



Zebra finch behaviour differs consistently between individuals but is not affected by early life adversity

Yoran H. Gerritsma^{*} , Merijn M. G. Driessen¹ , Simon Verhulst^{*}

University of Groningen, Groningen, the Netherlands

ARTICLE INFO

Article history:

Received 30 September 2022
Initial acceptance 13 December 2022
Final acceptance 9 January 2023
Available online 25 March 2023
MS. number: 22-00475

Keywords:

anxiety
asset protection
boldness
early life stress
novel environment

Individual variation in animal personality is ubiquitous, but little is known of its proximate causes. Theory predicts individuals will be more risk prone when their life expectancy is short, and adverse environmental conditions during development shorten life span. We modified developmental conditions of zebra finches, *Taeniopygia castanotis*, by manipulating parental foraging conditions to be either 'harsh' or 'benign' and studied offspring behaviour in adulthood using multiple tests: exploration, novel object, tonic immobility, dominance and sociality. Each test was done in duplicate, and repeatabilities of test scores ranged from 0.31 to 0.55 independent of treatment with one exception: tonic immobility was repeatable in individuals reared in harsh conditions only. Correlations between behaviours (i.e. behavioural syndromes) were generally weak and nonsignificant. Growing up in harsh conditions and in larger broods both negatively affected offspring growth and thereby presumably their life expectancy, but neither affected behaviour in standardized tests. This raised the question whether the origin of variation in personality was perhaps largely genetic, but heritability estimates were low to moderate. We conclude that variation in personality in our study can be attributed to environmental effects, but independent of early life adversity as manipulated here. To explain this finding, which runs counter to our expectation, we present two hypotheses. First, the adolescent social environment impacted the ontogeny of personality, overshadowing developmental effects. Second, the risk associated with a behaviour may be state dependent, with identical behaviours encompassing a greater risk in low-quality than in high-quality individuals. Thus, while individuals with different developmental backgrounds displayed similar behaviour, the perceived risk this entailed may have been very different and in accordance with theoretical predictions. Given that we found no behavioural syndromes, a theory that assumes behaviours to be correlated does not apply, and we discuss alternative functional explanations for individual differences in behaviour.

© 2023 The Author(s). Published by Elsevier Ltd on behalf of The Association for the Study of Animal Behaviour. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Consistent interindividual differences in behaviour, or personality (Dingemanse & Wright, 2020; Réale et al., 2007), have been well studied in recent years. In particular ecological and evolutionary consequences of personality variation have been elucidated (e.g. Dingemanse et al., 2004), revealing that personality variation can be maintained by natural and sexual selection (Carere & Eens, 2005; Dingemanse & Réale, 2005; Schuett et al., 2010; Wolf & Weissing, 2010). Furthermore, the genetic causes of personality variation have been well studied (Edwards et al., 2015; Korsten et al., 2010; Lesch et al., 1996; Mueller et al., 2013; Thys et al., 2021; Verhulst et al., 2016), and on average 13.7% of personality

variation is attributable to additive genetic effects (Dochtermann et al., 2015). Thus, the largest part of the proximate causes of personality variation remains unexplained. Yet the study of the proximate causes of variation in animal personality can provide insights into the evolutionary and functional constraints, and further our understanding of how selection acts on (the plasticity of) personality (Stamps & Groothuis, 2010).

On a functional level, life history theory predicts that individuals with higher fitness expectations, that is, a higher residual reproductive value (RRV), should be more risk averse (Wolf et al., 2007). The relationship between the benefits and costs of taking risks shifts towards taking less risk when the RRV is high because these individuals have more to lose (i.e. 'asset protection', Clark, 1994; Roff, 2002). A core assumption of this theory is that individuals with different life history strategies end up with similar fitness, but

^{*} Corresponding authors.

E-mail addresses: yorangerritsma@gmail.com (Y. H. Gerritsma), s.verhulst@rug.nl (S. Verhulst).

individuals with different personalities will not have similar fitness when the expression of personality is state or condition dependent. Such associations could, for example, arise when the risk associated with specific behaviours is state dependent, such as a better capability to escape predators (Stamps, 2007), where individuals in a better state face reduced risk. In this scenario, state-dependent behaviour would negate the functional prediction that individuals with higher fitness expectations behave as if more risk averse, and alternative predictions are not obvious.

Testing the relationship between fitness expectations and behaviour is preferably done in wild populations, where resources such as energy are limited, creating an incentive to take risks. However, results from such studies are mixed. For example, lower RRV was associated with more risk-prone behaviour in superb fairy wrens, *Malurus cyaneus* (Hall et al., 2015), in line with the functional framework. However, RRV was not associated with risk-taking behaviour in field crickets, *Gryllus campestris* (Fisher et al., 2018) or in yellow-bellied marmots, *Marmota flaviventris* (Petelle et al., 2019). Moreover, recent meta-analyses found ‘no directional relationship between riskier behaviour and greater mortality’ (Moiron et al., 2020, p. 399), and generally weak relationships between behaviour and ‘pace-of-life’ (Royauté et al., 2018). A potential explanation for the weak support is the correlational nature of most studies, making it difficult to separate genetic and environmental (e.g. developmental) effects on behaviour. This can potentially be resolved with an experimental approach, for example manipulating factors that modulate state and associated fitness prospects, followed by the assessment of individual variation in risk-taking behaviour.

In this study, we manipulated developmental conditions of captive zebra finches, *Taeniopygia castanotis*, and thereby their fitness expectations, and studied their behaviour in adulthood, to test for effects of life expectancy on risk-prone behaviour. Developmental conditions were manipulated by breeding zebra finches in aviaries where the parents experienced either benign or harsh foraging conditions (Briga & Verhulst, 2017; Jimeno et al., 2019; Koetsier & Verhulst, 2011). We previously showed that harsh foraging conditions during development negatively affected offspring growth (Gerritsma et al., 2022), indicating reduced life span expectations (Briga et al., 2017; Briga & Verhulst, 2021; De Kogel, 1997). We measured behaviour in multiple behavioural tests: (1) exploration in a novel environment (Fidler et al., 2007; Verbeek et al., 1994; Wuerz & Krüger, 2015), (2) boldness in a novel object setting (King et al., 2015; Spencer & Verhulst, 2007), (3) response in a tonic immobility test (Gallup, 1977, 1979; Peixoto et al., 2020), (4) dominance in a competition over a nonfood resource (Spencer & Verhulst, 2007) and (5) sociality in a preference setting (Kelly et al., 2011; Kelly & Goodson, 2015). All tests were first optimized to yield high repeatabilities, which is essential to characterize individual variation in behaviour. In the context of our study, risk-prone behaviour was assumed to be expressed as faster exploration of a novel environment, faster investigation of a novel object (boldness), less time spent immobile in the tonic immobility test and more dominant behaviour. For the sociality test we did not see a clear prediction with respect to the risk involved in being more or less social.

We further investigated whether behavioural syndromes, that is, correlations between different personality traits, emerged depending on developmental conditions (Bell & Sih, 2007; Dhellemmes et al., 2020; Dingemanse et al., 2007) or sex (Fresneau et al., 2014). This is of interest in itself but of particular relevance with respect to the prediction that individuals with higher fitness expectations should be more risk averse, because such a prediction is only tenable when individuals consistently differ in the risk they take, regardless of the nature of the behavioural test.

METHODS

Ethical Note

All methods and experiments in this research were performed with approval of the Central Committee for Animal Experiments (‘Centrale Commissie Dierproeven’) of the Netherlands, under licence AVD1050020174344. Minor stress from transporting individuals from and to test locations was unavoidable; however, test room and indoor aviaries were located close to each other which limited transport times. No injuries or mortality were observed during the behavioural tests. Animals were kept alive for further experiments after this study.

Animals and Developmental Treatment

We used 93 zebra finches (57 females, 36 males) raised in 58 different broods. The experimental setting in which the birds were reared is described in Gerritsma et al. (2022, cohorts 2018 and 2019). In brief, breeding took place in four outside aviaries (320 × 150 cm and 210 cm high) with up to 24 individuals each (sex ratio approximately 1:1). Two aviaries had low foraging costs and two had high foraging costs, as described in Koetsier and Verhulst (2011). Briefly, food, a commercial seed mixture, was offered in a single container (120 × 10 cm and 60 cm high), which was suspended from the ceiling. Food containers had five holes on each side (i.e. 10 in total), where the birds could access the seeds. In the benign foraging environment, there were perches beneath the holes, allowing the birds to perch while eating (low costs). In the harsh foraging environment, these perches were absent, forcing the individuals to fly back and forth between a distant perch and the food box for each seed.

The aviaries were equipped with nestboxes and ad libitum nest material (hay) from mid-April until mid-November each year. We measured offspring mass and biometry at the ages of 15 (±1) and 31 (±1) days, and growing up in harsh foraging conditions substantially reduced offspring growth (Gerritsma et al., 2022). Based on our earlier work, we assume that individuals growing up in harsh conditions had reduced fitness expectations (e.g. Briga et al., 2017) and refer to them as low-quality individuals. Offspring did not learn to eat from the foraging box in either easy or harsh foraging conditions. Consequently, they were nutritionally dependent on their parents until removal from the aviary to an identical aviary with food offered in containers on the floor, from which they would readily eat. At the age of 65 ± 4 days, the offspring were moved to single-sex groups. Between removal from the natal aviary and age 100 days, offspring were housed with unrelated adults, two males and two females, for sexual imprinting and song learning. After reaching the age of 100 days, individuals were moved to indoor aviaries (295 × 150 cm and 230 cm high, 14:10 h light:dark, lights on 0700–2100, ±20 °C, relative humidity ca. 50%, food (a seed mixture) and water ad libitum) with up to 24 individuals each. Individuals were colour ringed (size SP2, A.C. Hughes, Redruth, U.K.) for identification purposes and behavioural testing commenced afterwards (age at first test: mean ± SD: 296 ± 92 days).

Behavioural Tests

Behavioural tests were carried out in a fixed order to standardize potential sequence effects (Bell, 2013). The order of testing was: (1) novel environment, (2) novel object, (3) tonic immobility, (4) sociality and (5) dominance. Each individual was subjected to each behavioural test twice before the next test in the sequence. All tests were recorded using a video camera (Polaroid Cube+ for the

novel environment, JVC Everio R for all others). The novel environment, tonic immobility and dominance videos were analysed using Solomon coder (v19.08.02, <https://solomon.andraspeter.com/>). We tracked an individual's position in the novel object and sociality videos using EthoVision XT 11.5 (Noldus Information Technology, Wageningen, the Netherlands). Researchers were blind to the developmental manipulation of the individual when analysing the video recordings.

We performed the novel object, sociality and dominance tests in the same specially designed cage (104×52 cm and 52 cm high), with six perches, equally distributed over the length of the cage, which was permanently placed in a test room, in which we also performed the tonic immobility test. Prior to each behavioural test, birds were moved from the indoor aviaries to this designated test room. Hence, only the novel environment test was performed in a different room. To habituate the birds to the test cage and room, we moved each individual with three same-sex conspecifics to the testing cage for 1 h per day, for 4 consecutive days, before the first behavioural test. On the fifth day, each individual was individually habituated to the behavioural testing cage for 1 h, with sounds from an unfamiliar zebra finch colony playing in the background.

Novel environment

The novel environment was a room (200×150 cm and 200 cm high) in which were three artificial trees with four branches each (the same trees as used in Verbeek et al., 1994). The position and the colour (either wood or white) of the trees were changed between the first and second test to keep the environment novel. Individuals were transported in a small opaque wooden box (15×15 cm and 20 cm high) to the test room, where the box was put on a pedestal (1 m high) against a screen that was one wall of the test room. Immediately after leaving the novel environment, the experimenter opened the box with a pulley system from behind a screen. The individual was then recorded for 60 min. We measured the latency in leaving the box ($\log_{10}$ transformed), the number of unique locations explored (up to 12 'branches') and the number of movements made ($\log_{10}$ transformed). We combined these three variables into a single variable using principal component analysis (PCA; see Statistics for details) and used the first principal component (PC) as the dependent variable. The first PC explained 72.7% of the variation. Individuals with a higher score on this PC, that is, individuals that left the box early, explored more unique locations and made more movements were considered to be fast explorers (David et al., 2011; Dingemans et al., 2003).

