Research Article

Repeatability and Heritability of Behavioural Types in a Social Cichlid

Noémie Chervet, Markus Zöttl, Roger Schürch, Michael Taborsky, and Dik Heg

Department of Behavioural Ecology, Institute of Ecology and Evolution, University of Bern, 3032 Hinterkappelen, Switzerland

Correspondence should be addressed to Dik Heg, dik.heg@iee.unibe.ch

Received 11 December 2010; Accepted 28 February 2011

Academic Editor: Kristina M. Sefc

Copyright © 2011 Noémie Chervet et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aim. The quantitative genetics underlying correlated behavioural traits (“animal personality”) have hitherto been studied mainly in domesticated animals. Here we report the repeatability ($R$) and heritability ($h^2$) of behavioural types in the highly social cichlid fish Neolamprologus pulcher.

Methods. We tested 1779 individuals repeatedly and calculated the $h^2$ of behavioural types by variance components estimation (GLMM REML), using 1327 offspring from 162 broods from 74 pairs.

Results. Repeatability of behavioural types was significant and considerable (0.546), but declined from 0.83 between tests conducted on the same day, to 0.19 on tests conducted up to 1201 days apart. All $h^2$ estimates were significant but low (e.g., pair identity $h^2 = 0.15 \pm 0.03$ SE). Additionally, we found significant variation between broods nested within the parent(s), but these were not related to several environmental factors tested.

Conclusions. We conclude that despite a considerable $R$, $h^2$ in this cichlid species is low, and variability in behavioural type appears to be strongly affected by other (non)genetic effects.

1. Introduction

Individuals within animal populations often differ consistently in how they cope with environmental and social challenges, for instance with some individuals typically reacting shy and nonaggressively to such (novel) challenges and others reacting bold and aggressively [1–3]. Often, these individual differences are referred to as “behavioural types” or “coping styles” (e.g., shy, bold), and behavioural traits may covary amongst each other (e.g., shy individuals are also nonaggressive and nonexplorative, whereas bold individuals are more aggressive and explorative, e.g., [4]). Correlated behavioural trait values on the population level are commonly denoted as animal personalities, behavioural syndromes, or temperaments [5]. Central questions in animal personality research are (i) whether differences in behavioural types have a genetic basis, that is, whether they are heritable; (ii) whether and to what extent individuals remain consistent in their behavioural traits over time, that is, whether behavioural responses of individuals are repeatable [6–8].

The genetic components affecting the expression of different personalities have been well explored in humans (see reviews e.g., [9–15]), domesticated animals [16, 17], and animal model systems [18–20]. The evolutionary ecological factors responsible for the evolution of variation in behavioural types remains somewhat enigmatic and difficult to explain in natural animal populations, despite recent theoretical advances showing how life-history tradeoffs might generate and maintain such variation [21, 22]. In principle, consistent individual differences and covariation in behavioural traits are a paradox in evolutionary biology, particularly if such differences have a genetic basis. Standard theory would expect each individual to flexibly adjust their behaviour to the environment. For instance, in a predator rich environment each individual should devalue future fitness in favour of current fitness and adjust their behaviour accordingly (e.g., hide more and be less explorative).

In the human psychology literature it has been acknowledged that although personality differences have a genetic basis [9–15], other (social) factors might impinge upon and alter the expression of human behavioural types over a lifetime, including for instance life events (like the death of a partner, [23]). If personalities are truly fixed over life and can be accurately measured by using standardized
questionnaires (e.g., [24]), or by using standardized observational assessments (e.g., [25–28]), or behavioural tests (as used to determine personalities in many invertebrates and vertebrates, see [2]), the repeatability of behavioural types should approach one. Clearly, this is not the case in humans ([29]; where it increases from 0.31 in childhood to a plateau of 0.73 beyond 50 years of age: [30]) and in most animal studies [7], which led to the proposition to incorporate behavioural reaction norms into the animal personality concept [6]. By incorporating behavioural reaction norms, animal behaviour can be analysed using standard game theory, where behavioural strategies may be governed by internal “state” (e.g., body condition, sex, status, and reproductive activity), sometimes resulting in alternative strategies (e.g., depending on morphology or age), but where these strategies and behaviours are not necessarily fixed for life (i.e., without the need to assume the existence of animal personalities). In fact, this has been the major criticism of animal personality research: if the long-term stability (“repeatability”, $R$) is not proven in any study population, why not assume that the current variability between individuals reflects the current variability in “states” of the individuals tested? Using a meta-analysis, Bell and others showed that the repeatability estimate significantly declined with the time interval between the different tests [7]. As the number of long-term repeatability studies in animal personality research is still very low [31] and biased towards domesticated animals [32–42], our first target is to test for the long-term repeatability of behavioural types, using descendants of a wild population of a cooperatively breeding cichlid, *Neolamprologus pulcher*. We use the cichlid *N. pulcher* as model species, where individuals have been shown to differ consistently in behavioural traits (across the bold–shy continuum) in both the field [43] and the laboratory (stocks derived from the same field population, [4, 39, 44–46]). In addition, males and females have been shown to remain relatively stable in trait values from the juvenile stage (when they are the small, subordinate helpers of an adult pair) to early adulthood (when they are large, subordinate helpers and reproductively mature [39]). Behavioural types in this species may also influence sociality, reproduction [46], and helping behaviour [39, 44, 47]; and also alloparental brood care differs consistently between female helpers [47]. The major hypothesis proposed to explain this variability in behavioural types suggests that subordinates trade-off effort to gain social dominance inside versus outside their territory, which either selects for distinct life-history strategies (e.g., nonexplorative, helpful, and risk-averse individuals opt for dominance inside their group’s territory, whereas explorative, bold, and aggressive individuals opt for dominance outside their group’s territory, which involves early dispersal and independent breeding [39, 48]) or for diverse ontogenetic trajectories in behavioural types (e.g., young fish being risk averse whereas older fish being risk prone). Here we expand the time frame of standardised tests to encompass the lifetime of these fish.

