix elements (Table S1). Importantly, **G**-matrix elements were estimated with considerable uncertainty; inferences made on an element-by-element basis would thus also be uncertain. By contrast, SEMs tested multivariate (rather than bivariate) hypotheses utilising the entire **G**-matrix, resulting in greatly increased statistical power (Roff & Fairbairn 2001; Grace 2006); this explains why strong support for a particular SEM structure is compatible with uncertain matrix elements.

The specific set of SEMs considered for describing genetic (**G**) correlation structure (Figs 2 and 3) could not be fitted to residual (**R**) correlation matrices because treatment did not

vary within individuals. We thus applied an alternative approach: estimates of residual correlations were transformed into absolute values, and then averaged over all estimates for each of the four unique combinations of treatment and population.

RESULTS

Genetic integration

Behavioural genetic integration

For both populations, patterns of variation among the fifteen within- and among-treatment behavioural genetic correlations best supported Model 1E (Fig 2; model likelihoods: Cae Mawr: 0.999; Llyn Alaw: 1.000, Table 1). Model 1E was supported $0.731/0.181 = 4.03$ (Cae Mawr) and $0.714/0.194 = 3.68$ (Llyn Alaw) times more often than the next best model, Model 1C. Other models fitted poorly (Table 1a). The strength of path coefficients was similar between populations (Fig 4); overlapping 95% credible intervals (CIs) (Table S2) implied no statistical differences.

Model 1E implied genetic correlations between behaviours *within* low-risk and *within* high-risk treatments, and weaker cross-treatment genetic correlations, and two genetically correlated treatment-specific latent variables (labelled “Low risk” and “High risk” in Fig 4). Visual inspection of genetic correlation matrices (Table S1) confirmed this conclusion for two

Table 1 Comparisons of relative fit of alternative structural equation models (SEMs) explaining additive genetic correlation structure among three behavioural and three morphological traits as a function of ontogenetic experience with predators (low-risk versus high-risk) for each of two populations (Cae Mawr and Llyn Alaw). We present here the relative fit of five alternative hypotheses predicting integration of traits within behavioural and morphological domains (Models 1A-1E; visualised in Fig 2); we also present the relative fit of four alternative hypotheses predicting the combined integration of behavioural and morphological traits (Model 2A-2C; visualised in Fig 3). For each posterior ($n = 1000$) of the additive genetic correlation matrix, we calculated each model's Akaike's information criterion (AIC) value relative to the model that best fitted the data (Δ AIC), from which we calculated corresponding model likelihoods. From this, we report the posterior mode of the model likelihoods and the proportion of posteriors ($n = 1000$) where the focal model was the best-fitting model (Δ AIC = 0)

		Model 1A	Model 1B	Model 1C	Model 1D	Model 1E
(a) Population	Domain	model likelihood (proportion best fitting)				
Cae Mawr	Behaviour	0 (0)	0 (0.067)	0 (0.181)	0 (0.021)	0.999 (0.731)
	Morphology	0 (0)	0 (0)	0 (0)	0 (0.001)	1.000 (0.999)
Llyn Alaw	Behaviour	0 (0)	0 (0.071)	0 (0.194)	0 (0.021)	1.000 (0.714)
	Morphology	0 (0)	0 (0)	0 (0)	0 (0.064)	1.000 (0.936)
		Model 2A	Model 2A ⁰	Model 2B	Model 2C	
(b) Population	Domain	support value				
Cae Mawr*	Combined	1.000 (0.828)	0 (0.026)	0 (0.141)	0 (0.005)	
Llyn Alaw	Combined	1.000 (0.716)	0 (0.059)	0 (0.222)	0 (0.003)	

*Models could only be fit to 999 of the 1000 posteriors, proportions were calculated accordingly.