In 22 tests ($N = 18$ birds), the individual remained in the box the full 60 min. After observing this a few times, we wondered whether these birds were not exploratory or did not leave the box for a different reason. Thus, if an individual remained in the box for 60 min, the experimenter slightly moved the box through the screen, which stimulated the individual to exit the box after which it was recorded for another 10 min. From the 10 tests ($N = 9$ birds) in which this occurred, the individual visited only one perch and remained immobile there for 10 min in seven tests. In the other three tests, the individual visited three, seven or nine unique locations. The additional 10 min were excluded from the analysis, that is, if a bird remained in the box for the full 60 min it explored 0 unique locations and made 0 movements.

Novel object

Each test consisted of two runs which were performed on consecutive days, and there were at least 3 days between the second run of the first test and the first run of the second test. For each run, individuals were moved to the behavioural testing cage, and given a 30 min habituation period. The experimenter then returned and divided the behavioural testing cage into two compartments

with an opaque divider, leaving the bird in one half of the cage. The experimenter then put their hand in the empty cage-half, with or without placing a novel object there, after which the divider was removed. The novel object was placed on the second perch, counting from the far end from the perspective of the bird. The run without a novel object served as a within-individual control. Half of the birds received the novel object in the first run and the other half in the second run and the order was reversed for each individual in the second test. Two different novel objects (a green clothes peg and a black hair clamp) were used, balanced across the first and second test. We recorded the latency to first move and the latency to visit the perch with (novel object treatment) or without (control treatment) the novel object for up to 30 min after removal of the partition. We combined these variables using PCA, on the (1) latency to inspect the novel object ($\log_{10}$ transformed) and (2) the Δ latency to move between phase 2 (disturbance) and phase 3 (disturbance + novel object) ($\log_{10}$ transformed). The first PC explained 62.9% of the variation (for loadings, eigenvector values and directions of both PCAs, see Appendix Figs A1 and A2 and Table A1). Individuals with a higher score on this PC, i.e. those that took longer until their first move when a novel object was present and that took more time before perching next to the novel object, were considered to be shy (Wuerz & Krüger, 2015).

Tonic immobility

We measured tonic immobility in a quiet room with minimum disturbance, and the two tests were performed on consecutive days. For each test, an individual was fixed on its back, with wings pressed to its body, on a felt-lined aluminium cradle. Fixation took place by holding the index and middle finger gently on its chest for 5 s (mean $\pm$ SD = 5.5 ± 1.1 s) and then slowly removing them. Immobilization was deemed successful if the bird did not move for 5 s, and we measured the time spent immobile with a maximum score of 600 s. If immobilization was unsuccessful, the bird was fixed again for up to 10 times in total, using only the time spent immobile after successful immobilization. If immobilization was not successful 10 times, we assigned an immobility time of 1 s. The time spent immobile was used as the dependent variable ($\log_{10}$ transformed to normalize the distribution), and individuals that spent more time immobile were considered more anxious (Brust et al., 2013; Gallup, 1977; Wuerz & Krüger, 2015).

Sociality

For the sociality tests, side cages (52×52 cm and 52 cm high) were fitted to the sides of the behavioural testing cage, separated by metal bars, limiting physical contact, but otherwise allowing all communication. In each side cage, there was one perch, ca. 4 cm from the outer perches of the behavioural testing cage. Prior to the introduction of the focal individual, the experimenter placed 10 individuals in one side cage and two individuals in the other side cage, with sides balanced between first and second tests and individuals. These individuals had never previously interacted with the focal individual. For 10 min, we recorded the time spent on each of the six perches in the behavioural testing cage, where we took the time spent on the outer perches as time spent interacting with the conspecifics.

Group preference was taken to be the proportion of the 10 min spent on the outer perches, and group size preference was the time spent next to the group of 10 individuals divided by the total time spent interacting with either group. Group preference and group size preference were used as the dependent variables, where individuals that spent more time next to the groups, and those that preferred the large group, were considered more social. Both group preference and group size preference were square-root transformed to obtain a normal distribution.

Dominance

Social dominance score was the outcome of tournaments among single-sex groups of five or six individuals, and competition was over the single perch in the testing cage on which there was room for only one bird (Spencer & Verhulst, 2007). Each individual competed against every other individual in their group in dyadic interactions, with a maximum of one match per day for each individual. Thus, there were 10–15 dyadic matches per tournament, and each tournament was held twice, yielding two dominance scores per individual. Assignment to groups was such that the groups did not differ in size (based on their biometry at day 100, using the mean and SD of mass, tarsus and head+bill length) and birds from the benign and harsh treatments were equally represented in the groups. Unfortunately, we were not also able to assign groups in such a way that they did not differ in the mean and SD of age.

The behavioural testing cage was split into two compartments with an opaque divider, each with a low perch. Two birds were then moved to the behavioural cage, one in each compartment, and 30 min later the low perches and the divider were removed. Removal of the divider triggered a single, higher, perch to drop into place in the middle of the cage. We measured the time each individual spent on the perch for 30 min, assuming that the perch is the preferred position. To test this assumption, we included the single dominance perch in the solo habituation that was performed prior to the behavioural tests (ca. 1 h) in the test cage of the last 47 individuals. Of these 47 individuals, 41 perched, with 34 of them perching within 1 min. For the 41 individuals that perched, a mean $\pm$ SD of 3157 ± 916 s were spent on the perch. For the second run of each dyadic interaction, the starting positions were switched, with half of the individuals starting in the left half of the behavioural testing cage. We used the time spent on the perch to calculate the David's score (Gammell et al., 2003; Hemelrijk et al., 2005). The David's score was used as the dependent variable, and individuals that received a higher David's score were considered more dominant. Individuals that never sat on the perch, with their opponents also not perching, had a David's score of 0, and were not included in the analysis.

Statistics

All analyses were performed in R v 3.5.3 (R Development Core Team, 2008) in the RStudio IDE (v1.3.1073; Rstudio Team, 2020).

Principal components analysis

We employed a PCA on the novel environment and novel object tests to reduce the number of variables into a single component variable (Bell & Stamps, 2004; Kelly & Goodson, 2015; Rangassamy et al., 2015). We used the packages 'factoextra' (v1.0.7), 'FactoMineR' (v2.4), 'tibble' (v3.0.5) and 'tidyr' (v1.1.2) to calculate the PCA and extract the PC1 (Kassambara & Mundt, 2020; Lê et al., 2008; Müller & Wickham, 2021; Wickham, 2020).

Repeatability

Repeatability was calculated as the between-individual variance divided by the sum of the between- and within-individual variance. Repeatability estimates were obtained using the 'rpt' function from the 'rptR' (v0.9.22) package (Nakagawa & Schielzeth, 2010), with variance components extracted from mixed-effects models. Here, we included individual identity as a random effect, and the test number (1 or 2) as a fixed effect. We used a Gaussian distribution for all (transformed) metrics. Confidence intervals and significance were calculated by parametric bootstrapping, with 500 iterations and 0 permutations. Additional random effects were added per behavioural test: for the novel object, we included the phase the

individual was subjected to first (disturbance or novel object). For the dominance and sociality tests we included a group identifier, for the group of individuals the focal bird was competing against or the group of stimulus individuals present. For tonic immobility, we included an identity for the experimenter performing the immobilization.

Additionally, we tested whether repeatability estimates differed between individuals from benign or harsh developmental conditions by separately estimating repeatability for each group. If repeatability estimates were similar, this did not necessarily mean that individuals from one group behaved more consistently, as there may have been larger variance between individuals (Dochtermann & Royauté, 2019). Therefore, after estimating the repeatability for both groups (i.e. benign and harsh development), we tested whether within- or between-individual variance in behaviour differed significantly between these groups with a Levene's test ('car' package, v3.0-10, Fox & Weisberg, 2019). Within-individual variance was calculated by subtracting the value of the first test from the value of the second test.

Multivariate analyses

As we had up to five behavioural tests per individual, we employed multivariate response models, allowing us to estimate individual level pairwise residual correlations between behaviour in the different tests (i.e. behavioural syndromes). We fitted Bayesian multivariate response models with the 'brms' package (v2.14.4; Bürkner, 2017), interfaced with the MCMCsampler of RStan (v2.19.2; Stan Development Team, 2020). All response variables were transformed to a standard normal distribution (mean = 0, SD 1), resulting in informative main effects even when included in interaction terms. Furthermore, standardization facilitates comparison of effect sizes and increases the efficiency of the MCMCsampler. For fixed effects, we used weakly informative Gaussian priors (mean = 0, SD = 1, Lemoine, 2019). Default priors of 'brms' were used for random effects, a Student's *t* density with three degrees of freedom for standard deviations and an LKJ density (Lewandowski et al., 2009) for correlations.

As fixed effects, we included test number (1 or 2), developmental treatment, sex, and the interaction between sex and developmental treatment. We initially included an individual's age at the first test but found no significant effect thereof in any of the analyses and decided to exclude it from further models. For the tonic immobility test, we included the fixation time as a fixed effect. As Gaussian random intercepts, we included individual identity and the identity of the genetic parents. We further included the same random factors per behavioural test as in the repeatability analysis (Appendix Table A2). As we previously found that both the developmental treatment and brood size affected offspring growth (Gerritsma et al., 2022), and thereby their fitness expectations, we ran additional multivariate models where we replaced developmental treatment with either brood size or mass on day 15.

Multivariate models were run on three chains, with 1000 warm-up iterations each, followed by 3333 iterations per chain resulting in 9999 iterations in total. Adapt_delta was set to 0.999 to prevent divergent transitions. Trace plots were used to check proper mixing of chains and convergence of chains was checked by checking whether Rhat values were (close to) 1. Posterior predictive checks were inspected to evaluate model fits by using the pp_check function in the 'brms' package.

As multivariate models require complete data sets, we ran these models with ($N = 68$) and without ($N = 93$) the dominance test results (sample sizes are given in Table 1).

For hypothesis testing, we calculated the probability of direction (Pd), following Makowski, Ben-Shachar, Chen, et al. (2019). The Pd is defined as the portion of the posterior that has the same sign as

Table 1
Sample sizes for the multivariate models

Model	Two tests		One test	
Excluding dominance	88	♀ 54 (31/23) ♂ 34 (14/20)	93	♀ 57 (33/24) ♂ 36 (16/20)
Including dominance	50	♀ 18 (10/8) ♂ 32 (14/18)	68	♀ 32 (18/14) ♂ 36 (16/20)

Sample sizes are given per sex and number of successful tests per individual. Sample size per treatment is given between brackets, for benign/harsh developmental conditions. Within each model, only those individuals with a value for each test were included. Thus, excluding the dominance test results, where there were many missing cases, increased the sample size.

the mean and can be considered the Bayesian equivalent of the frequentist P value, following the formula $P = 2 \times (1 - \text{Pd})$ (Makowski, Ben-Shachar, Chen, et al., 2019). Following this formula, Pd values between 0.95 and 0.975 are described as 'weak evidence' (corresponding to $0.1 \geq P \geq 0.05$), Pd values between 0.975 and 0.99 as 'moderate evidence' ($0.05 \geq P \geq 0.02$) and Pd values above 0.99 as 'strong evidence' ($P \leq 0.02$). Effect sizes were calculated as the posterior means, with the highest density intervals (95% confidence interval, CI) of the posterior mean calculated with the 'hdi' function from the package 'bayestestR' (v0.8.2; Makowski, Ben-Shachar, & Lüdtke, 2019).