The second target of this study has been to estimate the genetic variance underlying phenotypic variation in personality traits in *N. pulcher* (“heritability” $h^2$ [49, 50]), and to compare estimates of $R$ and $h^2$. Repeatability $R$ often sets the upper limit to the heritability of a trait, and both measures are correlated in comparisons across species or populations, if the phenotypic variances in the compared units are governed by similar processes involving additive genetic variance. Any species or population showing a high heritability in a behavioural trait should also show a high repeatability in this trait (as genes are more likely to be involved in the determination of the behaviour). In contrast, a high repeatability can coincide with a low heritability, if the behavioural type is based, for instance, on the (current) internal state of individuals or on their (life-history) strategy [21] or social strategy [51], which may cause variation that is largely independent from genetic effects. Maternal or paternal effects, maternal additive genetic effects and genotype × environment interactions may also yield discrepancies between the heritability and the repeatability estimates (e.g., with heritability exceeding the repeatability, [52]). Both low repeatability and low heritability would indicate that the population exhibits no animal personalities, particularly if the repeatability diminishes over time.

In *Neolamprologus pulcher*, personality traits (such as boldness, aggressiveness, and propensity to explore) are consistently different between individuals (see references above) and related to two major life-history decisions: whether to help and whether to disperse. Furthermore, these traits are not related to growth rate in fish kept singly [39], but if living in groups consisting of members with divergent personality types, shy fish were found to grow quicker in body length than bold fish [53]. Studies of personalities are particularly interesting in social species like cichlids and primates [54], as they may bear similarities to the human personality axes which incorporate significant aspects of human behaviour and sociality (e.g., “OCEAN”: Openness, Conscientiousness, Extraversion, Agreeableness, but with the possible exception of Neuroticism [24, 55–58]). Human and animal personality differences may be governed by similar differences in (neuro)physiological responses to environmental challenges and stressful situations [23, 59–62]. In this study we first estimated the repeatability of behavioural types based on a combined score of boldness, aggressiveness, and exploration propensity by comparing test results of individuals obtained successively with intervals ranging from 0 to 1201 days (which approaches the maximum life span of this species: see [63]). In a second step we estimated the heritability of these same behavioural types using parent(s)-offspring regressions and variance component estimations from offspring derived from different broods of the same pair.

2. Methods

2.1. Study Species. Descendants of wild caught *Neolamprologus pulcher* (from wild animals collected in 1996, 2006, and 2009 near Kasalakawe, Zambia, so-called three different “stocks” and their crossings) were used for this study and tested in the years 2005–2008 and 2010. These fish are well-studied cooperative breeders [64], endemic to Lake
Tanganyika where they live in breeding groups composed usually of a dominant breeder male, one to several breeder females and some helpers [65]. All fish were fed twice a week with fresh food (JBL Cyclops spp., shrimp, Artemia spp., mosquito larvae) and the other five days with JBL De Novo Lake Tanganyika cichlid flake food (except for some missing days due to absence). This is the standard feeding regime for all cichlid fish at our laboratory. During 2010 we additionally fed all individuals on six to seven days per week with fresh small food items (Artemia freshly hatched eggs and JBL Cyclops spp., the latter replaced with Daphnia spp. when all offspring of all broods had grown beyond 10 mm standard length). This was done to ensure that all offspring received enough food to grow proficiently and to reach testing age. Cichlids were kept in tanks within climatised rooms (24–29°C, lights on between 08:00–21:00 h).

2.2. Experimental Setup. The experimental setup and tests performed in 2005 and 2006 were reported in [39] (repeated tests of the same 36 individuals up to 150 days apart by Roger Schürch), tests performed in 2007 were reported in [4] (repeated tests of the same 272 individuals up to 53 days apart by Susan Rothenberger). In 2008 (unpublished) repeated tests were performed of 32 individuals by Liana Lasut, Estée Bochud, and Sebastian Keller (using the same setup as [4]), and 10 individuals retested in 2010 by Noémie Chervet and Dirk Heg). These data were used to calculate the repeatability of behavioural types for time intervals of up to 1201 days after the first test was done. Individuals were marked between 2005 and 2008 with colour marks or fin clips and kept in breeding pairs, or fitted with PIT transponders and kept in aggregation tanks up to the moment of re-testing in 2010.

In September 2009 new breeding pairs were established and allowed to breed without monitoring their clutches (all “control pairs”), their offspring were tested without any information about the clutches (and thus offspring were up to 6 months of age). Between March and August 2010 these pairs were augmented with more new breeding pairs that were allowed to breed as follows.

Control pairs (n = 19 pairs) were kept in 89 or 93 litre tanks (length × breadth × height cm, height water level cm: 60 × 40 × 40 cm, 37 or 50 × 50 × 40 cm, 37 cm; resp.) and one clutch was left to hatch inside their tank (usually the first clutch), and these offspring remained there until personality testing (Figure 1(a), clutch treatment “with parents”, see below for more details). All other clutches were removed on the day of egg laying and put separately to hatch in a 24-litre tank (40 × 25 × 25 cm, 24 cm; clutch treatment “isolation”, see below for details). As pairs may produce a clutch about every other week [66], removal of the clutches ensured that we knew the identity of the offspring if they remained with their parents. However, some clutches remained undetected and were discovered after hatching and in such cases we removed all the offspring to an isolation tank as soon as possible. If the parents had very large offspring in their tank from a previous brood, we allowed them to hatch and keep a second clutch, as the offspring from the two broods could be easily distinguished according to their large size difference (this occurred only in 2 pairs). In 4 pairs a pair member died when already offspring were present and these offspring were removed and stored in a separate tank until personality testing, and the dead partner was replaced with a new partner.