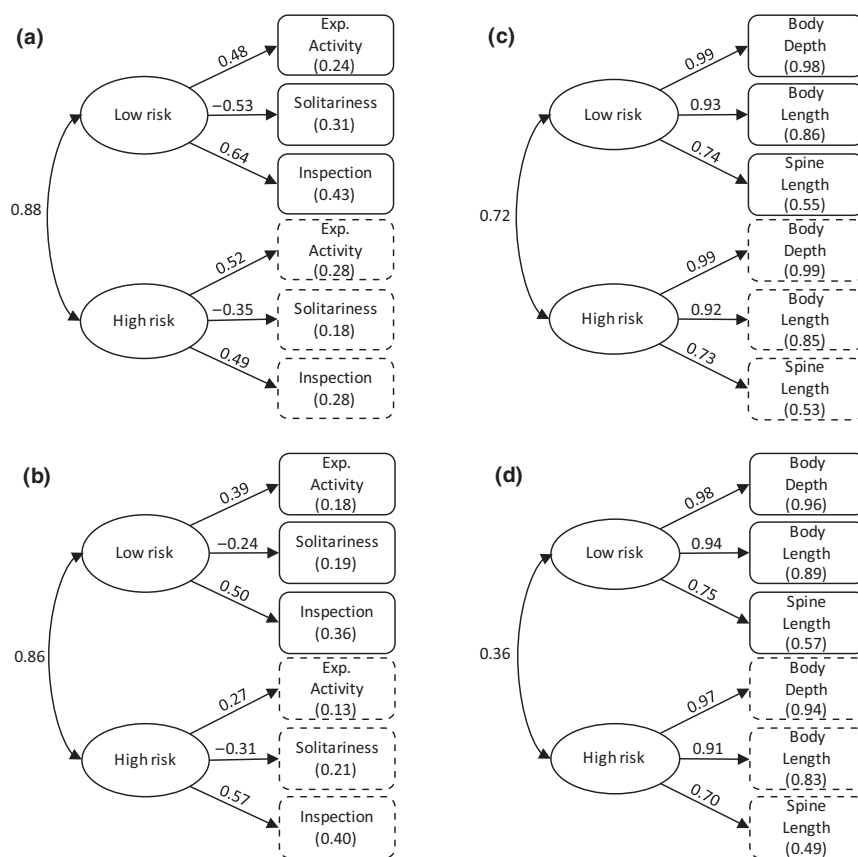


Figure 4 Parameter estimates of the structural equation model that best described the integration of behavioural traits for (a) Cae Mawr and (b) Llyn Alaw sticklebacks, and morphological traits for (c) Cae Mawr and (d) Llyn Alaw sticklebacks. Printed values are parameter estimates for factor loadings (single-headed arrows) and correlations (double-headed arrows). In all cases, traits were correlated within and across treatments but more tightly so within treatments, giving rise to support for a SEM with two correlated yet treatment-specific latent variables (Model 1E; Fig 2)

reasons: First, exploratory activity, solitariness, and predator-inspection behaviour were genetically correlated as individuals genetically predisposed towards expressing higher levels of

exploratory activity, were also genetically predisposed to inspect predators and to show a greater willingness to shoal. Second, cross-treatment genetic correlations were weaker

compared to within-treatment counterparts, thus supporting treatment-specific latent variables. The genetic correlation between the latent variables was positive, matching our finding of positive within-trait cross-treatment genetic correlations (Table S1).

Behavioural correlations are tighter in predator-sympatric populations (Dingemanse *et al.* 2007), which we hypothesised to result from genetic variation in plasticity, causing tighter genetic correlations in high-risk treatments (Fig 1a). The cross-treatment genetic correlation between the treatment-specific latent variables confirmed the existence of genetic variation in plasticity because it was below one (Fig 4; Table 1). Full support would also require tighter genetic correlations and tighter path coefficients connecting latent variables to observed traits in the high- vs low-risk treatment (single-headed arrows in Fig 4). However, neither genetic correlations (Table S1) nor path coefficients were weaker under low-risk (owing to overlapping 95% CIs; Table S2). Fig 1b conceptually illustrates why genetic variation in plasticity may not have resulted in treatment-specific genetic correlations.

Morphological genetic integration

Model 1E (Fig 2) also best described variation among all genetic correlations between morphological traits (model likelihoods: Cae Mawr: 1.000; Llyn Alaw: 1.000; Table 1a); all other models fitted poorly (Table 1a). As was the case for behaviour, Model 1E was also consistently ranked best (proportion of posteriors where $\Delta AIC = 0$: Cae Mawr: 0.999, Llyn Alaw: 0.936; Table 1a). Body length, body depth, and spine length were genetically more positively correlated within versus among treatments (Figure S1), leading to a positive genetic correlation between the two latent variables below one (Fig 4), implying genetic variation in plasticity. Notably, this genetic correlation was much stronger for Cae Mawr (mode, 95% CIs: 0.72, 0.39, 0.91) compared to Llyn Alaw (0.37, -0.06, 0.69) but 95% CIs overlapped, implying that Llyn Alaw did not exhibit significantly more $G \times E$. As found for behaviour, path coefficients were very similar between populations/treatments (Table S2), resulting in a failure to support the hypothesis that plasticity tightened correlations under high-risk.

