



Assessing Short- and Long-Term Repeatability and Stability of Personality in Captive Zebra Finches Using Longitudinal Data

Morgan David*, †, Yannick Auclair* & Frank Cézilly*, ‡

* Équipe Écologie Évolutive, UMR CNRS 6282 Biogéosciences, Université de Bourgogne, Dijon, France

† Groupe de Recherche en Écologie Comportementale et Animale, Département des Sciences Biologiques, Université du Québec à Montréal, Montréal, QC, Canada

‡ Institut Universitaire de France, Paris, France

Correspondence

Morgan David, Centre for Ecology & Conservation, School of Biosciences, College of Life & Environmental Sciences, University of Exeter, Cornwall Campus, Penryn, Cornwall TR10 9EZ, UK.
E-mail: Morgan.David@gmx.fr

Received: May 12, 2012

Initial acceptance: June 19, 2012

Final acceptance: June 26, 2012

(J. Wright)

doi: 10.1111/j.1439-0310.2012.02085.x

Abstract

Assessing behavioural consistency is crucial to understand the evolution of personality traits. In the present study, we examined the short- and long-term repeatability and stability of two unrelated personality traits – exploratory tendencies and struggling rate – using captive female zebra finches (*Taeniopygia guttata*). We performed two experimental sessions of behavioural tests with a 7-mo interval, which represents up to one quarter of a zebra finch's life expectancy. We showed that, overall, exploratory tendencies and struggling rate were significantly repeatable in the short term. However, only exploratory tendencies were repeatable in the long term. We found interindividual differences in short-term stability of exploratory tendencies, but not struggling rate, providing evidence for differences in intraindividual variability. In the long term, struggling rate significantly decreased between the two experimental sessions, whereas exploratory tendencies remained stable. Finally, the amount of interindividual variation measured at both sessions did not differ. Our results suggest that short- and long-term repeatability and stability of personality may vary between individuals, depending on the behavioural trait under scrutiny. As a consequence, deducing personality from measures realized earlier in a subject's life should be performed with caution. We discuss the implications of inter- and intraindividual variation in personality consistency on individual fitness.

Introduction

The study of animal personality is based on the characterization of interindividual behavioural variation (Bell 2007). Several experimental designs have been set up which allow researchers to measure individual personality in some standardized way (Réale et al. 2007). Yet, while studies investigating human personality usually take several broad dimensions into account (McCrae et al. 1998; Uher 2008), animal personality is commonly inferred from a few traits only (Butler et al. 2012; Hedrick & Kortet 2012). The discovery that several personality traits may be related within so-called 'behavioural syndromes' provides empirical grounding to this procedure. Yet, the extent to which the assessment of a

single trait is representative of global personality, and possibly several different dimensions, remains poorly investigated (but see David et al. 2011). More importantly, relationships between personality traits have been shown to be mediated by natural selection (Bell & Sih 2007; Dingemanse et al. 2007). The global structure of personality is thus thought to be dynamic over time. Investigating the properties of several different personality traits is thus of major concern to provide an account of their evolution in the face of natural selection.

The concept of animal personality has been developed around the premise that individuals of the same populations display consistent behavioural tendencies over time and across contexts (Réale et al. 2007). Repeatability and stability are commonly used as

measures of consistency yet provide different information (see Dingemanse et al. 2010). From a population genetics viewpoint, repeatability is an estimate of the proportion of phenotypic variation, which is explained by individual differences (Falconer & Mackay 1996). Stability refers to the absolute individual variation of the trait under study over an ecologically relevant time period. Assessing both repeatability and stability of personality traits is thus crucial to the characterization of within-population behavioural differences; and to appraise the biological relevance of this variation in an evolutionary framework (Dingemanse & Réale 2005).

Several studies have estimated the repeatability of personality traits using wild-caught animals (Ward et al. 2004; Quinn & Cresswell 2005; Montiglio et al. 2010; Jenkins 2011; Mafli et al. 2011) or individuals bred in captivity (Schuett & Dall 2009; Jones & Godin 2010; Schielzeth et al. 2011). Several factors have been found to influence behaviour repeatability and thus could potentially affect the accuracy of the measures recorded. For instance, repeatability may differ between sexes (Schuett & Dall 2009; Jenkins 2011) or across populations (Sinn et al. 2010; van Dongen et al. 2010). In a recent meta-analysis, Bell et al. (2009) found that experimental trials performed at short intervals of time generally yielded more repeatable estimates than behaviours recorded at longer intervals (Bell et al. 2009). However, few studies have investigated the repeatability of personality over both short- and long-time intervals using captive individuals (but see Svartberg et al. 2005; Koski 2011 and Kanda et al. 2012). As a consequence, less is known about whether testing captive animals provides accurate and stable estimates of an individual's personality and if such experiments can be used to infer the evolution of personality in the wild (see Weigensberg & Roff 1996 and Herborn et al. 2010). The question of the use of captive individuals within personality studies remains an issue (Archard & Braithwaite 2010).

In the present study, we used the zebra finch (*Taeniopygia guttata*) to estimate the repeatability and stability of two personality traits: exploration tendencies and struggling rate. The first trait has been found to be part of a behavioural syndrome, along with neophobia, risk-taking behaviour and activity, whereas the second one was found to be independent of other personality traits (David et al. 2011). Performing experimental trials over short- (7 d) and long-term intervals (~7 mo), we assessed to what extent intra-population variation in personality is repeatable over time, and whether different personality dimensions show similar patterns of stability.

