

**Within-individual canalization contributes to age-related increases in trait repeatability: a
longitudinal experiment in red knots**

Eva M. A. Kok¹, Joseph B. Burant^{1,2}, Anne Dekinga¹, Petra Manche^{1,3}, Darren Saintonge¹,
Theunis Piersma^{1,3}, Kimberley J. Mathot^{1,4}

1. NIOZ Royal Netherlands Institute for Sea Research, Department of Coastal Systems and Utrecht
University, P.O. Box 59, 1790 AB Den Burg, Texel, The Netherlands

2. Current address: Department of Integrative Biology, University of Guelph, Guelph, Ontario, N1G 2W1,
Canada

3. Rudi Drent Chair in Global Flyway Ecology, Conservation Ecology Group, Groningen Institute for
Evolutionary Life Sciences (GELIFES), University of Groningen, PO Box 11103, 9700 CC Groningen, The
Netherlands

4. Current address: Canada Research Chair in Integrative Ecology, Department of Biological Sciences,
University of Alberta, Edmonton, Alberta, T6G 2E9, Canada

*Corresponding author: E-mail: eva.kok@nioz.nl

MS type: article

Keywords: consistent among-individual differences, variance partitioning, within-individual
variation, state-behavior feedbacks

List of elements: Title page, English Abstract, Frisian Title and Abstract, Introduction, Materials
and Methods, Results, Discussion, Acknowledgments and Authorship Statement, Table A1,
References, Table 1-4, Figure Legends.

ABSTRACT

Age-related increases in the repeatable expression of labile phenotypic traits are often assumed to arise from an increase in among-individual variance due to differences in developmental plasticity or by means of state-behavior feedbacks. However, age-related increases in repeatability could also arise from a decrease in within-individual variance as a result of stabilizing trait expression, i.e. canalization. Here we describe age-related changes in within- and among-individual variance components in two correlated traits, gizzard mass and exploration behavior, in a medium-sized shorebird, the red knot (*Calidris canutus*). Increased repeatability of gizzard mass came about due to an increase in among-individual variance, unrelated to differences in developmental plasticity, together with decreases in within-individual variance, consistent with canalization. We also found canalization of exploration, but no age-related increase in overall repeatability, which suggests that showing predictable expression of exploration behavior may be advantageous from a very young age onward. Contrasts between juveniles and adults in the first year after their capture provide support for the idea that environmental conditions play a key role in generating among-individual variation in both gizzard mass and exploration behavior. Our study shows that stabilization of traits occurs under constant conditions: with increased exposure to predictable cues, individuals may become more certain in their assessment of the environment allowing traits to become canalized.

De foarspelbere ûntwikkeling fan lichems- en gedrachseigenskippen: in eksperiment oer welhelberhyd fan eigenskippen by mientsen

GEARFETTING

Alhoewol guon lichemseigenskippen fan jonge yndividuën hyltiten wer feroarje kinne, komme by it âlder wurden sokke eigenskippen ornaris dochs hyltiten mear fêst te lizzen. Soks kin komme troch it feroarjen fan yndividuele plastisiteit en troch weromkeppelingen tusken it gedrach en de steat fan sa'n bist. Yndividuele ferskillen yn it fêstlizzen fan eigenskippen fergrutsje de fariaasje tusken yndividuën. In technysk begrip om yndividuele fariaasje fan eigenskippen te kwantifisearjen is 'repeatability' ('werhelberhyd'), mar it euvel is dat in taname fan dizze statistyske maat sawol komme kin troch in taname yn 'e fariaasje *tusken* yndividuën en troch in ôfname fan de fariaasje *binnen* yndividuën; dit lêste neame wy 'kanalisaasje'. Yn dit artikel beskriuwe wy hoe't dizze twa boarnen fan fariaasje by it âlder wurden feroarje kinne by mientsen (*Calidris canutus*), en dat dogge wy oan 'e hân fan twa besibbe eigenskippen: (1) it gewicht fan de spiermage, en (2) de wize werop mientsen yn in eksperimentele romte lytse stikjes waad ferkenne (der binne fûgels dy't bot eksploraasje, en guon dy't ôfwachtsje). It die bliken dat de werhelberhyd fan it magegewicht feroare troch tanimmende ferskillen tusken yndividuën en in ôfname fan de fariaasje binnen yndividuën, in kombinaasje fan plastisiteit en kanalisaasje dus. By eksploraasje-gedrach fûnen wy by it âlder wurden oanwizings foar kanalisaasje. Dat soe betsjutte kinne dat der foardielen binne om al op jonge leeftyd in bepaalde yndividuele wize fan eksploraasje oan te hâlden. Út in fergeliking tusken jonge en âlde fûgels (dy't in ferskil yn ûntwikkeling yn it frije fjild wjerspegelje), blykte it bestean fan weromkeppelingen tusken de steat fan it lichem (magegewicht) en it gedrach (eksploraasje). Ús stúdzje lit lykwols foaral sjen dat eigenskippen, sels yn sitewaasjes dy't net feroarje, fêst komme te lizzen. Miskien makket it fenomeen dat fûgels har omjouwing hyltiten better foarspelle kinne sokke kanalisaasje mooglik.

Introduction

Individuals often differ consistently from one another in suites of behavioral, physiological, and morphological traits (Sih et al. 2004; Bell et al. 2009; Réale et al. 2010*a*; Dall et al. 2012; Carere and Maestriperi 2013). Although inter-individual variability was traditionally viewed as merely the substrate for natural selection, evidence is accumulating that among-individual variation has greater ecological and evolutionary implications (Bolnick et al. 2003; Sih et al. 2012; Wolf and Weissing 2012). It is increasingly recognized that such intraspecific variation may be adaptive (Wilson and Yoshimura 1994; Bolnick et al. 2003; Dall et al. 2004; Sih et al. 2004; Réale et al. 2007).

Individuals of the same population may differ in dispersal behavior (Cote et al. 2010*a*, 2010*b*), foraging behavior, resource use (Svanbäck and Bolnick 2007; Toscano et al. 2016; Sheppard et al. 2018), and aggression (Bell et al. 2009). In some cases, these among-individual differences in behavior are associated with differences in reproduction and survival (Wilson 1998; Smith and Blumstein 2008; Réale et al. 2010*b*). Variation between individuals can lead to non-random distributions of individuals, an increase in the range of resources that can be exploited, and higher resilience to change for populations as a whole (Bolnick et al. 2003; Wolf and Weissing 2010; Sih et al. 2012). Therefore, not surprisingly, a significant amount of work has aimed to address the causes of among-individual variation (Wolf et al. 2007; Dingemanse and Wolf 2010).

The ontogeny of repeatable among-individual variation, a potentially core process underlying individual development, has received markedly less attention (but see Sinn et al. 2008; Bell et al. 2009; Biro and Stamps 2015; Polverino et al. 2016). Although among individual

variance can in theory either increase or decrease with age, the majority of empirical studies have reported age-related increases in repeatable among-individual variation (Table 1). However, these studies often fail to evaluate whether these changes are driven by changes in the amount of among-individual variation, within-individual variation, or both (Table 1).

Theoretical considerations of the development of trait repeatability have focused on processes that affect the amount of variation among individuals (Fisher et al. 2018; Sih et al. 2015; Stamps and Frankenhuis 2016). One obvious explanation for age-related increases in among-individual trait expression is that it reflects among-individual differences in the conditions experienced during development (West-Eberhard 2003). However, among-individual variation can also arise when individuals are reared under near-identical conditions (Crabbe et al. 1999; Brust et al. 2015; Bierbach et al. 2017). This could be due to (epi-)genetic variation among individuals (Dall et al. 2012) or to individual differences in developmental plasticity (i.e., the effect of environment on phenotypic differences) (West-Eberhard 1989, 2003; Stamps and Krishnan 2014a; Stamps and Frankenhuis 2016).

Among-individual variance can also increase over time through positive-feedbacks between two traits (Sih et al. 2015). For example, foraging boldness (i.e., willingness to forage in the presence of predators), may allow individuals to acquire more resources and grow relatively more quickly compared to individuals that do not forage in the presence of predators (Luttbegg and Sih 2010). If, at the same time, being larger confers some safety advantage (e.g., because predators are gape limited and large prey are less accessible), then being larger will also favor higher boldness (Luttbegg and Sih 2010). The reciprocal effects of boldness on body

size, and body size on boldness, mean that initially very small differences between individuals can increase over time (Sih et al. 2015).

Though less often considered, increased repeatability during development may also result from decreasing within-individual variance, or canalization (Waddington 1942). A trait is considered canalized if phenotypic expression remains invariable under mildly differing developmental conditions. The term canalization was originally used to refer to the evolution of developmental stability (Waddington 1942). According to Waddington's (1942) metaphor of canalization, the individual development of traits can be visualized as the movement of balls rolling down alternative valleys in a landscape that has been shaped by natural selection. Environmental effects can either be implemented either as modifications to the width and depth of a single valley (Boonekamp et al. 2018) or as switches between alternative valleys (Waddington 1942). More recently the term canalization has been applied to refer to the reduction in residual phenotypic variance at the within-individual level (Westneat et al. 2015). A reduction in within-individual variance (i.e., residual variance) can occur if phenotypic variation decreases in the course of development (e.g., Stamps and Krishnan 2014*b*, 2017; Westneat et al. 2015; Stamps and Frankenhuis 2016). Following Waddington's metaphor, this is analogous to the valleys in the phenotypic landscape to deepen and/or narrow over time, producing more rigid and predictable trait expression across ontogeny (Boonekamp et al. 2018).

Thus, age-related increases in trait repeatability under identical conditions can be the outcome of at least three distinct developmental processes. Two of these affect the degree of among-individual variance (i.e. differences in developmental plasticity and state-behavior feedbacks) and one affects the degree of within-individual variance (within-individual

canalization) (See Table 2 and Figure A1 for a matrix of predictions). Recognizing that many organisms are only sensitive to environmental cues during particular stages of ontogeny (Bateson 1979; Fawcett and Frankenhuis 2015; Panchanathan and Frankenhuis 2016), systematic investigation of the development of trait expression requires longitudinal studies of development.

