

Chapter 2

REPEATABILITY AND HERITABILITY OF EXPLORATORY BEHAVIOUR IN GREAT TITS FROM THE WILD

Niels J. Dingmanse, Christiaan Both, Piet J. Drent, Kees van Oers
& Arie van Noordwijk

Animal Behaviour (2002) 64, 929-937

ABSTRACT

We investigated whether individual great tits, *Parus major*, vary consistently in their exploratory behaviour in a novel environment and measured the repeatability and heritability of this trait. Wild birds were caught in their natural habitat, tested in the laboratory in an open field test on the following morning, then released at the capture site. We measured individual consistency of exploratory behaviour for recaptured individuals (repeatability) and estimated the heritability with parent–offspring regressions and sibling analyses. Measures of exploratory behaviour of individuals at repeated captures were consistent in both sexes and study areas (repeatabilities ranged from 0.27 to 0.48). Exploration scores did not differ between the sexes, and were unrelated to age, condition at fledging or condition during measurement. Heritability estimates were 0.22–0.41 (parent–offspring regressions) and 0.37–0.40 (sibling analyses). We conclude that (1) consistent individual variation in open field behaviour exists in individuals from the wild, and (2) this behavioural variation is heritable. This is one of the first studies showing heritable variation in a behavioural trait in animals from the wild, and poses the question how of this variation is maintained under natural conditions.

INTRODUCTION

In a wide range of vertebrates, individuals show consistent behavioural differences in how they respond to novel situations (Wilson et al. 1994; Gosling 2001). These behavioural responses have been arranged along axes such as shyness/boldness (Clark & Ehlinger 1987; Wilson et al. 1994), aggression (Benus et al. 1992), neophobia (Greenberg & Mettke-Hofmann 2001) or exploration (Winkler & Leisler 1999). Many laboratory studies have shown heritable variation in such behaviour towards novel social or nonsocial situations in mammals (e.g. van Oortmerssen & Bakker 1981; Simmel & Bagwell 1983; Kagan et al. 1988; Magurran 1990; Benus et al. 1991; Sluyter et al. 1996).

Behavioural variation in reaction to novel situations may be adaptive and be maintained by natural selection (Clark & Ehlinger 1987; Wilson 1998), rather than reflecting nonadaptive variation around an adaptive mean. However, the processes responsible for maintaining this genetic variation in the wild are largely unexplored, in part because most study species have a long history of domestication.

If we want to understand the existence and maintenance of this behavioural variation in the wild, we need to know both its repeatability and heritability in wild-living animals. Phenotypic selection will have evolutionary consequences only when the phenotypic variation is heritable (Fisher 1930; Endler 1986). Some studies have demonstrated repeatable differences in reaction to novel situations in wild-caught animals (Armitage 1986; Wilson et al. 1993; Réale et al. 2000; Fraser et al. 2001) or wild-caught animals reared under standardized conditions (Greenberg 1992; Verbeek et al. 1994, 1996). However, despite the considerable evidence for heritable variation in animals reared under laboratory conditions, heritability estimates have, to our knowledge, not yet been published for wild populations. Heritability estimates may be substantial when measured for laboratory populations, but insignificant in the wild because of large effects of environmental factors (Falconer & Mackay 1996). For instance, behaviour towards novel situations can be strongly affected by experiences during early stages of ontogeny (e.g. Winkler & Leisler 1999). Such potential sources of environmental variation are usually controlled in laboratory studies, but are present in wild-caught individuals. Therefore, field heritabilities are more reliable indicators of the potential response to natural selection on behaviour towards novel situations.