Cross-breeding pairs (n = 38 pairs that produced at least one clutch, includes 17 pairs from [46]) were kept in 54 or 58 litre tanks (60 × 30 × 33 cm, 30 cm; or 60 × 30 × 35 cm, 32 cm). However, they included repeated measures of the same female with different males, and vice versa, so in total 19 different females and 20 different males were involved. The pairs were left together to breed for one and a half months. In between they were remeasured and reallocated to different mates. In total, we attempted to mate each individual with 5 different mates, but were only successful (i.e., the pair produced at least one clutch in 1.5 months) for a maximum of 4 different mates (males: 7 × 1, 8 × 2, 1 × 3 and 3 × 4 mates; males: 7 × 1, 9 × 2, 3 × 3 and 1 × 4 mates), partly because we lost and had to replace 4 individuals intermittently. All clutches were removed into 24-litre tanks (Figure 1(a), clutch treatment “isolation”), penultimate clutches were used for the clutch treatment “with parents” (in their 54 or 58 litre tanks) or “with foster parents” (see below), to increase the sample sizes for these latter two treatments.

Cross-fostering pairs (n = 14 pairs that produced at least one clutch, in one pair their single clutch did not hatch) were kept in 54 or 58 litre tanks (60 × 30 × 33 cm, 30 cm; or 60 × 30 × 35 cm, 32 cm). Their first clutch was removed into a different empty 54 or 58 litre tank (Figure 1(a), clutch treatment “cross-fostering”, see below for details), however, if this clutch did not hatch, the second clutch received the same treatment. All other clutches were removed into the treatment “isolation” (24 litre tank). Pairs kept producing clutches until they were transferred as foster parents to a foster clutch (see Clutch treatments below).

2.3. Clutch Treatments. All pairs were checked every day for new broods. Upon detection of a new clutch, we commenced with a 15 min brood care observation, counting the frequency of cleaning the eggs (each mouth movement counted) and fanning the eggs (aerating the eggs by vibrating with body and fins [67]) for the female and the male separately [47, 66, 68]. The minimum distance to the eggs in cm (with 0 indicating inside the pot(s) containing the eggs) was also noted for the female and male separately. Unfortunately, pairs could also lay clutches under/behind the filter or on the aquarium walls, and these clutches were sometimes not detected until the fry hatched (which occurs 2 to 3 days after egg laying). These cases account for some missing data on clutch size, average egg mass, hatching success (but not the number of hatched offspring, as fry were immediately counted), and brood care behaviour. After the brood care observation, the pot was removed and we counted the number of eggs (clutch size) and measured the average dry egg mass (by sampling up to 5 eggs per clutch and weighing them after 32 h drying in a 70°C oven). Data on brood care, clutch size, and egg mass will be treated in detail elsewhere.

The broods in the treatment “with parents” broods were then immediately placed back with their parents...
Figure 1: Treatment of the broods and experimental setup of the three behavioural tests (black fish show the starting position of the focal individual in each test). (a) Offspring remained with their parents (treatment “with parents”); or were isolated and raised only together with their siblings (treatment “isolation”); or were isolated, raised together with their siblings for 65 days, and from this day onwards received a foster pair (“with fosters”). Six offspring were removed for behavioural testing on days 120 and 150 each (or fewer offspring if less than 6 offspring were still alive on day 120, and fewer offspring if less than 6 not yet tested offspring were still alive on day 150). Offspring were measured and moved singly to a 40-litre tank depicted in (b, c). After two days acclimatization, each offspring was tested. Note that offspring were permanently removed to avoid confusion with previously tested offspring. (b) Setup of the aggression test, where aggressive displays/attacks were scored towards the mirror (either placed left or right), and hiding time inside the pot was measured. (c) Setup of the boldness test, where latency and shortest distance moved to the novel object was scored (object placed left or right). (d) Setup of the exploration test, the test started by removing the opaque partition, and latencies plus visits to 10 pots were measured. Note that the home compartment was either on the left or the right and pots were shifted accordingly, visits to the home compartment pot were not counted. Offspring and parents were similarly tested according to (b–d).

Figure 1(a)). The broods in the treatments “isolation” and “with foster-pair” broods were permanently removed (Figure 1(a), the pair received a new clean flower-pot halve) and placed into an isolation net inside a separate 24-litre tank (“isolation”) or 54/58 litre tank (“with foster-pair”), and the eggs were incubated using an air stone. Approximately five days after hatching they were released from their isolation net. “Isolation” broods were kept in their 24-litre tanks with their siblings until 55 to 114 days after egg laying, when they were transferred to a bigger tank (34 to 188 litre) to accommodate both their size and numbers (as numbers were highly variable at transfer, ranging between 2 and 67 siblings, we also used highly variable tank sizes). Two broods with a single offspring each and three broods with two offspring each remained in their original tanks, as transfer was not necessary due to their low number and limitations in the availability of bigger tanks. “With fosters” brood were kept together with their siblings in their 54/58-litre tank, until they received a foster pair from day 65 onwards (Figure 1(a)).