Genetic integration of behaviour and morphology

We performed comparisons of four *a posteriori* derived SEMs (Fig 3) to test alternative hypotheses compatible with our finding that Model 1E best fit both behavioural (Figs 4a and b) and morphological (Figs 4c and d) domains. Of these, Model 2A best fitted both populations (model likelihoods Cae Mawr: 1.000; Llyn Alaw: 1.000; Table 1b). While Model 2B had a posterior modal likelihood of 0 for both populations, it was occasionally the top model (Table 1b). Consistent with model likelihoods, this support for Model 2B was weak: Model 2A was supported $0.828/0.141 = 5.87$ (Cae Mawr) and $0.716/0.22 = 3.25$ times more often than Model 2B; other models fitted poorly (Table 1b).

Model 2A described a structure with two genetically positively correlated treatment-specific latent variables for behaviour (detailed above), two genetically positively correlated treatment-specific latent variables for morphology (detailed above), and an overarching genetic integration between behaviour and morphology independent of treatment (Fig 5). A

positive genetic correlation between behaviour and morphology also characterised both Cae Mawr ($r_A = 0.43$) and Llyn Alaw ($r_A = 0.39$). These estimates were relatively uncertain (their 95% CIs spanned the entire range; Table S2), but this connection between behaviour and morphology was strongly supported as a model where the correlation was set to zero (Model 2A⁰) fitted poorly (Table 1b). Individuals genetically predisposed towards higher exploratory activity, predator-inspection behaviour, and shoaling, were thus also genetically predisposed to be larger (Table S1).

Besides the same structure of integration being supported in both populations (Table 1b, Fig 5), the strength of connections among traits was also highly consistent between them. Specifically, the 95% CIs associated with path coefficients for Model 2A were highly concordant between populations (Table S2).

Residual integration

Genetic and residual correlations were generally of the same sign (Table S1), though genetic correlations were stronger (Figure S1). Comparison across populations/treatments showed that average absolute residual correlations did not differ across treatments (Figure S1).

Geometric mean heritabilities

Genetic correlations were tighter than residual ones (see above). This implied that treatment effects on geometric mean heritability ($\sqrt{h_A^2 h_B^2}$, where A and B represent two different traits, could potentially alter the strength of phenotypic correlations (Eqn. 1). This mechanism was not important for behaviour, as the three calculable geometric mean behavioural heritabilities were relatively similar for the low-risk (mean among these three values, range: 0.37, 0.32-0.40) versus high-risk treatment (0.39, 0.35-0.43) in Llyn Alaw, though somewhat lower for the low-risk (0.32, 0.31-0.33) versus high-risk treatment (0.36, 0.33-0.44) in Cae Mawr. As expected, the three calculable geometric mean heritabilities were substantially higher among morphological traits, and again similar across treatments for both Cae Mawr (low-risk: 0.72, 0.70-0.74; high-risk: 0.71, 0.69-0.75) and Llyn Alaw (low-risk: 0.77, 0.75-0.81; high-risk: 0.75, 0.72-0.79).

DISCUSSION

Behaviours associated with predation risk are integrated in wild-caught three-spined stickleback, but more tightly so within predator-sympatric populations (Bell 2005; Dingemanse *et al.* 2007, 2010). Our manipulations allowed studying non-consumptive mechanisms by which predators affect trait correlations. We applied quantitative genetics to distinguish between effects of perceived risk on phenotypic correlations resulting from treatment-specific i) genetic correlations, ii) residual correlations, and/or iii) geometric mean heritabilities.