Verbeek et al. (1994) showed consistent individual differences in exploratory behaviour in hand-reared great tits, *Parus major*. They studied behavioural reactions towards novel objects and novel environments, defining exploration as any behaviour that provides the animal with information about its environment (Birke & Archer 1983). These behavioural differences covaried: birds that quickly explored novel environments also quickly explored novel objects in their home cage, readily formed routines and were aggressive in pairwise confrontations (Verbeek et al. 1996; Drent & Marchetti 1999; Marchetti & Drent 2000). Furthermore, a five-generation artificial selection experiment yielded a significant response to both downward ('slow') and upward ('fast') selection for a combined measure of behaviour towards both novel objects and environments (Drent et al. 2003). These heritable and covarying behaviours towards novel situations can be viewed as evidence for the concept of temperament (Boissy 1995; Segal & MacDonald 1998), personality (Buss 1991) or coping strategies (Jensen 1995). Furthermore, these individual strategies may have major fitness consequences, particularly since they relate to foraging, exploration and aggression.

We measured phenotypic variation in exploratory behaviour in a novel laboratory environment in wild-caught great tits to estimate the repeatability and heritability of exploratory behaviour. The repeatability of exploratory behaviour was measured for recaptured individuals in two study populations. The heritability of exploratory behaviour was estimated in one population, using parent–offspring regressions and sibling analyses. We further tested whether an aspect of rearing conditions, fledging weight, or body weight at testing influenced exploratory behaviour.

METHODS

Study sites

We measured exploratory behaviour of great tits captured from two study areas between May 1998 and March 2001. The Westerheide (WH) study area near Arnhem, the Netherlands, covers 250 ha of mixed wood, with patches of predominantly birch, *Betula pendula*, pedunculate oak, *Quercus robur*, red oak, *Q. rubra*, pine, *Pinus sylvestris*, larch, *Larix decidua* and beech, *Fagus sylvatica*. Approximately 600 nestboxes were provided from 1995 onwards. The Oosterhout (OH) study area near Nijmegen, the Netherlands, covers about 12 ha of isolated deciduous wood with pedunculate oak as the predominant tree species. Here, about 150 nestboxes were provided from 1956 onwards as part of a long-term study on the population dynamics of the great tit (van Balen 1973). All data were analysed separately for Oosterhout and Westerheide, because both catching intensity and methodology differed between the two study areas (see below). Heritabilities were calculated for Westerheide only, because sample sizes to do so were insufficient for Oosterhout.

Catching methods

In the breeding season of 1998 (May–June) we captured 61 individuals from both study areas and used them to develop a suitable behavioural test in the laboratory (their offspring were removed and hand-reared for another experiment). To take birds from the two study areas into the laboratory and score their behaviour, we inspected nestboxes twice a year during winter nights in November/December and February/March. In Westerheide individuals were also captured in mistnets at feeding stations baited with sunflower seeds (about twice a week from July to April). To exclude juveniles that were still dependent on their parents, only individuals older than 50 days of age or with fully grown tail feathers were taken to the laboratory for testing (Drent 1984; Verhulst & Hut 1996). All the adult birds captured at feeding stations had stopped breeding, and their offspring had already been independent for over 3 weeks in all cases.

Within 1.5 h of catching, we transported the selected birds to the laboratory in darkened boxes, weighed them and housed them individually. The following morning, we measured individual exploratory behaviour in a novel environment test (see below), except in the breeding season of 1998, when few birds were measured 10 days after catching (61 of 1070 individuals). After the last bird had been tested, all birds were weighed to the nearest 0.1 g and released in the field near their individual place of capture. We measured exploratory behaviour in the laboratory for 878 (WH)

and 192 (OH) individuals. Between November 1998 and December 1999, some birds were recaptured ($n_{WH}=185$, $n_{OH}=33$) to score their behaviour in the laboratory for a second time. Another experiment on Westerheide between March 2000 and April 2001 provided another opportunity to enlarge this sample of recaptured birds. Birds were never tested twice within a week.

The age of the birds not ringed as nestlings was determined by the colour of their greater wing coverts, allowing distinction between juvenile and older birds (Jenni & Winkler 1994). Juveniles caught in July and August were sexed with molecular markers (Griffiths et al. 1998) and adult birds were sexed by plumage characteristics (Jenni & Winkler 1994). Blood samples (5–15 μ l) used for sexing were taken from the brachial vein in the wing. This procedure has been shown to have no adverse effects (Hoysak & Weatherhead 1991; Lubjuhn et al. 1998). Blood samples were always taken after the measurement of exploration behaviour to prevent any effect of the sampling procedure on the behaviour of the bird.