2.4. Body Measurements and Personality Testing. Before the personality test (on day −2 before testing) or mate exchange (on day 0 of release with the new mate), each individual was sexed (external papilla inspection under a dissecting microscope), measured (standard length SL and total length TL in 0.1 mm using a dissecting microscope) and weighed (body mass in mg), and fin clipped for a DNA sample and for male/female identification within pairs. All offspring produced in the period September 2009 to March 2010 were tested in 2010, their parents were tested in March 2010 (5 females had lost their mate before March 2010, so 5 pairs had missing data for the male parent; 1 male had lost his mate before March 2010, so 1 pair had missing data for the female parent). The reasoning for testing all offspring was that they had experienced a very prolonged time with their parents, so might be well suited to serve as a benchmark for future studies. Offspring produced in the period March to August 2010 were identified by their clutch identifier and clutch treatment and tested on day 120 since the clutch was laid
The personality testing procedure has been outlined in detail in [4, 39]. Briefly, boldness and aggressiveness tests were conducted for each single focal fish inside a 42-litre tank (Figures 1(b) and 1(c), 50 × 30 × 30 cm; 28 cm water level). In total, 34 of such tanks were available for testing. Focal fish were left for two days to acclimate and settle territory around a single flower pot halve (placed 30 cm from the front glass). In the aggressiveness test, a mirror was placed along one side (Figure 1(b)). Here the total time hiding (in seconds) and aggressive behaviours towards the mirror image were noted: restrained aggression (frequency of slow approach, fast approach, head down display, spread fins, s-bending) and overt aggression (frequency of contact with the mirror, again 5 min total test duration). In the boldness test, a novel object (Figure 1(c), plastic beetle, plastic funnel, plastic blue half moon, clay bird, or plastic white cross) was placed in a front corner (left or right), and the latency to approach this object (in seconds) plus the closest distance to the object (in cm) was recorded (5 min total test duration). The exploration test was conducted in a 400-litre tank, where the fish were left in a partitioned area with a flower pot half for ten minutes before testing (Figure 1(d), so-called “home compartment”). Then the partition wall was removed and the focal fish could start exploring the unknown part of the tank where ten other flower pot halves were placed (“exploration compartment”, Figure 1(d)). Here the time spent moving outside the pots in any compartment, the latency before leaving the home compartment (in seconds), and for the exploration compartment, the latency before entering the first pot, the number of pots approached, the number of pots entered, and the number of different pots entered (0–10) were recorded (again 5 min total test duration). The three tests (boldness, aggressiveness, and exploration propensity) were conducted in randomized order. The three tests were repeated one day later, or rarely on the same day or up to four days later (due to time constraints and tank constraints). Note that in 2005 and 2006 all three tests were conducted in the 400-litre tank and the exploration test lasted 10 min (instead of 5 min, observer Roger Schürch, see [39]). Moreover, due to severe time constraints, the exploration test could not be conducted for all focal animals in 2010, as it involved a lot of time lost in handling fish. See the description of statistical analyses for details on how we have dealt with these differences in procedures.

2.5. Statistical Analyses. We used Categorical Principal Components analyses CatPCA with two-knot spline transformations [69] to summarise the three different tests (boldness, aggressiveness, and exploration propensity) into a single measure of “behavioural type” (object scores, see also [4]). To account for the observer effects, each CatPCA was run separately for each observer (2005-2006: Roger Schürch n = 216 series of three tests, 2007: Susan Rothenberger n = 1042 series of three tests, 2008: Sebastian Keller/Estée Bochud/Liana Lasut n = 64 series of three tests, boldness/aggressiveness/exploration tests; 2010: Markus Zöttl n = 288 tests, including 264 tests where only aggressiveness was recorded; Noémie Chervet n = 1412 series of two tests and Dik Heg n = 1268 series of two tests: aggressiveness and boldness). Noémie Chervet and Dik Heg also conducted exploration tests, but these were excluded from the analyses to keep the data amongst the individuals consistent. This procedure has the advantage that first, all observers automatically scale to a mean “behavioural type” of zero; and second, the different test procedures are also scaled to a mean “behavioural type” of zero (i.e., in 2005-2006 all three tests were conducted inside a 400-litre tank and the exploration tests lasted 10 min versus in 2007–2010 the three tests were conducted according to Figures 1(b)–1(d) and always lasted 5 min). This procedure only assumes that all observers capture more or less the complete variation in behavioural types present in the population, which is a reasonable assumption considering the large number of tests each observer conducted, and that all original variables had very high correlations both before and after transformations in the CatPCA, for each observer separately.

Repeatability was estimated using the VARCOMP and RELIABILITY procedures in SPSS 17 [70], by extracting the variance components and the corresponding intraclass correlation coefficients (= repeatability) by using the Restricted Maximum Likelihood method (REML). The procedure was run once for the complete data set using VARCOMP (Restricted Maximum Likelihood Method REML) and once for each time difference between the first test and the next test(s) i, where individuals were tested up to 6 six different times using the RELIABILITY procedure (which was easier to use in the latter analyses for data management reasons). Time differences between test i and the first test was calculated in days and the repeatability was calculated for days = 0 (test i conducted on the same day as the first test), 1, 2, 3, 4, 15 (between 11 and 20 days), 25 (between 21 and 30 days), 35 (between 31 and 40 days), 45 (between 41 and 50 days), 55 (between 51 and 60 days), 90, 120, 150, 175 (between 151 and 200 days), 225 (between 201 and 250 days), and 930 days (between 732 and 1201 days). To estimate the change in repeatability over time, these 16 estimates of repeatability (from 0 to 732–1201 days) were regressed against ln (days + 1) weighted by their sample size (weighted linear regression). Similarly, change in the test scores (test i minus the first test) were analysed by regressing this difference against ln (days + 1).

Heritability was estimated using (1) the mid-parent versus mid-offspring weighted regression slope (weighted by the square root of the number of offspring tested [71]). (2) The intraclass correlation coefficients derived from the variance components extracted using the VARCOMP procedure in SPSS 17 for random effects of the pair identity, mother identity, and father identity, respectively, using the REML method. These estimates were verified by using the minimum norm unbiased estimator method, using both
the priors 0 or 1 (MINQUE(0) or MINQUE(1) method: see [70]), and since the MINQUE estimates were virtually identical to the REML estimates only the latter are given. Finally, fixed effects on the brood level were tested by General Linear Mixed Models, using as random effects pair identity and brood identity nested within pairs (and extracting the variance components accordingly). The following fixed effects were tested: female stock (1996, 2006, 2009), male stock (1996, 2006, 2009), treatment of the brood (with parents, with foster parents, isolated), volume of the tank in litres (both before and after transfer, if offspring were not transferred volumes were identical), temperature of the tank in degrees Celsius (both before and—if this applies after transfer), and body size of the focal offspring (SL mm). Note that the fixed effects were measured on the brood identity level, and therefore varied between broods within pairs, and that the offspring varied in body sizes (both within and between broods).