The first mechanism—non-consumptive effects on phenotypic correlations via effects of perceived risk on genetic correlations—requires genetic variation in plasticity. This could cause within-trait cross-treatment genetic correlations below

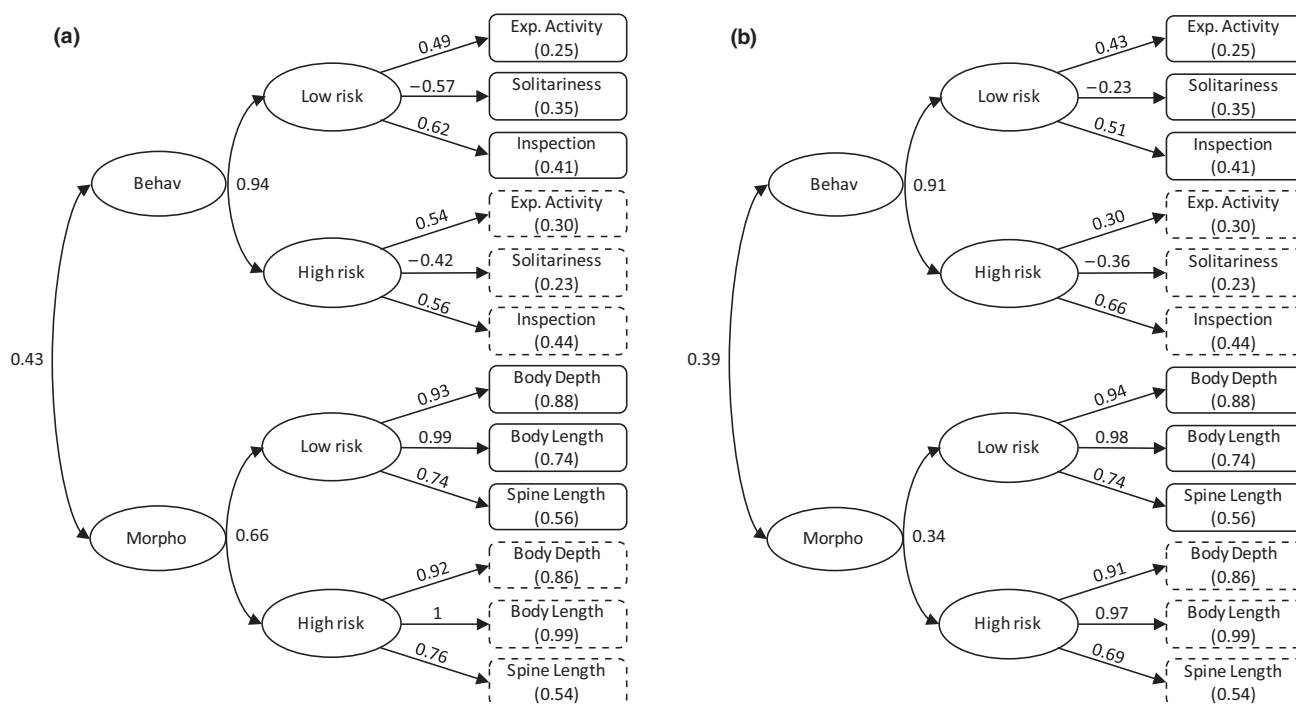


Figure 5 Parameter estimates of the structural equation model that best described the integration of behaviour and morphology for (a) Cae Mawr and (b) Llyn Alaw sticklebacks. Printed values are parameter estimates for factor loadings (single-headed arrows) and correlations (double-headed arrows). In both populations, within-treatment genetic correlations were tighter than cross-treatment genetic correlations for the behavioural versus morphological traits separately; this combined with an overall treatment-independent genetic correlation between behaviour and morphology supported Model 2A (Fig 3)

one, and tighter genetic correlations within the high-risk treatment (Fig 1a). While we experimentally confirmed a key component (i.e. genetic variation in plasticity), the hypothesis was not supported overall because genetic correlations did not tighten with increased risk.

The second mechanism—non-consumptive effects on phenotypic correlations via effects on residual correlations—would imply greater environmental heterogeneity in high-risk environments. This assumes—perhaps inappropriately—that residual correlations represented “environmental” correlations (*sensu* Searle 1961), capturing plasticity in response to temporary environmental effects (Gavrilets & Scheiner 1993; Whitman & Agrawal 2009; Berdal & Dochtermann 2019). Residual correlations might tighten with risk, for example, because predator presence generates spatiotemporal variation in risk (Lima & Bednekoff 1999), and because individuals perceiving higher risk up-regulate multiple anti-predator behaviours relative to individuals perceiving lower risk. Our analyses also did not support this mechanism: residual correlations did not differ between treatments (Figure S1).