Methods

Study Subjects

Twenty young, but sexually mature, captive-born zebra finches were purchased from a reliable local supplier. Only females which were initially 2 mo of age were used, as the present study was part of a larger research programme on the influence of personality on female sexual behaviour (David & Cézilly 2011). The zebra finch is a common tropical passerine bird species, widely used in behavioural ecology (Zann 1996) and specifically within personality studies (Schuett & Dall 2009; David et al. 2011, 2012). Birds were kept in a single experimental room one to 3 wks before the experiment started and were maintained in individual cages (60 × 33 × 30 cm), containing four perches at different heights and four feeders with food or water. Room temperature was maintained at 22 ± 2°C, and the photoperiod was 13:11 h light/dark cycle (0730 am–0830 pm), with a 30-min period simulating dawn and dusk through increasing and decreasing light intensity in the morning and evening. Each individual was identified by an orange numbered ring (size XF; AC Hughes, Hampton Hill, UK). Birds were provisioned with millet seeds, cuttlebones and water *ad libitum*. Vegetables were only given outside of the experiment periods. Birds were not given any additional vitamins. After the experiment, birds were kept in groups of two to be used in subsequent studies. The experimental procedures were in agreement with the ethical requirements of the Université de Bourgogne.

General Procedures

Two different personality traits were assessed: exploratory tendencies and struggling rate (David et al. 2011). Long-term consistency of each trait was assessed through two experimental sessions separated by 263 d [214, 306] (median [lower value, upper value]) for exploration and 209 d [155, 236] for struggling rate. Within each experimental session, the short-term consistency of each trait was assessed through two trials for each individual at a (mean ± standard deviation) 7 ± 0 d interval in the first session and a 3 ± 1 d interval in the second session. Struggling rate, then exploratory tendencies, were assessed from the same individuals with a 6-d interval during the first session and a ~2-mo interval during the second session. Whereas zebra finches' lifespan has been found to vary between different wild populations, individuals typically live between 2 to

3 yrs (Zann 1996). The lifespan of zebra finches kept in optimal conditions in captivity may reach 5 yrs (Burley 1985). The between-session interval therefore represents an important proportion (~23%) of zebra finches' life expectancy, at least compared to the highest lifespan recorded in the wild (Zann 1996). After the first experimental session, birds underwent two spatial association mate-choice trials with no physical male contact (M. David, unpubl. data). They were then kept in same-sex groups of two until the second session without being disturbed but by caretakers.

Assessment of Exploratory Tendencies

Individual exploratory tendencies were assessed in a large cage (140 × 140 × 70 cm) with which birds were unfamiliar. The cage consisted of opaque walls with a clear Plexiglas ceiling and contained five artificial trees each composed of four small branches (David et al. 2011). Individuals were deprived of food for 1 h before the trial and had no access to any further food source within the apparatus. Individual behaviour inside the apparatus was recorded for 30 min with a video camera placed one metre above it. During analysis, the cumulative number of movements between perches and between trees performed by the bird was considered as a measure of its exploratory tendencies.

Assessment of Struggling Rate

Individuals were caught from their home cages and gently put on a flat surface below a small landing net (20 × 15 cm). The numbers of hops and escape attempts under the net were recorded during 30 s (David et al. 2011), while the experimenter remained motionless one metre away from the bird. All these manipulations were performed by M.D. The cumulative number of hops and escape attempts were considered as a measure of individual struggling rate (also termed 'obstinacy' in David et al. 2011).

Statistical Analyses

The results obtained during the first session represent a subset of the data reported in David et al. (2011). Three individuals were not tested for their exploratory tendencies thus reducing the sample size to 17 for analyses with exploratory tendencies as a dependent variable. Repeatability of personality was assessed using the 'rptR' package (Nakagawa & Schielzeth 2010) developed in R software (R Development Core Team 2011). Thus, the repeatability of exploratory

tendencies was computed using link scale method for count data. Here, we report the estimates (repeatability coefficient [R] and 95% confidence interval) of multiplicative overdispersion glmm-based models. Differences between these repeatability estimates computed within and among the two sessions were assessed by comparing the corresponding 95% confidence intervals. Overlapping intervals indicated that the repeatability estimates were not significantly different, at the 5% significance threshold (Nakagawa & Cuthill 2007).

A preliminary analysis of struggling rate, using procedures for count data, showed that confidence interval estimates were unreliable as the overdispersion parameter of the model fell below 1. This would have led to erroneous statistical comparisons of repeatability estimates. As number of hops and escape attempts were normally distributed (Shapiro tests, $p > 0.05$ for all tests) and estimated overdispersion fell below 1, we decided to use ANOVA-based models to assess the repeatability of struggling rate. To that end, we used the command 'rpt.aov' of the 'rptR' R package (Nakagawa & Schielzeth 2010). Differences between repeatability estimates were thus appraised by comparing confidence intervals, with overlaps indicating no differences at a 5% threshold.

Between-session differences in personality, averaged within sessions for each bird and each trait, were analysed using generalized linear mixed models with either 'exploratory tendencies' or 'struggling rate' as a dependent variable, 'individual' as a random intercept and 'session' as a fixed parameter. Following Schielzeth & Forstmeier's (2009) recommendation, we included 'session' as a random slope factor to test for differences in slope (i.e., between-individual differences in between-session variation). Residual error distribution was set to Poisson. Between-session differences in personality and between-individual slope differences were thus determined by comparing models with or without the corresponding fixed, or random, factors.

Moreover, we investigated whether between-session variation in personality was related to average personality values for the two traits. To do this, we used Oldham's method (Oldham 1962) as recommended by Tu & Gilthorpe (2007). This method has been designed to estimate the association between change and initial value, while avoiding the pitfalls arising from relating a variable to any computation including it (for instance, 'x' to 'x-y'; see Tu & Gilthorpe 2007). We have then correlated between-session variation, computed as the absolute value of the difference in individual personality values – averaged within sessions – between the first and the second session $|m_{\text{first session}} - m_{\text{second session}}|$, to individual personality

values, averaged within sessions, then averaged across the two sessions ($m_{\text{first session}} + m_{\text{second session}}/2$).