We studied the development of two ecologically important phenotypic traits in a migratory shorebird: gizzard mass (Piersma et al. 2003; van Gils et al. 2003, 2005) and exploration behavior (Bijleveld et al. 2014, 2016; Oudman et al. 2016). In a longitudinal study spanning two consecutive years, we measured individual changes in gizzard mass and exploration behavior in red knots (*Calidris canutus*; hereafter called 'knot'). During the non-breeding season, knots forage on a diet of hard-shelled prey, primarily mollusks, that they crush in their muscular gizzards (Zwarts and Blomert 1992; Piersma et al. 1993; Battley and Piersma 2005), and gizzard mass is therefore a key trait (van Gils et al. 2005). Further, exploration behavior scored in standardized behavioral assays (see Methods) has been shown to be correlated with large-scale (100s of km) patterns of space use in the wild (Bijleveld et al. 2014). Notably, exploration and gizzard mass in free-living knots co-vary at the among individual level; individuals with large gizzards at the time of capture have lower exploration scores than individuals with small gizzards (Bijleveld et al. 2014, 2016). Previous laboratory experiments have shown that both gizzard mass and exploration behavior exhibit repeatable among-individual variation in knots (≥ 2 cy) (Bijleveld et al. 2014, Mathot et al. 2017). However, experimental manipulations of gizzard mass produced no changes in exploration behavior (Bijleveld et al. 2014). These series of observations led to the speculation that the among-

individual variation in gizzard mass and exploration behavior could be the result of state-behavior feedbacks between searching behavior and digestive quality of food during a limited window in early ontogeny (Bijleveld et al. 2014). Here we describe the development of age-related difference in trait repeatability in the light of three non-exclusive developmental processes to elucidate the developmental origin of among-individual variation in trait expression.

Material and methods

Study species and housing conditions

The knots (*islandica* subspecies; Piersma 2007) used in this study were captured with mist nets at two different high tide roosts in the Dutch Wadden Sea; Schiermonnikoog (53.29°N, 6.15°E) (n=53) and Griend (53.15°N, 5.16°E) (n=31) between 20 August and 20 October 2015. Birds were aged based on plumage characteristics as either juvenile birds (< 6 months), second calendar year birds (between 6 and 18 months), or older (i.e., adult birds, > 18 months) (Prater et al. 1977). Only juveniles and adults were selected for the study (N = 44 juveniles, N = 46 adults). We collected a small blood sample (< 75 µl) for molecular sexing (van der Velde et al. 2017). For simplicity, we refer to the birds caught as first year birds as “juveniles” throughout the paper, despite the fact that they changed from being juvenile to second calendar year to adults in the course of this two-year-long study.

Birds were housed in outdoor aviaries (4.0 m deep, 1.9 m wide, and 2.3 m high at one end, sloping down to a height of 1.9 across the depth of the aviary) at the Experimental Shorebird Facility of the NIOZ Royal Netherlands Institute for Sea Research on the island of Texel, The

Netherlands (53°00'N, 04°47'E). The aviaries had smoothly coated concrete floors that were constantly irrigated with running seawater. The back of each aviary had a basin with sand collected from the Wadden Sea and running seawater. Outside of experiments, birds had *ad libitum* access to Trout food pellets (Produits Trouw, Vervins, France) and a continuous source of fresh water for drinking and bathing in a separate tray. Every week, while the aviary floors were cleaned and disinfected with chlorine, the birds were weighed, their molt and plumage status scored and their bodies, especially their feet, checked for small wounds and *Staphylococcus* infection (Milot et al. 2014).

The focal birds (*islandica* subspecies) were kept together with knots of the *canutus* subspecies in mixed flocks (14-17 knots per aviary, randomized with stratification based on age and subspecies). Flock composition was largely constant throughout the first year, but before the start of the second year of experiments new birds of the *canutus* subspecies were caught (N = 22). Thus, to maintain constant flock sizes across the two study years, 24 *islandica* knots were released between year-1 and year-2 of the experiment.

Diet manipulations

To prevent circannual endogenous rhythms from unduly affecting our measurements (Battley and Piersma 2005), experiments were only carried out over two non-breeding periods, from late October 2015 to early April 2016 in year-1, and from early October 2016 to mid-March 2017 in year-2. During the experimental period, birds were fed *ad libitum* diets of either high or low digestive quality. The high digestive quality food (HQ) consisted of Trout food pellets and the low digestive quality (LQ) food was thawed mud snails, *Peringia ulvae*. Previous work has

shown that these two food types induce approximately two-fold variation in gizzard mass (Vézina et al. 2011; Mathot et al. 2017). We used a crossover design: birds in half of the aviaries (n=4) were fed HQ food first, while the other half received LQ food first. Previous studies showed that knots can fully adjust their gizzard mass to a new food type within approximately one week (Dekinga et al. 2001), but we allowed three weeks of acclimatization to the new diet to ensure that the general condition of the birds would be stable and equal between diets (Bijleveld et al. 2014; Mathot et al. 2017).

These three weeks of diet manipulation were followed by two weeks of behavioral observation, during which time the knots remained on the same *ad libitum* diet. When all behavioral observations were completed, a new replicate of diet manipulations commenced; aviaries previously assigned the HQ food treatment became LQ aviaries and *vice versa*. Four diet treatments were carried out per bird during each of the two experimental years. To prevent for systematic differences between the knots as a result of the order of testing, we randomized the sequence with which we tested individuals in each behavioral test. On average 43 days (ranging between 21-65 days) elapsed between successive behavioral tests in year-1 and 40 days (ranging between 24-57 days) between successive tests in year-2.

Gizzard mass measurements

After each diet treatment, and before behavioral observations, the gizzard mass of all birds was measured using ultrasonography (Dekinga et al. 2001; Dietz et al. 1999). To standardize the measurements, birds were deprived of food for at least one hour prior to measurement to ensure an empty gizzard. Subsequently birds were selected in a haphazard order for

measurement. The observer was blind to the age and diet treatment of the birds. Gizzard measurements were done following a standardized procedure developed by AD, see Mathot et al. (2017) for a detailed description of the method.

Exploration behavior

The exploration behavior of individual birds was quantified in an arena that was novel for the birds during first exposure. Studies on exploration traditionally focus on individual movements after introduction to a novel environment (Réale et al. 2007; Verbeek et al. 1994). Studies that assess repeatability of exploration typically re-use the same test arena for subsequent tests (e.g., Bijleveld et al. 2014; Dingemanse et al. 2002, 2012; Dubuc-Messier et al. 2017; McCowan et al. 2015; Minderman et al. 2010; Wuerz and Krüger 2015). Therefore, a decline in novelty with repeated exposure is an inherent feature of studies estimating repeatability in exploration. The exploration arena used here was identical to the one used in Bijleveld et al. (2014). It measured 7 x 7 m and was filled with a layer of 30 cm seawater and five 1m x 1m trays filled with wet sand (Figure A2).

Birds were caught from their holding aviaries two hours prior to their randomly assigned observation time and kept individually in holding crates in a semi-dark and quiet room. Birds were food-deprived during these two hours to standardize hunger levels between birds. Immediately prior to the test, each bird was moved to a small aviary adjacent to the arena. After five minutes of acclimatization, the door between the aviary and the arena was opened by means of a remote pulley system and the bird was gently herded into the arena. Exploration trials lasted 30 minutes, during which time the behavior of the bird was scored live through

one-way-glass using the behavioral observation software, JWatcher

(<http://www.jwatcher.ucla.edu/>) and recorded (GoPro HERO+ LCD) for future reference. The observations were done by five different observers (three observers in year-1, and four observers in year-2, one common observer across both years) who were randomly distributed over the trials and blind to the treatment and age of the birds being tested. Behaviors recorded were: flying, walking on patches, searching for food, preening, resting, vigilance, out of sight, or 'other'. After the trial ended, the bird was returned to its holding aviary.

Using the protocol developed by Bijleveld et al. (2014) for the same subspecies of red knot, exploration behavior was scored as the fraction of time spent searching or walking on patches. Proportions were logit transformed to meet normality assumptions (Warton and Hui, 2011). This measure of exploration may not be assaying information gathering by individuals in a broad sense, but more specifically information gathering related to the distribution of food.

All experiments described in this paper complied with Dutch law and regulation and were carried out under protocol number AVD802002016740.

Data selection and statistical analyses

We obtained 569 complete observations of exploration behavior and gizzard mass ($N_{\text{adult}} = 283$, $N_{\text{juvenile}} = 286$). We excluded replicates during which the exploration experiments were disturbed ($n = 2$), when birds swam in the arena for more than 10 consecutive min ($n = 11$) because of welfare concerns for the bird (i.e., the risk of drowning or inability to thermoregulate with wet feathers), or when the diet manipulation was unsuccessful ($n = 14$). Unsuccessful diet manipulations occurred when knots failed to switch to the experimentally

determined diet (as evidenced by significant body mass loss), or when knots had to be removed from the experiment to be treated for *Staphylococcus* infection.

We constructed univariate models for gizzard mass and exploration behavior to study the development of age-related difference in trait repeatability. To be able to compare both age groups between year-1 and year-2, as well as to contrast adults and juveniles, we constructed separate models for each age group in each year (i.e., 4 models per trait; juveniles year-1, juveniles year-2, adults year-1, adults year-2). Because we were explicitly interested in age- and year-specific estimates for both among- and within-individual variance components, we included a random intercept for individual ID.

Although contrasts between the among- and within- individual variance components for each age-cohort and year combination could have been carried out in a single analysis by modelling heterogeneous residual errors, such analyses have very low statistical power (Cleasby and Nakagawa 2011). Therefore, we split the data in 4 bins and estimated the variance components for each trait per age group and year. To be able to correctly calculate within-individual variance that was unrelated to diet we used two measures of gizzard mass and exploration on each diet per year following Araya-Ajoy et al. (2015). We obtained complete data records (i.e., two measures of gizzard mass and exploration on each diets in each year ($N_{\text{replicates_per_bird}} = 8$) for a total of 58 birds (30 “adults” and 28 “juveniles” ($N_{\text{adult_measurements}} = 240$, $N_{\text{juvenile_measurement}} = 224$). This data can be found in the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.dn28cn6> (Kok et al. 2019). We restricted our analyses to these birds, as any changes in variance components from year-1 to year-2 necessarily reflected within-individual, age- (or time in captivity-) related changes in variance components, as

opposed to changes resulting from comparing different cohorts of birds. However, our data selection criteria did not affect the estimates of either the fixed effects or the variance components (results not shown).

We modeled gizzard mass and logit exploration as a function of sex (2-level factor: M or F), diet (2-level factor: LQ or HQ), replicate (continuous factor: range 1-4) and the interaction between diet and replicate. The addition of replicate in the model allowed us to test for changes in the response variables over time. The interaction term between diet and replicate allowed for a comparison of diet-related differences in the effect of replicate. In the results section, we focus on the effects of diet, replicate (time), and their interaction, on gizzard mass and exploration. We did, however, also include a fixed effect for sex to control for potential differences due to structural size differences between the sexes, since female knots are larger than males (Tomkovich 1992), but we will not discuss this any further in the results. For the models of exploration behavior, we fitted an additional random intercept for observer ID, to control for potential among-observer differences in behavioral scoring that would otherwise introduce additional residual variance. Because the observer was blind to the age group and diet treatment of each experimental bird, and because birds were randomly assigned to each observer, observer effects are not biologically meaningful, and are not relevant for the hypotheses being tested. They are presented in Table 3 for completeness, but are not discussed further. Models were built using the *lmer* function from the *lme4* package (Bates et al. 2015) in the R (v3.4.3) statistical environment (R Core Development Team 2017).