To test for behavioural syndromes, we estimated an individual level pairwise residual correlation matrix from the multivariate models that included only effects related to the methodology of the tests, thereby yielding correlations between the 'raw' behavioural observations in the different tests and its 95% CI. In these models, fixation time was included for the tonic immobility trait, and all random effects that were used in the repeatability analysis were kept (Appendix Table A2). Also, here we ran separate multivariate models, excluding and including dominance, and used the latter for correlations between dominance and other traits only.

To test whether individual level behavioural syndromes differed between individuals from benign or harsh early life conditions, or between males and females, additional multivariate models were performed on subsets of the data set. More specifically, residual correlations were extracted from the multivariate models using the VarCorr function from the 'brms' package. We subtracted the residual correlations from the benign development model from the residual correlations from the harsh development model. For sex, we subtracted the residual correlations from the male model from the residual correlations from the female model. We calculated the mean and 95% CI of these differences, where an overlap of the CI with zero indicated that behavioural syndromes did not differ significantly between individuals with different developmental backgrounds or between the sexes. Sample sizes were not sufficient to determine whether behavioural syndromes differed between males and females depending on the developmental treatment.

Heritability

The heritability (h^2) of the behavioural observations was estimated following equation 8 (given below) from de Villemereuil et al. (2018). Heritability estimates were calculated for the genetic father and the genetic mother as the ratio of the additive genetic variance, V_a , which is the amount of variance explained by the identity of the genetic parent, to the total phenotypic variance. As suggested by de Villemereuil et al. (2018), we included only 'natural' fixed effects (i.e. sex, and not developmental treatment) and included the random effects that were also included in the repeatability analysis (Appendix Table A2). The total phenotypic

variance consisted of (1) V_a , (2) V_f , the variance from fixed effects, (3) V_r , the variance of additional random effects, and (4) V_{re} , the residual variance component.

$$h^2 = \frac{V_a}{V_a + V_f + V_r + V_{re}}$$

RESULTS

Sample Size per Behavioural Test

Some tests were unsuccessful (see Appendix Table A3 for sample sizes), in particular the dominance test, usually because neither of the two contestants used the perch (45/114 cases in females, 3/72 cases in males).

Developmental Manipulation

Offspring growth was reduced in harsh foraging conditions (Gerritsma et al., 2022). In brief, at age 15 days, offspring reared in harsh foraging conditions were 0.7 g lighter (mean $\pm$ SD = 8.57 ± 1.99 versus $7.88 \pm 1.80 \text{ g}$, benign versus harsh, respectively) and had shorter wings (39.84 ± 5.21 versus $38.47 \pm 6.30 \text{ mm}$) and head + bill lengths (18.77 ± 1.40 versus $18.28 \pm 1.51 \text{ mm}$), but tarsus length (13.62 ± 0.96 versus $13.41 \pm 1.03 \text{ mm}$) did not differ. At the age of 31 days, offspring born in the harsh foraging environment were still significantly lighter (11.57 ± 1.59 versus $11.14 \pm 1.45 \text{ g}$), but wing, tarsus and head + bill length were not significantly smaller. As the growth of the offspring was negatively affected, we refer to growing up in harsh foraging conditions as being faced with early life adversity.

Individual Repeatability

Behaviour in the standardized tests varied repeatably between individuals (Fig. 1), with repeatability estimates ranging from 0.31 to 0.55 (Table 2). An exception was group size preference in the sociality test ($R = 0.11$, $P = 0.17$), and we therefore excluded this variable from further analysis. Repeatability estimates did not depend on developmental conditions, except for the time spent immobile in the tonic immobility test, which was repeatable in individuals from harsh developmental conditions ($R = 0.62$, $P < 0.01$), but not in individuals from benign developmental conditions ($R = 0.17$, $P = 0.12$). This difference was due to significantly smaller within-individual variance in individuals from harsh developmental conditions, while between-individual variance was independent of treatment (Table 2).

Effects (standardized effect sizes) of developmental conditions on behaviour in the standardized tests ranged from -0.27 to 0.38 with CIs including zero in all cases (Appendix Table A4). Thus, we conclude that early life adversity as applied in the present study did not affect adult behaviour in the standardized tests. Neither did behaviour in the standardized tests differ significantly between the sexes (standardized effect sizes ranging from -0.45 to 0.20 with CIs overlapping zero). There was weak evidence for individuals to be more social in the second test (Pd = 0.954, posterior mean, 95% CI = 0.158 [-0.027 , 0.338]). There was no evidence for an interaction between development and sex (Fig. 2).

Behaviour in standardized tests was independent of mass on day 15, with standardized effect sizes ranging from -0.05 to 0.26 , but CIs always overlapping with zero (Appendix Table A5). However, dominance and tonic immobility were affected by an interaction between mass on day 15 and sex. To elucidate this effect, we

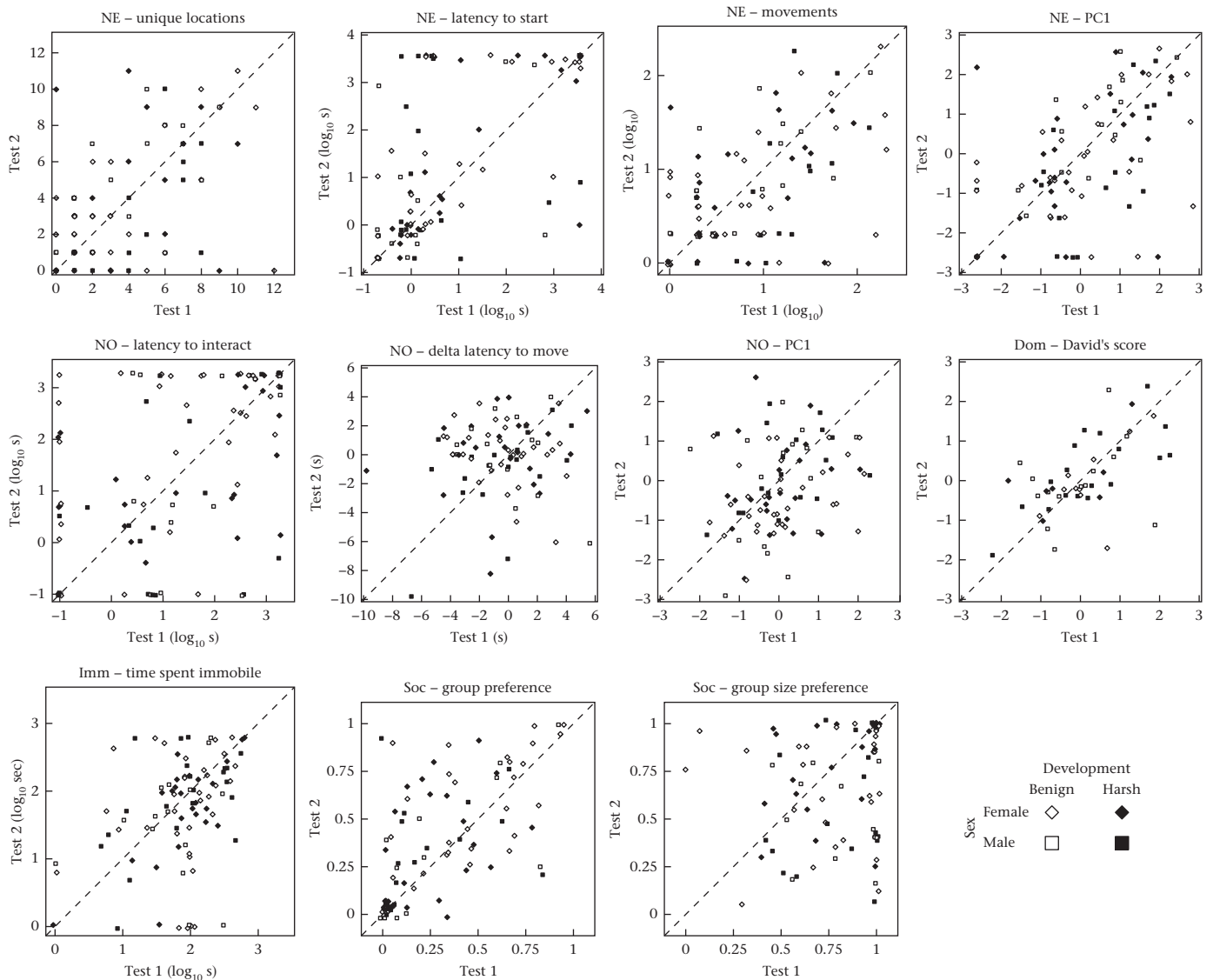


Figure 1. Behavioural scores in the first and second tests of males and females that were reared in either benign or harsh conditions. The dotted line ($y = x$) represents equal scores on tests 1 and 2. NE: novel environment test; Soc: sociality test; NO: novel object test; Imm: tonic immobility test; Dom: dominance test.

ran additional multivariate models for the sexes separately. In females, mass on day 15 did not predict dominance (effect size 0.40, CI $[-0.155, 0.922]$) or tonic immobility (0.08, CI $[-0.249, 0.408]$). However, males that were heavier at age 15 days were less dominant (-0.78 , CI $[-1.211, -0.347]$) and spent less time immobile (-0.347 , CI $[-0.663, -0.045]$) compared to lighter males.

Behaviour in the standardized tests was independent of brood size, with standardized effect sizes ranging from -0.20 to 0.14 , with CIs overlapping with zero (Appendix Table A6).

Behavioural Syndromes

Correlations between behavioural scores in the different standardized tests were low (Fig. 3, Appendix Table A7), with two exceptions: exploration was positively correlated with dominance and negatively with tonic immobility (Appendix Table A7). Developmental conditions may affect the development of behavioural syndromes, and we therefore repeated this analysis for birds reared in benign and harsh conditions separately. This yielded evidence for two behavioural syndromes in individuals reared in harsh conditions only. First, there was a correlation

between exploration and tonic immobility (Appendix Table A7), where individuals that explored faster spent more time immobile. These correlations differed significantly between benign and harsh developmental conditions (difference and 95% CI = 0.533 , $[0.130, 0.930]$). Second, there was a correlation between novel object and tonic immobility (Appendix Table A7), where individuals that were bolder when faced with a novel object spent more time immobile, but this difference was not statistically significant (difference and 95% CI = -0.14 , $[-0.53, 0.25]$). When examining behavioural syndromes in the sexes separately, we found exploration to be positively correlated with dominance in females, but not in males (Appendix Table A7), but the difference between the sexes was not statistically significant (estimate and 95% CI = -0.437 , $[-0.908, 0.069]$).

Heritability

Heritability estimates of the behavioural scores ranged from 0.08 to 0.12 , with the heritability of the behaviour towards a novel object being somewhat higher at 0.23 (Table 3).

Table 2
Repeatability of each metric per behavioural test

Test	Metric	Days between tests 1 and 2 ± SE	All birds [95% CI]	Benign	Harsh	Within-individual variance	Between-individual variance	Additional random effects
Dominance	David's score	10±1	0.55 [0.302, 0.711]	0.43 [0.022, 0.707]	0.67 [0.286, 0.814]	F = 0.00 P = 0.98	F = 1.01 P = 0.32	Identity & Dominance group
Sociality	Group preference	1±0	0.502 [0.343, 0.674]	0.651 [0.396, 0.846]	0.313 [0.066, 0.563]	F = 2.25 P = 0.14	F = 1.79 P = 0.18	Identity & Social stimulus group
Tonic immobility	Group size preference	1±0	0.11 [0.0325]	0 [0.0161]	0.217 [0.0494]	Excluded		
	Log ₁₀ time spent immobile	1±0	0.367 [0.171, 0.524]	0.17 [0.0415]	0.616 [0.418, 0.778]	F = 4.71 P = 0.03*	F = 0.14 P = 0.71	Identity & Experimenter
Novel object	PC1 novel object	20±2 ^a	0.314 [0.135, 0.48]	0.263 [0.0498]	0.354 [0.089, 0.599]	F = 0.15 P = 0.70	F = 0.03 P = 0.87	Identity & First phase
Novel environment	PC1 exploration	8±0	0.526 [0.368, 0.657]	0.55 [0.353, 0.707]	0.503 [0.26, 0.686]	F = 0.24 P = 0.63	F = 0.89 P = 0.35	Identity

Repeatabilities of behavioural test scores with 95% confidence interval (CI) in brackets, for all birds pooled and for individuals reared in benign or harsh conditions separately. Estimates from model with test number (1 or 2) included as fixed effect and random effects as shown. Columns denoting within- and between-individual variance report the results of a Levene's test comparison between benign and harsh treatments.

