Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*

Niels J. Dingemanse[1,2,3]*, Karen M. Bouwman[2], Martijn van de Pol[4], Thijs van Overveld[5], Samantha C. Patrick[6], Erik Matthysen[5] and John L. Quinn[6]

[1]*Department of Behavioural Ecology & Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany;* [2]*Animal Ecology Group, Centre for Ecology and Evolution, University of Groningen, The Netherlands;* [3]*Department of Animal Ecology, Netherlands Institute of Ecology, Wageningen, The Netherlands;* [4]*Ecology, Evolution & Genetics, Research School of Biology, Australian National University, Canberra, Australia;* [5]*Evolutionary Ecology Group, Department of Biology, University of Antwerp, Wilrijk, Belgium; and* [6]*Edward Grey Institute, Department of Zoology, University of Oxford, Oxford, UK*

Summary

1. Interest in the evolutionary origin and maintenance of individual behavioural variation and behavioural plasticity has increased in recent years.

2. Consistent individual behavioural differences imply limited behavioural plasticity, but the proximate causes and wider consequences of this potential constraint remain poorly understood. To date, few attempts have been made to explore whether individual variation in behavioural plasticity exists, either within or between populations.

3. We assayed 'exploration behaviour' among wild-caught individual great tits *Parus major* when exposed to a novel environment room in four populations across Europe. We quantified levels of individual variation within and between populations in average behaviour, and in behavioural plasticity with respect to (i) repeated exposure to the room (test sequence), (ii) the time of year in which the assays were conducted and (iii) the interval between successive tests, all of which indicate habituation to novelty and are therefore of functional significance.

4. Consistent individual differences ('I') in behaviour were present in all populations; repeatability (range: 0·34–0·42) did not vary between populations. Exploration behaviour was also plastic, increasing with test sequence – but less so when the interval between subsequent tests was relatively large – and time of year; populations differed in the magnitude of plasticity with respect to time of year and test interval. Finally, the between-individual variance in exploration behaviour increased significantly from first to repeat tests in all populations. Individuals with high initial scores showed greater increases in exploration score than individuals with low initial scores; individual by environment interaction ('I × E') with respect to test sequence did not vary between populations.

5. Our findings imply that individual variation in both average level of behaviour and behavioural plasticity may generally characterize wild great tit populations and may largely be shaped by mechanisms acting within populations. Experimental approaches are now needed to confirm that individual differences in behavioural plasticity (habituation) – not other hidden biological factors – caused the observed patterns of I × E. Establishing the evolutionary causes and consequences of this variation in habituation to novelty constitutes an exciting future challenge.

Key-words: behavioural reaction norm, habituation, individual by environment interaction, local adaptation, personality, temperament

Introduction

There has been increasing interest in consistent individual differences in single behaviours, or suites of correlated behav-

iours, across time or contexts. This variation has been referred to as 'animal personality' (Gosling 2001; Réale *et al.* 2007) or 'behavioural syndromes' (Sih, Bell & Johnson 2004; Sih *et al.* 2004). One form of variation in personality exists when rank order differences between individuals in their behaviour are maintained over time or contexts (Sih *et al.*

*Correspondence author. E-mail: ndingemanse@orn.mpg.de

2004; Réale *et al.* 2007), implying limits to behavioural plasticity (Sih *et al.* 2004). From an adaptive perspective, limited plasticity is unexpected (DeWitt, Sih & Wilson 1998; Dall, Houston & McNamara 2004) because heterogeneous environments should favour the evolution of behavioural plasticity rather than behavioural consistency (Via & Lande 1985; de Jong 1995; Via *et al.* 1995). Theoreticians have therefore begun to address variation in personality within an evolutionary framework and started to develop adaptive explanations for animal personality (reviewed by Dingemanse & Wolf 2010; Wolf & Weissing 2010).

Animal personality studies typically document repeatability of behaviours within populations (Réale *et al.* 2007; Bell, Hankison & Laskowski 2009), a measure providing insight into whether individuals –on average– differ in behavioural profile. Such an analysis, however, does not provide information on the level of individual plasticity within this population (Dingemanse *et al.* 2010; Réale & Dingemanse 2010). Individual variation in plasticity is important because it implies that the potential amount of additive genetic variance that selection can act upon is not stable across environments, which would affect predictions of evolutionary change in response to selection (Roff 1997; see Brommer, Rattiste & Wilson 2008 for a worked example). Furthermore, recent research has demonstrated that individual plasticity itself can have a genetic basis and is under selection (Brommer *et al.* 2005; Nussey *et al.* 2005; Nussey, Wilson & Brommer 2007). Despite its importance, evolutionary ecologists have only recently begun to quantify individual variation in plasticity using a reaction norm approach (Nussey, Wilson & Brommer 2007), a widely adopted framework in quantitative genetics (Via *et al.* 1995). This approach considers the phenotype of an individual expressed in two (or more) environments as a line, described by an elevation ('intercept' in statistical terms; the individual's level of behaviour in the average environment) and a slope (the individual's plasticity over an environmental gradient), and specifically aims to estimate individual variation in reaction norm elevation (individual variation or 'I') and slope (individual × environment interaction or 'I × E').

Few studies have quantified variation in plasticity in wild populations, even at the level of the phenotype (reviewed in Nussey, Wilson & Brommer 2007), and those that have generally focussed on the response of life history or morphological traits to changing environmental conditions (e.g., Nussey *et al.* 2005; Brommer, Rattiste & Wilson 2008). Few attempts have been made to quantify levels of variation in individual plasticity in the context of behaviour specifically (Dingemanse *et al.* 2010). Examples of contexts in which such attempts have been made include links between provisioning rate and offspring begging intensity (Smiseth, Wright & Kölliker 2008), between dispersal and wind velocity (Bonte, Bossuyt & Lens 2007), or between crypsis and predation risk (Quinn & Cresswell 2005).

Here, we investigate individual variation in phenotypic plasticity (i.e. I × E) of a behavioural trait often used in personality research (exploration behaviour) based on large numbers (1007) of wild individual great tits *Parus major* sampled from four populations across Western Europe. Previous work has shown that repeatable and heritable variation exists in exploration behaviour, both in wild populations (Dingemanse *et al.* 2002; Quinn *et al.* 2009) and under controlled laboratory conditions (Drent, van Oers & van Noordwijk 2003). Exploration behaviour is also plastic, as it increases within individuals from autumn to spring (Dingemanse *et al.* 2002; Quinn *et al.* 2009) and, independently, with repeated exposure to the test procedure (Dingemanse *et al.* 2002) because of habituation to repeated testing or acclimation to repeated capture and handling (Romero 2004). Although we have yet to reveal the functional significance of the time of year effects, habituation to repeated tests was expected because responses to novelty normally decrease with exposure (Réale *et al.* 2007; Martin & Réale 2008). Moreover, while the experimental 'novel' room used in these studies is the same physical environment across repeated measures, habituation can be viewed as a specific type of behavioural plasticity because the individual's perception of the room's novelty and risk changes over time, representing an environmental axis of variation in the broad sense (Réale *et al.* 2007; Martin & Réale 2008).

The existence of individual variation in plasticity in response to experience with challenging stimuli (e.g. experience with our test procedure), or time of year, is important because individuals are continuously exposed, and have to habituate, to challenging environmental conditions as part of everyday activities (e.g. during foraging, or encounters with novel conspecifics). The presence of individual variation in plasticity would thus provide insight into the potential for the heritability of exploration behaviour to be a function of these environmental gradients (Nussey, Wilson & Brommer 2007) and imply that the evolutionary consequences of selection on exploratory behaviour would depend on novelty, and experience with, challenging stimuli, or time of year in which selection might be acting.