We report adjusted repeatabilities (i.e. after correcting for the fixed effects in the model) that were calculated following Nakagawa and Schielzeth (2010, 2013). To study the age-

dependent changes in repeatabilities, we first compare the changes in repeatability of gizzard mass and exploration for juveniles and adults between year-1 and year-2 in a longitudinal analyses. Subsequently we report a cross-sectional comparisons (e.g., comparing juveniles in year-1 with adults in year-1 as well as juveniles in year-2 with adults in year-2) to separate age-dependent effects from effects resulting from free-ranging experience or time in captivity. In all cases, we report how both within- (i.e., residual) and among-individual variance components contributed to the overall repeatability (Cleasby and Nakagawa 2011).

In cases where we found a value of zero for the among-individual variance, we verified that this was not a false negative result (e.g., singularity due to model overfitting), by rerunning the model with bird ID fitted as a fixed effect rather than a random effect and evaluating its effect size. We found no evidence that any of our estimates of zero among-individual variance were due to model overfitting (results not shown).

We used the *sim* function of the *arm* package (Gelman and Su 2016), to simulate values of the posterior distribution of the model parameters (Gelman and Hill 2007). 95% credible intervals (CIs) around the mean (β) were extracted based on 1000 simulations using the *MCMCglmm* package (Hadfield 2010). The 95% CI indicates a *margin of error* in terms of a range of plausible values for β . With this CI we indicate that we are 95% confident that our CI includes the actual effect size (Cumming and Finch 2005). To evaluate the differences between means and CIs of the fixed effects, as well as the variance components and the repeatability estimates, we followed Cumming and Finch (2005). Independent 95% CIs were deemed to indicate significant differences between averages when they did not overlap. This corresponds to a traditional p value < 0.006. In cases that the 95% CIs did overlap, we used the proportion of

overlap of the 95% CIs (the overlap between the two independent CIs divided by the average CI) to calculate the exact probability (p) that two CIs overlap (Cumming 2009) and discuss the level of support for a given effect as a continuous, rather than binary, characteristic.

Results

Changes in mean trait values

In each age group in both years, knots developed heavier gizzards on the LQ diet treatment than on the HQ diet treatment (95% CIs for fixed effect diet do not overlap with 0, except for a 0.15 proportion of overlap in juveniles in year-2, Table 4). Within each year, diet effects were similar between juveniles and adults, indicating similar gizzard mass plasticity across age cohorts. With time, the relatively light gizzard mass as a result of the HQ diet increased for both juveniles and adults (in year-2 the 95% CIs for replicate do not overlap with 0, Table 4). The negative interaction effect between diet and replicate found for adults in year-1 showed a decrease in gizzard masses on the LQ diet treatment across replicates (Table 4). These contrasting effects of replicate on HQ versus LQ diet meant that the overall diet effect decreased across replicates (Figure 1).

Exploration did not differ significantly between years or across age groups (all 95% CIs overlap), but exploration behavior decreased across replicates in adults in year-1 (95% CI for replicate does not overlap with 0, Table 3). Contrary to expectation, we found no effect of diet on exploration in any of the age groups (all 95% CI overlap with 0, Table 3, Figure 2).

Age-dependent changes in trait repeatability

Juvenile knots showed no repeatability in gizzard mass in year-1 (r and 95% CI centered on 0) and a small but significant repeatability in year-2 (95% CI does not overlap with 0, Table 4, Figure 3). In contrast, juvenile knots already showed significant repeatability in exploration behavior in year-1 (95% CI does not overlap with 0) and there was no support for change in the repeatability between year-1 and year-2 ($p = 0.29$, Table 3, Figure 3).

In adults we observed significant repeatability in gizzard mass in both year-1 and year-2 (95% CIs do not overlap with 0). However, the repeatability in year-2 was significantly lower (95% CIs do not overlap, Table 4, Figure 3). For exploration behavior in adults we found significant (95% CIs do not overlap with 0), and near identical repeatabilities in year-1 and year-2 (proportion overlap between years: $p = 0.89$, Table 3, Figure 3).

When comparing gizzard mass repeatabilities between age groups, repeatability was lower in juveniles than in adults in year-1 (95% CIs do not overlap). In year-2, both repeatability estimates were very small but juvenile repeatability was higher when compared to adults ($p < 0.01$, Table 4, Figure 3). For exploration behavior, we found no differences in repeatability between adults and juveniles in year-1 ($p = 0.54$) nor in year-2 ($p = 0.90$, Table 3, Figure 3).

Changes in among- and within-individual variance components

In juveniles, the increase in gizzard mass repeatability between year-1 and year-2 was the result of an increase in among-individual variance (95% CIs do not overlap) together with a decrease in within-individual variance of 32% between year-1 and year-2 ($p = 0.03$, Table 4, Figure 3). For exploration behavior we found no significant change in repeatability between year-1 and year-

2. However, in the absence of a change in among-individual variance in juveniles ($p = 0.90$, Table 3, Figure 3), a significant decrease in within-individual variance ($p < 0.01$) led to a (non-significant) increase in repeatability.

In adults, the significant decrease of gizzard mass repeatability between year-1 and year-2 was the result of a decrease in among-individual variation (95% CIs do not overlap), while the within-individual variance did not change ($p = 0.90$, Table 4, Figure 3). We found no change in repeatability in exploration behavior for adults from year-1 to year-2. However, when considering the changes in among- and within-individual variance components separately, we found a decrease in the within-individual variance ($p = 0.01$) together with a non-significant concomitant decrease in among-individual variance ($p = 0.15$, Table 3, Figure 3)

The age-related difference in gizzard mass repeatability between juveniles and adults in year-1 was due to lower among-individual variance and higher within-individual variance in juveniles compared with adults (among-individual variance; 95% CIs do not overlap, within-individual variance; $p = 0.01$, Table 4, Figure 3). In year-2, the small but significantly higher repeatability in gizzard mass of juveniles compared to adults was the result of higher among-individual variance in juveniles than in adults ($p = 0.01$). The within-individual variance did not differ between juveniles and adults ($p = 0.85$, Table 4, Figure 3).

As for repeatability, the within- ($p = 0.73$) and among-individual variance ($p = 0.11$) in exploration behavior did not differ between juveniles and adults in year-1 (Table 3, Figure 3). However, the limited overlap in 95% CIs in among-individual variance in adults and juveniles suggests that adults showed higher among-individual variance than juveniles in year-1 (Table 3,

Figure 3). In year-2, there was no difference between adults and juveniles in either the within- ($p = 0.63$) or among-individual ($p = 0.92$) variance in exploration behavior (Table 3, Figure 3).

Discussion

In this study, we evaluated support for three non-exclusive developmental processes that may underlie age-related changes in repeatable trait expression in knots (individual difference in phenotypic plasticity, state-behavior feedbacks, and within-individual canalization, Table 2). From year-1 to year-2 gizzard mass repeatability increased in juveniles and decreased in adults. This increase in gizzard mass repeatability in juveniles was the result of an increase in among-individual variance and a decrease in within-individual variance. In adults the decrease in repeatability was due to a decrease in among-individual variance alone. Initially (in year-1), juveniles showed lower among-individual variance and higher within-individual variance in gizzard mass than adults. In year-2 within-individual variance in juveniles declined to levels similar to older individuals (i.e. adults in year-1 and year-2). We found no linear age-related changes in among-individual variance in gizzard mass. Although exploration repeatability did not differ between age groups and years, we found a significant decrease in within-individual variance for both juveniles and adults between year-1 and year-2.

The observed age-related differences in gizzard mass repeatability resulted from changes in both the among- and within-individual variance components (Table 4). We can exclude the possibility that feedbacks led to an increase in gizzard mass repeatability, because state-behavior feedbacks would have presented themselves as within-individual correlations between gizzard mass and exploration (Luttbegg and Sih 2010; Sih et al. 2015), for which we

found no support (Table 2 process 2, Table A1). We considered the possibility that the observed increase in among individual difference in gizzard mass in juveniles between year-1 and year-2 came about as a result of individual differences in developmental plasticity (Table 2 process 1). However, since we found a decrease in among-individual variance in gizzard mass in adults between year-1 and year-2, we do not interpret these, apparently reversible, changes in among-individual variance in gizzard mass as the outcome of individual differences in developmental plasticity (Table 2 process 1; Figure A1, A; West-Eberhard 1989, 2003).

A reduction of within-individual variance contributed to increased repeatability in gizzard mass in juveniles between year-1 and year-2. In year-1, the within-individual variance in gizzard mass was higher in juveniles than in adults. However, between year-1 and year-2 within-individual variance in juveniles decreased, while we found no year-related differences in within-individual variance in adults (Table 4). Taken together, we interpret the decrease in within-individual variance found in juveniles as canalization of gizzard mass during ontogeny (Table 2 process 3).

The absence of age-related differences in repeatability of exploration behavior concealed underlying changes in variance components in both juveniles and adults. Within-individual variance in exploration decreased significantly between year-1 and year-2 for both juveniles and adults. Concomitant (non-significant) decreases in among-individual variance meant that there was no overall change in repeatability of exploration between year-1 and year-2 (Table 3). As we found no systematic change in average exploration behavior between year-1 and year-2 (Table 3, Figure 2), we rule out the possibility that the decrease in within-individual variance in exploration from year-1 to year-2 (Table 3) was the result of habituation

to the experimental arena. Instead, the decrease in within-individual variance is consistent with the idea of canalization (Table 2, process 3). Because this decrease in within-individual variance was quantitatively similar for juveniles and adults, we consider the possibility that this decrease in within-individual variation may reflect a “time in captivity” effect, as opposed to a strictly developmental process (that would result in changes in the juvenile cohort alone). Since we found no increase in among-individual variance in exploration behavior in either juveniles and adults, we suggest that among-individual differences in developmental plasticity in exploration were not at play during our experiments (Table 2, process 1).

Taken together, our results suggest that canalization may play an important role in the development of among individual differences for both gizzard mass and exploration. Importantly, our results also demonstrate that studying age-related differences in repeatability alone, without considering the differences in among- and within-individual variance components separately, is insufficient for studying developmental processes. This is because no overall change in repeatability can occur even when there are significant changes in the underlying variance components if these exhibit changes in the same direction.