Due to replacement of dead mates, we ended with a total sample size of 74 pairs, 162 broods, and 1327 offspring tested for the heritability analyses (49 individual females and 50 individual males), in one pair the mother was not tested and in three pairs the father was not tested.

3. Results

3.1. Categorical Principal Component Extraction of Behavioural Types. In total 1779 individuals were tested two to six times for their behavioural traits (on average 2.41 tests per individual or 4290 tests in total). Categorical Principal Components analyses were run for each observer separately, and in each case a single factor was extracted with an Eigenvalue higher than 1 explaining a high proportion of the correlated behaviours in the one to three tests conducted (Table 1). The extracted factor scores were saved as the “behavioural type” of the individual in each test (for the repeatability analyses) or averaged per individual over all their tests as “behavioural type” (for the heritability analyses).

3.2. Repeatability of the Behavioural Types. VARCOMP analysis (REML) showed significant repeatabilities of the behavioural types (n = 4290 tests of 1779 individuals): the variance attributed to the individuals was 0.5495, the error variance was 0.4576, which gives a repeatability (intraclass correlation coefficient) of 0.5495/(0.5495 + 0.4576) = 0.5457 (+0.0149 SE, standard error of the estimate). However, by comparing the test results pairwise with the first test result over time (from 0 days between two tests up to 732–1201 days between two tests) it became clear that the repeatability R significantly declined over time (Figure 2(a); regression analysis weighted by sample size: R = 0.830 ± 0.002 – 0.093 ± 0.001 ln(days between tests + 1); F = 5864.7, P < .001, R² = 0.72, n = 16 pairwise R estimates, ± SE). Both the intercept and the slope of this regression line (depicted in Figure 2(a)) were significantly different from zero (t = 340.1 and −76.6, resp., both P < .001).

Although the repeatability changed over time, the actual behavioural type test scores changed very little over time (Figure 2(b)). On a short-term basis, individuals became bolder, more aggressive, and explorative compared to their first test score, but this difference to the first test score rapidly diminished over time and approached the “no difference” (marked by the red line in Figure 2(b): y = 0). These changes were modeled with a regression analysis (black line in Figure 2(b); change in score = 0.442 ± 0.027 – 0.065 ± 0.012 × ln(days between tests + 1); F = 29.1, P < .001,

![Figure 2: Repeatability of behavioural types significantly declined over time. (a) Pairwise repeatability from the test results (behavioural type test i versus behavioural type first test), from two test series conducted on the same day 0 (n = 44), 1 (n = 1435), 2 (n = 277), 3 (n = 18), 4 (n = 45), 11–20 (n = 7), 21–30 (n = 101), 31–40 (n = 101), 41–50 (n = 63), 51–60 (n = 48), 90 (n = 36), 120 (n = 36), 150 (n = 36), 151–200 days (n = 7), 201–250 days (n = 15) and 732–1201 days apart (n = 18), respectively. The two low sample sizes of 7 are indicated by open circles. (b) Pairwise changes in the test series results over time (behavioural type test i minus behavioural type first test, n = 2501). Note the maximum time difference between two tests of 1201 days. See the text for the regression line.]
Table 1: Categorical Principal Component results for the behavioural testing, for each observer separately (in brackets the year(s) when the tests were conducted). The variance accounted for is represented by Cronbach’s alpha, and the Eigenvalue is given (% explained variance in brackets). In each case a single factor score was extracted, and used to characterize the behavioural type of the focal individuals on a testing day (one test series), used for the repeatability analyses. Scores were then averaged per individual for the heritability analyses (scores from two to six test series averaged). Tests used for Categorical Principal Component extraction: B = boldness, A = aggression, E = exploration.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>n test series</th>
<th>Tests used per series</th>
<th>Cronbach’s alpha</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. Schürch (2005-2006)</td>
<td>216</td>
<td>B, A, E</td>
<td>0.91</td>
<td>4.95 (61.9%)</td>
</tr>
<tr>
<td>S. Rothenberger (2007)</td>
<td>1042</td>
<td>B, A, E</td>
<td>0.94</td>
<td>6.65 (60.5%)</td>
</tr>
<tr>
<td>E. Bochud/S. Keller/L. Lasut (2008)</td>
<td>64</td>
<td>B, A, E</td>
<td>0.94</td>
<td>6.93 (63.0%)</td>
</tr>
<tr>
<td>M. Zottl (2010)</td>
<td>288</td>
<td>A</td>
<td>0.85</td>
<td>2.30 (76.7%)</td>
</tr>
<tr>
<td>N. Chervet (2010)</td>
<td>1412</td>
<td>B, A</td>
<td>0.90</td>
<td>3.53 (70.4%)</td>
</tr>
<tr>
<td>D. Heg (2010)</td>
<td>1268</td>
<td>B, A</td>
<td>0.89</td>
<td>3.49 (69.7%)</td>
</tr>
</tbody>
</table>

\(^{a}\)Includes 72 individuals that were subjected to an additional test series following the procedure of Schürch and Heg [39] in the 400-litre tank, analysed separately with CatPCA: Cronbach’s alpha = 0.93, Eigenvalue = 6.48 (38.9%).

\(^{b}\)Includes 64 individuals for which also the boldness test was conducted, analysed separately with CatPCA: Cronbach’s alpha = 0.92, Eigenvalue = 3.75 (75.0%).