Study replication is generally viewed as an essential part of modern biology (Kelly 2006). For example, the comparison of patterns of behavioural variation across different populations has proven useful because it can provide insight into whether behaviour has been subject to locally varying processes (Bell 2005; Dingemanse *et al.* 2007). Here, we investigated phenotypic plasticity of exploration behaviour measured among four West-European great tit populations, which have previously been shown to differ both in average level, and in the association between a gene polymorphism and exploration (Korsten *et al.* 2010). These findings suggest that local circumstances, for example selective regime, differ between these populations but whether they also differ with respect to phenotypic plasticity of exploration behaviour is unknown. We partition here the phenotypic variation in behaviour into individual components (I; random effect), population-average level of plasticity (E; fixed effect) and individual variation in plasticity (I × E; random effect) using procedures advocated by Nussey, Wilson & Brommer (2007)

and Dingemanse *et al.* (2010). First, we calculated average levels of plasticity in response to (i) experience with the test procedure, (ii) interval between tests and (iii) time of the year. Second, we estimated standardized indices (repeatability) of individual variation in exploratory behaviour. Third, we assessed whether individual variation in plasticity existed with regard to experience with the test procedure (habituation) and time of the year. Fourth, we compared patterns of individual variation across populations.

Materials and methods

STUDY SITES

Exploration behaviour was measured for wild-caught great tits from four populations: Boshoek (BH; Belgium), Lauwersmeer (LM; The Netherlands), Westerheide (WH; The Netherlands) and Wytham Woods (WW; Great Britain). Detailed descriptions of the study sites are given elsewhere (BH: Matthysen 2002; LM: Nicolaus *et al.* 2009; WH: Dingemanse *et al.* 2002; WW: Elton 1966). In short, BH consists of 17 fragmented woodlands within a 1000-ha area, intersected by roads, residential and agricultural areas (51°08′N, 04°32′E). All forest patches were provided with *c.* nine nest boxes per ha from 1993 onwards. LM consists of five woodlots surrounded by grassland areas within a 1700-ha area (53°20′N, 06°12′E). Twelve nest box plots (between 9 and 12 ha) with 50 nest boxes each have been created in the area since 2005. WH consists of mixed pine and deciduous forest of about 250 ha (52°00′N, 05°50′E). Approximately 600 nest boxes have been provided from 1995 onwards. Finally, WW consists of a single *c.* 352-ha area fitted with over 1000 nest boxes situated in mixed woodland dominated by oak trees (51°47′N, 1°20′W).

CATCHING METHODS

Similar capture methodology was used in all study areas (see details in Dingemanse *et al.* 2002). Outside the breeding season, individuals were caught using mist nets at feeding stations (baited with sunflower seeds or peanuts), or when roosting in nest boxes during winter. At capture, all birds were weighed and transported to the laboratory within 1·5 h. Re-captured birds were taken to the laboratory for a repeat test. Inter-test interval was on average about half a year: mean interval in days (range): BH: 198 (5–740); LM: 288 (1–889); WH: 130 (1–832); WW: 132 (2–763).

EXPLORATION TESTS

Each study area had its own laboratory located near the study site. All 'novel environment rooms' were constructed and furbished following Dingemanse *et al.* (2002). Subtle differences nevertheless existed between the laboratories: home cage size differed between study areas (L × W × H: BH: 0·8 × 0·4 × 0·5 m; LM: 0·5 × 0·4 × 0·5 m; WH: 0·9 × 0·4 × 0·5 m; WW: 0·7 × 0·5 × 0·5 m), as well as the number of cages connected to an observation room (BH: 16; LM: 48; WH 16; WW: 16), in part dictated by the laboratory's ethical committees. Consequently, the number of perches in the novel room differed between the laboratories, because some birds readily use the edges of sliding doors that connect each cage to the room as perches (Dingemanse *et al.* 2002).

In the laboratory, birds were housed individually under natural daylight regime in cages with a solid bottom and top. Birds had *ad libitum* access to food and water, and human disturbance was kept to a minimum. The morning following capture, exploration behaviour was measured of each bird alone placed in a sealed room (BH, LM, WH: 4·0 × 2·4 × 2·3 m; WW: 4·0 × 3·3 × 2·5 m) containing five artificial trees, following procedures detailed by Dingemanse *et al.* (2002). In BH, LM and WH, exploration score was calculated as the total number of flights and hops within the first 2 min after arrival. Flights were defined as movements that were made between trees, walls or other perches using flight. Hops were defined as movements that were made without using flight either between different branches on the same tree or over a distance of 30 cm (e.g. when hopping on the floor or across a ledge on the wall). Exploration scores of WW also included hops within a branch and within a distance of 30 cm. These additional hops could not be excluded from the exploration score retrospectively because of the methodology of data collection (see also Quinn *et al.* 2009).

All birds were released near their site of capture directly after the last bird had been tested (always within 24 h after capture). The age class of birds not ringed as nestlings was determined by the colour of their greater wing coverts, allowing distinction between juveniles and adults (Jenni & Winkler 1994).

DATA SELECTION

We selected data of birds that had undergone two (*n* = 719 birds), three (*n* = 222 birds) or four (*n* = 66 birds) exploration tests. To increase comparability between data sets, only records from the months July through March were included in the analyses, because data from other months were available for WH only. This meant including data collected from January 2006 (BH), September 2005 (LM) or February 2005 (WW; the onset of these studies) through February 2009 (BH), March 2009 (LM) or March 2007 (WW).The WH data set includes a selection of the data published by Dingemanse *et al.* (2002; July 1998 through March 2001) supplemented with additional data collected in 2002. Tests of 1007 birds were included (total number of tests: 2368), with individuals in each study site tested on average 2·42 (BH; 2×: 154, 3×: 46, 4×: 24), 2·37 (LM; 2×: 244, 3×: 83, 4×: 23), 2·37 (WH; 2×: 153, 3×: 62, 4×: 11) and 2·23 (WW; 2×: 168, 3×: 31, 4×: 8) times.