We acknowledge that these trait-specific processes may be partially caused by aspects of our experimental design. Notably, our diet manipulation did not have any effect on mean exploration, even in juveniles (Figure 2). This indicates that, if there is a sensitive window during early development in which knots adjust their exploration behavior directly in response to changes in diet quality, it occurs earlier in development than the timescale of our experiments (i.e., 4 months of age). Consequently, we repeatedly induced within-individual shifts in gizzard mass, but not in exploration. This may have allowed for greater reduction in within-individual

variance in exploration in the course of the study compared with gizzard mass. As with exploration behavior, there was no evidence that juveniles were more sensitive to diet manipulations than adults in terms of adjustment in gizzard mass (Table 4), because overall diet effects were similar for adults and juveniles within each study year.

By offering standardized aviary conditions, we deliberately separated the individual from its natural environment as a way of explicitly controlling for factors that might influence physiology and behavior (Gibbons Jr. et al. 1994). Contrary to free-living knots, the birds in the context of our experiments lived in absence of competition for food and actual predation danger; these are potentially important factors known to contribute to among-individual variation (e.g., Bengston et al. 2014; Urszán et al. 2015; but see Horváth et al. 2017). The lower among-individual variance in gizzard mass and exploration in juveniles when compared to adults in year-1, indicates that among-individual variation increases over time in free-living knots. Our results therefore suggest that environmental conditions play a key role in shaping patterns of phenotypic (co-)variation in knots.

In addition, whereas co-variation between gizzard mass and exploration behavior has been reported in free-living knots (Bijleveld et al. 2014), we replicated the earlier finding that there is no co-variation between the gizzard mass and exploration in captivity (Bijleveld et al. 2014). At the same time, among-individual variation in both gizzard mass and exploration behaviour between year-1 and year-2 eroded in adults, and not in juveniles, suggesting that our captive conditions indeed removed a variable that generates and maintains among-individual variation (e.g. Archard and Braithwaite 2010; Fisher et al. 2015). However, our results are consistent with Bijleveld et al. (2014) and Oudman et al. (2017), who both found experimental

manipulations of gizzard mass not to affect the expression of exploration behavior. Thus, repeatable variation in field gizzard mass may be considered ‘pseudo-repeatable’ (Niemelä and Dingemanse 2017; i.e., it is the outcome of repeatable exposure to different conditions such as habitat or diet). Indeed, the observed among-individual variance in gizzard mass was greater in two earlier studies where the choice for prey quality was greater (Bijleveld et al. 2014) and the time in captivity was shorter (Mathot et al. 2017).

At least two other studies have reported how within-individual canalization can contribute to increased repeatability. In sea anemones (*Actinia equina*), within-individual variation in the startle response decreased over time (Osborn and Briffa 2017). Similarly, a reduction of within-individual variance explained the strong age-related increase in repeatability in multiple behavioral traits in mosquitofish (*Gambusia holbrooki*) (Polverino et al. 2016; Table 1). We suggest that reduction in within-individual variation may be the outcome of Bayesian updating (Stamps and Krishnan 2017). With increased exposure to environmental cues, individuals may be more certain in their assessment of the environment. As their ‘estimate’ becomes more accurate, smaller phenotypic adjustments are needed (Stamps and Krishnan 2014b).

Returning to Waddington’s (1942) metaphor of canalization, it is likely that the exact canalization process, and the adaptive value of within-individual canalization, varies between traits and that some traits are shaped more rigidly than others. The limited level of canalization of gizzard mass found here may be explained by the fact that there is strong selection to retain plasticity in gizzard mass. Red knots benefit by being able to fine-tune gizzard mass to seasonally changing diets and highly variable food conditions at the nonbreeding grounds

(Zwarts and Blomert 1992; Piersma et al. 1993; Battley and Piersma 2005). The strong within-individual canalization of exploration on the other hand, might be the result of a few environmental switches during early development (Waddington 1942); it suggests that having a predictable expression of exploration behavior (which varies between individuals) is beneficial even early in life. The individual exploration behavior in red knots may well start as the hatchling chicks begin to forage and self-explore their tundra birthplaces.

Acknowledgments

We thank Luc de Monte, Baptiste Garde, Demy Mollink and Evelien Witte for help during experiments and animal care. Bram Fey, Ewout Adriaanse and the crew of the RV *Navicula* and RV *Stern* for help transporting birds and fishing. Sander Holthuijsen, Job ten Horn, Olivier Kramer, Roeland Bom, Kelly Duro, Hebo Peng, Maureen Sikkema and Cheyenna de Wit for help with fishing. Job ten Horn and the ringing group 'Calidris' (Schiermonikoog) for catching birds and Evelien Witte, Kasper van Kraaij and Jan Wijmenga for animal care. We thank members of the Department of Coastal Systems (NIOZ) for discussions, and especially Allert Bijleveld, for comments on earlier drafts on the manuscript. We are grateful for the great feedback from two anonymous referees and we thank Dick Visser for preparing the figures. This work was supported by a VENI fellowship from the Netherlands Organization for Scientific research (NWO grant no. 863.14.021) to KJM. The field work, AD and TP were supported by grants to TP from NWO-ALW (TOP-grant 'Shorebirds in space', no. 854.11.004) and Waddenfonds (project 'Metawad', WF 209925).

Authorship statement: EMAK and KJM conceived of and designed the experiment. EMAK, JBB, PM, DS, AD, and KJM collected the data. EMAK analyzed the data and wrote the manuscript with KJM and TP. All authors gave final approval for publication.

Appendix: Supplementary material

Table A1. Outcome of the multivariate model on the covariance between gizzard mass and (logit-) exploration behavior.

	Year-1		Year-2	
	among	within	among	within
Juvenile	-0.19 (-0.92,0.79)	0.07 (-0.12, 0.26)	0.09 (-0.81,0.87)	-0.02 (-0.21,0.18)
Adult	0.19 (-0.77, 0.92)	0.03 (-0.15,0.23)	0.16 (-0.75, 0.87)	0.06 (-0.13, 0.23)

Note. – Means (β) and 95% Credible Intervals (CI) of within-individual correlation estimates between gizzard mass (g) and (logit-)exploration behavior. To test whether individual differences in gizzard mass and exploration behavior could have been the result of state-behavior feedbacks between gizzard mass and exploration behavior during early ontogeny, we compared the within-individual correlation between gizzard mass and exploration between age groups and years. Within-individual correlation between gizzard mass and exploration in juveniles in year-1, but not in year-2, nor in adults or years, would have been indicative for state-behavior feedbacks during early ontogeny.

We constructed four separate bivariate models with gizzard mass and (logit) exploration behavior as response variables. These models were constructed with a Gaussian error distributions. As we were only interested in the covariance structure, no fixed effects were included. A random effect bird ID was included to be able to separate within- and among-individual variance components of gizzard mass and exploration. Models were run for 1003000 simulations with a burnin of 3000 and a thinning interval of 1000, yielding a total of 1000 estimates from which to construct posterior distributions.

Literature Cited

- Araya-Ajoy, Y. G., K. J. Mathot, and N. J. Dingemanse. 2015. An approach to estimate short-term, long-term and reaction norm repeatability. *Methods in Ecology and Evolution* 6:1462–1473.
- Archard, G. A., and V. A. Braithwaite. 2010. The importance of wild populations in studies of animal temperament. *Journal of Zoology* 281:149–160.
- Bates, D., Maechler, M., Bolker, B., and S. Walker. 2015. Fitting Linear Mixed-Effects models using lme4. *Journal of Statistical Software* 67:1–48.
- Bateson, P. 1979. How do sensitive periods arise and what are they for? *Animal Behaviour* 27:470–486.
- Battley, P. F., and T. Piersma. 2005. Adaptive interplay between feeding ecology and features of the digestive tract in birds. Pages 201–228 in J. M. Starck and T. Wang, eds. *Physiological and Ecological Adaptations to Feeding in Vertebrates*. USA: Science Publishers.
- Bell, A. M., S. J. Hankison, and K. L. Laskowski. 2009. The repeatability of behaviour: a meta-analysis. *Animal Behaviour* 77:771–783.
- Bengston, S. E., J. N. Pruitt, and S. E. Riechert. 2014. Differences in environmental enrichment generate contrasting behavioural syndromes in a basal spider lineage. *Animal Behaviour* 93:105–110.
- Bierbach, D., K. L. Laskowski, and M. Wolf. 2017. Behavioural individuality in clonal fish arises despite near-identical rearing conditions. *Nature Communications* 8:15361.
- Bijleveld, A. I., R. B. MacCurdy, Y.-C. Chan, E. Penning, R. M. Gabrielson, J. Cluderay, E. L. Spaulding, et al. 2016. Understanding spatial distributions: negative density-dependence in prey causes predators to trade-off prey quantity with quality. *Proceedings of the Royal Society B* 283:20151557.
- Bijleveld, A. I., G. Massourakis, A. van der Marel, A. Dekinga, B. Spaans, J. A. van Gils, and T. Piersma. 2014. Personality drives physiological adjustments and is not related to survival. *Proceedings of the Royal Society B* 281:20133135.

- Biro, P. A., B. Adriaenssens, and P. Sampson. Individual and sex-specific differences in intrinsic growth rate covary with consistent individual differences in behaviour. 2014. *Journal of Animal Ecology* 83:1186-1195.
- Biro, P. A., and J. A. Stamps. 2015. Using repeatability to study physiological and behavioural traits: ignoring time-related change at your peril. *Animal Behaviour* 105:223-230.
- Bolnick, D. I., R. Svanbäck, J. A. Fordyce, L. H. Yang, J. M. Davis, C. D. Hulsey, and M. L. Forister. 2003. The ecology of individuals: incidence and implications of individual specialization. *American Naturalist* 161:1–28.
- Bosco, J. M., S. E. Riechert, and B. C. O'Meara. 2017. The ontogeny of personality traits in the desert funnel-web spider, *Agelenopsis lisa* (Araneae: Agelenidae). *Ethology* 123:648-658.
- Boonekamp, J. J., E. Mulder, and S. Verhulst. 2018. Canalisation in the wild: effects of developmental conditions on physiological traits are inversely linked to their association with fitness. *Ecology Letters* 21:857–864.
- Brust, V., P. M. Schindler, and L. Lewejohann. 2015. Lifetime development of behavioural phenotype in the house mouse (*Mus musculus*). *Frontiers in Zoology* 12:S17.
- Carere, C., and D. Maestriperi. 2013. *Animal Personalities. Behaviour, Physiology and Evolution.* University of Chicago Press, Chicago.
- Cleasby, I. R., and S. Nakagawa. 2011. Neglected biological patterns in the residuals. A behavioural ecologist's guide to co-operating with heteroscedasticity. *Behavioral Ecology and Sociobiology* 65:2361-2372.
- Cote, J., J. Clobert, T. Brodin, S. Fogarty, and A. Sih. 2010a. Personality-dependent dispersal: characterization, ontogeny and consequences for spatially structured populations. *Philosophical Transactions of the Royal Society B* 365:4065–4076.
- Cote, J., S. Fogarty, K. Weinersmith, T. Brodin, and A. Sih. 2010b. Personality traits and dispersal tendency in the invasive mosquitofish (*Gambusia affinis*). *Proceedings of the Royal Society B* 277:1571–9.