\(^{c}\)Includes 8 test series were only aggressiveness was scored and analysed separately with CatPCA Cronbach’s alpha = 0.86, Eigenvalue = 2.36 (78.6%).

Table 2: Descriptive statistics of the offspring tested.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment of the brood</td>
<td>1327</td>
<td>n with parents: 419, with foster parent: 91, isolation: 817</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tank volume litre(^a)</td>
<td>1327</td>
<td>47.87</td>
<td>28.63</td>
<td>24 (^d)</td>
<td>93</td>
</tr>
<tr>
<td>Tank volume litre(^b)</td>
<td>1327</td>
<td>88.45</td>
<td>48.13</td>
<td>24 (^d)</td>
<td>188</td>
</tr>
<tr>
<td>Tank temperature(^a)</td>
<td>1219</td>
<td>27.17</td>
<td>1.22</td>
<td>24.25</td>
<td>29.50</td>
</tr>
<tr>
<td>Tank temperature(^b)</td>
<td>1219</td>
<td>26.58</td>
<td>0.87</td>
<td>24.25</td>
<td>28.30</td>
</tr>
<tr>
<td>Offspring body size SL (mm)(^c)</td>
<td>1327</td>
<td>24.11</td>
<td>0.87857</td>
<td>7.0</td>
<td>77.0</td>
</tr>
<tr>
<td>Offspring behavioural type(^c)</td>
<td>1327</td>
<td>−0.01565</td>
<td>1.49879</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\)Volume and water temperature of the offspring raising tanks: \(^{b}\)before and \(^{b}\)after transfer.

\(^{c}\)Body size at personality testing.

\(^{d}\)All clutches were produced by pairs in 54-to 93-litre tanks, but eggs and offspring from the treatment “isolation” were incubated and raised from the day of egg laying onwards inside 24 litre tanks and later transferred to larger tanks (see information given on tank volumes after transfer).

\(^{c}\)Offspring behavioural type computed as the average of the average score per offspring—behavioural type from the Categorical Principal Component analysis. Note that each offspring was scored at least twice (at least two test series of boldness test, exploration test and aggression test).

\(R^2 = 0.012, n = 2501, \pm SE of the estimates, both the intercept and the slope of this regression line were significantly different from zero: t = 16.3 and −5.4, resp., both P < .001).\)

3.3. Heritability of Behavioural Types. The raw data of the offspring behavioural type scores are given in Figure 3 and descriptive data are provided in Table 2. First, we estimated heritability using weighted regression equations (weighing the regression analysis by offspring number; Figure 4, Table 2): that is, the mid-parent versus mid-offspring behavioural types (Figure 4(a)), mother versus mid-offspring behavioural types (Figure 4(b)), and father versus mid-offspring behavioural types (Figure 4(c)). Heritabilities were significant but low for the mid-parent and mother versus offspring regressions, and nonsignificant for the father-offspring regression (Table 3). Moreover, the heritability estimated from sibling comparisons was high and significant (last row in Table 3).

We should like to point out that the regression approach is inferior to the variance components method: first, because the regression approach assumes the behavioural type of the parents is measured without error; second, because the mixed nature of the data can be better accommodated by the variance components (using the REML method) and this method can be extended to estimate fixed effects (GLMM REML method, see below); third, because the above analyses suggest strong brood identity effects (last row of Table 3), which should be estimated as a random effect nested within pair identity effects (again using the GLMM REML method). We therefore recalculated heritability using the GLMM REML, once with the pair identity or parent’s identity (female or male) as random effects (first three rows in Table 4) and once by adding also a brood identity nested within pair identity as random effects (last three rows in Table 4). Heritability estimates were all significant and now slightly larger than the previous estimates (cf. Table 4 with Table 3).
Table 3: Heritabilities $h^2$ of offspring behavioural type using the weighted regression equation approach for parents versus offspring (weighing the mid-offspring behavioural type by the square root of the number of offspring tested) and the one-way ANOVA approach for siblings (with brood identifier as a random factor).

<table>
<thead>
<tr>
<th>Comparison</th>
<th>$n$ pairs or parent, $n$ offspring</th>
<th>MS</th>
<th>$df$, error $df$</th>
<th>$F$</th>
<th>Estimates</th>
<th>$h^2 \pm SE$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intercept ± SE</td>
<td>Slope ± SE</td>
</tr>
<tr>
<td>Midparent-offspring</td>
<td>70, 1272</td>
<td>0.95</td>
<td>1, 276</td>
<td>8.58***</td>
<td>−0.010 ± 0.024 ns</td>
<td>0.12 ± 0.04</td>
</tr>
<tr>
<td>Mother-offspring</td>
<td>49, 1318</td>
<td>2.10</td>
<td>1, 283</td>
<td>12.46***</td>
<td>0.008 ± 0.025 ns</td>
<td>0.19 ± 0.03</td>
</tr>
<tr>
<td>Father-offspring</td>
<td>50, 1281</td>
<td>0.07</td>
<td>1, 279</td>
<td>0.44 ns</td>
<td>−0.007 ± 0.025 ns</td>
<td>0.00 ± 0.03</td>
</tr>
<tr>
<td>Siblings</td>
<td>162*, 1327</td>
<td>1.90</td>
<td>161, 1165</td>
<td>3.08***</td>
<td>−0.004 ± 0.040 ns</td>
<td>0.41 ± 0.03</td>
</tr>
</tbody>
</table>

*Number of broods.

ns: nonsignificant. *P < .05, **P < .01, ***P < .001.

In total 74 pairs with 162 broods were tested (in one pair the mother was untested for behavioural type and in three pairs the father was untested for behavioural type). MS: mean square.

Heritability estimates are twice the slopes for mother-offspring and father-offspring regressions, and twice the intraclass correlation coefficient for siblings.

Table 4: Heritabilities $h^2$ of offspring behavioural type using the variance components approach (GLMM REML).