* $P < 0.05$.

^a Increased time difference due to covid-19 delays.

DISCUSSION

Causes of Individual Variation

We successfully manipulated early life adversity in zebra finches, as evidenced from their growth, and subjected these birds to multiple behavioural tests. However, early life adversity did not affect the behaviour in any of these tests. This lack of effect can in principle be due to imprecise measurements, but repeatabilities of the behavioural scores were substantial (mean 0.45, range 0.31–0.55), and in the range expected for behavioural traits (Appendix Table A8). Moreover, each individual was subjected to each test twice, and the repeatability of the mean of two observations is higher, ranging in this case from 0.48 to 0.71 (mean 0.62; calculated using equation 37 in Nakagawa & Schielzeth, 2010). Given that the sample size was substantial at 93 birds, we assume that any undetected effects of early adversity must have been small relative to the total variance.

The absence of a behavioural response to adversity in our study is consistent with a meta-analysis that found no effect of maternal undernutrition during pregnancy on offspring behaviour in mammals (Besson et al., 2016). It is also consistent with the results of a meta-analysis of associations between 'state' and personality traits, which concluded that such associations explain at most 8% of the variation (Niemelä & Dingemanse, 2018). Furthermore, high heterogeneity in these meta-analyses is reported (44.1–85.6%, Besson et al., 2016; 80%, Niemelä & Dingemanse, 2018), indicating a need for more studies to unravel the source of this variation.

High heritability of behaviour could in principle explain the lack of experimental effects in our study, because variation in personality is known to have a heritable component (Dochtermann et al., 2015). However, on average only 13.7% of the variation in animal personality can be attributed to additive genetic variation (Dochtermann et al., 2015), and our mean estimate ($\pm$ SE: $12.2 \pm 1.2\%$) is close to this estimate. We were limited in estimating heritability as variation in relatedness was limited to full-siblings and half-siblings but this limitation likely generated an over-estimation due to shared environments, and hence is unlikely to affect our conclusion that a genetic constraint does not explain the absence of experimental effects.

The finding that neither early life adversity nor additive genetic effects explains the substantial individual variation in behaviour indicates that the causes of this variation must be found elsewhere. Zebra finches live in flocks in the wild (Zann, 1996), and the quantity and quality of interactions within flocks vary widely (Boogert et al., 2014; Brandl et al., 2019). Most studies of social environment effects on later behaviour investigated the pre- or postnatal environment (Caldji et al., 1998; Kaiser & Sachser, 2005; Kemme et al., 2008), and attention has only recently shifted to the adolescent stage (Sachser et al., 2011). In zebra finches, the adolescent social environment affects adult affiliative and aggressive behaviour (Bölting & von Engelhardt, 2017; Ruploh et al., 2013, 2014, 2015), suggesting it to be an important period in shaping zebra finch personality. In our study, birds were always in groups, and interactions in these groups at any point before the behavioural testing may therefore have shaped their personality as expressed in our behavioural assays, perhaps through processes as described by reinforcement sensitivity theory (Corr, 2008) but this remains to be investigated.

Functional Aspects

Life history theory predicts individuals with reduced life expectancy will be more risk prone (Wolf et al., 2007). The results from our study and from Krause et al. (2017), who studied long-

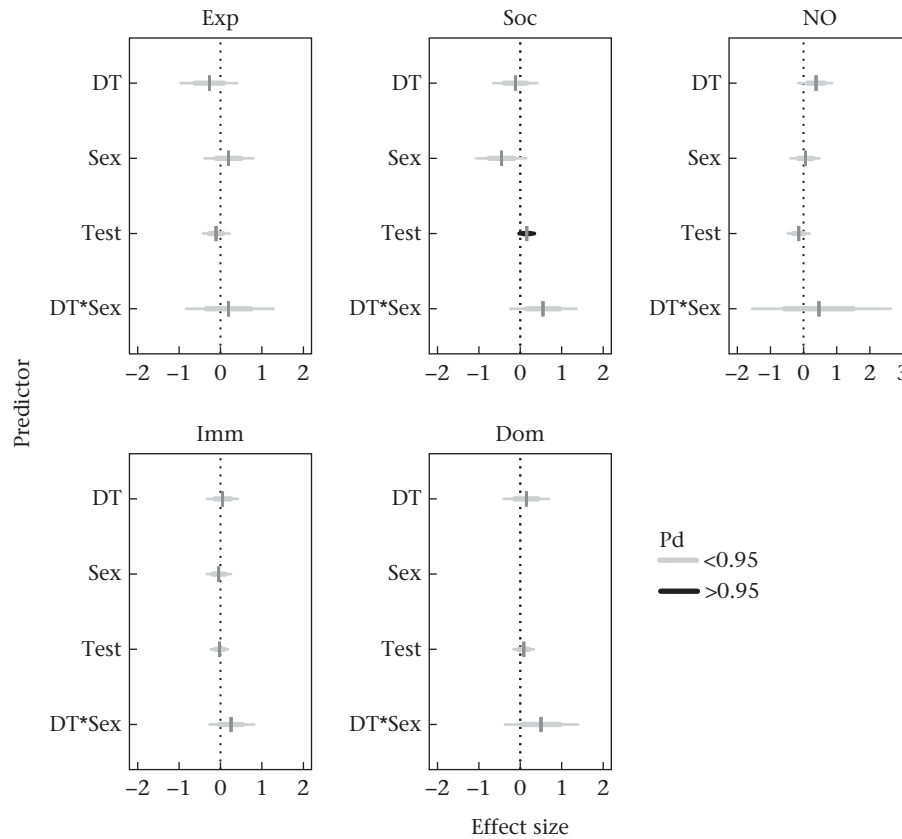


Figure 2. Effects of developmental treatment, sex and test sequence on behavioural test results, expressed as standardized effect size (vertical line) with SE (thick whisker) and 95% confidence interval (thin whisker). Pd: the probability of direction, where values below 0.95 are not significant and values above 0.95 are significant. Exp: exploration in novel environment; Soc: group preference in the sociality test; NO: novel object; Imm: tonic immobility; Dom: dominance. Dominance tests were always intrasex, so there can be no differences between the sexes. Negative effect sizes indicate lower scores for individuals reared in harsh conditions (DT), males (Sex) or the second test. Note that the x-axis for novel object is on a longer scale, and that variables were mean-centred, rendering the main effect estimates informative, even in the presence of the interaction term in the model.

term effects of early life nutrition, do not agree with this prediction, since neither study found effects of developmental conditions on behaviour in adulthood. However, a core assumption of this theory is that individuals with different life history strategies ultimately achieve similar fitness. Given that early life adversity negatively affected growth in our study, fitness prospects are likely to differ between these individuals, and this complicates predictions. Specifically, the risk associated with behaviour could be state dependent (Luttbeg & Sih, 2010), where low-quality individuals may be at greater risk in identical contexts, for example if they are less able to escape a predator. This implies that while early life adversity did not affect behaviour, it may well have had an undetected effect on the risk that they took in expressing it. Testing this hypothesis requires different tests, for example quantifying escape behaviour.

Behavioural syndromes potentially limit the behavioural expression of an individual, and can be the result of specific selection pressures, when combinations of behavioural traits are favoured by selection. For example, in sticklebacks, *Gasterosteus aculeatus*, and lemon sharks, *Negaprion brevirostris*, behavioural syndromes only formed when individuals were exposed to predation (Bell & Sih, 2007; Dhellemmes et al., 2020; Dingemanse et al., 2007). However, greater plasticity can also be selected for, as early life adversity disrupted the emergence of a behavioural syndrome in cichlids (Hope et al., 2020). We therefore tested whether early life adversity affected the development of behavioural syndromes in our study. However, except for a negative correlation ('syndrome') between exploration and tonic immobility in birds that

experienced early life adversity, we found little evidence for behavioural syndromes.

The ability to detect behavioural syndromes depends on the repeatability of the behavioural scores. For example, when the repeatability of two behavioural tests is 0.5, the maximum correlation that can be observed between the two behavioural scores is $0.5 \times 0.5 = 0.25$. Thus, even if the two behaviours in this example were perfectly correlated ($r = 1$), when averaged over many measurements in each individual, we would observe a correlation of only 0.25 in a data set where each behaviour is measured once and the difference is due to the within-individual variability of the behaviour. In our study, the (extrapolated) repeatabilities ranged from 0.48 to 0.71, and the upper limits of the correlations we could find (i.e. the products) therefore ranged from 0.26 to 0.49 across all combinations. We were very likely to detect correlations at this level, that is, perfect correlations ($r = 0.26$, power = 0.73; $r = 0.49$, power > 0.99 when using Pearson correlation), but suppose the true correlation was 0.5 then our power becomes substantially lower (range 0.24–0.66). It remains difficult to exclude the possibility, therefore, that behavioural syndromes were present but undetected in our study. On the other hand, due to the within-individual variability in behaviour, the expression of undetected behavioural syndromes is nevertheless likely to have been relatively weak. This reduces their relevance when trying to understand causes and consequences of variation in behaviour, because selection will act on the expression of the phenotype.

A starting point of this study was the assumption that individuals can be distinguished on the basis of the amount of risk