- Crabbe, J. C., D. Wahlsten, and B. C. Dudek. 1999. Genetics of mouse behavior: interactions with laboratory environment. *Science* 284:1670–1672.
- Cumming, G. 2009. Inference by eye: reading the overlap of independent confidence intervals. *Statistics in Medicine* 28:205–220.
- Cumming, G., and S. Finch. 2005. Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist* 60:170-180.
- Dall, S. R. X., A. M. Bell, D. I. Bolnick, and F. L. Ratnieks. 2012. An evolutionary ecology of individual differences. *Ecology Letters* 15:1189–1198.
- Dall, S. R. X., A. I. Houston, and J. M. McNamara. 2004. The behavioural ecology of personality: consistent individual differences from an adaptive perspective. *Ecology Letters* 7:734–739.
- Debeffe, L., J. F. Lemaître, U. A. Bergvall, A. J. M. Hewison, J. M. Gaillard, N. Morellet, P. Kjellander, et al. 2015. Short- and long-term repeatability of docility in the roe deer: sex and age matter. *Animal Behaviour* 109:53–63.
- Dekinga, A., M. W. Dietz, A. Koolhaas, and T. Piersma. 2001. Time course and reversibility of changes in the gizzards of red knots alternately eating hard and soft food. *Journal of Experimental Biology* 204:2167-2173.
- Dietz, M. W., A. Dekinga, T. Piersma, and S. Verhulst. 1999. Estimating organ size in small migrating shorebirds with ultrasonography: An intercalibration exercise. *Physiological and Biochemical Zoology* 72:28-37.
- Dingemanse, N. J., C. Both, P. J. Drent, K. van Oers, and A. J. van Noordwijk, 2002. Repeatability and heritability of exploratory behavior in great tits from the wild. *Animal Behaviour* 64:929–938.
- Dingemanse, N. J., K. M. Bouwman, M. van de Pol, T. van Overveld, S. C. Patrick, E. Matthysen, and J. L. Quinn. 2012. Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. *Journal of Animal Ecology* 81:116–126.

- Dingemanse, N. J., and M. Wolf. 2010. Recent models for adaptive personality differences: a review. *Philosophical Transactions of the Royal Society B* 365:3947–3958.
- Dubuc-Messier, G., D. Réale, P. Perret, and A. Charmantier. 2017. Environmental heterogeneity and population differences in blue tits personality traits. *Behavioral Ecology* 28:448–459.
- Fawcett, T. W., and W. E. Frankenhuis. 2015. Adaptive explanations for sensitive windows in development. *Frontiers in Zoology* 12:S3.
- Fisher, D. N., M. Brachmann, and J. B. Burant. 2018. Complex dynamics and the development of behavioural individuality. *Animal Behaviour* 38:e1-e6.
- Fisher, D. N., A. James, R. Rodríguez-Muñoz, and T. Tregenza. 2015. Behaviour in captivity predicts some aspects of natural behaviour, but not others, in a wild cricket population. *Proceedings of the Royal Society B* 282:20150708.
- Gelman, A., and J. Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- Gelman, A., and Y.-S. Su. 2016. *arm: Data analysis using regression and multilevel/hierarchical models*. Retrieved from <https://cran.r-project.org/package=arm>
- Gibbons Jr., E. F., E. J. Wyers, E. Waters, and E. W. Menzel Jr. 1994. *Naturalistic Environments in Captivity for Animal Behavior Research*. State University of New York Press, Albany.
- Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software* 33:1-22.
- Horváth, G., B. Mészáros, T. J. Urszán, K. Bajer, O. Molnár, L. Z. Garamszegi, and G. Herczeg. 2017. Environment-dependence of behavioural consistency in adult male European green lizards (*Lacerta viridis*). *PLoS ONE* 12: e0187657.
- Kaiser, A., T. Merckx, and H. Van Dyck. 2018. Urbanisation and sex affect the consistency of butterfly personality across metamorphosis. *Behavioral Ecology and Sociobiology* 72:1-11.

- Kim, S. L., M. T. Tinker, J. A. Estes, and P. L. Koch. 2012. Ontogenetic and among-individual variation in foraging strategies of northeast pacific white sharks based on stable isotope analysis. *PLoS ONE* 7:e45068.
- Kok, E. M. A., J. B. Burant, A. Dekinga, P. Manche, D. Saintogne, T. Piersma, and K. J. Mathot. 2019. Data from: Within-individual canalization contributes to age-related increases in trait repeatability: a longitudinal experiment in red knots. *American Naturalist*, Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.dn28cn6>
- Lendvai, A. Z., M. Giraudeau, V. Bókony, F. Angelier, and O. Chastel. 2015. Within-individual plasticity explains age-related decrease in stress response in a short-lived bird. *Biology Letters* 11:20150272.
- Luttbeg, B., and A. Sih. 2010. Risk, resources and state-dependent adaptive behavioural syndromes. *Philosophical Transactions of the Royal Society B* 365:3977–3990.
- Mathot, K. J., A. Dekinga, and T. Piersma. 2017. An experimental test of state-behaviour feedbacks: gizzard mass and foraging behaviour in red knots. *Functional Ecology* 31:1111–1121.
- McCowan, L. S. C., M. C. Mainwaring, N. H. Prior, and S. C. Griffith. 2015. Personality in the wild zebra finch: Exploration, sociality, and reproduction. *Behavioral Ecology* 26:735–746.
- Milot, E., A. A. Cohen, F. Vézina, D. M. Buehler, K. D. Matson, and T. Piersma. 2014. A novel integrative method for measuring body conditions in ecological studies based on physiological dysregulation. *Methods in Ecology and Evolution* 5:146-155.
- Minderman, J., J. M. Reid, M. Hughes, M. J. H. Denny, S. Hogg, P. G. H. Evans, and M. J. Whittingham. 2010. Novel environment exploration and home range size in starlings *Sturnus vulgaris*. *Behavioral Ecology* 21:1321–1329.
- Nakagawa, S., and H. Schielzeth. 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews* 85:935-956.
- Nakagawa, S., and H. Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* 4:133-142.

- Niemelä, P. T., and N. J. Dingemanse. 2017. Individual versus pseudo-repeatability in behaviour: lessons from translocation experiments in a wild insect. *Journal of Animal Ecology* 86:1033–1043.
- Osborn, A., and M. Briffa. 2017. Does repeatable behaviour in the laboratory represent behaviour under natural conditions? A formal comparison in sea anemones. *Animal Behaviour* 123:197–206.
- Oudman, T., A. I. Bijleveld, M. M. Kavelaars, A. Dekinga, J. Cluderay, T. Piersma, and J. A. van Gils. 2016. Diet preferences as the cause of individual differences rather than the consequence. *Journal of Animal Ecology* 85:1378–1388.
- Panchanathan, K., and W. E. Frankenhuis. 2016. The evolution of sensitive periods in a model of incremental development. *Proceedings of the Royal Society B* 283: 20152439.
- Petelle, M. B., D. E. McCoy, V. Alejandro, J. G. A. Martin, and D. T. Blumstein. 2013. Development of boldness and docility in yellow-bellied marmots. *Animal Behaviour*, 86:1147–1154.
- Piersma, T. 2007. Using the power of comparison to explain habitat use and migration strategies of shorebirds worldwide. *Journal of Ornithology* 148:S45-S59.
- Piersma, T., A. Dekinga, J. A. van Gils, B. Achterkamp, and G. H. Visser. 2003. Cost–benefit analysis of mollusc eating in a shorebird I. Foraging and processing costs estimated by the doubly labelled water method. *Journal of Experimental Biology* 206:3361–3368.
- Piersma, T., A. Koolhaas, and A. Dekinga. 1993. Interactions between stomach structure and diet choice in shorebirds. *Auk* 110:552–564.
- Polverino, G., C. Cigliano, S. Nakayama, and T. Mehner. 2016. Emergence and development of personality over the ontogeny of fish in absence of environmental stress factors. *Behavioral Ecology and Sociobiology* 70:2027–2037.
- Prater, A. J., J. H. Marchant, and J. Vuorinen. 1977. *Guide to the Identification & Aging of Holarctic Waders*. BTO Guide 17. Tring: British Trust for Ornithology.

- R Core Team. 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Réale, D., N. J. Dingemanse, A. J. N. Kazem, and J. Wright. 2010*a*. Evolutionary and ecological approaches to the study of personality. *Philosophical Transactions of the Royal Society B* 365:3937–3946.
- Réale, D., D. Garant, M. M. Humphries, P. Bergeron, V. Careau, and P. O. Montiglio. 2010*b*. Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philosophical Transactions of the Royal Society B* 365:4051–4063.
- Réale, D., S. M. Reader, D. Sol, P. T. McDougall, and N. J. Dingemanse. 2007. Integrating animal temperament within ecology and evolution. *Biological Reviews* 82:291–318.
- Sakai, O. 2018. Comparison of personality between juveniles and adults in clonal gecko species. *Journal of Ethology* 36:221–228.
- Schuster, A. C., T. Carl, and K. Foerster. 2017. Repeatability and consistency of individual behaviour in juvenile and adult Eurasian harvest mice. *Sci Nat* 104:10.
- Sheppard, C. E., R. Inger, R. A. McDonald, S. Barker, A. L. Jackson, F. J. Thompson, E. I. K. Vitikainen, et al. 2018. Intragroup competition predicts individual foraging specialisation in a group-living mammal. *Ecology Letters* 21:665:673.
- Sih, A., A. M. Bell, and J. C. Johnson. 2004. Behavioral syndromes: an integrative overview. *Quarterly Review of Biology* 81:241–277.
- Sih, A., J. Cote, M. Evans, S. Fogarty, and J. Pruitt. 2012. Ecological implications of behavioural syndromes. *Ecology Letters* 15:278–289.
- Sih, A., K. J. Mathot, M. Moirón, P. O. Montiglio, M. Wolf, and N. J. Dingemanse. 2015. Animal personality and state-behaviour feedbacks: a review and guide for empiricists. *Trends in Ecology & Evolution* 30:50–60.
- Sinn, D. L., S. D. Gosling, and N. A. Moltischniwsykyj. 2008. Development of shy/bold behaviour in squid: context-specific phenotypes associated with developmental plasticity. *Animal Behaviour* 75:433–442.