STATISTICAL ANALYSES

We used General Linear Mixed Models (GLMM) with a normal error distribution to assess sources of variation in exploration score. The dependent variable exploration score, and continuous explanatory variables, were centred around their grand mean. Individual was entered as a random effect (i.e. random intercepts for individual). Three variables, previously shown to explain within-individual variation in exploratory behaviour (Dingemanse *et al.* 2002), were included as fixed effects. Test number ('sequence') was included as a categorical factor to investigate the change in exploration behaviour in a completely novel environment (first test; coded as 0) vs. a more familiar environment (second, third and fourth test; all coded as 1). Preliminary analyses showed that sequence effects were adequately modelled in this way, because exploration behaviour typically changed most from first to second tests and changed less from second to further tests (Fig. 1a–d). 'Interval' in days between two consecutive tests was log-transformed (interval + 1) and included as a linear covariate, because sequence effects decay over time (Dingemanse *et al.* 2002). Furthermore, we entered log-transformed 'time of year' (here defined as days from July 1st, covariate) as a third fixed effect. Exploration scores have been shown to change linearly from autumn to spring (see Fig. 2 in Dingemanse *et al.* 2002). However,
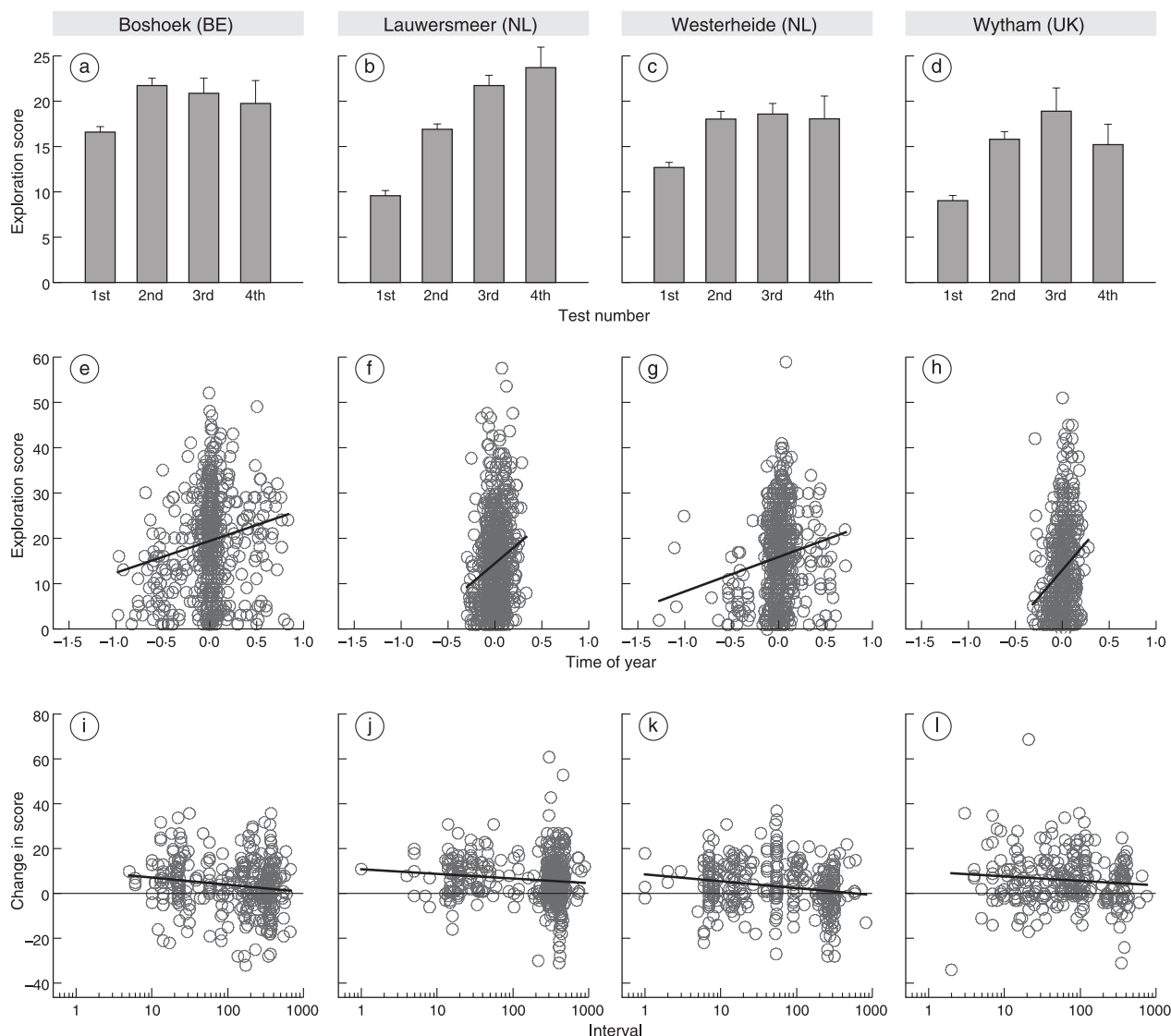
Fig. 1. Exploration score as a function of test sequence (a–d), time of year (e–h) and change in score (the score of a focal test minus the score of the previous one) as a function of interval between subsequent tests (i–l). Panels a–d show that exploration scores (± SE) increased from first (1st) to repeat (2nd–4th) tests within individuals for all populations. Panels e–h show that exploration scores changed with Julian date ('time of year') for all populations. The *x*-axis gives values for log-transformed time of year (Julian date) that are expressed in deviations from the individual's mean (i.e. 'centred' values, see Materials and methods). Panels e–h therefore capture the within-individual relationship between behaviour and time of year. For example, an individual with three observations with the values 1·959, 2·369 and 2·230 for log-transformed time of year would have an average value of 2·186 and plotted centred values of −0·227, 0·183 and 0·044, respectively. Panels i–l show that the increase in exploration score between subsequent tests was generally more pronounced for shorter intervals (note that interval is plotted on a log-axis; data points above the thin, reference line represent individuals that increased their exploration score from one test to the next; the solid line fits a linear regression through the raw data shown here). Statistical analyses for all data are shown in Table 1. Samples sizes (grand total number of tests) were 542 (Boshoek), 829 (Lauwersmeer), 536 (Westerheide) and 461 (Wytham Woods).

preliminary exploration of our data set showed that the model fit was better when using the log-transformation of time of year. In agreement with Dingemanse *et al.* (2002), we found that other fixed effects that could potentially bias the data (age class, body mass, catching method, time of day) did not significantly affect exploration score (Results not shown).

Relationships between exploration score and time of year or interval can exist both at the between-individual level (e.g. if certain types of bird are more likely to be caught at certain times of the year) and the within-individual level (e.g. if individuals change their behaviour over the year, i.e. individual plasticity). We used within-individual centring to separate the within-individual effects from the between-

individual effects (e.g., van de Pol & Verhulst 2006; van de Pol & Wright 2009), a methodology recently advocated for the analysis of reaction norms based on observational data (see Box 2 in Dingemanse *et al.* 2010). For each individual, we thus calculated, first, the mean value for interval and time of year, and second, for all observations of an individual, the deviation from this mean value. Because we were interested in modelling within-individual effects here, we only included the within-individual effects (deviations) in our models (Tables 1 and 2).

Individual repeatability (*r*; the *intraclass correlation coefficient*) of exploration score was calculated as the ratio of the between-individual variance over the sum of the between- and within-individual

variance (Rasbash *et al.* 2005), with standard errors and confidence intervals calculated following Fisher (1925). Repeatability is best interpreted when the timing of the measurements is standardized between individuals (Falconer & Mackay 1996), and we therefore calculated *r* from models where sequence, interval and time of year were included as fixed effects, also called 'adjusted repeatability' (Nakagawa & Schielzeth 2010). We statistically compared adjusted repeatability across study sites by fitting a GLMM that included data from all four study sites and estimated individual and residual variance components for each study site simultaneously, which we compared with a reduced model where individual variances of the different study sites were constrained to the same value. For this comparison, we standardized values across our study sites by rescaling the total variance within study sites not accounted for by fixed effects to one, thereby ensuring that we tested for differences in repeatability rather than individual variance across populations.