We investigated the importance of intraindividual variability (IIV) in our sample (Herczeg & Garamszegi 2012; Stamps et al. 2012). Intraindividual variability refers to between-individual differences in intraindividual variation when animals are tested several times in the same situation and the same conditions. We thus computed individual stability statistic (ISS) scores for each individual in each of the two sessions (Asendorpf 1990; Sinn et al. 2008). As the scores were skewed, they were transformed to reach normality (Asendorpf 1990; Sinn et al. 2008). Individual stability statistic scores' repeatabilities were tested using ANOVA-based methods (Nakagawa & Schielzeth 2010). Between-session variation in ISS scores was assessed using linear mixed models fitted with a Gaussian distribution with 'transformed ISS score' as the dependent variable, 'session' as the independent variable, 'individual' as a random intercept factor and 'session' as a random slope factor (Schielzeth & Forstmeier 2009).

We assessed the correlation linking between-session variation in personality, as previously computed, and the time interval between the two experimental sessions for the two personality traits. This procedure was aimed at testing whether differences in the length of time between the two sessions could explain individual between-session variation in personality.

Finally, the correlation between exploratory tendencies and struggling rate was tested, for both the first and the second experimental session, using the individual average of the two within-session measures for each variable.

For each analysis, we provided Pearson's correlation coefficients with its associated 95% Confidence Interval (CI) (Anderson et al. 2001; Cassey et al. 2004; Nakagawa & Cuthill 2007). When data were not normally distributed, we did not transform them in accordance with previous recommendations (O'Hara & Kotze 2010; Zuur et al. 2010). Instead, we performed Spearman's correlation analyses. We then computed 95%CI for Spearman's coefficients using bootstrapping method with 20 000 repetitions. Statistical analyses were performed using JMP 5.0.1 (SAS Institute, Cary, NC, USA) and R 2.11.1 (R Development Core Team 2011).

Results

Within- and Between-Session Repeatabilities

Exploratory tendencies were significantly repeatable between the two trials of the first session ($R = 0.81$,

95%CI [0.59, 0.95], $p = 0.001$) and across the two sessions ($R = 0.76$, 95%CI [0.46, 0.92], $p = 0.001$), but not between the two trials of the second session ($R = 0.35$, 95%CI [0, 0.73], $p = 0.14$). However, the confidence intervals' overlap indicated that these three repeatability estimates did not significantly differ from each other.

Struggling rate was significantly repeatable between the two trials of the first ($R = 0.54$, 95%CI [0.21, 0.88], $p = 0.005$) and the second session ($R = 0.88$, 95%CI [0.78, 0.99], $p < 0.0001$). However, struggling rate was not repeatable across the two sessions ($R = 0.15$, 95%CI [-0.31, 0.61], $p = 0.25$). The overlap of confidence intervals indicated that the repeatability estimates of the first and second session did not differ from each other.

Individual stability statistic scores of exploratory tendencies were significantly repeatable between the two sessions ($R = 0.51$, 95%CI [0.11, 0.91], $p = 0.016$). This indicates that individuals differed in their short-term stability. Thus, a given individual showed similar levels of IIV during both the first and the second sessions. Conversely, ISS scores of struggling rate were not repeatable ($R = 0.05$, 95%CI [-0.42, 0.53], $p = 0.41$). Repeatability analyses are summed up in Fig. 1.

Within- and Between-Session Stability

Exploratory tendencies, averaged within sessions, did not vary between the first and the second session (1st session: median [interquartile range], 8 [3, 62]; 2nd session, 12 [2, 53]; $\chi^2 = 0.24$; $df = 1$; $p = 0.63$; Fig. 2a). Within-session variance did not significantly differ between the two sessions (1st session [log-transformed]: 0.46; 2nd session [log-transformed], 0.50; Levene's test, $F_{1,32} = 0.14$; $p = 0.71$). There were significant between-individual differences in between-session variation in exploratory tendencies (random slope effect, $\chi^2 = 117.8$; $df = 2$; $p < 0.0001$). Also, between-session variation was positively related to exploratory tendencies averaged across sessions (Spearman's $r_s = 0.69$, 95%CI [0.23, 0.96], $n = 20$, $p = 0.001$; Fig. 3a). This indicates that those individuals with higher exploratory tendencies demonstrated more variation in their exploratory tendencies between sessions. Between-session variation was not related to the time interval between the two experimental sessions (Spearman's $r_s = -0.21$, 95%CI [-0.67, 0.37], $n = 17$, $p = 0.39$). Individual stability statistic scores of exploratory tendencies did not significantly differ between the two sessions ($\chi^2 = 2.60$, $df = 1$, $p = 0.11$; Fig. 4a). However, there were significant between-individual differences in between-session

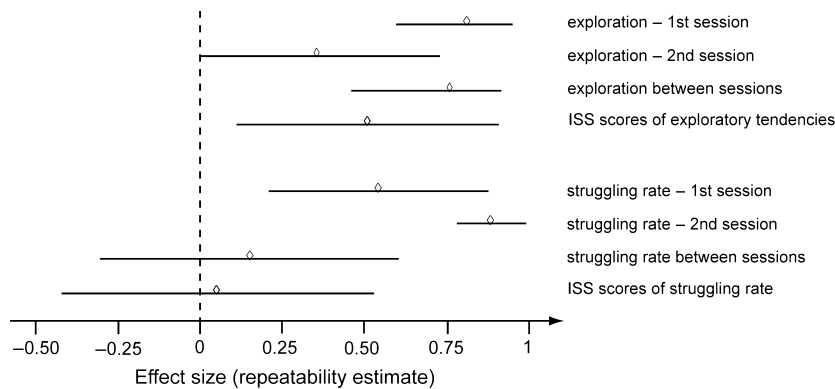


Fig. 1: Summary of each personality trait's repeatability estimate and individual stability statistic (ISS) scores with their 95% confidence interval. ISS is an index of intraindividual variability (see Methods). The x-axis represents the effect size. Open diamonds are the computed estimates of repeatability and horizontal lines represent their 95% confidence intervals.