- Smith, B. R., and D. T. Blumstein. 2008. Fitness consequences of personality: A meta-analysis. *Behavioral Ecology* 19:448–455.
- Stamps, J. A., and W. E. Frankenhuis. 2016. Bayesian models of development. *Trends in Ecology & Evolution* 31:260-268.
- Stamps, J. A., and V. V. Krishnan. 2014a. Combining information from ancestors and personal experiences to predict individual differences in developmental trajectories. *American Naturalist* 184:647-657.
- Stamps, J. A., and V. V. Krishnan. 2014b. Individual differences in the potential and realized developmental plasticity of personality traits. *Frontiers in Ecology and Evolution* 2:1–15.
- Stamps, J. A., and V. V. Krishnan. 2017. Age-dependent changes in behavioural plasticity: insights from Bayesian models of development. *Animal Behaviour* 126:53–67.
- Svanbäck, R., and D. I. Bolnick. 2007. Intraspecific competition drives increased resource use diversity within a natural population. *Proceedings of the Royal Society B* 274:839–844.
- Tomkovich, P. 1992. An analysis of the geographic variability in knots *Calidris canutus* based on museum skins. *Wader Study Group Bulletin* 64:17-23.
- Toscano, B. J., N. J. Gownaris, S. M. Heerhartz, and C. J. Monaco. 2016. Personality, foraging behavior and specialization: integrating behavioral and food web ecology at the individual level. *Oecologia* 182:55-69.
- Urszán, T. J., L. Z. Garamszegi, G. Nagy, A. Hettyey, J. Török, and G. Herczeg. 2015. No personality without experience? A test on *Rana dalmatina* tadpoles. *Ecology and Evolution* 5:5847–5856.
- van Berkum, F. H., R. B. Huey, J. S. Tsuji, and T. Garland Jr. 1989. Repeatability of individual differences in locomotor performance and body size during early ontogeny of the lizard *Sceloporus occidentalis* (Baird & Girard). *Functional Ecology* 3:97–105.
- van der Velde, M., O. Haddrath, Y. I. Verkuil, A. Baker, and T. Piersma. 2017. New primers for molecular sex identification of waders. *Wader Study* 124:147-151.

- van Gils, J. A., S. R. de Rooij, J. van Belle, J. van der Meer, A. Dekinga, T. Piersma, and R. Drent. 2005. Digestive bottleneck affects foraging decisions in red knots *Calidris canutus*. I. Prey choice. *Journal of Animal Ecology* 74:105–119.
- van Gils, J. A., T. Piersma, A. Dekinga, and M. W. Dietz. 2003. Cost-benefit analysis of mollusc-eating in a shorebird. II. Optimizing gizzard size in the face of seasonal demands. *The Journal of Experimental Biology* 206:3369–3380.
- Verbeek, M. E. M., P. J. Drent, and P. R. Wiepkema. 1994. Consistent individual differences in early exploratory behaviour of male great tits. *Animal Behaviour* 48:1113–1121.
- Vézina, F., A. Dekinga, and T. Piersma. 2011. Shorebirds' seasonal adjustments in thermogenic capacity are reflected by changes in body mass: How perprogrammed and instantaneous acclimation work together. *Integrative and Comparative Biology* 51:394-408.
- Waddington, C. H. 1942. Canalization of development and the inheritance of acquired characters. *Nature* 150:563–565.
- Warton, D. I., F. K. C. Hui. 2011. The arcsine is asinine: the analysis of proportions in ecology. *Ecology* 92:3-10.
- West-Eberhard, M. J. 1989. Phenotypic plasticity and the origins of diversity. *Annual Review of Ecology and Systematics* 20:249–278.
- West-Eberhard, 2003. *Developmental Plasticity and Evolution*. Oxford University Press, Oxford.
- Westneat, D. F., J. Wright, and N. J. Dingemanse. 2015. The biology hidden inside residual within-individual phenotypic variation. *Biological Reviews* 90:729–743.
- Wexler, Y., A. Subach, J. N. Pruitt, and I. Scharf. 2016. Behavioral repeatability of flour beetles before and after metamorphosis and throughout aging. *Behavioral Ecology and Sociobiology* 70:745–753.
- Wilson, D. S. 1998. Adaptive individual differences within single populations. *Philosophical Transactions of the Royal Society B* 353:199–205.

- Wilson, D. S., and J. Yoshimura. 1994. On the coexistence of specialists and generalists. *American Naturalist* 144:692–707.
- Wolf, M., G. S. van Doorn, O. Leimar, and F. J. Weissing. 2007. Life-history trade-offs favour the evolution of animal personalities. *Nature* 447:581–584.
- Wolf, M., and F. J. Weissing. 2010. An explanatory framework for adaptive personality differences. *Philosophical Transactions of the Royal Society B* 365:3959–3968.
- Wolf, M., and F. J. Weissing. 2012. Animal personalities: Consequences for ecology and evolution. *Trends in Ecology & Evolution* 27:452–461.
- Wuerz, Y., and O. Krüger. 2015. Personality over ontogeny in zebra finches: Long-term repeatable traits but unstable behavioural syndromes. *Frontiers in Zoology*, 12:S9.
- Zwarts, L., and A.-M. Blomert. 1992. Why knot *Calidris canutus* take medium-sized *Macoma balthica* when six prey species are available. *Marine Ecology Progress Series* 83:113–128.

Tables

Table 1. A non-exhaustive review of papers reporting age-related changes in trait repeatability.

Species	Age group	Trait	Age-related effect	R/VC	Ref
Insects					
Red flour beetle (<i>Tribolium castaneum</i>)	Sub-adult → Young Adult → Adults	B: Movement B: Edge-preference	– –	R R	Wexler et al. 2016
Wild cricket (<i>Gryllus campestris</i>)	Juveniles → Adults	B: Flight initiation distance	–	VC	Niemelä and Dingemanse 2017
Speckled wood butterfly (<i>Pararge aegeria</i>)	Juvenile → Adult	B: Activity ¹	– ²	VC	Kaiser et al. 2018
Arachnids					
Desert funnel-web spider (<i>Agelenopsis lisa</i>)	Juvenile → penultimate molt → sexually mature	B: Foraging B: Exploration B: Neophobia B: Boldness	↑ ↑/↓ ³ – (M), ↑ (F) ↑ (M), – (F)	R R R R	Bosco et al. 2017
Basal tarantula (<i>Brachypelma smithi</i>)	Juvenile (1-year old) → immature (2-year old)	B: Boldness B: Latency to attack B: Conspecific	– ³ – ³	R R	Bengston et al. 2014

		tolerance	↑ ³	R	
		B: Neophobia	- ³	R	
Fish					
Eastern mosquitofish (<i>Gambusia holbrooki</i>)	Juvenile→Sub-adult→Adult	B: Distance moved	↑	R	Polverino et al. 2016
		B: Freezing time	↑	R	
		B: Hiding time	↑	R	
		B: Latency to emerge	↑	R	
Amazon molly (<i>Poecilia formosa</i>)	Hatchling→Juvenile	B: Movement	↑	VC	Bierbach et al. 2017
White shark (<i>Carcharodon carcharis</i>)	Sub-adult→Adult	B: Prey choice	↑	VC	Kim et al. 2012
Crayfish (<i>Cherax destructor</i>)	0→4 months	P: Growth rate	↑	R	Biro et al. 2014
		B: Boldness	↑		
Reptiles					
Western fence lizard (<i>Sceloporus occidentalis</i>)	Hatchling→Juvenile→ Sub-adult	B: Locomotion performance	-	R	Van Berkum et al. 1989
		M: Body size	-	R	
Gecko (<i>Lepidodactylus lugubris</i>)	Juveniles → Adult	B: Exploration	- ²	VC	Sakai 2018
		B: Boldness	- ²	VC	

Birds

Zebra finch (<i>Taeniopygia guttata</i>)	Sub-adult→Young adult→Adult	B: Fearlessness	–	R	Wuerz and Krüger 2015 ²
		B: Exploration	–	R	
		B: Activity	–	R	
		B: Aggression	–	R	
		B: Boldness	–	R	
House sparrow (<i>Passer domesticus</i>)	Nestling→Adult	P: Corticosterone levels	–	R	Lendvai et al. 2015

Mammals

Yellow bellied marmot (<i>Marmot flaviventris</i>)	Juvenile→/ Yearling→/ Adult	B: Boldness	↑/↓ ⁴	R	Petelle et al. 2013
		B: Docility	↑	R	
Eurasian harvest mice (<i>Micromys minutus</i>)	Juvenile→Adult	B: Exploration	–	R	Schuster et al. 2017
		B: Activity	↑	R	
		B: Boldness	–/↓ ⁵	R	
		C: Spatial Recognition	↑	R	
European roe deer (<i>Capreolus capreolus</i>)	Juvenile/Adult	B: Docility	–	R	Debeffe et al. 2015

Note. - In each study age groups were compared using either a longitudinal approach (→; i.e. measuring the same individual over different life stages) or a cross-sectional approach (/ ; i.e. comparing individuals belonging to different life stages) or a combination of both approaches. The type of trait is indicated as B (behavioral), M (morphological), P (physiological) or C (cognitive), with a brief description. Age related effects are

summarized as increasing (\uparrow), no change ($-$) or decreasing repeatability (\downarrow). In situations that effect differed between sexes this is indicated with M (males) and F (females). The column labeled R/VC indicates whether only changes in repeatability were considered (R), or whether changes in within- and among-individual variance components were assessed (or provided) separately (VC).

¹Note that other behaviors were also scored in this study, but the only common assay across age cohorts was activity.

²Did not formally compare the repeatability across the age classes, but point estimates were presented and could be compared.

³Repeatability increased from juvenile to penultimate molt, decreased from penultimate molt to mature adult

⁴ Juvenile and adult marmots showed no repeatability in boldness but yearlings did.

⁵ Boldness is repeatable in juveniles and in adult males but not in females. This implies an age-related decrease in repeatability for females. But this is not explicitly discussed in the paper.

Table 2. Predicted alternative processes that may generate age-related increases in trait repeatability.

Process	Among-individual variance	Within-individual variance	Within-individual trait covariance
Individual differences in developmental plasticity	↑	–	N ¹
State-behavior feedbacks	↑ ²	–	Y ³
Within-individual canalization	–	↓	N

Note. - Predicted age-related changes in among- and within-individual variance (↑ = increase, ↓ = decrease, - = no prediction), and the presence of within-individual trait covariances (Y = Yes, N = No) for three alternative processes that may generate age-related increases in trait repeatability. Predicted age-related changes in among-individual variance and within-individual variance apply both within-individuals over time, as well as across age categories (e.g., between juveniles versus adults).

1. Within-individual trait covariance is not explicitly predicted from individual differences in developmental plasticity. However, when developmental plasticity of two traits covaries, this could generate within-individual trait covariance.
2. Our experimental protocol interrupts potential state-behavior feedbacks (see methods). Thus, we would not predict to generate increased among-individual variance within-individuals over time. However, we would expect to find lower among individual variance in juveniles than in adults in year-1, as knots captured as adults would have experienced state-behavior feedbacks prior to capture.
3. Within-individual trait covariance present during ontogeny, absent after ontogeny.