Subsequently, we analysed whether there was individual by environment interaction (I × E) in the effect of sequence, interval and

Table 1. Sources of within-individual variation in exploration behaviour based on repeated exploration tests of wild-caught great tits in four West-European populations. We used GLMMs with random intercepts fitted for individual, and test sequence (1st vs. repeat tests), interval between tests (in days; log-transformed) and time of year (days from July 1st; log-transformed) fitted as fixed effects. For interval and time of year, we fitted for all observations of an individual the deviation from its mean value (see Materials and methods). We further give estimates of between- and within-individual variances and adjusted repeatability. Values are reported with 95% credible intervals (CI), and ΔDIC values refer to the change in DIC when the specific parameter was included vs. excluded

Fixed effects	Boshoek (Belgium) β (95%CI)	ΔDIC	Lauwersmeer (the Netherlands) β (95%CI)	ΔDIC	Westerheide (the Netherlands) β (95%CI)	ΔDIC	Wytham Woods (United Kingdom) β (95%CI)	ΔDIC
Sequence	8·21 (4·09, 12·33)	−22·9	9·75 (6·54, 13·16)	−31·2	9·44 (6·41, 12·42)	−45·6	7·1 (2·72, 11·48)	−11·3
Interval	−1·91 (−3·84, −0·02)	−2·9	−0·74 (−2·18, 0·65)	1·3	−2·77 (−4·37, −1·24)	−13·7	−0·59 (−2·76, 1·61)	1·8
Time of year	4·72 (2·18, 7·27)	−15·7	10·81 (6·00, 15·81)	−25·7	4·87 (1·86, 7·91)	−12·0	10·34 (2·23, 18·7)	−6·6
Intercept	−0·98 (−3·52, 1·69)		−6·78 (−8·84, −4·75)		−5·42 (−7·37, −3·42)		−6·65 (−9·3, −4·11)	
Variances	σ^2 (95%CI)	ΔDIC	σ^2 (95%CI)	ΔDIC	σ^2 (95%CI)	ΔDIC	σ^2 (95%CI)	ΔDIC
Between-individual	46·8 (33·5, 62·6)	−153·1	33·0 (25·0, 41·8)	−217·2	33·1 (23·2, 44·8)	−30·0	30·1 (18·6, 43·3)	−71·5
Within-individual	65·9 (56·4, 76·9)		49·0 (43·1, 55·4)		53·4 (45·5, 62·5)		59·6 (49·9, 70·9)	
Repeatability	*r* (95%CI)		*r* (95%CI)		*r* (95%CI)		*r* (95%CI)	
	0·42 (0·37, 0·46)		0·40 (0·37, 0·44)		0·38 (0·34, 0·43)		0·34 (0·30, 0·39)	

DIC, deviance information criterion; GLMM, general linear mixed model.

Table 2. Between-individual variation in effects of test sequence (first vs. repeat tests), log-transformed interval between tests (days) and log-transformed time of year (days from July 1st) on exploration score of great tits for four populations. Results are from a GLMM with individual fitted as random effect and sequence, interval and time of year fitted as fixed effects. For interval and time of year, we fitted for all observations of an individual the deviation from its mean value (see Materials and methods). For each fixed effect, we report the between-individual variance in elevation ($\sigma^2_{elevation}$) and slope (σ^2_{slope}), and their covariance ($\sigma_{elevation,slope}$) for a model where the magnitude of the fixed effect was allowed to vary randomly between individuals (random slopes model). ΔDIC values of the variance in slope refer to the change in DIC when both the slope-variance and elevation-slope covariance term were included into the model compared with a model that only contained variance in elevation. Note that 95% credible intervals (CI) of variances can only take positive values

(Co)variances	Boshoek Estimate (95%CI)	ΔDIC	Lauwersmeer Estimate (95%CI)	ΔDIC	Westerheide Estimate (95%CI)	ΔDIC	Wytham Woods Estimate (95%CI)	ΔDIC
Individual × Sequence								
$\sigma^2_{elevation}$	26·7 (14·7, 41·8)		14·7 (8·7, 21·8)		16·2 (7·8, 27·0)		11·5 (4·8, 20·1)	
σ^2_{slope}	11·1 (2·9, 23·7)	−24·2	15·3 (8·0, 24·6)	−85·3	12·1 (3·8, 24·4)	−30·0	26·0 (12·0, 44·5)	−71·5
$\sigma_{elevation,slope}$	14·8 (7·9, 21·9)	−9·9	13·6 (10·0, 17·3)	−82·6	11·9 (6·9, 16·8)	−37·6	18·0 (9·5, 20·7)	−88·3
Individual × Time of year								
$\sigma^2_{elevation}$	47·2 (34·0, 63·4)		33·3 (25·3, 42·2)		33·6 (23·5, 45·3)		30·8 (19·7, 43·6)	
σ^2_{slope}	0·01 (0·00, 0·06)	0·9	0·05 (0·00, 0·06)	1·8	0·01 (0·00, 0·05)	1·2	0·09 (0·01, 0·16)	0·6
$\sigma_{elevation,slope}$	0·04 (−0·57, 0·78)	0·3	0·08 (−0·50, 0·61)	1·3	0·02 (−0·49, 0·69)	1·5	0·04 (−0·49, 0·60)	0·3
Individual × Interval								
$\sigma^2_{elevation}$	47·6 (34·1, 63·5)		31·9 (24·2, 40·9)		33·8 (23·6, 45·5)		30·4 (19·6, 43·2)	
σ^2_{slope}	0·05 (0·00, 0·50)	1·8	1·15 (0·05, 3·23)	0·0	0·07 (0·00, 0·27)	0·9	0·12 (0·01, 0·44)	1·5
$\sigma_{elevation,slope}$	0·30 (−0·41, 3·50)	0·1	1·86 (0·21, 8·25)	0·7	0·27 (−0·34, 3·47)	1·4	0·08 (−0·48, 0·58)	1·7

DIC, deviance information criterion; GLMM, general linear mixed model.

time of year by including random coefficients at the individual level for these fixed effects (so-called 'random regression' or 'random slopes' models; Data S1, Supporting Information). These models thus estimated the amount of between-individual variance in the slope of the (within-individual) relationship between exploration score and an explanatory variable (either sequence, interval or time of the year; for detailed explanation of the approach, see van de Pol & Wright 2009). Random slope models describe the pattern of between-individual variation over an environmental gradient by estimating three random parameters: (i) the variance in elevation, (ii) the variance in slope and (iii) the covariance between elevation and slope. Evidence for the presence of random slopes (I × E) was assessed by comparing a model that included all three above parameters against a model that included only variance in elevation. All random intercepts and slopes were modelled as normally distributed random variables with zero mean and variance (σ^2).

We used a two-step analytical approach. First, we present models for each study site separately. Second, we present a joint model, where we test for differences between the four study sites in the above-mentioned fixed and random effects by comparing models in which these effects were allowed to differ between populations or not (see Data S1, Supporting Information for details).

Model parameters were estimated using Markov Chain Monte Carlo (MCMC) methods within a Bayesian framework (see Data S1, Supporting Information for details). We compared model support using the Deviance Information Criterion (DIC). The DIC is a hierarchical modelling generalization of the Akaike and Bayesian information criterions and is used in Bayesian model selection analyses whenever posterior distributions of the parameters have been derived from MCMC simulation (Spiegelhalter *et al.* 2002). The model with the lowest DIC value is considered best supported. Support for the presence of relationships between exploration behaviour and other fixed effects (Table 1), or for presence of between-individual variation (Tables 1 and 2), was based upon comparisons of DIC values between models where the effect of interest was included vs. excluded. Throughout the Results section, we considered values of ΔDIC ($DIC_{\text{effect included}}$ minus $DIC_{\text{effect excluded}}$) below −2 as 'support for the presence', and values of ΔDIC above + 2 as 'support for the absence', of a focal (random or fixed) effect.

General Linear Mixed Models were performed with MLwiN 2.02 and its winBUGS interface (Rasbash *et al.* 2005). Unless stated otherwise, parameter estimates are reported with 95% credible intervals, which reflect the degree of belief that the parameter has a 95% probability of being within this interval.