Housing

In the laboratory, we housed the birds individually under normal daylight in cages of 0.9×0.4 m and 0.5 m high, with a solid bottom and top, side and rear walls, a wire mesh front and three perches. Each cage was connected to an observation room via a 20×20-cm sliding door in the rear wall. The birds received mealworms and unrestricted access to water, sunflower seeds and commercial seed mixture (Verbeek et al. 1994). Human exposure was kept to a minimum.

Novel environment test

We used the novel environment test (Verbeek et al. 1994), which is a variant of the classical open field test of animal psychologists (Walsh & Cummins 1976). Each bird was tested alone between 0800 and 1400 hours in a sealed room under artificial light. Trials began at least 1 h after sunrise, allowing the birds to eat before the first trials started. The room contained five artificial wooden trees (figure 1). We introduced each bird into the room without handling by darkening the cage with a curtain, opening the sliding door, turning on the light in the test room and briefly lifting the curtain, after which all birds flew into the room. We observed the behaviour of the birds in the observation room for the first 2 min after arrival. If the bird had not visited all five trees within this period, the observation time was extended until it had reached all trees, or up to a maximum of after 10 minutes, after which observation was stopped.

Besides the trees, birds could land on the edges of sliding doors, the main door and the one-way observation screen, and most birds explored the environment by flying between these and the trees, hopping between branches within the trees and/or actively scanning the room from a perch. During the observation period, we noted where a bird was at each second in time, and we noted details of movements, including whether the bird flew or hopped between locations.

We used the total number of flights and hops within the first 2 min as an index of exploratory behaviour ('exploration score'). Earlier laboratory studies of hand-reared great tits used the time of first arrival on the fourth (Drent et al. 2003) or fifth tree (Verbeek et al. 1994) as a measure of exploratory behaviour. We did not use these variables, because in our study many individuals (48%) did not reach the fifth tree

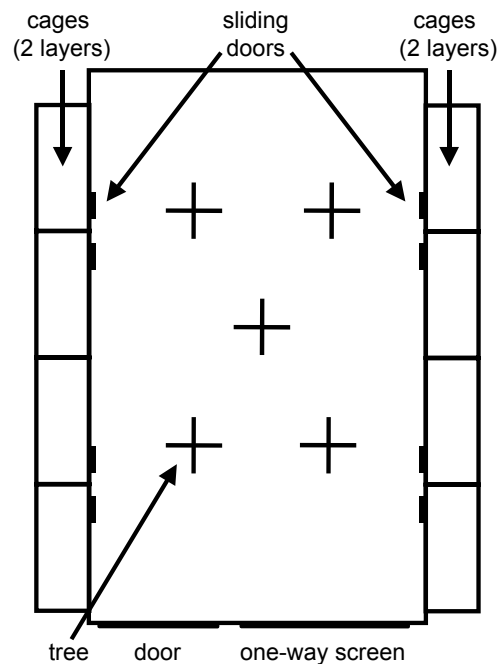


Figure 1 Plan of the observation room (4.0×2.4×2.3 m) in which we tested the exploratory behaviour of wild-caught great tits. Along each 4.0-m wall were eight sliding doors (in two rows of four above each other), connecting the holding cages to the room. The front 2.4-m wall had a 0.9×2.0-m door at the left side and a 1.1×0.16-m one-way screen through which the birds could be observed. The room contained five artificial trees made of wood with a trunk of 4×4 cm and a height of 1.5 m. Each tree had four cylindrical branches 20 cm long. The upper two branches (5 cm below the top) were on opposite sides of the trunk, perpendicular to the lower branches (25 cm below the top). Birds entered the room through one of the sliding doors.

within the 10-min observation period, and thus had identical censored scores, masking the behavioural variation in which we were interested. However, our measure of exploratory behaviour was highly correlated with first arrival time in our data (fourth tree: $r_s=0.78$; fifth tree: $r_s=0.72$; $n=1070$ individuals, both $p<0.0001$).