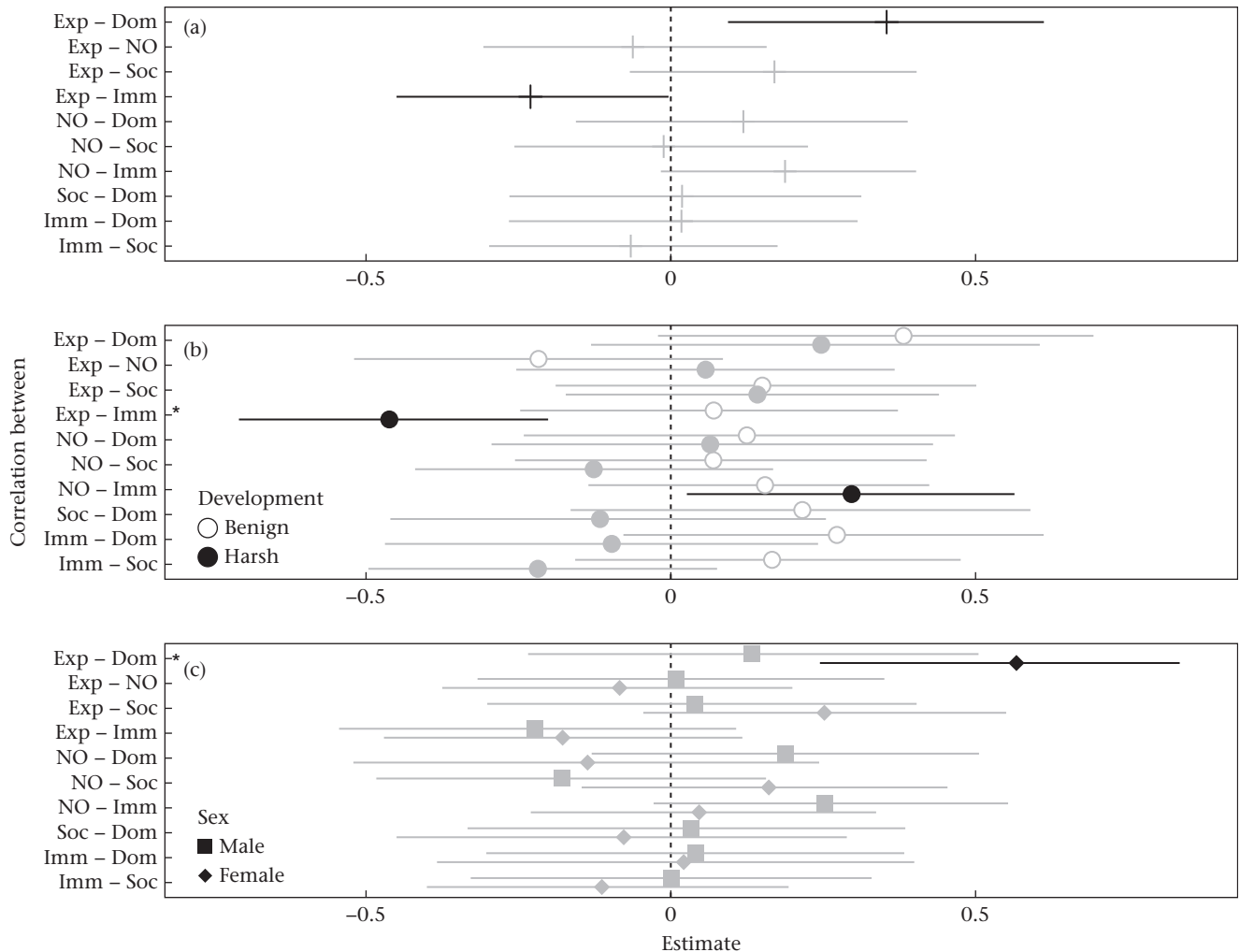


Figure 3. Correlations are given as the estimate (point) and the 95% confidence interval (CI, whiskers). (a) All individuals combined, (b) individuals separated by developmental treatment and (c) individuals separated by sex. The colour of the point and line corresponds to whether the 95% CI overlaps with zero (grey/white = yes, black = no), and denotes whether behavioural traits were correlated within that group (all, development or sex). Exp: exploration in novel environment; Soc: group preference in the sociality test; NO: novel object; Imm: tonic immobility; Dom: dominance. A small asterisk on the y-axis denotes whether the correlations differed significantly between the groups.

Table 3
Heritability estimates of behaviour in each test

Behavioural test	Mother	Father	Cumulative
Novel environment	0.051 [0, 0.179]	0.058 [0, 0.199]	0.109
Novel object	0.127 [0, 0.437]	0.089 [0, 0.341]	0.216
Tonic immobility	0.03 [0, 0.131]	0.047 [0, 0.178]	0.077
Sociality	0.063 [0, 0.183]	0.061 [0, 0.186]	0.124
Dominance	0.045 [0, 0.171]	0.029 [0, 0.107]	0.074

Estimates are given as the additive genetic variance explained by the genetic mother or genetic father with the 95% confidence interval between brackets.

they take, combined with the prediction based on life history theory that individuals with lower fitness prospects would take more risk (Clark, 1994; Roff, 2002; Wolf et al., 2007). However, zebra finches in our study did not differ consistently in the amount of risk they took, as evidenced by the weak correlations between behaviour in the different tests, despite the consistent individual differences in behaviour. In this setting, theories that assume consistent individual differences in risk taking are of limited relevance to functionally explain the individual variation in behaviour. This raises the question how to explain the consistent individual variation in behaviour we observed from a functional perspective. If

fitness payoffs are frequency dependent (e.g. being aggressive or not; Duckworth, 2008), this could lead to consistent interindividual differences in behavioural traits through diversifying selection, which can act on multiple behavioural traits, yet not necessarily in a similar direction (e.g. towards risk taking; Dall et al., 2004). In this scenario, there is no reason to anticipate differences in fitness between individuals with different behavioural scores. Such a diversifying process may be at the phenotypic level only, in accordance with low to modest heritability of behavioural scores in our study. Furthermore, such a diversifying process could, for example, take place during adolescence, as discussed above, with individuals continuing to behave in the way that proved to be beneficial. In this scenario, small initial differences in behaviour, potentially with a largely stochastic origin, may become amplified through reinforcement, resulting in consistent interindividual differences in adulthood.

Author Contributions

Conceptualization: Y.H.G., S.V.; Data curation: Y.H.G.; Formal analysis: Y.H.G. in discussion with S.V.; Funding acquisition: S.V.; Investigation: Y.H.G.; Methodology: Y.H.G., M.M.G.D.; Project

administration: Y.H.G., M.M.G.D.; Resources: Y.H.G., S.V.; Software: Y.H.G.; Supervision: S.V.; Validation: Y.H.G., S.V.; Visualization: Y.H.G., S.V.; Writing – original draft: Y.H.G.; Writing – review & editing: M.M.G.D., S.V.

Data Availability

Data are stored on 10.5281/zenodo.6406723 and will be made available by the authors upon request.

Declaration of Interest

None.

Acknowledgments

We thank the animal caretakers at the University of Groningen for their efforts that made this project possible. We further thank Katharina Riebel for insightful discussion, Owen Clements for designing and building the dominance perch, and Anne Meurs and Bo Westerdijk for their help in performing and scoring all behavioural tests. We thank Nadia van Eekelen and Juultje Ennink for initial pilots of the novel environment test, and Douwe Mul, Jade van Dam, Jeanette Jansma and Lynn Eizinga for initial pilots of the sociality test. Y.H.G. and M.M.G.D. were supported by Adaptive Life grants awarded by the University of Groningen.

References

- Bell, A. M. (2013). Randomized or fixed order for studies of behavioral syndromes? *Behavioral Ecology*, 24(1), 16–20. <https://doi.org/10.1093/beheco/ars148>
- Bell, A. M., & Sih, A. (2007). Exposure to predation generates personality in threespined sticklebacks (*Gasterosteus aculeatus*). *Ecology Letters*, 10(9), 828–834. <https://doi.org/10.1111/j.1461-0248.2007.01081.x>
- Bell, A. M., & Stamps, J. A. (2004). Development of behavioural differences between individuals and populations of sticklebacks, *Gasterosteus aculeatus*. *Animal Behaviour*, 68(6), 1339–1348. <https://doi.org/10.1016/j.anbehav.2004.05.007>
- Besson, A. A., Lagisz, M., Senior, A. M., Hector, K. L., & Nakagawa, S. (2016). Effect of maternal diet on offspring coping styles in rodents: A systematic review and meta-analysis. *Biological Reviews*, 91(4), 1065–1080. <https://doi.org/10.1111/brv.12210>
- Bölting, S., & von Engelhardt, N. (2017). Effects of the social environment during adolescence on the development of social behaviour, hormones and morphology in male zebra finches (*Taeniopygia guttata*). *Frontiers in Zoology*, 14(1), 1–18. <https://doi.org/10.1186/s12983-017-0190-4>
- Boogert, N. J., Farine, D. R., & Spencer, K. A. (2014). Developmental stress predicts social network position. *Biology Letters*, 10(10), 20140561. <https://doi.org/10.1098/rsbl.2014.0561>
- Brandl, H. B., Farine, D. R., Funghi, C., Schuett, W., & Griffith, S. C. (2019). Early-life social environment predicts social network position in wild zebra finches. *Proceedings of the Royal Society B: Biological Sciences*, (1897), 286. <https://doi.org/10.1098/rspb.2018.2579>
- Briga, M., Koetsier, E., Boonekamp, J. J., Jimeno, B., & Verhulst, S. (2017). Food availability affects adult survival trajectories depending on early developmental conditions. *Proceedings of the Royal Society B: Biological Sciences*, 284(1846). <https://doi.org/10.1098/rspb.2016.2287>
- Briga, M., & Verhulst, S. (2017). Individual variation in metabolic reaction norms over ambient temperature causes low correlation between basal and standard metabolic rate. *Journal of Experimental Biology*, 220(18), 3280–3289. <https://doi.org/10.1242/jeb.160069>
- Briga, M., & Verhulst, S. (2021). Mosaic metabolic ageing: Basal and standard metabolic rates age in opposite directions and independent of environmental quality, sex and life span in a passerine. *Functional Ecology*, 35(5), 1055–1068. <https://doi.org/10.1111/1365-2435.13785>
- Brust, V., Wuerz, Y., & Krüger, O. (2013). Behavioural flexibility and personality in zebra finches. *Ethology*, 119(7), 559–569. <https://doi.org/10.1111/eth.12095>
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01.v2.14.4>
- Caldji, C., Tannenbaum, B., Sharma, S., Francis, D., Plotsky, P. M., & Meaney, M. J. (1998). Maternal care during infancy regulates the development of neural systems mediating the expression of fearfulness in the rat. *Proceedings of the National Academy of Sciences*, 95(April), 5335–5340.
- Carere, C., & Eens, M. (2005). Unravelling animal personalities: How and why individuals consistently differ. *Behaviour*, 142(9), 1149–1157. <https://doi.org/10.1163/156853905774539436>
- Clark, C. W. (1994). Antipredator behavior and the asset-protection principle. *Behavioral Ecology*, 5(2), 159–170. <https://doi.org/10.1093/beheco/5.2.159>
- Corr, P. I. (2008). *The reinforcement sensitivity theory of personality*. Cambridge University Press.
- Dall, S. R. X., Houston, A. I., & McNamara, J. M. (2004). The behavioural ecology of personality: Consistent individual differences from an adaptive perspective. *Ecology Letters*, 7(8), 734–739. <https://doi.org/10.1111/j.1461-0248.2004.00618.x>
- David, M., Auclair, Y., & Cézilly, F. (2011). Personality predicts social dominance in female zebra finches, *Taeniopygia guttata*, in a feeding context. *Animal Behaviour*, 81(1), 219–224. <https://doi.org/10.1016/j.anbehav.2010.10.008>
- David, M., Auclair, Y., & Cézilly, F. (2012). Assessing short- and long-term repeatability and stability of personality in captive zebra finches using longitudinal data. *Ethology*, 118(10), 932–942. <https://doi.org/10.1111/j.1439-0310.2012.02085.x>
- De Kogel, C. H. (1997). Long-term effects of brood size manipulation on morphological development and sex-specific mortality of offspring. *Journal of Animal Ecology*, 66(2), 167. <https://doi.org/10.2307/6019>
- Dhellemmes, F., Finger, J. S., Laskowski, K. L., Guttridge, T. L., & Krause, J. (2020). Comparing behavioural syndromes across time and ecological conditions in a free-ranging predator. *Animal Behaviour*, 162, 23–33. <https://doi.org/10.1016/j.anbehav.2020.01.009>
- Dingemanse, N. J., Both, C., Drent, P. J., & Tinbergen, J. M. (2004). Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society B: Biological Sciences*, 271(1541), 847–852. <https://doi.org/10.1098/rspb.2004.2680>
- Dingemanse, N. J., Both, C., van Noordwijk, A. J., Rutten, A. L., & Drent, P. J. (2003). Natal dispersal and personalities in great tits (*Parus major*). *Proceedings of the Royal Society B: Biological Sciences*, 270(1516), 741–747. <https://doi.org/10.1098/rspb.2002.2300>
- Dingemanse, N. J., & Réale, D. (2005). Natural selection and animal personality. *Behavior*, 142(9), 1159–1184.
- Dingemanse, N. J., & Wright, J. (2020). Criteria for acceptable studies of animal personality and behavioural syndromes. *Ethology*, 126(9), 865–869. <https://doi.org/10.1111/eth.13082>
- 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 three-spined stickleback. *Journal of Animal Ecology*, 76(6), 1128–1138. <https://doi.org/10.1111/j.1365-2656.2007.01284.x>
- Dochtermann, N. A., & Royauté, R. (2019). The mean matters: Going beyond repeatability to interpret behavioural variation. *Animal Behaviour*, 153, 147–150. <https://doi.org/10.1016/j.anbehav.2019.05.012>
- Dochtermann, N. A., Schwab, T., & Sih, A. (2015). The contribution of additive genetic variation to personality variation: Heritability of personality. *Proceedings of the Royal Society B: Biological Sciences*, 282(1798), 20142201. <https://doi.org/10.1098/rspb.2014.2201>
- Duckworth, R. A. (2008). Adaptive dispersal strategies and the dynamics of a range expansion. *American Naturalist*, 172(S1), S4–S17. <https://doi.org/10.1086/588289>
- Edwards, H. A., Hajduk, G. K., Durieux, G., Burke, T., & Dugdale, H. L. (2015). No association between personality and candidate gene polymorphisms in a wild bird population. *PLoS One*, 10(10), 1–13. <https://doi.org/10.1371/journal.pone.0138439>
- Fidler, A. E., van Oers, K., Drent, P. J., Kuhn, S., Mueller, J. C., & Kempenaers, B. (2007). Drd4 gene polymorphisms are associated with personality variation in a passerine bird. *Proceedings of the Royal Society B: Biological Sciences*, 274(1619), 1685–1691. <https://doi.org/10.1098/rspb.2007.0337>
- Fisher, D. N., David, M., Rodríguez-Muñoz, R., & Tregenza, T. (2018). Lifespan and age, but not residual reproductive value or condition, are related to behaviour in wild field crickets. *Ethology*, 124(5), 338–346. <https://doi.org/10.1111/eth.12735>
- Fox, J., & Weisberg, S. (2019). *An {R} companion to applied regression* (3rd ed.). Sage. Retrieved from <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
- Fresneau, N., Klue, E., & Brommer, J. E. (2014). A sex-specific behavioral syndrome in a wild passerine. *Behavioral Ecology*, 25(2), 359–367. <https://doi.org/10.1093/beheco/aru008>
- Gallup, G. G. (1977). Tonic immobility: The role of fear and predation. *Psychological Record*, 27(S1), 41–61. <https://doi.org/10.1007/bf03394432>
- Gallup, G. G. (1979). Tonic Immobility as a Measure of Fear in Domestic Fowl. *Animal Behaviour*, 27(1), 316–317. [https://doi.org/10.1016/0003-3472\(79\)90159-3](https://doi.org/10.1016/0003-3472(79)90159-3)
- Gammell, M. P., De Vries, H., Jennings, D. J., Carlin, C. M., & Hayden, T. J. (2003). David's score: A more appropriate dominance ranking method than Clutton-Brock et al.'s index. *Animal Behaviour*, 66(3), 601–605. <https://doi.org/10.1006/anbe.2003.2226>
- Gerritsma, Y. H., Driessen, M. M. G., Tangili, M., de Boer, S. F., & Verhulst, S. (2022). Experimentally manipulated food availability affects offspring quality but not quantity in zebra finch meso- populations. *Oecologia*, 199, 769–783. <https://doi.org/10.1007/s00442-022-05183-y>
- Hall, M. L., van Asten, T., Katsis, A. C., Dingemanse, N. J., Magrath, M. J. L., & Mulder, R. A. (2015). Animal personality and pace-of-life syndromes: Do fast-exploring fairy-wrens die young? *Frontiers in Ecology and Evolution*, 3(Mar), 1–14. <https://doi.org/10.3389/fevo.2015.00028>
- Hemelrijk, C. K., Wantia, J., & Gyga, L. (2005). The construction of dominance order: Comparing performance of five methods using an individual-based model. *Behaviour*, 142(8), 1043–1064.