ETHICAL NOTE

Possible adverse effects of the novel environment test have been studied elsewhere but were not detected (Dingemanse *et al.* 2002; Hollander *et al.* 2008). Formal permission for short-term housing (maximum of 24 h) and personality testing was granted by an ethical committee for each study site, or in the case of WW, under license granted by Natural England.

Results

POPULATION-AVERAGE BEHAVIOURAL PLASTICITY

For all four populations, we found strong support for plasticity in exploration behaviour with respect to test sequence (ΔDIC ranged between −45·6 and −11·3; Table 1): on aver-

age individuals explored the room faster during repeat compared to initial tests (Fig. 1a–d). There was strong support for lack of variation between populations in the magnitude of behavioural change with sequence (study site × sequence: ΔDIC = + 5·3).

For all four populations, we found further strong support for plasticity in exploration score with respect to time of year (ΔDIC ranged between −25·7 and −6·6; Table 1): exploration scores increased with Julian date in all populations (Fig. 1e–h). We also found support for between-population differences in the magnitude of population-average plasticity (interaction study site × time of year: ΔDIC = −2·6). The magnitude of the time of year effect differed most strongly between BH and WH vs. LM and WW.

There was also evidence for plasticity with respect to interval because longer test intervals reduced exploration scores in repeat tests within the average individual for population BH (ΔDIC = −2·9; Table 1; Fig. 1i) and WH (ΔDIC = −13·7; Table 1; Fig. 1k). Analysis for the other two populations resulted in slightly better support for the absence as opposed to the presence of interval effects (LM: ΔDIC = + 1·3; WW: ΔDIC = + 1·8; Table 1). Consequently, we found support for between-population differences in the magnitude of population-average plasticity (interaction study site × interval: ΔDIC = −2·4; Fig. 1i–l).

VARIATION IN REPEATABILITY BETWEEN POPULATIONS

We found strong support for the presence of consistent individual variation in exploration behaviour within each population (ΔDIC ranged between −217·2 and −30·0 for the random effect individual; Table 1), with adjusted repeatabilities ranging from 0·34 to 0·42 (Table 1). We also found strong support for the absence of population differences in adjusted repeatability (ΔDIC = + 4·2).

INDIVIDUAL VARIATION IN PLASTICITY

For all four study sites, there was strong model support for the presence of between-individual variation in the change in exploration score from the first to repeat tests within individuals (ΔDIC ranged between −85·3 and −24·2 for individual × sequence (σ^2_{slope}); Table 2), implying that individuals differed systematically in how quickly they habituated to the novel environment and/or other aspects of the experimental procedure (Fig. 2) (see the Discussion for alternative interpretations). In all populations, individuals typically explored the room faster during their repeat tests compared to their first one (see above, Table 1), but individuals that were initially fast explorers became disproportionally faster in repeat tests compared to birds that initially explored the room less quickly (Fig. 2). As a consequence, the between-individual variance in exploration score increased from first to repeat tests (Fig. 2). This link between personality and plasticity was present in all populations and was shown statistically by high support for a positive covariance between the elevation and slope
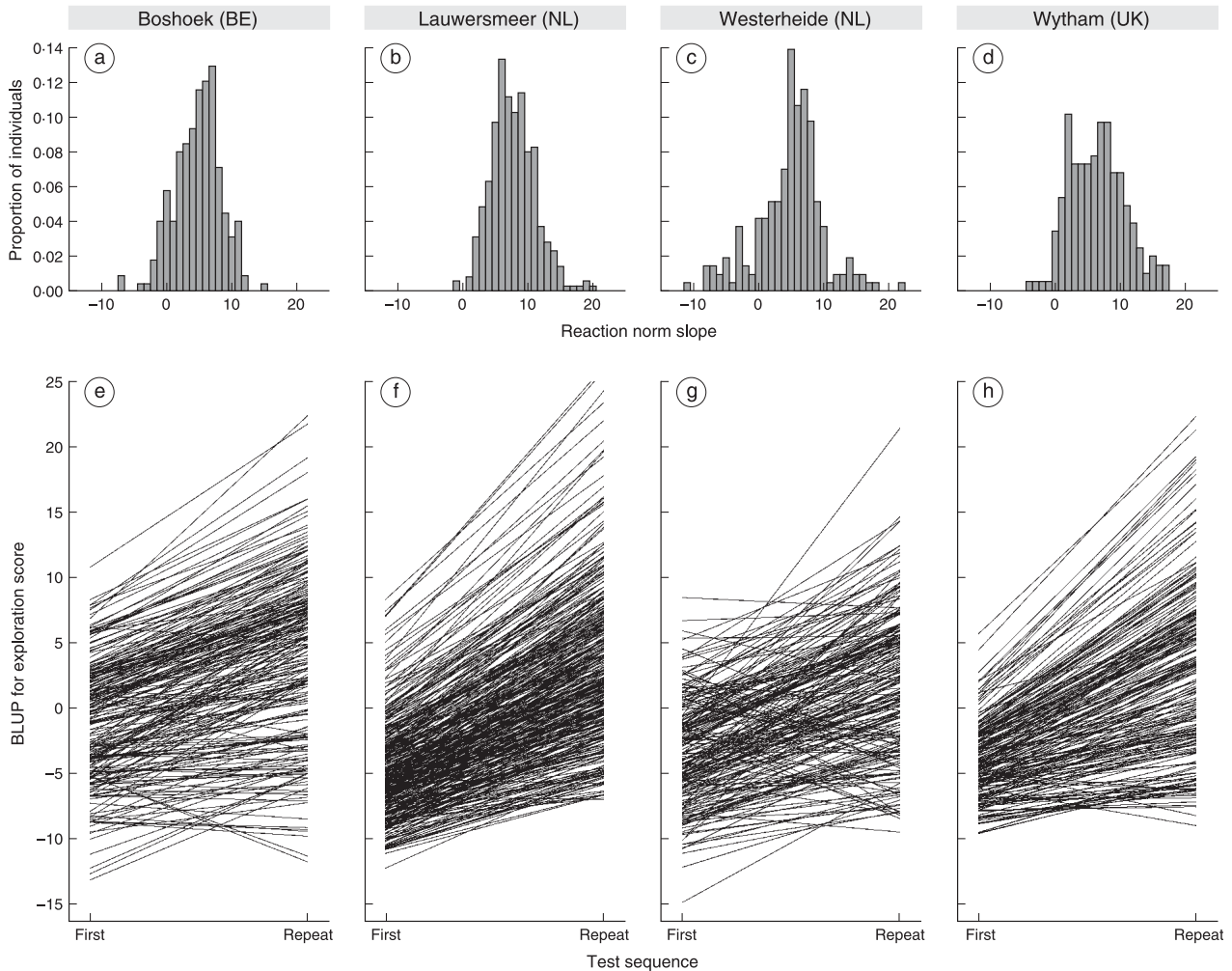
Fig. 2. Frequency distributions (a–d) and plots (e–g) of the best linear unbiased predictors of the individual reaction norm slopes (I × E: Individual × Sequence) between first ('novel environment') and repeated (second to fourth; 'familiar environment') exploration tests for wild-caught great tits in four different study areas; (a) & (e): Boshoek (*n* = 224 birds); (b) & (f): Lauwersmeer (*n* = 350 birds); (c) & (g): Westerheide (*n* = 226 birds); (d) & (h): Wytham Woods (*n* = 207 birds). Exploration scores were corrected for the population-specific effects of interval and time of year between tests (given in Table 2).