The third mechanism—non-consumptive effects on phenotypic correlations via effects on geometric mean heritabilities—could act, provided that genetic and residual correlations differed (Equation 1). This was indeed the case: the former were stronger than the latter (Figure S1). Yet, we also found no support for this explanation because geometric mean heritabilities did not differ between treatments.

Lack of differences in phenotypic integration across treatments was unexpected, as we previously demonstrated greater

phenotypic correlations among predator-sympatric populations (Dingemans *et al.* 2007, 2010). Our current analyses imply that genetic integration among both behavioural and morphological traits does not vary greatly with major ecological differences, such as those created by risk treatments. Importantly, our study was replicated, documenting the same treatment effect in two populations with very different ecological backgrounds (see Introduction); this implies that **G**-matrix stability may generally characterise the Anglesey stickleback populations, rather than presenting a result conditional on the ecology of each population (Wilson 1998; Kelly 2006; Nakagawa & Parker 2015). Future research should address whether **G** is also stable with respect to other ecological gradients and across larger geographical scales. For example, Siren *et al.* (2017) have shown that **G**-matrix structure can be somewhat labile in this species.

We therefore conclude that differences in phenotypic integration among wild-caught stickleback from predator-naïve and predator-sympatric populations must stem from plasticity integration, arising from variation in environmental conditions that are unique to predator-sympatric environments, yet not considered in our experiment. For example, spatiotemporal variation in predation risk experienced by wild predator-sympatric stickleback populations may favour alternative strategies, such as solitary monopolisation of a safe yet unproductive foraging patch versus joining a roaming shoal taking risks to seek out rich food patches (Krause & Ruxton 2002). Individuals may plastically adjust multiple traits simultaneously, as solitary versus group living requires different

combinations of behaviours (Ward *et al.* 2004), a situation not made possible by our experimental setup. Alternatively, long-term population differences in predation may have led to population-specific genetic correlations (Armbruster & Schwaegerle 1996; Bell 2005). A firm test of this idea would require analyses statistically demonstrating differences in **G** between *types* of population, requiring greater replication than we currently have. Assuming that differences between our single predator-naïve (Cae Mawr) and predator-sympatric (Llyn Alaw) populations are representative, the lack of statistical evidence for population differences in SEM-based genetic correlation structure revealed in this paper, would suggest that this explanation is unlikely. If **G**-matrices indeed do not differ between types of population, predators may perhaps have direct consumptive effects by inducing correlational survival selection on the environmental components of these traits (Rauscher 1992), thereby tightening phenotypic correlations among parents that are not expressed in laboratory-bred offspring.

Our experiment demonstrated genetic variation in plasticity in response to perceived risk, as an SEM hypothesising tighter within- than cross-treatment genetic correlations (i.e. treatment-specific latent variables; Model 1E, Fig 2) was statistically supported both for behaviour and morphology, and across both populations (Table 1a; Fig 4). Genetic variation in plasticity, however, never produced an outcome of tighter genetic correlations in the high-risk treatment: path coefficients connecting each treatment-specific latent variable (ovals labelled 'low-risk' and 'high-risk' in Fig 4) to its treatment-specific traits (rectangles in Fig 4) were not different in the high-risk treatment (Table S2). This implies the expression of genes with pleiotropic effects specific to each treatment (Pavlicev & Wagner 2012), causing tighter within- versus cross-treatment genetic correlations. Importantly, strongly positive cross-treatment genetic correlations (double-headed arrows between the latent variables 'low-risk' and 'high-risk' in Fig 4) demonstrated a largely stable genetic correlation matrix, though perhaps less so for morphology (Fig 4c,d). Thus, while multivariate gene–environment interactions were important (Sgro & Hoffmann 2004), they explained little genetic variation overall and failed to cause changes in phenotypic and genetic correlation structures with perceived risk treatment (conceptually visualised in Fig 1b).

We further demonstrated the integration of behaviour and morphology, finding results in line with meta-analyses showing that individuals expressing risky phenotypes are also relatively large (Niemelä & Dingemanse 2018). Our analyses demonstrated an overarching positive genetic correlation between behaviour and morphology (ovals labelled "Behav" and "Morpho" in Fig 5), implying that relatively active genotypes were relatively bold towards predators, relatively social, and relatively large. Models hypothesising structures where the genetic integration between behaviour and morphology varied with treatment (Models 2B,C; Fig 3) were poorly supported (Table 1b), implying that the multivariate gene–environment interactions causing stronger behavioural and morphological genetic correlations within versus among risk-environments (Fig 2b) did not also influence genetic correlations between behaviour and morphology.