- Hope, B. V., Fjellner, K. L., Renn, S. C. P., & Hurd, P. L. (2020). Juvenile stress disrupts the development of an exploration–boldness behavioural syndrome in convict cichlid fish. *Animal Behaviour*, 161, 95–102. <https://doi.org/10.1016/j.anbehav.2020.01.004>
- Jimeno, B., Hau, M., Gómez-Díaz, E., & Verhulst, S. (2019). Developmental conditions modulate DNA methylation at the glucocorticoid receptor gene with cascading effects on expression and corticosterone levels in zebra finches. *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-019-52203-8>
- Kaiser, S., & Sachser, N. (2005). The effects of prenatal social stress on behaviour: Mechanisms and function. *Neuroscience and Biobehavioral Reviews*, 29(2), 283–294. <https://doi.org/10.1016/j.neubiorev.2004.09.015>
- Kassambara, A., & Mundt, F. (2020). factoextra: Extract and visualize the results of multivariate data analyses. R package. Retrieved from <https://cran.r-project.org/package=factoextra>.
- Kelly, A. M., & Goodson, J. L. (2015). Functional interactions of dopamine cell groups reflect personality, sex, and social context in highly social finches. *Behavioural Brain Research*, 280, 101–112. <https://doi.org/10.1016/j.bbr.2014.12.004>
- Kelly, A. M., Kingsbury, M. A., Hoffbuhr, K., Schrock, S. E., Waxman, B., Kabelik, D., Thompson, R. R., & Goodson, J. L. (2011). Vasotocin neurons and septal V1a-like receptors potentially modulate songbird flocking and responses to novelty. *Hormones and Behavior*, 60(1), 12–21. <https://doi.org/10.1016/j.yhbeh.2011.01.012>
- Kemme, K., Kaiser, S., & Sachser, N. (2008). Prenatal stress does not impair coping with challenge later in life. *Physiology and Behavior*, 93(1–2), 68–75. <https://doi.org/10.1016/j.physbeh.2007.07.013>
- King, A. J., Williams, L. J., & Mettke-Hofmann, C. (2015). The effects of social conformity on Gouldian finch personality. *Animal Behaviour*, 99, 25–31. <https://doi.org/10.1016/j.anbehav.2014.10.016>
- Koetsier, E., & Verhulst, S. (2011). A simple technique to manipulate foraging costs in seed-eating birds. *Journal of Experimental Biology*, 214(Pt 8), 1225–1229. <https://doi.org/10.1242/jeb.050336>
- Korsten, P., Mueller, J. C., Hermannstädter, C., Bouwman, K. M., Dingemanse, N. J., Drent, P. J., Liedvogel, M., Matthysen, E., van Oers, K., van Overveld, T., Patrick, S. C., Quinn, J. L., Sheldon, B. C., Tinbergen, J. M., & Kempenaers, B. (2010). Association between DRD4 gene polymorphism and personality variation in great tits: A test across four wild populations. *Molecular Ecology*, 19(4), 832–e843. <https://doi.org/10.1111/j.1365-294X.2009.04518.x>
- Krause, E. T., Krüger, O., & Schielzeth, H. (2017). Long-term effects of early nutrition and environmental matching on developmental and personality traits in zebra finches. *Animal Behaviour*, 128, 103–115. <https://doi.org/10.1016/j.anbehav.2017.04.003>
- Krause, E. T., & Naguib, M. (2011). Compensatory growth affects exploratory behaviour in zebra finches, *Taeniopygia guttata*. *Animal Behaviour*, 81(6), 1295–1300. <https://doi.org/10.1016/j.anbehav.2011.03.021>
- Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, 25(1), 1–18. <https://doi.org/10.18637/jss.v025.i01>
- Lemoine, N. P. (2019). Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, 128(7), 912–928. <https://doi.org/10.1111/oik.05985>
- Lesch, K., Bengel, D., Heils, A., Sabol, S. Z., Benjamin, D., Petri, S., Benjamin, J., Müller, C. R., Hamer, D. H., & Dennis, L. (1996). Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science*, 274(5292), 1527–1531. <https://doi.org/10.1016/B978-0-12-385897-9.00003-3>
- Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9), 1989–2001. <https://doi.org/10.1016/j.jmva.2009.04.008>
- Luttbeg, B., & Sih, A. (2010). Risk, resources and state-dependent adaptive behavioural syndromes. *Philosophical Transactions of the Royal Society B*, 365(1560), 3977–3990. <https://doi.org/10.1098/rstb.2010.0207>
- Mainwaring, M. C., Beal, J. L., & Hartley, I. R. (2011). Zebra finches are bolder in an asocial, rather than social, context. *Behavioural Processes*, 87(2), 171–175. <https://doi.org/10.1016/j.beproc.2011.03.005>
- Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdtke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, 10(December), 1–14. <https://doi.org/10.3389/fpsyg.2019.02767>
- Makowski, D., Ben-Shachar, M. S., & Lüdtke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, 4(40), 1541. <https://doi.org/10.21105/joss.01541>
- Moiron, M., Laskowski, K. L., & Niemelä, P. T. (2020). Individual differences in behaviour explain variation in survival: A meta-analysis. *Ecology Letters*, 23(2), 399–408. <https://doi.org/10.1111/ele.13438>
- Mueller, J. C., Partecke, J., Hatchwell, B. J., Gaston, K. J., & Evans, K. L. (2013). Candidate gene polymorphisms for behavioural adaptations during urbanization in blackbirds. *Molecular Ecology*, 22(13), 3629–3637. <https://doi.org/10.1111/mec.12288>
- Müller, K., & Wickham, H. (2021). tibble: Simple Data Frames. Retrieved from <https://cran.r-project.org/package=tibble>.
- Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews*, 85(4), 935–956. <https://doi.org/10.1111/j.1469-185X.2010.00141.x>
- Niemelä, P. T., & Dingemanse, N. J. (2018). Meta-analysis reveals weak associations between intrinsic state and personality. *Proceedings of the Royal Society B: Biological Sciences*, 285(1873). <https://doi.org/10.1098/rspb.2017.2823>
- Peixoto, M. R. L. V., Karrow, N. A., Newman, A., & Widowski, T. M. (2020). Effects of maternal stress on measures of anxiety and fearfulness in different strains of laying hens. *Frontiers in Veterinary Science*, 7(March). <https://doi.org/10.3389/fvets.2020.00128>
- Petelle, M. B., Martin, J. G. A., & Blumstein, D. T. (2019). Mixed support for state maintaining risky personality traits in yellow-bellied marmots. *Animal Behaviour*, 150, 177–188. <https://doi.org/10.1016/j.anbehav.2019.02.008>
- R Development Core Team. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org>.
- Rangassamy, M., Dalmas, M., Féron, C., Gouat, P., & Rödel, H. G. (2015). Similarity of personalities speeds up reproduction in pairs of a monogamous rodent. *Animal Behaviour*, 103, 7–15. <https://doi.org/10.1016/j.anbehav.2015.02.007>
- Réale, D., Reader, S. M., Sol, D., McDougall, P. T., & Dingemanse, N. J. (2007). Integrating animal temperament within ecology and evolution. *Biological Reviews*, 82(2), 291–318. <https://doi.org/10.1111/j.1469-185X.2007.00010.x>
- Roff, D. A. (2002). *Life history evolution*. Sinauer.
- Royauté, R., Berdal, M. A., Garrison, C. R., & Dochtermann, N. A. (2018). Paced life? A meta-analysis of the pace-of-life syndrome hypothesis. *Behavioral Ecology and Sociobiology*, 72(3), 64. <https://doi.org/10.1007/s00265-018-2472-z>
- Rstudio Team. (2020). Rstudio: Integrated Development for R. Retrieved from Rstudio, PBC, Boston, MA website: www.rstudio.com.
- Ruploh, T., Bischof, H. J., & von Engelhardt, N. (2013). Adolescent social environment shapes sexual and aggressive behaviour of adult male zebra finches (*Taeniopygia guttata*). *Behavioral Ecology and Sociobiology*, 67(2), 175–184. <https://doi.org/10.1007/s00265-012-1436-y>
- Ruploh, T., Bischof, H. J., & von Engelhardt, N. (2014). Social experience during adolescence influences how male zebra finches (*Taeniopygia guttata*) group with conspecifics. *Behavioral Ecology and Sociobiology*, 68(4), 537–549. <https://doi.org/10.1007/s00265-013-1668-5>
- Ruploh, T., Henning, M., Bischof, H. J., & von Engelhardt, N. (2015). Effects of social conditions during adolescence on courtship and aggressive behavior are not abolished by adult social experience. *Developmental Psychobiology*, 57(1), 73–82. <https://doi.org/10.1002/dev.21262>
- Sachser, N., Hennessy, M. B., & Kaiser, S. (2011). Adaptive modulation of behavioural profiles by social stress during early phases of life and adolescence. *Neuroscience and Biobehavioral Reviews*, 35(7), 1518–1533. <https://doi.org/10.1016/j.neubiorev.2010.09.002>
- Schuett, W., & Dall, S. R. X. (2009). Sex differences, social context and personality in zebra finches, *Taeniopygia guttata*. *Animal Behaviour*, 77(5), 1041–1050. <https://doi.org/10.1016/j.anbehav.2008.12.024>
- Schuett, W., Dall, S. R. X., & Royle, N. J. (2011). Pairs of zebra finches with similar 'personalities' make better parents. *Animal Behaviour*, 81(3), 609–618. <https://doi.org/10.1016/j.anbehav.2010.12.006>
- Schuett, W., Tregenza, T., & Dall, S. R. X. (2010). Sexual selection and animal personality. *Biological Reviews*, 85(2), 217–246. <https://doi.org/10.1111/j.1469-185X.2009.00101.x>
- Spencer, K. A., & Verhulst, S. (2007). Delayed behavioral effects of postnatal exposure to corticosterone in the zebra finch (*Taeniopygia guttata*). *Hormones and Behavior*, 51(2), 273–280. <https://doi.org/10.1016/j.yhbeh.2006.11.001>
- Stamps, J. A. (2007). Growth-mortality tradeoffs and 'personality traits' in animals. *Ecology Letters*, 10(5), 355–363. <https://doi.org/10.1111/j.1461-0248.2007.01034.x>
- Stamps, J. A., & Groothuis, T. G. G. (2010). The development of animal personality: Relevance, concepts and perspectives. *Biological Reviews*, 85(2), 301–325. <https://doi.org/10.1111/j.1469-185X.2009.00103.x>
- Stan Development Team. (2020). RStan: the R interface to Stan. Retrieved from <http://mc-stan.org/>.
- Thys, B., Grunst, A. S., Staes, N., Pinxten, R., Eens, M., & Grunst, M. L. (2021). The serotonin transporter gene and female personality variation in a free-living passerine. *Scientific Reports*, 1–9. <https://doi.org/10.1038/s41598-021-88225-4>
- Verbeek, M. E. M., Drent, P. J., & Wiepkema, P. R. (1994). Consistent individual differences in early exploratory behaviour of male great tits. *Animal Behaviour*, 48, 1113–1121. <https://doi.org/10.1006/anbe.1994.1344>
- Verhulst, E. C., Mateman, A. C., Zwier, M. V., Caro, S. P., Verhoeven, K. J. F., & van Oers, K. (2016). Evidence from pyrosequencing indicates that natural variation in animal personality is associated with DRD4 DNA methylation. *Molecular Ecology*, 25(8), 1801–1811. <https://doi.org/10.1111/mec.13519>
- de Villemereuil, P., Morrissey, M. B., Nakagawa, S., & Schielzeth, H. (2018). Fixed-effect variance and the estimation of repeatabilities and heritabilities: Issues and solutions. *Journal of Evolutionary Biology*, 31(4), 621–632. <https://doi.org/10.1111/jeb.13232>
- Wickham, H. (2020). tidy: Tidy Messy Data. Retrieved from <https://cran.r-project.org/package=tidy>.
- Wolf, M., van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs favour the evolution of animal personalities. *Nature*, 447(7144), 581–584. <https://doi.org/10.1038/nature05835>
- Wolf, M., & Weissing, F. J. (2010). An explanatory framework for adaptive personality differences. *Philosophical Transactions of the Royal Society B*, 365(1560), 3959–3968. <https://doi.org/10.1098/rstb.2010.0215>
- Wuerz, Y., & Krüger, O. (2015). Personality over ontogeny in zebra finches: Long-term repeatable traits but unstable behavioural syndromes. *Frontiers in Zoology*, 12(Suppl 1), S9. <https://doi.org/10.1186/1742-9994-12-S1-S9>
- Zann, R. A. (1996). *The zebra finch: A synthesis of field and laboratory studies*. Oxford University Press.