($\sigma_{\text{elevation,slope}}$) of exploration-sequence reaction norms (ΔDIC ranged between −88·3 and −9·9; Table 2).

At the same time, we found more support for the absence than the presence of population variation in the amount of between-individual variation in these reaction norm slopes for sequence (study site × individual × sequence: ΔDIC = +1·2). Similar estimates and statistics were obtained when we only considered the 288 individuals tested more than twice (study site × individual × sequence: ΔDIC = +1·0), implying that the inclusion of these individuals did not cause bias. Similarly, the absence of variation among the four populations in the covariance between the elevation and sequence slope was better supported than its presence (ΔDIC = +2·1). In other words, the same 'fanning-out' pattern of between-individual variance with test sequence characterized all populations (Fig. 2).

For individual variation in behavioural plasticity with respect to interval or time of year, the absence of I × E was generally better supported than its presence (ΔDIC ranged

between +0·6 and +1·8 for individual × time of year (σ^2_{slope}); ΔDIC ranged between +0·0 and +1·8 for Individual × Interval (σ^2_{slope}); Table 2). Therefore, given the data at hand, it appears that individual variation in plasticity with regard to interval or time of year was either absent (as suggested by the observation that values of ΔDIC were always positive), or that we had insufficient power to detect it.

Discussion

Our results demonstrate considerable complexity in the patterns of between-individual and between-population variation underlying assays of an easily measured avian personality trait, which we summarize here in four main findings. (i) Significant individual differences ('I') in exploration behaviour were found in all populations, and this component of variation was similar across all populations. (ii) Analyses of the fixed effects suggested that exploration behaviour increased with sequence and time of year but decreased with

interval between successive tests within individuals; the average patterns were similar across populations for sequence but differed quantitatively across populations for both interval and time of year. (iii) Significant individual by environment interaction (I × E) was present with respect to sequence (individual × sequence), but not with respect to time of year or test interval, and populations did not appear to differ from one another in these patterns. (iv) Finally, the covariance between elevation and slope of individual exploration–habituation reaction norms showed that individuals that were fast explorers on the first test became disproportionally faster in repeat tests compared to birds that initially explored the room less quickly; this pattern was present in all populations, and populations did not appear to differ from one another in this link between exploration behaviour on the first test and subsequent plasticity over, or habituation to, subsequent tests.

Consistent individual differences in average level of exploration behaviour have been discussed extensively elsewhere, and here, we simply point out that this finding supports the assertion that exploration behaviour may be a ubiquitous, heritable personality trait in this species (Réale *et al.* 2007). We focus the remainder of our discussion primarily on patterns of plasticity variation because few attempts have previously been made to estimate differences in behavioural plasticity between individuals using variance components, let alone compare variation in behavioural plasticity between populations.

When confronted with a novel environment repeatedly over time, individuals become familiar with the situation that usually leads to decreased activity (e.g., Elliott & Grunberg 2005; Mettke-Hofmann *et al.* 2006; Martin & Réale 2008). The direction of the effect is likely to be system dependent, and exploration scores for our great tits instead increased over successive tests. This effect (i) was most pronounced when the interval between successive tests was relatively short, suggesting a memory component to the habituation process and (ii) varied between individuals (I × E; Fig. 2). Individual variation in habituation to a novel situation is commonly associated with individual variation in learning ability (e.g., File 2001), with information processing (Elliott & Grunberg 2005) and risk aversion (McSweeney, Murphy & Kowal 2004), all of which are potentially linked and involved in the patterns described here. Exploration behaviour in the great tit is often associated with fearfulness (reviewed by Groothuis & Carere 2005), and individual differences in the change in exploration behaviour over repeated tests may thus reflect differences in the speed of overcoming initial fear. Furthermore, Verbeek, Drent & Wiepkema (1994) suggested that the initial exploration of the environment is relatively thorough and cautious ('slow' exploring) in great tits, while exploration is more superficial in repeat tests ('fast' exploring). This process agrees with the cognitive map theory of habituation (O'Keefe & Nadel 1978) in which individuals construct a representation of the novel environment in their hippocampus: as the map becomes complete, exploration is reduced.

In our system, although birds became faster, paradoxically slow-exploring birds are known to explore more thoroughly compared with fast-exploring ones (reviewed by Groothuis & Carere 2005). We suspect that a combination of overcoming fear, learning and memory all contributed to these complex patterns, the functional significance of which is likely to be substantial.

Changes in exploration behaviour with time of year were observed for all populations, but the proximate and functional significance of this form of plasticity remains unknown. Interpretation of this pattern is complicated by the fact that our measure of exploration behaviour likely reflects a combination of behavioural tendencies, including exploration *per se*, risk taking behaviour and response to capture and handling (e.g. van Oers *et al.* 2004). Adaptive behavioural changes to seasonal environmental variation are widespread (e.g. migration, singing), and thus the functional significance of exploratory behaviour may be dependent on time of year. Increases in explorative behaviour from autumn to spring, as observed in our current and previous studies (Dingemanse *et al.* 2002; Quinn *et al.* 2009), might have evolved because the energetic or behavioural costs of fast exploration outweigh any functional benefits. For example, fast exploration behaviour in early spring may be particularly important if it facilitates information acquisition in the early stages of reproduction (Semenova *et al.* 2001; Prendergast & Nelson 2005; Mettke-Hofmann 2007), but in the non-breeding season, it may only be significant if food availability is scarce. Towards spring, an increase in hormones, such as testosterone, has been reported (e.g., Dittami & Gwinner 1985), and these may play a role in regulating changes in exploratory behaviour (but see Mutzel *et al.* 2011) and require further investigation.

We could not detect individual differences in how the phenotype of individuals changed across the season (individual × time of year), and we therefore suggest that this ubiquity indicates the seasonal trend is likely linked to energetic costs associated with high exploration scores. The absence of I × E over the season within populations also implies that the behavioural types of individuals captured at different times of the season can be compared as long as the population-average plasticity is known (Dingemanse *et al.* 2002), and that population-specific sampling bias is unlikely to bias our estimates of individual exploration phenotypes (Martin & Réale 2008; Dingemanse *et al.* 2010). The evolutionary implication of this similarity across populations is that there is probably no additive genetic variance for this form of plasticity within these populations, and that it therefore cannot be the target of selection (Nussey, Wilson & Brommer 2007). At the same time, heritable variation in this form of plasticity might exist at larger spatial scales, because our data suggested that populations differed in the magnitude of population-average plasticity with respect to time of year. The non-experimental nature of our data set, however, also provides scope for alternative explanations. For example, exploration behaviour may have been a function of other unmeasured environmental axes (e.g. population density,

food availability or risk of predation) that covaried with time of year in some, but not all populations.

The most novel contribution of our study to the field of animal personality research comes from comparisons in the patterns observed among populations. There was remarkable consistency across populations in the presence or absence of the effects observed. This finding points to the potential ubiquitous significance of exploration behaviour. However, the magnitude of two out of three fixed effects varied between populations: effects of both interval and time of year on the behaviour of individuals differed between study sites. These effects suggest that adaptive population differentiation in behavioural plasticity might be population or species-specific, which are likely to arise from evolutionary factors as well as non-genetic factors (e.g. parental effects). Elsewhere, studies on hermit crabs *Pagurus bernhardus* (Briffa, Rundle & Fryer 2008) and salt marsh wolf spiders *Pardosa purbeckensis* (Bonte, Bossuyt & Lens 2007) have recently documented population variation in average levels of plasticity. Together with these studies, our findings imply that population differences in average level of plasticity are likely to be common.