Table 3. Sources of variation in logit transformed exploration behavior.

	Juveniles		Adults	
	Year-1	Year-2	Year-1	Year-2
Fixed effects	$\beta \pm 95\% \text{ CI}$	$\beta \pm 95\% \text{ CI}$	$\beta \pm 95\% \text{ CI}$	$\beta \pm 95\% \text{ CI}$
Intercept ¹	-0.92 (-1.86,0.04)	-1.67 (-2.89,-0.27)	-0.98 (-1.87,0.35)	-1.37 (-2.99,0.18)
Diet (LQ)	-0.08 (-1.18,0.76)	0.30 (-1.63,2.10)	0.39 (-0.55,1.44)	0.68 (-1.54,2.32)
Sex (M)	0.14 (-0.57,0.97)	0.56 (-0.20,1.27)	-0.14 (-1.11,0.75)	-0.61 (-1.19,0.35)
Replicate	-0.14 (-0.40,0.14)	0.00 (-0.12,0.26)	-0.34 (-0.60,-0.09)*	0.12 (-0.13,0.26)
Diet (LQ):Replicate	0.14 (-0.21,0.50)	-0.01 (-0.30,0.27)	-0.20 (-0.53,0.20)	-0.07 (-0.37,0.22)
Random effects	$\sigma \pm 95\% \text{ CI}$	$\sigma \pm 95\% \text{ CI}$	$\sigma \pm 95\% \text{ CI}$	$\sigma \pm 95\% \text{ CI}$
Bird ID	0.86 (0.57,1.18)	0.70 (0.55,1.11)	1.23 (0.91,1.70)	0.90 (0.63,1.20)
Observer	0.02 (0.00,0.13)	0.00 (0.00,0.00)	0.15 (0.01,0.13)	0.00 (0.00,0.00)
Residual variance	1.02 (0.75,1.26)	0.61 (0.46,0.80)	1.12 (0.84,1.42)	0.69 (0.54,0.92)
Repeatability	$r \pm 95\% \text{ CI}$	$r \pm 95\% \text{ CI}$	$r \pm 95\% \text{ CI}$	$r \pm 95\% \text{ CI}$
Bird ID	0.48 (0.37,0.57)	0.60 (0.46,0.66)	0.53 (0.44,0.62)	0.55 (0.46,0.64)

Note. – Exploration behaviour is defined as the fraction of time searching.

¹ Intercept estimated for females (F) on high quality (HQ diet) during their first replicate.

* $p < 0.006$ (i.e., 95% CI does not overlap zero)

Table 4. Sources of variation gizzard mass (g).

	Juveniles		Adults	
	Year-1	Year-2	Year-1	Year-2
Fixed effects	$\beta \pm 95\% \text{ CI}$	$\beta \pm 95\% \text{ CI}$	$\beta \pm 95\% \text{ CI}$	$\beta \pm 95\% \text{ CI}$
Intercept ¹	4.50 (3.22,5.42)	2.42 (0.81,5.06)	4.04 (3.12,4.91)	2.03 (-0.20,3.85)
Diet (LQ)	4.13 (2.84,5.72)*	1.97 (-0.88,4.96)	4.86 (3.78,6.06)*	4.23 (1.44,7.21)*
Sex (M)	-0.56 (-1.16,-0.07)*	-1.02 (-1.53,-0.56)*	-0.64 (-1.20,-0.13)*	-0.90 (-1.31,-0.31)*
Replicate	0.38 (-0.06,0.70)	0.46 (0.10,0.73)*	0.26 (0.00,0.58)	0.44 (0.16,0.76)*
Diet (LQ):Replicate	-0.29 (-0.88,0.17)	0.05 (-0.44,0.44)	-0.90 (-1.19,-0.34)*	-0.44 (-0.74,0.13)
Random effects	$\sigma \pm 95\% \text{ CI}$	$\sigma \pm 95\% \text{ CI}$	$\sigma \pm 95\% \text{ CI}$	$\sigma \pm 95\% \text{ CI}$
Bird ID	0.00 (0.00,0.00)	0.08 (0.04,0.13)	0.11 (0.07,0.20)	0.03 (0.02,0.05)
Residual	2.60 (1.91,3.14)	1.77 (1.30,2.14)	1.67 (1.27,2.08)	1.82 (1.38,2.27)
Repeatability	$r \pm 95\% \text{ CI}$	$r \pm 95\% \text{ CI}$	$r \pm 95\% \text{ CI}$	$r \pm 95\% \text{ CI}$
Bird ID	0.00 (0.00,0.00)	0.05 (0.03,0.07)	0.07 (0.04,0.11)	0.02 (0.01,0.03)

Note. - ¹ Intercept estimated for females (F) on high quality (HQ diet) during their first replicate.

* $p < 0.006$ (i.e., 95% CI does not overlap zero)

Figure legends

Figure 1. Gizzard mass as a function of manipulated diet and time. Boxes represent the median, quartiles and interquartile outliers in within-individual centered gizzard mass (g) for juveniles (top) and adults (bottom). Gizzard mass was centered within individual by subtracting the individuals mean from each measurement (n=8). Dark grey indicate birds on high quality food (pellets) and the light grey indicate birds on low quality food (mudsnails). Means (black circles) are connected with grey lines for each group of individuals receiving the similar treatment order in the crossover design. The horizontal black lines represent the mean gizzard mass for juveniles (continuous line) and for adults (interrupted line). Measurement taken during the non-breeding season in year-1 (i.e. replicate 1-4) are separated from measurements taken during the non-breeding season in year-2 (i.e. replicate 5-8) by a breeding summer when no measurements were taken.

Figure 2. Exploration behavior scores as a function of manipulated diet and time. Boxes represent the median, quartiles and interquartile outliers in within-individual centered score for exploration behavior for juveniles (top) and adults (bottom). Exploration was centered within individual by subtracting the individuals mean from each measurement (n=8). Dark grey indicate birds on high quality food (pellets) and the light grey indicate birds on low quality food (mudsnails). Means (black circles) are connected with grey lines for each group of individuals receiving the similar treatment order in the crossover design. The horizontal black lines represent the mean exploration behavior for juveniles (continuous line) and for adults

(interrupted line). Measurement taken during the non-breeding season in year-1 (i.e. replicate 1-4) are separated from measurements taken during the non-breeding season in year-2 (i.e. replicate 5-8) by a breeding summer when no measurements were taken.

Figure 3. Development of variance components of gizzard mass and exploration behavior.

Adjusted repeatabilities (top), among-individual variation (middle) and within-individual variation (bottom) for gizzard mass (left) and exploration behavior (right) for juveniles (black) and adults (grey) in year-1 and year-2. Dots and bars represent $\beta + 95\%$ CI for each age group per year.

Figure A1. Visual representation of the three processes underlying age related increase in repeatability. Visual representation of three mechanisms underlying age-related increase in repeatability. Each color represents a different individual. Solid lines indicate mean trait values per individual, and dotted lines represent within-individual variance in trait values (i.e., residual variance). Panels A and B differ only in that state-behavior feedbacks require some initial among-individual differences, though these can be small, and may arise due to a chance events.

Figure A2. Layout of the experimental arena. Experimental birds were released individually from the holding aviary. Behavior was observed and digitally scored by one observer from behind a one-way-mirror next to island 5.

Figure 1

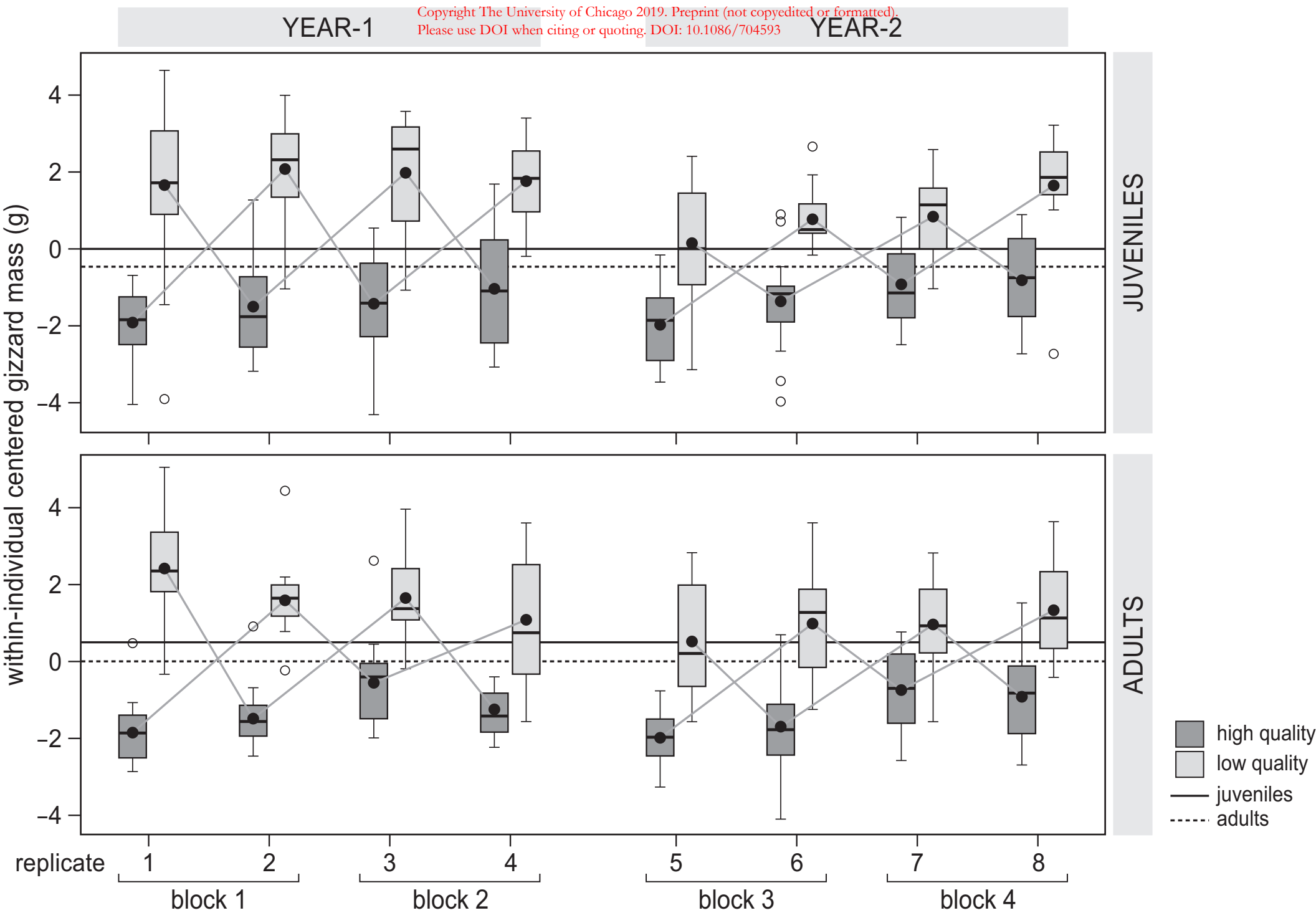


Figure 2

Copyright The University of Chicago 2019. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/704593