APPENDIX

Table A1

PCA eigenvalues and loadings for the novel environment and the novel object tests

Behavioural test	% Variance PC1	Eigenvalue	Loadings
Novel environment	72.7	2.18	Locations: 0.65 Latency to start: −0.42 Movements: 0.64
Novel object	62.9	1.26	Latency to inspect novel environment: 0.71 Δ first move between phase 2 and 3: 0.71

As higher scores on the first PC of the novel object corresponded with higher latency to inspect the novel object, and a lower latency to first move when a novel object was present, indicating lower levels of boldness, we multiplied it by −1.

Table A2

Model formulas for the multivariate analysis of the personality traits

Response variable	Personality trait	Fixed effects	Random effects	Additional random effects
PC1 exploration	Exploration	DT*Sex+	(1 BirdId)	–
PC1 novel object	Boldness	Test	+	+(1 FirstPhase)
David's score	Dominance		(1 GenDad)	+(1 DomGroup)
Group preference	Sociality		+	+(1 StimGroup)
Immobility time	Anxiety	DT*Sex+	(1 GenMom)	+(1 Experimenter)
		Test+FixTime		

Identical fixed effects are included for all response variables, with identity of the individual and the identity of its genetic parents included as random Gaussian intercepts. DT: developmental treatment. Additional random Gaussian intercepts included are specified per response variable.

Table A3

Sample sizes per behavioural test, separated by sex and developmental treatment

Behavioural test	Two tests		One test	
Novel environment	89	♀ 55 (32 – 23) ♂ 34 (14 – 20)	93	♀ 57 (33 – 24 ^a) ♂ 36 (16 – 20 ^b)
Novel object	93	♀ 57 ^a ♂ 36 ^b	93	♀ 57 ^a ♂ 36 ^b
Tonic immobility	93	♀ 57 ^a ♂ 36 ^b	93	♀ 57 ^a ♂ 36 ^b
Sociality	92	♀ 56 (32 – 24) ♂ 36 ^b	93	♀ 57 ^a ♂ 36 ^b
Dominance	51	♀ 18 (10 – 8) ♂ 33 (15 – 18)	68	♀ 32 (18 – 14) ♂ 36 ^b

The sample size per developmental treatment is given between brackets: the first number is for the benign and the second for the harsh treatment. Sample sizes are given for all individuals for which we have data for both tests or for at least one test.

^a All females, 33 from a benign and 24 from a harsh developmental treatment.

^b All males, 16 from a benign and 20 from a harsh development.

Table A4

Marginal effects of all predictors and interactions on behaviour for the model including the developmental manipulation

Predictor	Exploration		Novel object		Sociality		Tonic immobility		Dominance	
	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean
DT	0.787	−0.269 (−0.962, 0.404)	0.925	0.378 (−0.165, 0.865)	0.666	−0.115 (−0.657, 0.413)	0.618	0.053 (−0.322, 0.427)	0.699	0.146 (−0.405, 0.7)
Sex	0.746	0.197 (−0.387, 0.797)	0.607	0.06 (−0.389, 0.486)	0.934	−0.452 (−1.071, 0.146)	0.613	−0.042 (−0.332, 0.257)	–	–
Test	0.739	−0.107 (−0.42, 0.226)	0.798	−0.144 (−0.48, 0.186)	0.954	0.158 (−0.027, 0.338)	0.593	−0.025 (−0.241, 0.183)	0.753	0.086 (−0.167, 0.333)
DT*Sex	0.643	0.194 (−0.836, 1.28)	0.731	0.465 (−1.55, 2.641)	0.913	0.549 (−0.239, 1.355)	0.826	0.256 (−0.265, 0.818)	0.868	0.5 (−0.369, 1.402)

For each posterior mean estimate, the other predictors were kept averaged over levels. The variable not shown is fixation time, concerning tonic immobility, which was kept at its mean value when calculating posterior means. DT: developmental treatment.

Table A5

Marginal effects of all predictors and interactions on behaviour for the model including the mass on day 15

Predictor	Exploration		Novel object		Sociality		Tonic immobility		Dominance	
	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean
Mass on day 15	0.584	−0.049 (−0.497, 0.427)	0.943	0.256 (−0.068, 0.562)	0.876	0.186 (−0.12, 0.508)	0.855	0.108 (−0.093, 0.311)	0.94	0.255 (−0.073, 0.575)
Sex	0.795	0.294 (−0.391, 1.039)	0.661	−0.107 (−0.622, 0.397)	0.97	−0.666 (−1.398, 0.034)	0.918	−0.227 (−0.55, 0.098)	—	—
Test	0.753	−0.119 (−0.457, 0.217)	0.825	−0.164 (−0.521, 0.164)	0.952	0.181 (−0.04, 0.392)	0.622	−0.031 (−0.23, 0.174)	0.77	0.097 (−0.17, 0.355)
Mass on day 15*Sex	0.718	−0.257 (−1.118, 0.645)	0.694	−0.396 (−2.738, 1.959)	0.831	0.302 (−0.333, 0.927)	0.979	0.437 (0.01, 0.839)	1	1.105 (0.496, 1.779)

For each posterior mean estimate, the other predictors were kept averaged over levels (factors) or at their median (continuous predictor, i.e. mass). The variable not shown is fixation time, concerning tonic immobility, which was kept at its mean value when calculating posterior means.

Table A6

Marginal effects of all predictors and interactions on behaviour for the model including brood size

Predictor	Exploration		Novel object		Sociality		Tonic immobility		Dominance	
	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean	Pd	Posterior mean
Brood size	0.897	−0.197 (−0.51, 0.098)	0.522	0.006 (−0.212, 0.231)	0.558	0.014 (−0.199, 0.218)	0.701	0.039 (−0.103, 0.183)	0.859	0.139 (−0.119, 0.381)
Sex	0.626	0.105 (−0.508, 0.728)	0.574	0.042 (−0.409, 0.524)	0.95	−0.522 (−1.165, 0.086)	0.576	−0.032 (−0.35, 0.296)	—	—
Test	0.732	−0.106 (−0.446, 0.225)	0.825	−0.156 (−0.471, 0.195)	0.96	0.186 (−0.034, 0.383)	0.519	−0.004 (−0.218, 0.207)	0.739	0.087 (−0.176, 0.352)
Brood size*Sex	0.632	0.076 (−0.372, 0.516)	0.523	0.036 (−1.995, 2.031)	0.864	0.194 (−0.162, 0.541)	0.717	0.075 (−0.183, 0.339)	0.513	0.006 (−0.411, 0.415)

For each posterior mean estimate, the other predictors were kept averaged over levels (factor) or at their median (continuous predictor, i.e. brood size). The variable not shown is fixation time, concerning tonic immobility, which was kept at its mean value when calculating posterior means.