In conclusion, our study demonstrated that behavioural traits were generally integrated, but more tightly so within than across low- and high-risk environments. The same hierarchical structuring characterised morphology. Importantly, the non-consumptive effects of perceived predation risk did not tighten trait correlations under high-risk. A particularly intriguing finding is the lack of conditionality of the integration of stickleback behaviour and morphology, suggesting that the functional integration of these two trait domains is independent of predation risk. Importantly, despite major differences in ecology (predation risk, lake dimensions and depth, etc.) the effect of perceived risk on genetic correlation structures was the same for both populations. Further studies comparing **G**-matrix structures among populations and ecological regimes are now required to infer whether our findings are broadly generalizable (Kelly 2006).

ACKNOWLEDGEMENTS

We are grateful to K. Boeke, R. Geoghegan, K. Meijer, C. Oosterhof, M. Schrama, T. Sparrow and F. van der Plas for assistance, and Welsh Water for permission to capture stickleback on their property. The experiment was designed by NJD and IB within the guidelines of the Association for the Study of Animal Behaviour, and undertaken under a U.K. Home Office project licence (PPL 40.2969, held by IB). We thank J. Wright and B.-E. Saether for discussion and support. NJD was supported by the Netherlands Organisation for Scientific Research (grant 863.05.002) and by the Research Council of Norway (STORFORSK project 'Population genetics in an ecological context', held by B.-E.Saether). NAD was supported by the US National Science Foundation (grant IOS 1557951). The manuscript was prepared by NJD and NAD and co-written by IB.

DATA AVAILABILITY STATEMENT

Data used in this work are available from Dryad Digital Repository: <https://doi.org/10.5061/dryad.wdbrv15j5> (Dingemanse *et al.* 2019).

AUTHORSHIP

The study was conceived and data collected by NJD and IB; NJD compiled the database. Analytical strategies were designed by NJD and NAD, and statistical analyses performed by NAD. The manuscript was drafted by NJD and NAD and co-written by IB.

REFERENCES

- Abbey-Lee, R.N. & Dingemanse, N.J. (2019). Adaptive individual variation in phenological responses to perceived predation levels. *Nat. Commun.*, 10, 1601.
- Adriaenssens, B. & Johnsson, J.I. (2013). Natural selection, plasticity and the emergence of a behavioural syndrome in the wild. *Ecol. Lett.*, 16, 47–55.
- Aguirre, J.D., Hine, E., McGuigan, K. & Blows, M.W. (2014). Comparing **G**: multivariate analysis of genetic variation in multiple populations. *Heredity*, 112, 21–29.