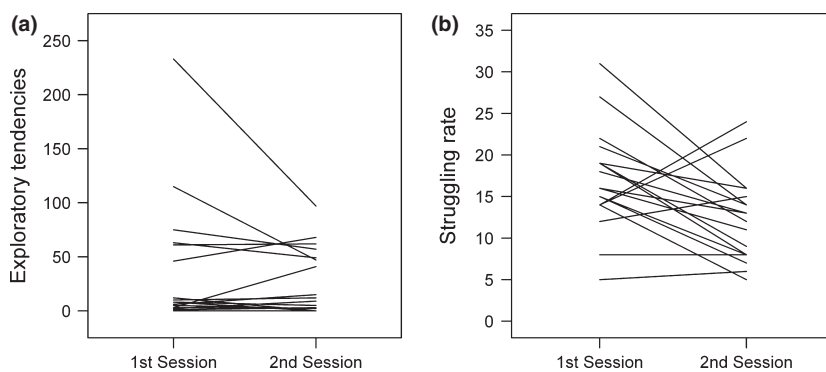


Fig. 2: (a) Between-session variation in exploratory tendencies, averaged within each session. Each line represents a single individual. (b) Between-session variation in struggling rate, averaged within each session. Each line represents one individual.

variation in ISS scores (random slope effect, $\chi^2 = 7.08$; $df = 2$; $p = 0.029$; Fig. 4a). The results of these analyses remained substantially similar when the most explorative bird was removed from the sample.

Average struggling rate decreased between the two sessions (1st session, mean \pm SEM, 16.0 ± 1.4 ; 2nd session: 12.0 ± 1.2 ; $\chi^2 = 11.6$; $p < 0.001$; Fig. 2b). However, within-session variance did not significantly differ between the two sessions (1st session, 38.6; 2nd session, 26.9; $F_{1,36} = 0.03$, $p = 0.86$). There were no significant between-individual differences in between-session variation of struggling rate (random slope effect, $\chi^2 = 1.71$; $df = 2$; $p = 0.42$). However, between-session variation was positively correlated to struggling rate averaged across sessions (Pearson's $r = 0.68$, 95%CI [0.34, 0.86], $t_{18} = 3.96$, $p < 0.001$; Fig. 3b). Thus, individuals with higher struggling rate demonstrated more variation between sessions. Between-session variation tended to be related to the time interval between the two experimental sessions (Spearman's $r_s = 0.41$; 95%CI [-0.02, 0.75]; $n = 20$; $p = 0.068$). Individual stability statistic scores of struggling rate tended to be higher in the second session than in the first, although the difference did not reach significance ($\chi^2 = 3.70$, $df = 1$, $p = 0.054$;

Fig. 4b). This indicates that struggling rate tended to be more stable in the short term throughout the first session than during the second session. There were no significant between-individual differences in between-session variation in ISS scores (random slope effect, $\chi^2 = 1.40$; $df = 2$; $p = 0.50$; Fig. 4b).

Correlation Between Exploratory Tendencies and Struggling Rate

There was no significant correlation between exploratory tendencies and struggling rate either during the first (Spearman's $r_s = -0.32$, 95%CI [-0.77, 0.19], $n = 17$, $p = 0.19$) or the second session (Spearman's $r_s = -0.10$, 95%CI [-0.55, 0.36], $n = 17$, $p = 0.68$).

Discussion

In the present study, we have shown that exploratory tendencies and struggling rate are overall repeatable in captive female zebra finches, at two and 9 mo of age. Our results demonstrate that a commonly studied personality trait such as exploration may be repeatable over a significant portion of zebra finches' life expectancy. However, struggling rate was not

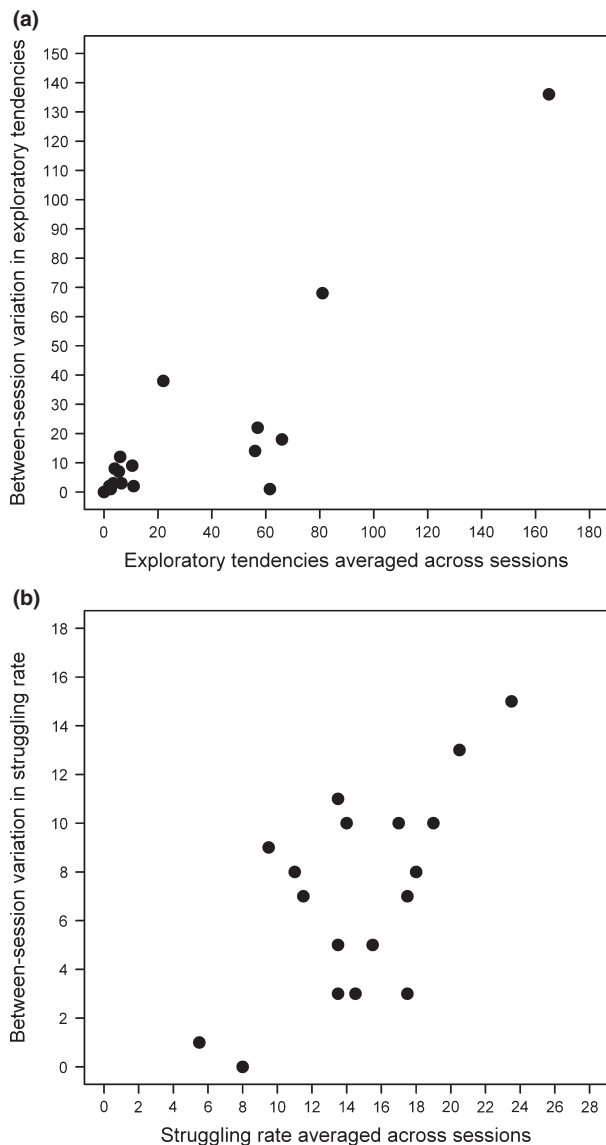


Fig. 3: (a) Relationship between between-session variation in individual exploratory tendencies and individual exploratory tendencies averaged across both sessions. (b) Relationship between between-session variation in individual struggling rate and individual struggling rate averaged across both sessions. For both personality traits, between-session variation was computed as the absolute value of the difference between individual average value in first session and individual average value in second session (see Methods for more details).

repeatable in the long term. Exploratory tendencies and struggling rate varied in their stability through time. Indeed, exploratory tendencies remained rather stable in the long term, whereas birds decreased their level of struggling rate over the 7-mo period. Here, we discuss each of these findings in turn and explore the importance of methodology in an evolutionary approach to animal personality.