Table A7

Correlations between behavioural traits

Trait combination	Group	Estimate	Low	High
Exploration–Novel object	All	−0.06236	−0.30752	0.157219
Exploration–Novel object	Benign	−0.21709	−0.51962	0.085546
Exploration–Novel object	Harsh	0.057124	−0.25356	0.367259
Exploration–Novel object	Male	0.00874	−0.31706	0.350349
Exploration–Novel object	Female	−0.08396	−0.37503	0.199457
Exploration–Immobility	All	−0.23024	−0.44995	−0.00371
Exploration–Immobility	Benign	0.070353	−0.2472	0.372662
Exploration–Immobility	Harsh	−0.46183	−0.70855	−0.20129
Exploration–Immobility	Male	−0.22283	−0.54424	0.107184
Exploration–Immobility	Female	−0.17752	−0.47062	0.117362
Exploration–Sociality	All	0.170006	−0.06723	0.403276
Exploration–Sociality	Benign	0.150086	−0.18897	0.501152
Exploration–Sociality	Harsh	0.1421	−0.17233	0.439972
Exploration–Sociality	Male	0.039942	−0.30114	0.403292
Exploration–Sociality	Female	0.251936	−0.04515	0.550231
Exploration–Dominance	All	0.3541	0.093847	0.612071
Exploration–Dominance	Benign	0.381778	−0.02101	0.693453
Exploration–Dominance	Harsh	0.246786	−0.13088	0.605352
Exploration–Dominance	Male	0.132357	−0.23413	0.504813
Exploration–Dominance	Female	0.56703	0.24465	0.83488
Novel object–Immobility	All	0.187423	−0.01637	0.402487
Novel object–Immobility	Benign	0.154583	−0.13517	0.424022
Novel object–Immobility	Harsh	0.296628	0.026344	0.563829
Novel object–Immobility	Male	0.252333	−0.02827	0.55335
Novel object–Immobility	Female	0.046514	−0.22974	0.336856
Novel object–Sociality	All	−0.01174	−0.25662	0.225082
Novel object–Sociality	Benign	0.069775	−0.25561	0.419715
Novel object–Sociality	Harsh	−0.12642	−0.41951	0.167872
Novel object–Sociality	Male	−0.17845	−0.48315	0.156185
Novel object–Sociality	Female	0.160823	−0.14622	0.453949
Novel object–Dominance	All	0.11904	−0.15569	0.388649
Novel object–Dominance	Benign	0.124799	−0.24125	0.466167
Novel object–Dominance	Harsh	0.064659	−0.29412	0.430256
Novel object–Dominance	Male	0.187484	−0.12978	0.505697

(continued on next page)

Table A7 (continued)

Trait combination	Group	Estimate	Low	High
Novel object–Dominance	Female	−0.13652	−0.52076	0.243385
Immobility–Sociality	All	−0.06571	−0.29801	0.175189
Immobility–Sociality	Benign	0.166172	−0.1569	0.475564
Immobility–Sociality	Harsh	−0.21807	−0.49615	0.075905
Immobility–Sociality	Male	0.000481	−0.32829	0.329451
Immobility–Sociality	Female	−0.11309	−0.40023	0.193283
Immobility–Dominance	All	0.017459	−0.26561	0.306507
Immobility–Dominance	Benign	0.272524	−0.07777	0.611622
Immobility–Dominance	Harsh	−0.09689	−0.46897	0.241623
Immobility–Dominance	Male	0.041548	−0.3027	0.382969
Immobility–Dominance	Female	0.021165	−0.3836	0.399416
Sociality–Dominance	All	0.018454	−0.2648	0.31257
Sociality–Dominance	Benign	0.215787	−0.16441	0.590023
Sociality–Dominance	Harsh	−0.11611	−0.46012	0.254496
Sociality–Dominance	Male	0.032914	−0.3334	0.384696
Sociality–Dominance	Female	−0.07719	−0.44994	0.288747

Correlations between traits are given for all individuals and are stratified over (1) developmental conditions and (2) sex.

Table A8

Repeatability estimates of behavioural tests in captive zebra finches

Repeatability	[95% CI], <i>F</i> statistic or $\pm$ SE	Test	Days between tests (mean $\pm$ SE)	<i>N</i>	Life stage	Sex	Source
0.44	[0.11, 0.64]	Activity	148 $\pm$ 67	24	Adult	M/F	Brust et al. (2013)
0.44	[0.21, 0.56]	Tonic immobility	48 $\pm$ 17				
0.40	$F_{40,80} = 2.15$	Activity	7 $\pm$ 0	41	Adult ^b	F	David et al. (2011)
0.43	$F_{40,80} = 2.51$	Neophobia					
0.37	$F_{40,80} = 2.19$	Startle test					
0.43	$F_{40,80} = 2.48$	Struggle rate					
0.81	$F_{40,80} = 9.67$	Exploration					
0.81	[0.59, 0.95]	Exploration	7 $\pm$ 0	20	Adult ^b	F	David et al. (2012)
0.35	[0, 0.73]		3 $\pm$ 0				
0.76	[0.46, 0.92]		263 ^{a,c}				
0.54	[0.21, 0.88]	Struggle rate	7 $\pm$ 0				
0.88	[0.78, 0.99]		3 $\pm$ 0				
0.15	[−0.31, 0.61]		209 ^{a,c}				
>0.37 ^c	— ^c	Exploration	1 $\pm$ 0	50 ^d	Adult	M/F	Krause and Naguib (2011)
0.44 ^e	$\pm$ 0.07	Exploration	72 ^c	147	Adult	M/F	Krause et al. (2017)
0.38	$\pm$ 0.07						
0.36 ^e	$\pm$ 0.06		1532 ^c				
0.32	$\pm$ 0.06						
0.52 ^e	$\pm$ 0.06	Neophobia	72 ^c				
0.38	$\pm$ 0.06						
0.50 ^e	$\pm$ 0.06		1532 ^c				
0.39	$\pm$ 0.06						
0.21	$F_{2,90} = 4.22$	Exploration	21 ^c	31	Adult	M	Mainwaring et al. (2011)
0.06	$F_{2,48} = 2.05$			17		F	
0.03	$F_{2,90} = 2.30$	Neophobia		31		M	
0.02	$F_{2,48} = 1.42$			17		F	
0.32	$\pm$ 0.11, $F_{23,69} = 2.83$	Exploration	7 $\pm$ 0	24	Adult	M/F	Schuett and Dall (2009)
0.60	$\pm$ 0.14, $F_{11,35} = 6.74$			12		M	
−0.06	$\pm$ 0.11, $F_{11,34} = 0.79$			12		F	
0.29	$\pm$ 0.09	Exploration	7 $\pm$ 0	37	Adult	M	Schuett et al. (2011)
0.41	$\pm$ 0.09			40		F	
0.37	— ^c	Aggression	11 $\pm$ 0.07	37		M	
0.64	— ^c			36		F	

F: female; M: male.

^a Median.

^b Young, but sexually mature.

^c No further information provided.

^d Number of broods.

^e Multiple behaviours within a test.

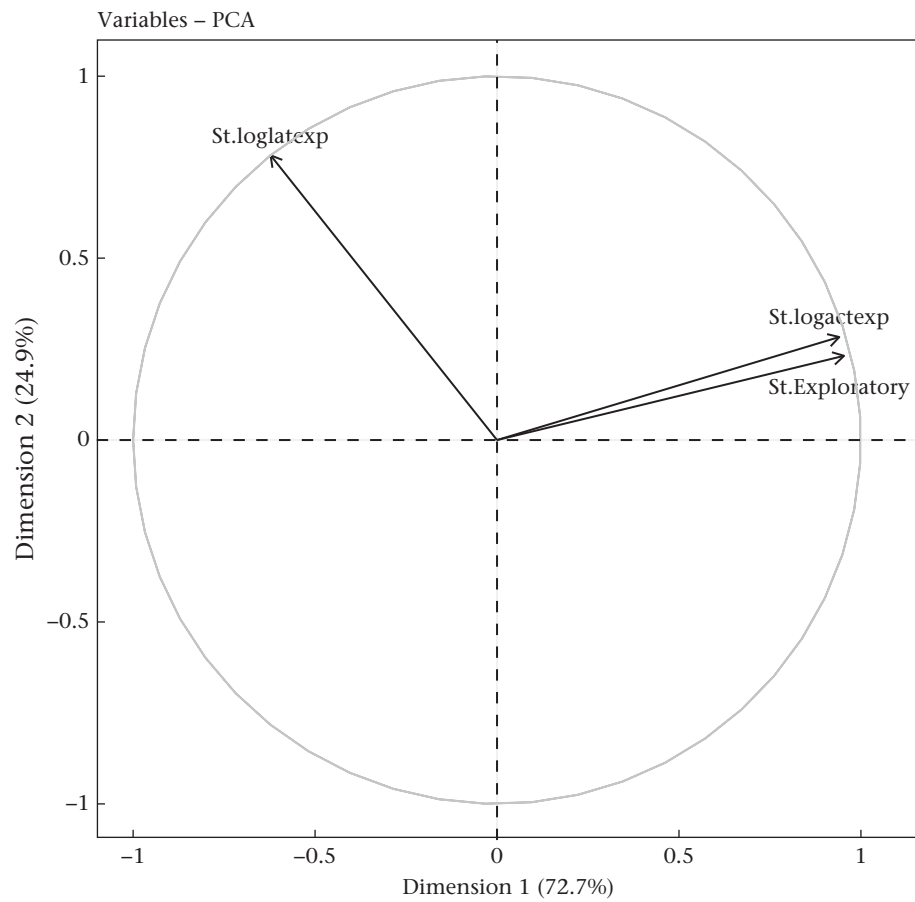


Figure A1. PCA on the novel environment test. The direction of each metric is given for the first two principal components, including the percentage of variance explained by each. All variables were standardized prior to the PCA (referred to as 'St.') and the latency in starting exploring (latexp) and the activity during the test (actexp) were $\log_{10}$ transformed.

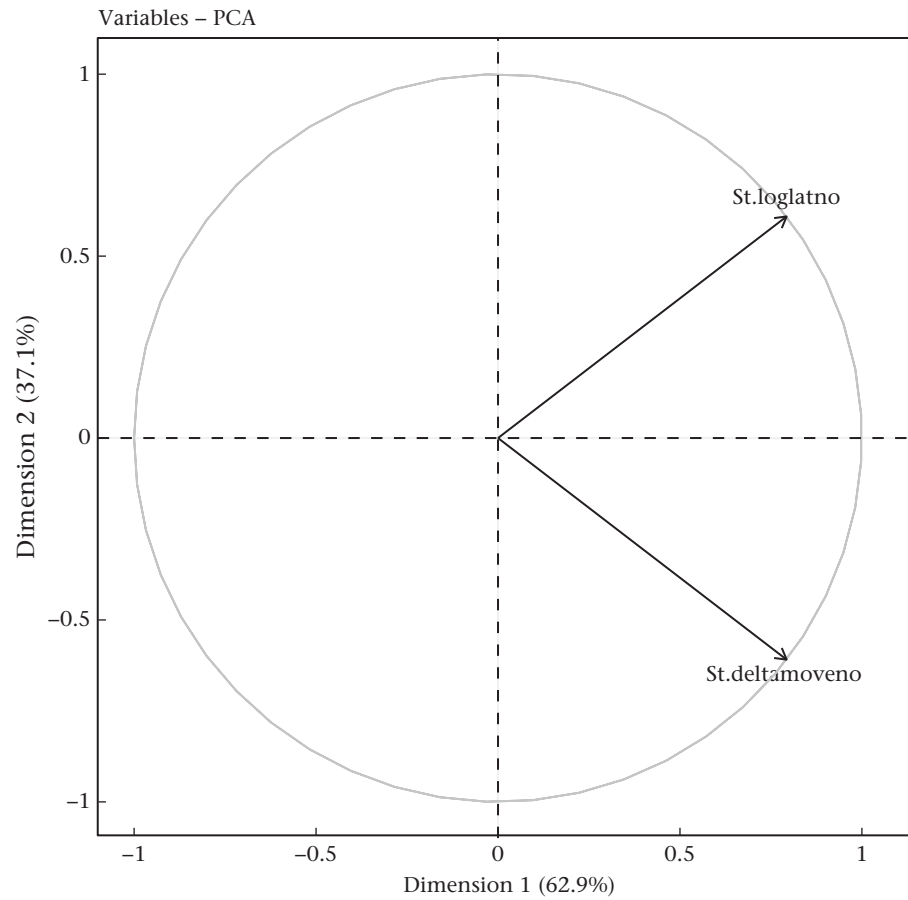


Figure A2. PCA on the novel object test. The direction of each metric is given for the first two principal components, including the percentage of variance explained by each. All variables were standardized prior to the PCA (referred to as 'St.') and the latency to inspect the novel object (latno) was $\log_{10}$ transformed. Deltamoveno refers to the difference in the latency to the first move between phase 2 and phase 3.