- Araya-Ajoy, Y.G. & Dingemanse, N.J. (2014). Characterizing behavioural 'characters': a conceptual and statistical framework. *Proc. R. Soc. Lond. B*, 281, 20132645.
- Armbruster, W.S. & Schwaegerle, K.E. (1996). Causes of covariation of phenotypic traits among populations. *J. Evol. Biol.*, 9, 261–276.
- Barber, I. & Arnott, S.A. (2000). Split-clutch IVF: a technique to examine indirect fitness consequences of mate preferences in sticklebacks. *Behaviour*, 137, 1129–1140.
- Bell, A.M. (2005). Behavioral differences between individuals and two populations of stickleback (*Gasterosteus aculeatus*). *J. Evol. Biol.*, 18, 464–473.
- Bell, A.M. & Sih, A. (2007). Exposure to predation generates personality in threespined sticklebacks (*Gasterosteus aculeatus*). *Ecol. Lett.*, 10, 828–834.
- Berdal, M.A. & Dochtermann, N.A. (2019). Adaptive alignment of plasticity with genetic variation and selection. *J. Hered.*, 110, 514–521.
- Brodie, E.D. (1992). Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. *Evolution*, 46, 1284–1298.
- Brodie, E.D. (1993). Homogeneity of the genetic variance-covariance matrix for antipredator traits in 2 natural populations of the garter snake *Thamnophis ordinoides*. *Evolution*, 47, 844–854.
- Burnham, K.P. & Anderson, D.R. (1998). *Model Selection and Inference*. Springer-Verlag, New York.
- Coleman, K. & Wilson, D.S. (1998). Shyness and boldness in pumpkinseed sunfish: individual differences are context specific. *Anim. Behav.*, 56, 927–936.
- Dingemanse, N.J. & Dochtermann, N.A. (2013). Quantifying individual variation in behaviour: mixed-effect modelling approaches. *J. Anim. Ecol.*, 82, 39–54.
- Dingemanse, N.J., Wright, J., Kazem, A.J.N., Thomas, D.K., Hickling, R. & Dawnay, N. (2007). Behavioural syndromes differ predictably between 12 populations of stickleback. *J. Anim. Ecol.*, 76, 1128–1138.
- Dingemanse, N.J., Oosterhof, C., van der Plas, F. & Barber, I. (2009a). Parasitism, head morphology and the strength of morphometric correlations in sticklebacks. *Biol. J. Linn. Soc.*, 96, 759–768.
- Dingemanse, N.J., van der Plas, F., Wright, J., Réale, D., Schrama, M., Roff, D.A., et al. (2009b). Individual experience and evolutionary history of predation affect expression of heritable variation in fish personality and morphology. *Proc. R. Soc. Lond. B*, 276, 1285–1293.
- Dingemanse, N.J., Dochtermann, N.A. & Wright, J. (2010). A method for exploring the structure of behavioural syndromes to allow formal comparison within and between datasets. *Anim. Behav.*, 79, 439–450.
- Dingemanse, N.J., Barber, I. & Dochtermann, N.A. (2019). Data from: non-consumptive effects of predation: does perceived risk strengthen the genetic integration of behaviour and morphology in stickleback? *Dryad Digital Repository*, <https://doi.org/10.5061/dryad.wdbrv15j5>.
- Dochtermann, N.A. (2011). Testing Cheverud's conjecture for behavioral correlations and behavioral syndromes. *Evolution*, 65, 1814–1820.
- Dochtermann, N.A. & Jenkins, S.H. (2007). Behavioural syndromes in Merriam's kangaroo rats (*Dipodomys merriami*): a test of competing hypotheses. *Proc. R. Soc. Lond. B*, 274, 2343–2349.
- Gavrilets, S. & Scheiner, S.M. (1993). The genetics of phenotypic plasticity. 5. Evolution of reaction norm shape. *J. Evol. Biol.*, 6, 31–48.
- Grace, J.B. (2006). *Structural Equation Modeling and Natural Systems*. Cambridge University Press, Cambridge.
- Hadfield, J.D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.*, 33, 1–22.
- Herczeg, G., Gonda, A. & Merila, J. (2009). Predation mediated population divergence in complex behaviour of nine-spined stickleback (*Pungitius pungitius*). *J. Evol. Biol.*, 22, 544–552.
- Kelly, C.D. (2006). Replicating empirical research in behavioural ecology: how and why it should be done but rarely ever is. *Q Rev Biol*, 81, 221–236.
- Kim, S.J. & Velando, A. (2015). Phenotypic integration between antipredator behavior and camouflage pattern in juvenile sticklebacks. *Evolution*, 69, 830–838.
- Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., et al. (2001). The strength of phenotypic selection in natural populations. *Am. Nat.*, 157, 245–261.
- Klingenberg, C.P. (2008). Morphological integration and developmental modularity. *Annu. Rev. Ecol. Evol. Syst.*, 39, 115–132.
- Klingenberg, C.P. (2014). Studying morphological integration and modularity at multiple levels: concepts and analysis. *Philos T R Soc B*, 369, 20130249.
- Kraft, P.G., Wilson, R.S., Franklin, C.E. & Blows, M.W. (2006). Substantial changes in the genetic basis of tadpole morphology of *Rana lessonae* in the presence of predators. *J. Evol. Biol.*, 19, 1813–1818.
- Krause, J. & Ruxton, G.D. (2002). *Living in groups*. Oxford University Press, Oxford.
- Kruuk, L.E.B. (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Proc. R. Soc. Lond. B*, 359, 873–890.
- Lande, R. & Arnold, S.J. (1983). The measurement of selection on correlated characters. *Evolution*, 37, 1210–1226.
- Lima, S.L. (1998). Stress and decision making under the risk of predation: Recent developments from behavioral, reproductive, and ecological perspectives. *Adv Stud Behav*, 27, 215–290.
- Lima, S.L. & Bednekoff, P.A. (1999). Temporal variation in danger drives antipredator behavior: the predation risk allocation hypothesis. *Am. Nat.*, 153, 649–659.
- Lynch, M. & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer, Sunderland, MA.
- Moiron, M., Araya-Ajoy, Y.G., Mathot, K.J., Mouchet, A. & Dingemanse, N.J. (2019). Functional relations between body mass and risk-taking behavior in wild great tits. *Behav. Ecol.*, 30, 617–623.
- Mutzel, A., Dingemanse, N.J., Araya-Ajoy, Y.G. & Kempenaers, B. (2013). Parental provisioning behaviour plays a key role in linking personality with reproductive success. *Proc. R. Soc. Lond. B*, 280, 20131019.
- Nakagawa, S. & Parker, T.H. (2015). Replicating research in ecology and evolution: feasibility, incentives, and the cost-benefit conundrum. *BMC Biol.*, 13.
- Niemelä, P.T. & Dingemanse, N.J. (2018). Meta-analysis reveals weak associations between intrinsic state and personality. *Proc. R. Soc. B Biol. Sci.*, 1873, 20172823.
- Pavlicev, M. & Wagner, G. (2012). A model of development evolution: selection, pleiotropy and compensation. *Trends Ecol. Evol.*, 27, 316–322.
- Pigliucci, M. & Preston, K. (2004). Phenotypic integration: studying the ecology and evolution of complex phenotypes. *Ecol. Lett.*, 6, 265–272.
- Rausher, M.D. (1992). The measurement of selection on quantitative traits - biases due to environmental covariances between traits and fitness. *Evolution*, 46, 616–626.
- Roff, D.A. & Fairbairn, D.J. (2001). The genetic basis of dispersal and migration, and its consequences for the evolution of correlated traits. In *Dispersal*. (eds Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D.). Oxford University Press, New York, pp. 191–202.
- Roff, D. & Fairbairn, D. (2011). Path analysis of the genetic integration of traits in the sand cricket: a novel use of BLUPs. *J. Evol. Biol.*, 24, 1857–1869.
- Roff, D.A. & Fairbairn, D.J. (2012). A test of the hypothesis that correlational selection generates genetic correlations. *Evolution*, 66, 2953–2960.
- Rossee, Y. (2012). Lavaan: an R package for structural equation modeling. *J. Stat. Softw.*, 48, 1–36.
- Searle, S.R. (1961). Phenotypic, genetic and environmental correlations. *Biometrics*, 17, 474–480.
- Sgro, C.M. & Hoffmann, A.A. (2004). Genetic correlations, tradeoffs and environmental variation. *Heredity*, 93, 241–248.
- Sih, A., Bell, A. & Johnson, J.C. (2004). Behavioral syndromes: an ecological and evolutionary overview. *Trends Ecol. Evol.*, 19, 372–378.
- Sinervo, B. & Svensson, E. (2002). Correlational selection and the evolution of genomic architecture. *Heredity*, 89, 329–338.
- Siren, J., Ovaskainen, O. & Merila, J. (2017). Structure and stability of genetic variance-covariance matrices: a Bayesian sparse factor analysis