Behavioural Repeatability

Similar short-term repeatability estimates were computed during the two experimental sessions, both for exploratory tendencies and struggling rate. Although the time interval between the two within-session trials differed between the two sessions (7 d for the first session vs. 3 d for the second session), the magnitude of repeatability did not vary as expected (Bell et al. 2009). Our findings yet differ from a previous study by Schuett & Dall (2009) which showed that, unlike males, exploratory tendencies of captive female zebra finches were not repeatable (see also Jenkins 2011). Taken together, these studies and the present one suggest that individuals from different populations may display varying patterns in personality trait repeatability, as has been found in another bird species (van Dongen et al. 2010).

Exploratory tendencies were repeatable in the long term. This contrasts with previous findings that exploratory tendencies were not repeatable across juvenile and adult stages in wild great tits (*Parus major*) (Carere et al. 2005). However, our results are in line with primate (Koski 2011) and mammal studies (Ray & Hansen 2005; Svartberg et al. 2005). Conversely, we found no evidence that struggling rate was repeatable across the two experimental sessions. In accordance with this finding Bell et al. (2009) provided evidence that behaviour repeatability generally decreased with increased intervals between successive observations (but see Gabriel & Black 2010). However, because our sample size was moderate, any individual variation occurring between the two sessions may have had a disproportionate effect on computed intraindividual variation compared to interindividual variation. This could cause repeatability estimate to decrease and even lose significance, although this was not true for exploration tendencies where the most explorative individual drastically decreased its exploration score between the two sessions (see Fig. 2a). The rather large confidence intervals surrounding repeatability estimates can also be interpreted as an outcome of our modest sample size. The possibility then remains that increasing sample size may reduce confidence intervals and reveal a significant, while low, repeatability estimate for struggling rate's long-term repeatability. This possibility deserves further experimental investigation.

Our results suggest that unrelated personality traits may exhibit different patterns of repeatability in the same individuals when recorded over short or long time intervals. Variation in repeatability may translate into some lability in the response of these traits to

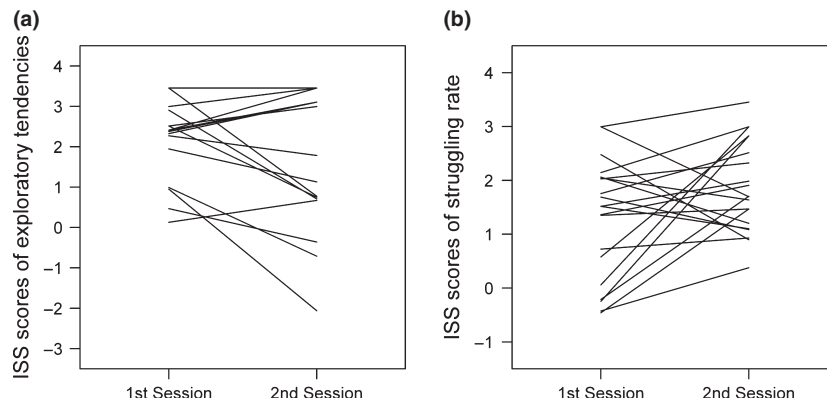


Fig. 4: (a) Between-session variation in individual stability statistic (ISS) scores of exploratory tendencies. (b) Between-session variation in ISS scores of struggling rate. Each line represents one individual. Individual stability statistic scores are indicative of intraindividual variability (see Methods) and thus quantify individual short-term behavioural stability.

natural selection. Assessing the repeatability or personality, and behavioural traits in general, is then crucial to understand its relevance over an individual's lifetime and their importance from an evolutionary perspective. The present study provides evidence that accurate estimates of personality can be obtained from captive studies but that these cannot be generalized to account for other behavioural traits, without a sense of caution.

Behavioural Stability

Individual exploratory tendencies did not significantly vary between the two experimental sessions. The number of movements performed within the exploration apparatus remained stable over the 7-mo interval. This finding contrasts with previous studies conducted with dogs (*Canis familiaris*) (Svartberg et al. 2005) or great tits (Dingemanse et al. 2002) which found that individual exploratory tendencies increased over time, whereas exploration has been found to decrease with age in rats (*Rattus norvegicus*) (Ray & Hansen 2005). However, we found that individual stability, as reflected in the computation of between-session variation, varied as a function of individual personality. Over a 7-mo period, the number of movements birds performed in the exploration apparatus decreased in those with higher exploratory tendencies to a greater extent than birds with lower exploratory tendencies. These results suggest that personality stability is likely to vary between individuals, at least in zebra finches. This has been documented in the present study through short-term stability of exploration (i.e., ISS scores) and between-session variation between individuals. This suggests that different individuals may exhibit varying levels of stability. As a consequence, deducing personality from measures realized earlier in a subject's life should be per-

formed with caution. Nonetheless, these results may have arisen because of a 'floor' effect as birds displaying very low exploratory tendencies during the first session could not exhibit lower tendencies during the second one. It remains possible that between-individual differences in long-term stability may be an artefact of our experimental design.

Struggling rate decreased overall between sessions, but showed a pattern similar to exploratory tendencies at the individual level: birds with higher rates displayed a greater decrease across the two sessions. Individual habituation to captive conditions and experimental situations may explain our findings. Indeed, a lack of important diverse stimulations, such as those limited field conditions provide, might considerably reduce individual general reactivity on a long-term basis (see Meehan & Mench 2002). In addition, individuals may become habituated to the presence of experimenters in captive conditions, thus reducing their docility and struggling rate. However, it remains to be explained why exploration and struggling rate showed a different pattern of variation across the sessions. Contrary to exploration, we did not find any evidence of interindividual differences in IIV in struggling rate. As struggling rate was estimated based on a shorter timescale than exploration (30 s vs. 30 min), the possibility remains that its assessment may be more influenced by small environmental perturbations, which could impact the true intensity of struggling and therefore individual measurements. Alternatively, an interesting but more hypothetical possibility is that different personality traits exhibit different levels of consistency in the same individuals (Sinn et al. 2008).