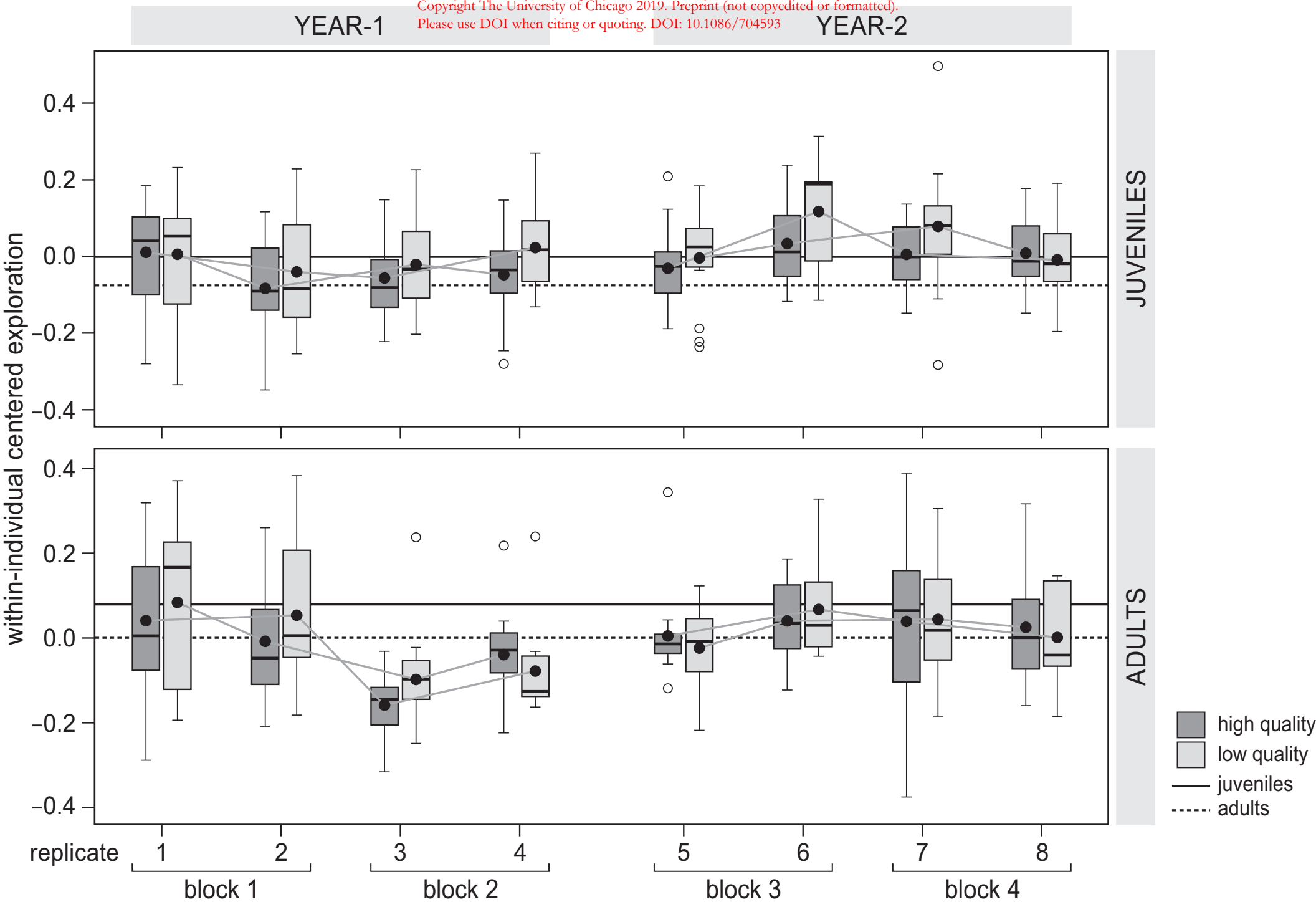


Figure 3

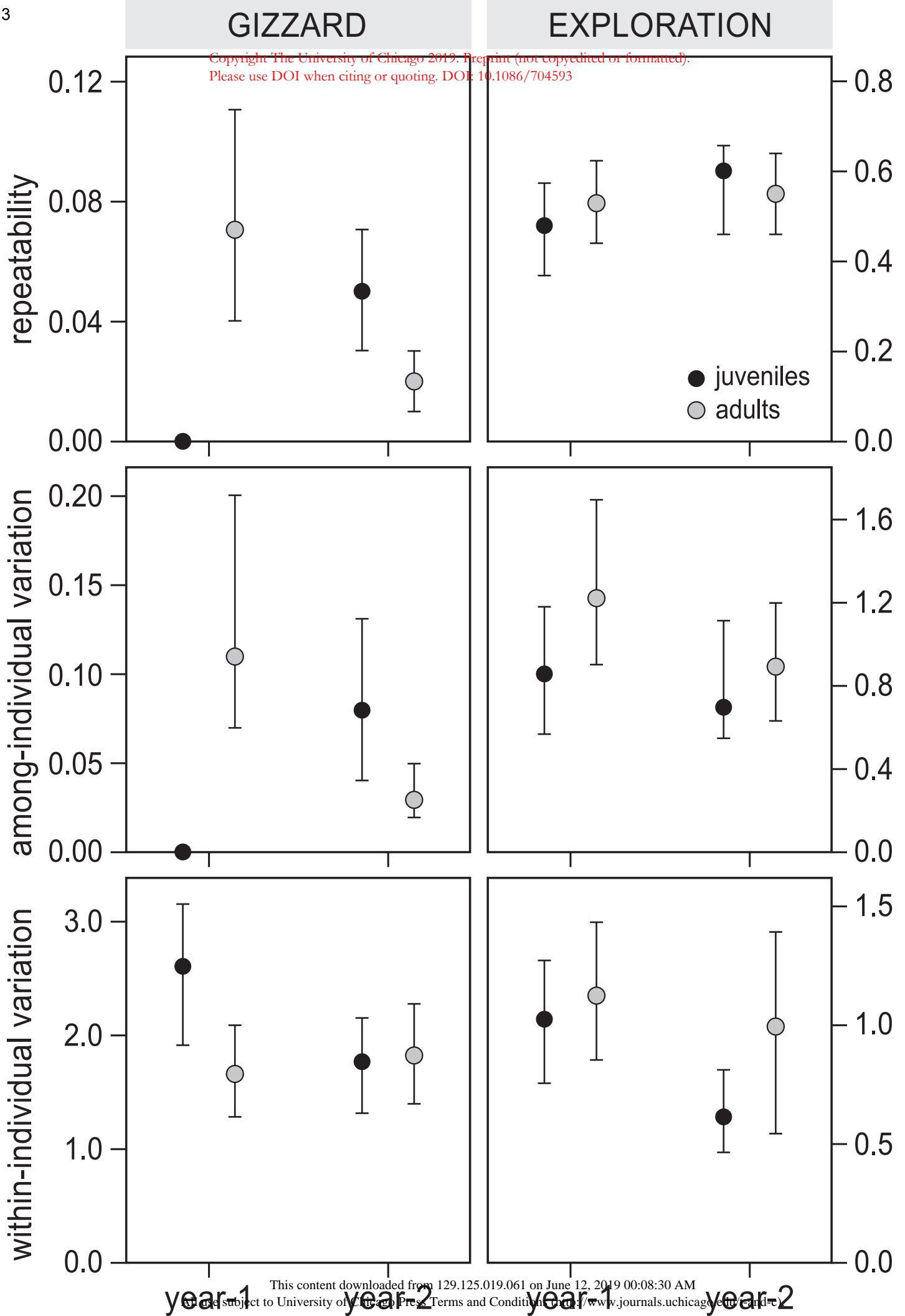


Figure A1

Copyright The University of Chicago 2019. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/704593

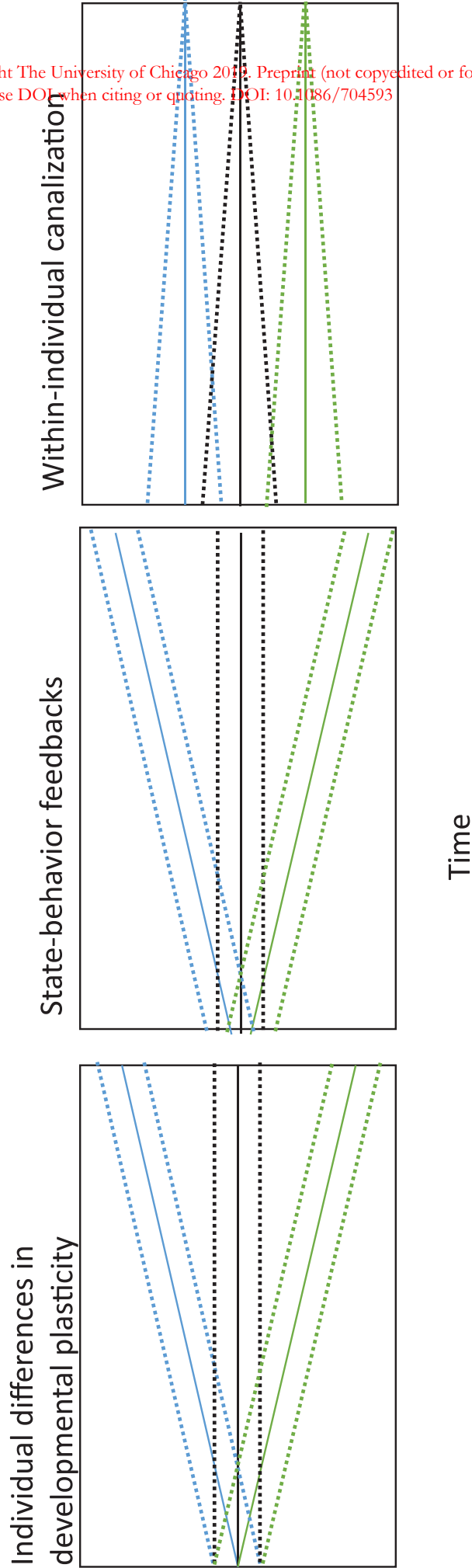


Figure A2

**HOLDING
AVIARY**

Copyright The University of Chicago 2019. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/704593