Statistical analyses

Within-individual sources of variation

Since part of the phenotypic variation may be due to factors with non-permanent effects (e.g. seasonality), which will lower our ability to compare the behaviour of different individuals, especially when individuals are measured consistently under different conditions (Falconer & Mackay 1996; Griffiths & Sheldon 2001), we tested whether within-individual variation in exploration score was related to within-individual variation in a number of explanatory variables with potentially non-permanent effects (table 1), to correct for them. We used a General Linear Model (GLM) and type III sums of squares to evaluate the within-individual relationship between exploration score and a range of explanatory variables. Only data for those individuals measured twice were included in this analysis. ‘Individual’ was entered as a factor in each model, treating the repeated exploration scores within subjects as independent cases. We constructed a model with all explanatory variables and all two-

Table 1 Within-individual sources of variation in exploratory behaviour in repeated tests of wild-caught great tits for the study areas Westerheide and Oosterhout. The results are from a GLM with normal errors and type III sums of squares with individual as a factor after backward elimination of nonsignificant terms. *F* values given are for the inclusion of the variable in the final model.

Variable	Westerheide			Oosterhout		
	<i>df</i>	<i>F</i>	<i>p</i>	<i>df</i>	<i>F</i>	<i>p</i>
Age*	1,181	0.09	0.76	1,29	0.37	0.55
Body weight	1,181	0.55	0.46	1,29	1.14	0.30
Catching method†	2,180	0.20	0.82	-	-	-
Season	1,182	14.00	0.0002	1,29	2.99	0.095
Time of day	1,181	0.15	0.70	<i>na</i> ‡		
Sequence‡	1,182	50.20	<0.0001	1,30	5.62	0.024
Interval	1,182	14.01	0.0002	1,29	3.50	0.072
Time of day × sex	1,180	3.12	0.08	1,30	5.55	0.025

*Juvenile or older.

†Breeding season 1998, net, roost.

‡First or second measurement.

§Not applicable; higher order interaction is significant.

way interactions with sex separately for each study area, using backward elimination of nonsignificant terms. Explanatory variables included the following factors (f) and covariates (c): individual (f), catching method (f; caught in (1) nestbox during breeding season 1998, (2) nestbox during roosting inspection, or (3) mistnet), time of the start of the novel room test (c; in minutes after sunrise), age (f; juvenile versus older), body weight (c; at the time of entering the laboratory), sequence (f; first or second measurement), and between-test interval (c; tested by entering exact age in days as a continuous variable; mean $x \pm SD$ interval = 131.2 ± 130.3 days, $n_{WH} = 185$ and 168.6 ± 120.0 days, $n_{OH} = 33$). We also included an explanatory variable for season, because exploration score was significantly related to month in first tests (ANOVA: $F_{10,1059} = 24.04$, $p < 0.00001$, $r^2 = 0.18$), with exploration score increasing from July to the following June (figure 2). Season was defined as the number of days from 30 June (range 1–365), because the coefficient of determination between exploration score and season ($r^2_{1068} = 0.17$, $p < 0.00001$) was not significantly different from that derived from the model with month as a factor ($Z = 0.39$, $p = 0.70$).

Based on the analyses described above, a corrected exploration score (*Y*) was calculated as $Y = \text{measured exploration score} - \sum b_i \times x_i + 10$ with b_i = parameter estimate of explanatory variable *i* and x_i = value of explanatory variable (categorical variables were transformed into dummy variables, and parameter estimates were calculated separately for each category). For both study areas, the data were corrected for explanatory variables that were significant ($p < 0.05$) in the final model (table 1) and used for all other analyses.