Our findings that struggling rate is neither repeatable nor stable in the long term have several implications. First, it suggests that the characterization of personality at a given time may prove unreliable to

infer its stability over long-time scales – depending on which trait is under scrutiny. This is even more astonishing for captive individuals who are assumed to live a less challenging lifestyle compared with wild animals (but see D'Eath 2004; Brown et al. 2007). One would expect captive individuals to experience a lower diversity of situations, in a neutral and predictable environment, and thus to show more stability in their behaviour than those in the wild (Archard & Braithwaite 2010; but see Weigensberg & Roff 1996 and McPhee 2003). Comparative designs based on both wild and captive individuals would be needed to test this assumption. In contrast to these expectations, Bell et al. (2009) found that behaviours measured in the field are generally more repeatable than behaviours measured in captive conditions. Natural conditions could thus play an important role in the selective maintenance of intraindividual consistency (Archard & Braithwaite 2010; Schuett et al. 2010). Second, insofar as selective pressures on personality traits are stable over time, a given population may display varying responses to selection with respect to variance in the expression of individuals' personality. In the present study, we have documented that individuals differed in their IIV – as assessed from short-term stability (ISS scores) – in exploratory tendencies but not in struggling rate. The notion of IIV is reminiscent of the notion of behavioural flexibility, which is known to affect some components of fitness (Sol et al. 2002; Nicolakakis et al. 2003). Exploring variation in IIV is thus of great importance to understand and predict the evolution of variation in personality and determine the evolutionary foundations of behavioural consistency and flexibility.

To conclude, we have provided evidence that two unrelated personality traits were repeatable in the short term in captive zebra finches, while only exploratory tendencies were repeatable in the long term. Our findings are yet based on a modest sample size. The possibility remains that significant differences between estimates could have been revealed by substantially increasing the number of individuals used or the number of trials performed with the same individuals (Martin et al. 2011). Exploratory tendencies proved to be stable across a significant portion of an individual's lifetime, whereas struggling rate decreased between the two sessions. Moreover, IIV, as assessed from short-term stability, was repeatable for exploratory tendencies, but not for struggling rate. These distinctions are particularly sensible as exploratory tendencies and struggling rate belong to two different personality dimensions in zebra finches (David

et al. 2011). Thus, the present study confirms David et al. (2011)'s assumption that overall personality can be composed of several different dimensions which are worth investigating simultaneously. Indeed, these dimensions may display varying patterns of development with varying stability over an individual's life. Taken together, these results argue in favour of a more extensive characterization of alternative personality dimensions within populations (Gosling 1998; Sinn & Moltschanivskyj 2005; Weiss et al. 2012).

The present study is an appropriate illustration of the extent to which different personality dimensions can differ by nature, by their relationship with life-history traits (David et al. 2011) and by their stability through an individual's life. Our findings emphasize the risk of erroneously inferring individual personality from non-contemporaneous observations. We thus recommend that personality should, when possible, be systematically assessed when it is related to any other variable under study, rather than relying on measures examined exclusively earlier in life (Boon et al. 2007; Höjesjö et al. 2011), at least for traits that are neither repeatable nor stable in the long term (Bell & Sih 2007). It must be kept in mind that varying selection pressures may act on a given individual's personality over its lifetime. We think it could be misleading to predict individual fitness as a function of personality in a selective environment when personality itself can vary over an individual's lifetime. We thus encourage the study of the simultaneous influences of genetical, epigenetical, environmental or experiential causes on the expression of personality (Brown et al. 2007; Frost et al. 2007; Dingemanse et al. 2009; Stamps & Groothuis 2010; Webster & Ward 2011) to provide an account of the selective pressures that apply over an individual's lifetime. Despite potential flaws, we found that interindividual variation in personality did not decrease over time, suggesting that, in natural populations, a substantial amount of variation may permanently remain available for natural selection to act upon it.

Acknowledgements

We thank François-Xavier Dechaume-Moncharmont, Pierre-Olivier Montiglio, Shinichi Nakagawa, Emma Parfitt and Denis Réale for their advice and comments on previous drafts, Jonathan Wright and two anonymous referees for their help in improving the manuscript. We thank Sébastien Motreuil for his assistance during bird care and the building of experimental devices, and Elise Melloul for her assistance during

the experiment. This work was funded with a Ph.D. grant to M.D. and a master honorific grant to Y.A., both provided by the Conseil Régional de Bourgogne.