<table>
<thead>
<tr>
<th>Random effect(s)</th>
<th>$n$ pairs or parent, $n$ offspring</th>
<th>Variance pair or parent</th>
<th>Variance brood</th>
<th>Variance error</th>
<th>Asymptotic covariance matrix†</th>
<th>$h^2 \pm SE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair</td>
<td>74, 1327</td>
<td>0.117***</td>
<td>0.653***</td>
<td>0.000787</td>
<td>0.000677</td>
<td>−0.000046</td>
</tr>
<tr>
<td>Mother</td>
<td>49, 1318</td>
<td>0.121***</td>
<td>0.665***</td>
<td>0.001043</td>
<td>0.000697</td>
<td>−0.000035</td>
</tr>
<tr>
<td>Father</td>
<td>50, 1281</td>
<td>0.097***</td>
<td>0.664***</td>
<td>0.000724</td>
<td>0.000716</td>
<td>−0.000036</td>
</tr>
<tr>
<td>Pair + Brood within paira</td>
<td>74, 1327</td>
<td>0.091**</td>
<td>0.069**</td>
<td>0.614***</td>
<td>0.000935</td>
<td>0.000641</td>
</tr>
<tr>
<td>Mother + Brood within mothera</td>
<td>49, 1318</td>
<td>0.096**</td>
<td>0.074***</td>
<td>0.616***</td>
<td>0.001099</td>
<td>0.000649</td>
</tr>
<tr>
<td>Father + Brood within fathera</td>
<td>50, 1281</td>
<td>0.076*</td>
<td>0.085***</td>
<td>0.605***</td>
<td>0.000871</td>
<td>0.000647</td>
</tr>
</tbody>
</table>

†For brevity, covariances involving the random brood effects are not given ($n = 162$ broods for 74 pairs, $n = 160$ broods for 49 mothers, $n = 156$ broods for 50 fathers).

*P < .05, **P < .01, ***P < .001.

Heritability estimates are twice the intraclass correlations for mother and father effects, standard errors calculated according to [70].

Interestingly, brood identity random effects remained significant when nested within their parent(s) (last three rows in Table 4). This suggests siblings from the same brood shared a common (environmental) effect on their behavioural type. To explore potential shared effects, we added single fixed effects to a base model containing random effects of pair identity and brood identity nested within pair (Table 5). However, none of these effects significantly affected the behavioural types of the siblings: treatment of the brood, stocks of their parents, tank volumes and water temperatures during raising, and also offspring body size at testing were all clearly nonsignificant (Table 5).

4. Discussion

There are three main results. First, repeatability of behavioural type significantly declined over time, that is, for two tests conducted on the same day repeatability was 0.83, and for two tests conducted up to 1201 days apart repeatability was only 0.19. This time period spans the entire expected maximum life-span of this species, which has been estimated to be ca. 1000 days [63]. Second, heritability was low but significant and depended on the random effects fitted (i.e., pair, mother, father, or brood identity). Third, a significant random effect of brood identity within pair identity suggests a shared effect on the behavioural types of the broods, which did not depend on (i) the treatment of the broods, (ii) origins of female and male stocks, (iii) volumes and temperatures of tanks used to raise the offspring, and (iv) sizes of the offspring at testing. These three main findings are discussed in more detail below.

Temporal and systematic changes in behavioural type have been reported for various animal populations and for humans (e.g., [20, 33, 37, 40, 72–80]), including our
Table 5: Fixed effects on offspring behavioural type using the variance components approach (GLMM REML). In the base model only random effects of pair identity and brood identity nested within pair identity were added. Fixed effects were then tested stepwise for entry into this base model. See Table 2 for offspring sample sizes. Treatments of the brood were “with parents”, “with foster parents”, or in “isolation”. Stocks were from 1996, 2006 or 2009. Volumes, temperatures, and offspring size were continuous (covariate) effects.

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>df</th>
<th>Error df</th>
<th>$F$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment of the brood</td>
<td>2</td>
<td>100.2</td>
<td>0.54</td>
<td>.59</td>
</tr>
<tr>
<td>Mother stock</td>
<td>2</td>
<td>72.8</td>
<td>0.21</td>
<td>.81</td>
</tr>
<tr>
<td>Father stock</td>
<td>2</td>
<td>65.5</td>
<td>0.18</td>
<td>.83</td>
</tr>
<tr>
<td>Tank volume (litre)$^a$</td>
<td>1</td>
<td>123.7</td>
<td>0.25</td>
<td>.62</td>
</tr>
<tr>
<td>Tank volume (litre)$^b$</td>
<td>1</td>
<td>121.3</td>
<td>0.02</td>
<td>.90</td>
</tr>
<tr>
<td>Tank temperature (°C)$^a$</td>
<td>1</td>
<td>112.5</td>
<td>0.30</td>
<td>.58</td>
</tr>
<tr>
<td>Tank temperature (°C)$^b$</td>
<td>1</td>
<td>112.8</td>
<td>0.21</td>
<td>.65</td>
</tr>
<tr>
<td>Offspring body size (SL mm)$^c$</td>
<td>1</td>
<td>440.0</td>
<td>0.29</td>
<td>.59</td>
</tr>
</tbody>
</table>

Volume and water temperature of the offspring raising tanks: $^a$ before and $^b$ after transfer.
$^c$ Body size at personality testing.