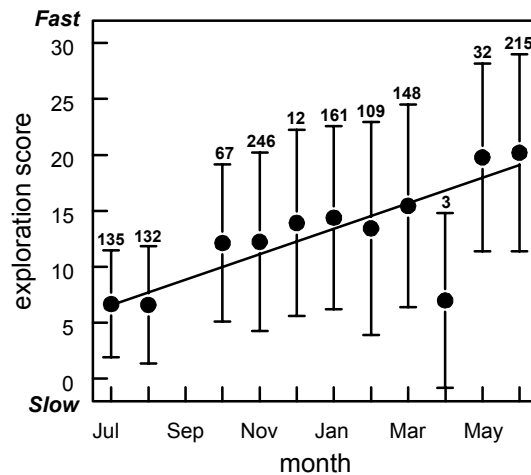


Figure 2 The relation between mean±SD exploration score and month for first captures of wild-caught great tits ($n=1070$ individuals; Westerheide and Oosterhout study areas combined). The solid line depicts the slope of the regression coefficient for season, defined as the number of days from 30 June (see Methods). Numbers represent sample sizes per month.

Repeatability and heritability

We calculated repeatability of exploration score for the recaptured individuals for which we obtained multiple measurements. Repeatability, the proportion of the phenotypic variance explained by the individual (Falconer & Mackay 1996), was calculated following Lessells & Boag (1987) and its standard errors following Becker (1984). The repeatability of exploration score was analysed separately for each combination of sex and area, because repeatabilities could differ between subgroups owing to differences in gene frequencies, large-scale environmental effects or sex-specific expression (Falconer & Mackay 1996).

We estimated narrow-sense heritabilities (h^2) of exploratory behaviour using parent–offspring regressions, where heritability is the proportion of total variance that is attributable to the additive effect of genes (Falconer & Mackay 1996). Because exploration scores of the male and female parent were not correlated ($r^2_{40}=0.03$, $p=0.30$), we did not correct the estimates for assortative mating. Unequal sample sizes in the number of offspring measured were taken into account by weighting by the square root of the number of offspring in the nest (Sokal & Rohlf 1995). Relatedness between offspring and parents was determined where the chicks were ringed in the nest and their parents were captured during the breeding cycle. Females were identified while incubating eggs, and both parents were caught when their young were 8–10 days old (see van Balen 1973 for a full description of nest box inspection methods). An alternative heritability estimate was derived from a one–way ANOVA with nest as a factor and calculated as twice the intraclass correlation coefficient, assuming that all nestlings are full siblings (sibling analysis, Falconer & Mackay 1996). Both heritability estimates include common environment effects and may be inflated when those effects explain significant variation, in particular sibling analyses (Falconer & Mackay 1996).

To test for ontogenetic effects on exploration score, we weighed nestlings 15 days after hatching (van Balen 1973), and tested whether fledging weight explained

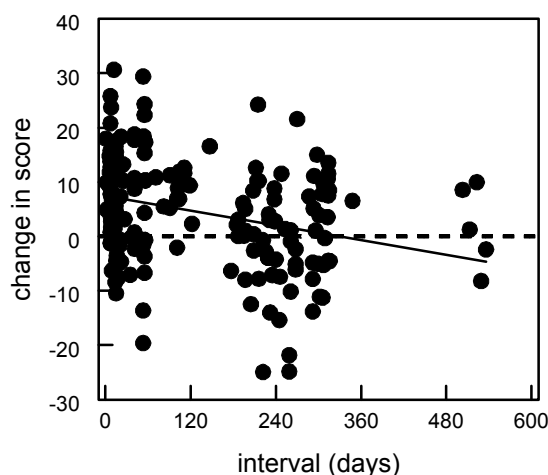


Figure 3 Within individual change in exploration score (second minus first measure, corrected for season) in wild-caught great tits in relation to the interval between the two measures in the Westerheide study area. The dotted line depicts equal scores for both measures. The slope of the solid line is the regression coefficient for interval from the model in table 1.

significant within- and between-brood variation in nests with sampled offspring of known origin.

The data were analysed using with SPSS v.10.1. Kolmogorov–Smirnov tests for normality confirmed that the assumptions underlying analyses of variance were fulfilled in all cases. *P* values are two tailed throughout.