Literature Cited

- Anderson, D. R., Link, W. A., Johnson, D. H. & Burnham, K. P. 2001: Suggestions for presenting the results of data analyses. *J. Wildl. Manage.* **65**, 373–378.
- Archard, G. A. & Braithwaite, V. A. 2010: The importance of wild populations in studies of animal temperament. *J. Zool.* **281**, 149–160.
- Asendorpf, J. B. 1990: The measurement of individual consistency. *Methodika* **4**, 1–23.
- Bell, A. M. 2007: Future directions in behavioural syndromes research. *Proc. Roy. Soc. Lond. B* **274**, 755–761.
- Bell, A. M. & Sih, A. 2007: Exposure to predation generates personality in threespined sticklebacks (*Gasterosteus aculeatus*). *Ecol. Lett.* **10**, 828–834.
- Bell, A. M., Hankison, S. J. & Laskowski, K. L. 2009: The repeatability of behaviour: a meta-analysis. *Anim. Behav.* **77**, 771–783.
- Boon, A. K., Réale, D. & Boutin, S. 2007: The interaction between personality, offspring fitness and food abundance in North American red squirrels. *Ecol. Lett.* **10**, 1094–1104.
- Brown, C., Burgess, F. & Braithwaite, V. A. 2007: Heritable and experiential effects on boldness in a tropical poeciliid. *Behav. Ecol. Sociobiol.* **62**, 237–243.
- Burley, N. 1985: Leg-band color and mortality patterns in captive breeding populations of zebra finches. *Auk* **102**, 647–651.
- Butler, M. W., Toomey, M. B., McGraw, K. J. & Rowe, M. 2012: Ontogenetic immune challenges shape adult personality in mallard ducks. *Proc. Roy. Soc. Lond. B* **279**, 326–333.
- Carere, C., Drent, P. J., Privitera, L., Koolhaas, J. M. & Groothuis, T. G. G. 2005: Personalities in great tits, *Parus major*: stability and consistency. *Anim. Behav.* **70**, 795–805.
- Cassey, P., Ewen, J. G., Blackburn, T. M. & Moller, A. P. 2004: A survey of publication bias within evolutionary ecology. *Proc. Roy. Soc. Lond. B* **271**, S451–S454.
- David, M. & Cézilly, F. 2011: Personality may confound common measures of mate-choice. *PLoS ONE* **6**, e24778.
- David, M., Auclair, Y. & Cézilly, F. 2011: Personality predicts social dominance in female zebra finches, *Taeniopygia guttata*, in a feeding context. *Anim. Behav.* **81**, 219–224.
- David, M., Auclair, Y., Giraldeau, L.-A. & Cézilly, F. 2012: Personality and body condition have additive effects on motivation to feed in Zebra finches *Taeniopygia guttata*. *Ibis* **154**, 372–378.
- D'Eath, R. B. 2004: Consistency of aggressive temperament in domestic pigs: the effects of social experience and social disruption. *Aggress. Behav.* **30**, 435–448.
- Dingemanse, N. J. & Réale, D. 2005: Natural selection and animal personality. *Behaviour* **142**, 1159–1184.
- Dingemanse, N. J., Both, C., Drent, P. J., Van Oers, K. & Van Noordwijk, A. J. 2002: Repeatability and heritability of exploratory behaviour in great tits from the wild. *Anim. Behav.* **64**, 929–938.
- Dingemanse, N. J., Wright, J., Kazem, A. J. N., Thomas, D. K., Hickling, R. & Dawnay, N. 2007: Behavioural syndromes differ predictably between 12 populations of three-spined stickleback. *J. Anim. Ecol.* **76**, 1128–1138.
- Dingemanse, N. J., Van der Plas, F., Wright, J., Réale, D., Schrama, M., Roff, D. A., Van der Zee, E. & Barber, I. 2009: Individual experience and evolutionary history of predation affect expression of heritable variation in fish personality and morphology. *Proc. Roy. Soc. Lond. B* **276**, 1285–1293.
- Dingemanse, N. J., Kazem, A. J. N., Réale, D. & Wright, J. 2010: Behavioural reaction norms: animal personality meets individual plasticity. *Trends Ecol. Evol.* **25**, 81–89.
- van Dongen, W. F. D., Maldonado, K., Sabat, P. & Vasquez, R. A. 2010: Geographic variation in the repeatability of a personality trait. *Behav. Ecol.* **21**, 1243–1250.
- Falconer, D. S. & Mackay, T. F. 1996: Introduction to Quantitative Genetics, 4th edn. Longman, Harlow.
- Frost, A. J., Winrow-Giffen, A., Ashley, P. J. & Sneddon, L. U. 2007: Plasticity in animal personality traits: does prior experience alter the degree of boldness. *Proc. Roy. Soc. Lond. B* **274**, 333–339.
- Gabriel, P. O. & Black, J. M. 2010: Behavioural syndromes in Steller's jays: the role of time frames in the assessment of behavioural traits. *Anim. Behav.* **80**, 689–697.
- Gosling, S. D. 1998: Personality dimensions in spotted hyenas (*Crocuta crocuta*). *J. Comp. Psychol.* **112**, 107–118.
- Hedrick, A. V. & Kortet, R. 2012: Sex differences in the repeatability of boldness over metamorphosis. *Behav. Ecol. Sociobiol.* **66**, 407–412.
- Herborn, K. A., Macleod, R., Miles, W. T. S., Schofield, A. N. B., Alexander, L. & Arnold, K. E. 2010: Personality in captivity reflects personality in the wild. *Anim. Behav.* **79**, 835–843.
- Herczeg, G. & Garamszegi, L. Z. 2012: Individual deviation from behavioural correlations: a simple approach to study the evolution of behavioural syndromes. *Behav. Ecol. Sociobiol.* **66**, 161–169.
- Höjesjö, J., Adriaenssens, B., Bohlin, T., Jönsson, C., Hellström, I. & Johnsson, J. I. 2011: Behavioural syndromes in juvenile brown trout (*Salmo trutta*); life history, family variation and performance in the wild. *Behav. Ecol. Sociobiol.* **65**, 1801–1810.