- of transcriptional variation in the three-spined stickleback. *Mol. Ecol.*, 26, 5099–5113.
- Stearns, S., Dejong, G. & Newman, B. (1991). The effects of phenotypic plasticity on genetic correlations. *Trends Ecol. Evol.*, 6, 122–126.
- Ward, A.J.W., Thomas, P., Hart, P.J.B. & Krause, J. (2004). Correlates of boldness in three-spined sticklebacks (*Gasterosteus aculeatus*). *Behav. Ecol. Sociobiol.*, 55, 561–568.
- Whitman, D. & Agrawal, A.A. (2009). What is phenotypic plasticity and why is it important? In: Phenotypic plasticity of insects: mechanisms and consequences (eds. Whitman, D & Anathakrishnan, TN). CRC Press Boca Raton (FL), pp. 1–63.
- Wilson, E.O. (1998). *Consilience: the unity of knowledge*. Little, Brown & Company, London.
- Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A., et al. (2010). An ecologist's guide to the animal model. *J. Anim. Ecol.*, 79, 13–26.
- Wood, C.W. & Brodie, E.D. (2015). Environmental effects on the structure of the G-matrix. *Evolution*, 69, 2927–2940.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Editor, Jean-Michel Gaillard

Manuscript received 15 July 2019

First decision made 23 August 2019

Manuscript accepted 29 September 2019