Ethical note

Apart from the 61 individuals kept in the laboratory for 10 days to develop a suitable behavioural test, our test procedure involved a 14–24-h period in which birds were removed from their natural environment. Our data showed no adverse effect of the test procedure. First, changes in body weight between arrival and departure from the laboratory (males: mean±*SE* change = -0.87±0.38 g; females: -0.96±0.39 g) were well within the natural range of overnight weight loss (of about 1 g; van Balen 1967). Second, only one of 1070 individuals died between capture and release from the laboratory, which is well within the natural range of this species (annual mortality rates are ca. 0.5; Bauchau & van Noordwijk 1995). Third, male great tits usually do not lose their territory when removed for less than 48 hours (Krebs 1982; N.J. Dingemans, unpublished data). Permission for short-term transport of great tits to the laboratory was granted by KNAW Dierexperimenten Commissie.

RESULTS

Within-individual sources of variation

In Westerheide, within-individual variation in exploration score was significantly related to season but was not related to age, catching method or time of start of the

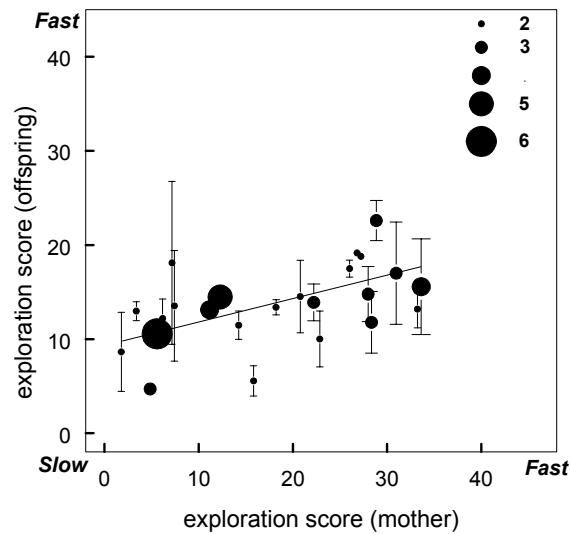


Figure 4 The relationship between mean \pm SE exploration score of offspring and mother in wild-caught great tits in the Westerheide study area, for broods with more than one offspring sampled. Heritability estimates are $h^2 \pm SE = 0.40 \pm 0.20$ (sibling analysis) and 0.41 ± 0.14 (mother–offspring regression; see Results). The line is the fitted regression equation weighted for the square root of the number of offspring sampled per mother ($0.21 \times \text{mother's score} + 1.93$). Dot sizes increase with the number of offspring sampled per mother.

novel room test (table 1). Within-individual variation in exploration score was unrelated to within-individual variation in body weight, indicating that the exploration score was not affected by changes in body condition. The birds explored the room significantly faster during their second measurement, but the effect diminished as the interval between subsequent measurements increased (table 1; figure 3). The analysis for Oosterhout confirmed that within-individual variation in exploratory behaviour was unrelated to age or body weight (table 1). Explanatory variables that were significant in the Westerheide model were either also significant or showed similar nonsignificant trends in the Oosterhout model, with qualitatively similar parameter estimates. The much smaller sample sizes for Oosterhout probably accounted for the lack of significance in some analyses.

Repeatability

Exploratory behaviour was repeatable for males and females in both study areas (range $r=0.27$ – 0.66 ; table 2) and the estimates did not differ significantly from each other ($F_{3,432}=1.09$, $p=0.35$; calculation of the F statistic after D.L. Thomson, N.J. Dingemanse & A.J. van Noordwijk, unpublished data). Furthermore, the effect of individual remained significant when all other (nonsignificant) variables in table 1 were also controlled for (Westerheide: $F_{184,177}=2.08$, $n=185$, $p<0.0001$; Oosterhout: $F_{32,27}=2.41$, $n=33$, $p=0.01$).

Table 2 Repeatability estimates (r) of exploratory behaviour in two populations of wild-caught great tits for study area and sex. Repeatability was calculated after Lessells & Boag (1987) from a one-way ANOVA with individual as a factor (n = number of individuals).