- Jenkins, S. H. 2011: Sex differences in repeatability of food-hoarding behaviour of kangaroo rats. *Anim. Behav.* **81**, 1155–1162.
- Jones, K. A. & Godin, J. G. J. 2010: Are fast explorers slow reactors? Linking personality type and anti-predator behaviour. *Proc. Roy. Soc. Lond. B* **277**, 625–632.
- Kanda, L. L., Louon, L. & Straley, K. 2012: Stability in activity and boldness across time and context in captive siberian dwarf hamsters. *Ethology* **118**, 518–533.
- Koski, S. E. 2011: Social personality traits in chimpanzees: temporal stability and structure of behaviourally assessed personality traits in three captive populations. *Behav. Ecol. Sociobiol.* **65**, 2161–2174.
- Mafli, A., Wakamatsu, K. & Roulin, A. 2011: Melanin-based coloration predicts aggressiveness and boldness in captive eastern Hermann's tortoises. *Anim. Behav.* **81**, 859–863.
- Martin, J. G. A., Nussey, D. H., Wilson, A. J. & Réale, D. 2011: Measuring individual differences in reaction norms in field and experimental studies: a power analysis of random regression models. *Meth. Ecol. Evol.* **2**, 362–374.
- McCrae, R. R., Costa, P. T., del Pilar, G. H., Rolland, J. P. & Parker, W. D. 1998: Cross-cultural assessment of the five-factor model – The revised NEO personality inventory. *J. Cross Cult. Psychol.* **29**, 171–188.
- McPhee, E. 2003: Generations in captivity increases behavioral variance: considerations for captive breeding and reintroduction programs. *Biol. Cons.* **115**, 71–77.
- Meehan, C. L. & Mench, J. A. 2002: Environmental enrichment affects the fear and exploratory responses to novelty of young Amazon parrots. *Appl. Anim. Behav. Sci.* **79**, 75–88.
- Montiglio, P. O., Garant, D., Thomas, D. & Réale, D. 2010: Individual variation in temporal activity patterns in open-field tests. *Anim. Behav.* **80**, 905–912.
- Nakagawa, S. & Cuthill, I. C. 2007: Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev.* **82**, 591–605.
- Nakagawa, S. & Schielzeth, H. 2010: Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol. Rev.* **85**, 935–956.
- Nicolakakis, N., Sol, D. & Lefebvre, L. 2003: Behavioural flexibility predicts species richness in birds, but not extinction risk. *Anim. Behav.* **65**, 445–452.
- O'Hara, R. B. & Kotze, D. J. 2010: Do not log-transform count data. *Meth. Ecol. Evol.* **1**, 118–122.
- Oldham, P. D. 1962: A note on analysis of repeated measurements of same subjects. *J. Chron. Dis.* **15**, 969–977.
- Quinn, J. L. & Cresswell, W. 2005: Personality, anti-predation behaviour and behavioural plasticity in the chaffinch *Fringilla coelebs*. *Behaviour* **142**, 1377–1402.
- R Development Core Team. 2011: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.
- Ray, J. & Hansen, S. 2005: Temperamental development in the rat: the first year. *Dev. Psychobiol.* **47**, 136–144.
- Réale, D., Reader, S. M., Sol, D., McDougall, P. T. & Dingemanse, N. J. 2007: Integrating animal temperament within ecology and evolution. *Biol. Rev.* **82**, 291–318.
- Schielzeth, H. & Forstmeier, W. 2009: Conclusion beyond support: overconfident estimates in mixed models. *Behav. Ecol.* **20**, 416–420.
- Schielzeth, H., Bolund, E., Kempnaers, B. & Forstmeier, W. 2011: Quantitative genetics and fitness consequences of neophilia in zebra finches. *Behav. Ecol.* **22**, 126–134.
- Schuett, W. & Dall, S. R. X. 2009: Sex differences, social context and personality in zebra finches, *Taeniopygia guttata*. *Anim. Behav.* **77**, 1041–1050.
- Schuett, W., Tregenza, T. & Dall, S. R. X. 2010: Sexual selection and animal personality. *Biol. Rev.* **85**, 217–246.
- Sinn, D. L. & Moltchanivskyj, N. A. 2005: Personality traits in dumpling squid (*Euprymna tasmanica*): context-specific traits and their correlation with biological characteristics. *J. Comp. Psychol.* **119**, 99–110.
- Sinn, D. L., Gosling, S. D. & Moltchanivskyj, N. A. 2008: Development of shy/bold behaviour in squid: context-specific phenotypes associated with developmental plasticity. *Anim. Behav.* **75**, 433–442.
- Sinn, D. L., Moltchanivskyj, N. A., Wapstra, E. & Dall, S. R. X. 2010: Are behavioral syndromes invariant? Spatiotemporal variation in shy/bold behavior in squid. *Behav. Ecol. Sociobiol.* **64**, 693–702.
- Sol, D., Timmermans, S. & Lefebvre, L. 2002: Behavioural flexibility and invasion success in birds. *Anim. Behav.* **63**, 495–502.
- Stamps, J. & Groothuis, T. G. G. 2010: The development of animal personality: relevance, concepts and perspectives. *Biol. Rev.* **85**, 301–325.
- Stamps, J. A., Briffa, M. & Biro, P. A. 2012: Unpredictable animals: individual differences in intraindividual variability (IIV). *Anim. Behav.* **83**, 1325–1334.
- Svartberg, K., Tapper, I., Temrin, H., Radesater, T. & Thorman, S. 2005: Consistency of personality traits in dogs. *Anim. Behav.* **69**, 283–291.
- Tu, Y. K. & Gilthorpe, M. S. 2007: Revisiting the relation between change and initial value: a review and evaluation. *Statist. Med.* **26**, 443–457.
- Uher, J. 2008: Comparative personality research: methodological approaches. *Eur. J. Pers.* **22**, 427–455.
- Ward, A. J. W., Thomas, P., Hart, P. J. B. & Krause, J. 2004: Correlates of boldness in three-spined sticklebacks (*Gasterosteus aculeatus*). *Behav. Ecol. Sociobiol.* **55**, 561–568.

- Webster, M. M. & Ward, A. J. W. 2011: Personality and social context. *Biol. Rev.* **86**, 759–773.
- Weigensberg, I. & Roff, D. A. 1996: Natural heritabilities: can they be reliably estimated in the laboratory? *Evolution* **50**, 2149–2157.
- Weiss, A., Inoue-Murayama, M., King, J. E., Adams, M. J. & Matsuzawa, T. 2012: All too human? Chimpanzee and orang-utan personalities are not anthropomorphic projections. *Anim. Behav.* **83**, 1355–1365.
- Zann, R. A. 1996: *The Zebra Finch: A Synthesis of Field and Laboratory Studies*. Oxford Univ. Press, Oxford.
- Zuur, A. F., Ieno, E. N. & Elphick, C. S. 2010: A protocol for data exploration to avoid common statistical problems. *Meth. Ecol. Evol.* **1**, 3–14.