The evolutionary significance of the observed patterns of variation in aspects of behavioural phenotypes within and across populations remains to be elucidated by focussing on three key questions. First, is the observed pattern of I × E because of individual variation in behavioural plasticity? This question is particularly valid because of the observational nature of our data. Perhaps all individuals had the same exploration – habituation reaction norm slope, but unmeasured environmental axes affecting the birds' behaviour might have harboured greater variance among individuals during repeat compared to initial tests. Such bias might, for example, be caused by age-related increases in between-individual variance in breeding conditions (Charmantier & Garant 2005) or accumulated experiences (Stamps & Groothuis 2010). Approaches where exploration behaviour of each individual would be essayed repeatedly over a range of environments differing in novelty *per se* may provide future experimental confirmation of our findings of I × E with respect to 'test sequence'.

Second, assuming that I × E was indeed reflecting individual variation in habituation, what are its proximate causes; specifically, does this variance component have an underlying additive genetic basis (G × E)? Work in two of our populations (WH and WW) has already revealed additive genetic (Dingemanse *et al.* 2002; Quinn *et al.* 2009) and permanent environmental (Quinn *et al.* 2009) sources underlying individual variation (I) in exploratory behaviour. G × E has not yet been estimated in these populations but could be carried out so using pedigree-based REML analysis (see Brommer, Rattiste & Wilson (2008) for a worked example). We note that evidence for G × E would by default imply that I × E did indeed partly reflect individual variation in behavioural plasticity.

Third, does selection act on plasticity in exploration behaviour (i.e. slopes of behavioural reaction norms)? Recently, we documented spatial and temporal variation in selection pressures on exploratory behaviour (Dingemanse *et al.* 2004; Quinn *et al.* 2009), but selection on the phenotypic, or genotypic, component of its plasticity awaits quantification. Notably, estimates of selection acting on behavioural plasticity in natural populations are generally missing from the empirical literature (Duckworth & Kruuk 2009; Dingemanse *et al.* 2010). Although the analyses associated with these questions are technically and empirically challenging (Hadfield *et al.* 2009), the differences and similarities in the patterns of variation observed across our populations suggest that they could be equally rewarding.

An additional challenge will be the identification of the mechanisms underlying selection on plasticity (Reed *et al.* 2006; Nussey, Wilson & Brommer 2007; Hadfield *et al.* 2009). Wolf, van Doorn & Weissing (2008) recently proposed that individual differences in plasticity might be maintained by a combination of negative frequency-dependent selection (maintaining variation in plasticity) and positive feedback mechanisms reducing the costs of plasticity (thereby generating consistent individual differences in plasticity). Our finding of positive correlations between the elevation and slope of exploration–habituation reaction norms in all four populations (Fig. 2) therefore warrants further investigation with regard to whether such associations are favoured and maintained by selection (Dingemanse *et al.* 2010).

Acknowledgements

References

Bell, A.M. (2005) Behavioral differences between individuals and two populations of stickleback (*Gasterosteus aculeatus*). *Journal of Evolutionary Biology*, **18**, 464–473.

Bell, A.M., Hankison, S.J. & Laskowski, K.L. (2009) The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, **77**, 771–783.

Bonte, D., Bossuyt, B. & Lens, L. (2007) Aerial dispersal plasticity under different wind velocities in a salt marsh wolf spider. *Behavioral Ecology*, **18**, 438–443.

Briffa, M., Rundle, S.D. & Fryer, A. (2008) Comparing the strength of behavioural plasticity and consistency across situations: animal personalities

in the hermit crab *Pagurus bernhardus. Proceedings of the Royal Society London Series B*, **275**, 1305–1311.

Brommer, J.E., Rattiste, K. & Wilson, A.J. (2008) Exploring plasticity in the wild: laying date-temperature reaction norms in the common gull *Larus canus. Proceedings of the Royal Society London Series B*, **275**, 687–693.

Brommer, J.E., Merila, J., Sheldon, B.C. & Gustafsson, L. (2005) Natural selection and genetic variation for reproductive reaction norms in a wild bird population. *Evolution*, **59**, 1362–1371.

Charmantier, A. & Garant, D. (2005) Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings of the Royal Society London Series B*, **272**, 1415–1425.

Dall, S.R.X., Houston, A.I. & McNamara, J.M. (2004) The behavioural ecology of personality: consistent individual differences from an adaptive perspective. *Ecology Letters*, **7**, 734–739.

DeWitt, T.J., Sih, A. & Wilson, D.S. (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution*, **13**, 77–81.

Dingemanse, N.J. & Wolf, M. (2010) A review of recent models for adaptive personality differences. *Philosophical Transactions of the Royal Society of London Series B*, **365**, 3947–3958.

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. *Animal Behaviour*, **64**, 929–937.

Dingemanse, N.J., Both, C., Drent, P.J. & Tinbergen, J.M. (2004) Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society London Series B*, **271**, 847–852.

Dingemanse, N.J., Wright, J., Kazem, A.J.N., Thomas, D.K., Hickling, R. & Dawnay, N. (2007) Behavioural syndromes differ predictably between 12 populations of stickleback. *Journal of Animal Ecology*, **76**, 1128–1138.

Dingemanse, N.J., Kazem, A.J.N., Réale, D. & Wright, J. (2010) Behavioural reaction norms: where animal personality meets individual plasticity. *Trends in Ecology and Evolution*, **25**, 81–89.

Dittami, J.P. & Gwinner, E. (1985) Annual cycles in the African Stonechat *Saxicola torquata axillaris* and their relationship to environmental factors. *Journal of Zoology*, **207**, 357–370.

Drent, P.J., van Oers, K. & van Noordwijk, A.J. (2003) Realized heritability of personalities in the great tit (*Parus major*). *Proceedings of the Royal Society London Series B*, **270**, 45–51.

Duckworth, R.A. & Kruuk, L.E.B. (2009) Evolution of genetic integration between dispersal and colonization ability in a bird. *Evolution*, **63**, 968–977.

Elliott, B.M. & Grunberg, N.E. (2005) Effects of social and physical enrichment on open field activity differ in male and female Sprague-Dawley rats. *Behavioural Brain Research*, **165**, 187–196.

Elton, C. (1966) *The Pattern of Animal Communities*. Methuen, London.

Falconer, D.S. & Mackay, T.F.C. (1996) *Introduction to Quantitative Genetics*. Longman, New York.

File, S.E. (2001) Factors controlling measures of anxiety and responses to novelty in the mouse. *Behavioural Brain Research*, **125**, 151–157.

Fisher, R.A. (1925) *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.

Gosling, S.D. (2001) From mice to men: what can we learn about personality from animal research? *Psychological Bulletin*, **127**, 45–86.

Groothuis, T.G.G. & Carere, C. (2005) Avian personalities: characterization and epigenesis. *Neuroscience and Biobehavioral Reviews*, **29**, 137–150.

Hadfield, J.D., Wilson, A.J., Garant, D., Sheldon, B.C. & Kruuk, L.E.B. (2009) The misuse of BLUP in ecology and evolution. *The American Naturalist*, **175**, 116–125.

Hollander, F.A., van Overveld, T., Tokka, I. & Matthysen, E. (2008) Personality and nest defence in the great tit (*Parus major*). *Ethology*, **114**, 405–412.