In our study, repeatability diminished strongly over time (Figure 2(a)), but nevertheless the behavioural type scores between two tests were quite comparable (Figure 2(b)). If two tests were conducted only a short time apart (e.g., on the same day), individuals were typically bolder, more aggressive, and explorative during the second test. This strongly suggests a training effect on the test scores of the individuals: a habituation effect which diminishes over time (see [36] for a similar example). Accordingly, if the two tests were conducted widely apart in time, test scores of individuals were on average very similar to each other. We note as a point of criticism that the change in repeatability over time was modeled by us using a simple regression model. However, it is quite likely (as the data in Figure 2(a) suggest) that the repeatability actually stabilizes to a level of 0.4 to 0.5 after 150 days between two tests. A "break-point regression" approach would be a way forward to study this, but would also need more repeatability data from day 150 onwards. We urge scientists to study the repeatability of behavioural types in more depth, as it plays a critical role in the concept of animal personality research, and we concur that these studies are particularly lacking in fish [39, 74, 75, 81–85].

We found that the heritability of behavioural type was ca. 0.15 in N. pulcher, which is a rather low estimate compared to other studies (see meta-analysis in [16]: mean = 0.31). However, heritability estimates may critically depend on the variance components which were actually tested, that is, whether the study design allowed testing for (permanent) environmental effects, maternal and paternal effects, and maternal additive genetic effects. Studies in domesticated animals have shown that many genetic and nongenetic factors may contribute to behavioural phenotype of the offspring (e.g., [86–91]). Similarly, we found strong evidence for shared sibling environmental effects (brood with pair effects in Table 3) and maternal/paternal effects on behavioural type (discrepancies between pair versus mother versus father effects in Table 3). This suggests that an animal
model statistical analysis of the behavioural types in *N. pulcher* might be a worthwhile enterprise in the future, to disentangle these variance components. We found no effects of some important aspects of the offspring’s rearing environment on their behavioural type (like temperatures, tank sizes, and clutch treatments).

It is yet unknown how behavioural type (e.g., aggressive propensity) matches the behaviour shown under natural conditions (e.g., regarding dominance [92–95], territory acquisition [94, 96, 97], mate acquisition [59, 98–102], and mating performance [103]; for review see [85, 104]). It would be of particular interest to know how the behavioural type affects fitness and therefore is subject to natural selection. Field studies show that fitness effects of behavioural type may vary over time [105–107] and space [37, 94, 108–119], and behavioural type scores may not match one to one with actual behaviour shown in nature (e.g., due to context dependence, [120, 121]). Similarly, although exploration propensity is part of the behavioural syndrome in *N. pulcher* under laboratory conditions, it appears decoupled from the syndrome under natural conditions (i.e., actual distances moved in field settings: [43]). Under seminatural settings, shy fish have more socially positive interactions with their neighbourhood than bold fish, which is contrary

---

**Figure 4:** Heritability of offspring behavioural type using the regression approach. Symbol sizes represent the number of offspring tested per brood (pairs: 1 to 54, *n* = 70, excludes broods were one parent was not tested; mothers: 3 to 100, *n* = 49; fathers: 3 to 94, *n* = 50). Note that multiple broods tested from the same pair or parent, have the same x-axis value in each panel.
to expectation [4]. However, as expected, bold fish are the hotspots of an aggressiveness network [4]. The relevance of all these effects for fitness in \textit{N. pulcher} remain unclear, as (i) the effects of behavioural type are always small compared to other effects known to affect the behaviour and fitness of \textit{N. pulcher} (e.g., social status, body size, and sex); (ii) frequencies of aggression, affiliation, and submission do not scale one to one on the behavioural type of the focal individual (Rothenberger et al. manuscript in preparation); (iii) effects of behavioural type on fitness (survival and reproduction) have not yet been measured in the field.

In domesticated and laboratory animals behavioural types (e.g., aggressiveness) are directly subjected to artificial selection by the experimenters [60, 122–136], for instance in types (e.g., aggressiveness) are directly subjected to artificial contests), and (ii) obtain estimates about how different behavioural types relate directly (or indirectly) to fitness in the natural situation. Finally, our results suggest other (non)genetic effects affecting the behavioural type in \textit{N. pulcher}, which should be analysed in more detail in the future.

Acknowledgments

The authors thank the editors for inviting us to contribute to this special issue. They thank Susan Rothenberger for behavioural tests performed in 2007, Liana Lasut, Esté Bochud and Sebastian Keller for behavioural tests performed in 2008, and Kristina Sefc and an anonymous reviewer for their comments. This study was supported by SNSF grants 301000A0-122511 to M. Taborsky and 3100A0-108473 to D. Heg. They declare that the results presented in this paper have not been published previously.

References


divergence in novelty-seeking traits: heritability and enrichment in


[130] L. F. Sundström, E. Petersson, J. Höjesjö, J. I. Johnsson,
and T. Järvi, “Hatchery selection promotes boldness in newly hatched brown trout (*Salmo trutta*): implications for
dominance,” *Behavioral Ecology*, vol. 15, no. 2, pp. 192–198,
2004.

[131] J. Nyberg, K. Sandnabba, L. Schalkwyk, and F. Sluyter,
“Genetic and environmental (inter)actions in male mouse
lines selected for aggressive and nonaggressive behavior,”

[132] G. Su, J. B. Kjaer, and P. Sørensen, “Variance components
and selection response for feather-pecking behavior in laying

[133] J. Schjolden, T. Backström, K. G. T. Pulman, T. G. Pottinger,
and S. Winberg, “Divergence in behavioural responses
to stress in two strains of rainbow trout (*Oncorhynchus
mykiss*) with contrasting stress responsiveness,” *Hormones

[134] M. Harri, J. Mononen, L. Ahola, I. Plyusnina, and T. Rekilä,
“Behavioural and physiological differences between silver
foxes selected and not selected for domestic behaviour,”

heritability of personalities in the great tit (*Parus major*),”
*Proceedings of the Royal Society B*, vol. 270, no. 1510, pp. 45–

and B. G. Bohus, “Differences in basal and stress-induced
HPA regulation of wild house mice selected for high and low
aggression,” *Hormones and Behavior*, vol. 43, no. 1, pp. 197–
Submit your manuscripts at
http://www.hindawi.com