Jenni, L. & Winkler, R. (1994) *Moult and Ageing of European Passerines*, Academic Press, London.

de Jong, G. (1995) Phenotypic plasticity as a product of selection in a variable environment. *The American Naturalist*, **145**, 493–512.

Kelly, C.D. (2006) Replicating empirical research in behavioural ecology: how and why it should be done but rarely ever is. *Quarterly Review of Biology*, **81**, 221–236.

Korsten, P., Mueller, J.C., Hermannstädter, C., Bouwman, K.M., Dingemanse, N.J., Drent, P.J., Liedvogel, M., Matthysen, E., van Oers, K., van Overveld, T., Patrick, S.C., Quinn, J.L., Sheldon, B.C., Tinbergen, J.M. & Kempenaers, B. (2010) Association between DRD4 gene polymorphism and personality in great tits: a test across four wild populations. *Molecular Ecology*, **19**, 832–843.

Martin, J.G.A. & Réale, D. (2008) Temperament, risk assessment and habituation to novelty in eastern chipmunks, *Tamias striatus. Animal Behaviour*, **75**, 309–318.

Matthysen, E. (2002) Boundary effects on dispersal between habitat patches by forest birds (*Parus major, P.caeruleus*). *Landscape Ecology*, **17**, 509–515.

McSweeney, F.K., Murphy, E.S. & Kowal, B.P. (2004) Varying reinforcer duration produces behavioral interactions during multiple schedules. *Behavioural Processes*, **66**, 83–100.

Mettke-Hofmann, C. (2007) Object exploration of garden and sardinian warblers peaks in spring. *Ethology*, **113**, 174–182.

Mettke-Hofmann, C., Rowe, K.C., Hayden, T.J. & Canoine, V. (2006) Effects of experience and object complexity on exploration in garden warblers (*Sylvia borin*). *Journal of Zoology*, **268**, 405–413.

Mutzel, A., Kempenaers, B., Laucht, S., Dingemanse, N.J. & Dale, J. (2011) Circulating testosterone levels do not affect exploratory behaviour in house sparrows: observational and experimental tests. *Animal Behaviour*, **81**, 731–739.

Nakagawa, S. & Schielzeth, H. (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews*, **85**, 935–956.

Nicolaus, M., Both, C., Ubels, R., Edelaar, P. & Tinbergen, J.M. (2009) No experimental evidence for local competition in the nestling phase as a driving force for density-dependent avian clutch size. *Journal of Animal Ecology*, **78**, 828–838.

Nussey, D.H., Wilson, A.J. & Brommer, J.E. (2007) The evolutionary ecology of individual phenotypic plasticity in wild populations. *Journal of Evolutionary Biology*, **20**, 831–844.

Nussey, D.H., Postma, E., Gienapp, P. & Visser, M.E. (2005) Selection on heritable phenotypic plasticity in a wild bird population. *Science*, **310**, 304–306.

van Oers, K., Drent, P.J., de Goede, P. & van Noordwijk, A.J. (2004) Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. *Proceedings of the Royal Society London Series B*, **271**, 65–73.

O'Keefe, J. & Nadel, L. (1978) *The Hippocumpus as a Cognitive Map*. Oxford University Press, Oxford.

van de Pol, M. & Verhulst, S. (2006) Age-dependent traits: a new statistical model to separate within- and between-individual effects. *The American Naturalist*, **167**, 766–773.

van de Pol, M. & Wright, J. (2009) A simple method for distinguishing within- versus between-subject effects using mixed models. *Animal Behaviour*, **77**, 753–758.

Prendergast, B.J. & Nelson, R.J. (2005) Affective responses to changes in day length in Siberian hamsters (*Phodopus sungorus*). *Psychoneuroendocrinology*, **30**, 438–452.

Quinn, J.L. & Cresswell, W. (2005) Chaffinch *Fringilla coelebs* personalities: correlations between activity level, apparent stress and anti-predation behaviour. *Behaviour*, **142**, 1383–1408.

Quinn, J.L., Patrick, S.C., Bouwhuis, S., Wilkin, T.A. & Sheldon, B.C. (2009) Heterogeneous selection on a heritable temperament trait in a variable environment. *Journal of Animal Ecology*, **78**, 1203–1215.

Rasbash, J., Steele, F., Browne, W. & Prosser, B. (2005) *A User's Guide to MLwiN – Version 2.0*. Centre for Multilevel Modelling, University of Bristol, London.

Réale, D. & Dingemanse, N.J. (2010) Personality and individual social specialisation. *Social Behaviour: Genes, Ecology and Evolution* (eds T. Szekely, A. Moore & J. Komdeur). pp. 527–557, Cambridge University Press, Cambridge.

Réale, D., Reader, S.M., Sol, D., McDougall, P. & Dingemanse, N.J. (2007) Integrating temperament in ecology and evolutionary biology. *Biological Reviews*, **82**, 291–318.

Reed, T.E., Wanless, S., Harris, M.P., Frederiksen, M., Kruuk, L.E.B. & Cunningham, E.J.A. (2006) Responding to environmental change: plastic responses vary little in a synchronous breeder. *Proceedings of the Royal Society London Series B*, **273**, 2713–2719.

Roff, D.A. (1997) *Evolutionary Quantitative Genetics*. Chapman and Hall, New York.

Romero, L.M. (2004) Physiological stress in ecology: lessons from biomedical research. *Trends in Ecology and Evolution*, **19**, 249–255.

Semenova, T.P., Anoshkina, I.A., Khomut, B.M. & Kolaeva, S.G. (2001) Seasonal peculiarities of behavior of ground squirrel *Citellus undulatus* in hole-board and open field tests. *Behavioural Processes*, **56**, 195–200.

Sih, A., Bell, A. & Johnson, J.C. (2004) Behavioral syndromes: an ecological and evolutionary overview. *Trends in Ecology and Evolution*, **19**, 372–378.

Sih, A., Bell, A.M., Johnson, J.C. & Ziemba, R.E. (2004) Behavioural syndromes: an integrative overview. *Quarterly Review of Biology*, **79**, 241–277.

Smiseth, P.T., Wright, J. & Kölliker, M. (2008) Parent-offspring conflict and co-adaptation: behavioural ecology meets quantitative genetics. *Proceedings of the Royal Society London Series B*, **275**, 1823–1830.

Spiegelhalter, D.J., Best, N.G., Carlin, B.R. & van der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B, Statistical Methodology*, **64**, 583–616.

Stamps, J.A. & Groothuis, T.G.G. (2010) Developmental perspectives on personality: implications for ecological and evolutionary studies of individual differences. *Philosophical Transactions of the Royal Society of London Series B*, **365**, 4029–4041.

Verbeek, M.E.M., Drent, P.J. & Wiepkema, P.R. (1994) Consistent individual differences in early exploratory behaviour of male great tits. *Animal Behaviour*, **48**, 1113–1121.

Via, S. & Lande, R. (1985) Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution*, **39**, 505–523.

Via, S., Gomulkiewicz, R., de Jong, G., Scheiner, S.M., Schlichting, C.D. & Van Tienderen, P.H. (1995) Adaptive phenotypic plasticity – consensus and controversy. *Trends in Ecology and Evolution*, **10**, 212–217.

Wolf, M., van Doorn, G.S. & Weissing, F.J. (2008) Evolutionary emergence of responsive and unresponsive personalities. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 15825–15830.

Wolf, M. & Weissing, F.J. (2010) An explanatory framework for adaptive personality differences. *Philosophical Transactions of the Royal Society of London Series B*, **365**, 3959–3968.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Data S1. Statistical Modelling Approach.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.