Study area/Sex	n	Mean score	V_p	$r \pm SE$	df	F	p
Westerheide							
Male	111	17.74	55.21	0.48±0.07	110,111	2.84	<0.0001
Female	74	17.50	56.27	0.27±0.11	73,74	1.72	0.010
Oosterhout							
Male	22	20.37	69.83	0.46±0.17	22,23	2.73	0.012
Female	11	24.01	107.58	0.66±0.17	10,11	4.95	0.007

Table 3 Heritability estimates of exploratory behaviour of wild-caught great tits in the Westerheide study area. Heritability estimates were calculated as (1) the slope (midparent–midoffspring) or twice the slope (single parent–midoffspring) of the regression for offspring and parents after weighing for the square root of offspring number per nest; (2) twice the intraclass correlation coefficient from a one-way ANOVA with nest as a factor (sibling analysis); n = number of nests, n_{off} = mean number of offspring per nest.

Method	n	n_{off}	Mean Score	V_p	$h^2 \pm SE$	df	F	p
Midparent–offspring	42	1.67	20.06	86.76	0.22±0.14	1,40	2.52	0.120
Father–offspring	59	1.78	20.96	72.52	0.31±0.19	1,57	2.53	0.117
Mother–offspring	63	1.65	18.96	95.41	0.34±0.16	1,61	4.30	0.042
Sibling analysis	50	2.64	15.13	90.72	0.61±0.20	49,82	2.16	0.001
Sibling analysis*	33	2.75	13.84	67.18	0.37±0.24	32,58	1.62	0.055

*Broods with parent(s) of known exploratory score.

Heritability

The heritability estimates of exploratory behaviour derived from parent–offspring regressions ranged from 0.22 to 0.34 (table 3) and did not differ significantly from each other (Z tests; all $p > 0.36$). Furthermore, based on the data from nests where more than one offspring was sampled, a sibling analysis confirmed that exploration scores differed significantly between nests (table 3). The heritability estimate derived from this sibling analysis was higher ($h^2 = 0.61$) but not significantly different from the estimates derived from parent–offspring regressions (Z tests: all $p > 0.11$). However, in 34% (11 of 50) of nests included in the sibling analysis, the parents' behaviour had not been scored. The parent–offspring estimate may be derived from a biased sample if parents of known and unknown score differ in their exploratory behaviour. To check this, we excluded nests from parents with unknown exploration score and reran

the sibling analysis. This indeed gave a heritability estimate that was much closer to that derived from parent–offspring regressions ($h^2=0.37$; table 3).

Heritability estimates derived from sibling analyses may be inflated by common environment effects (Falconer & Mackay 1996). However, the heritability estimates based on a sibling analysis and a mother–offspring regression for all nests with known maternal exploration score were very similar (respectively $h^2 \pm SE = 0.40 \pm 0.20$, $F_{23,41} = 1.67$, $p = 0.074$ and $h^2 = 0.41 \pm 0.14$, $F_{1,22} = 8.73$, $p = 0.007$; figure 4), suggesting that this estimate was not strongly determined by ontogenetic effects. Furthermore, fledging weight did not explain significant between-brood variation (linear regression: $F_{1,97} = 0.002$, $n = 37$ nests, $p = 0.97$) or within-brood variation (ANCOVA: $F_{1,61} = 0.80$, $n = 37$ nests, $p = 0.37$) in exploration score, although fledging weight differed significantly between broods (ANCOVA: $F_{36,62} = 3.95$, $n = 37$ nests, $p < 0.0001$). This latter result, combined with the similarity between heritabilities calculated from parent–offspring regression and sibling analyses, indicates that ontogenetic effects on exploration behaviour are relatively small.

DISCUSSION

This is one of the first studies showing heritable variation in a behavioural reaction towards a novel situation in nonhuman animals from the wild. Our heritability estimates of exploration score ranged from 0.22 to 0.61 and confirm the moderate heritabilities of comparable psychological traits in humans (e.g. Kagan et al. 1988; Bouchard 1994; Segal & MacDonald 1998) and other mammals (van Oortmerssen & Bakker 1981; Simmel & Bagwell 1983; Magurran 1990; Benus et al. 1991; Sluyter et al. 1996; Weiss et al. 2000). Furthermore, our results support the findings of two earlier studies in hand-reared great tits that showed (1) resemblance between siblings (Verbeek et al. 1994) and (2) a significant response to both upward and downward

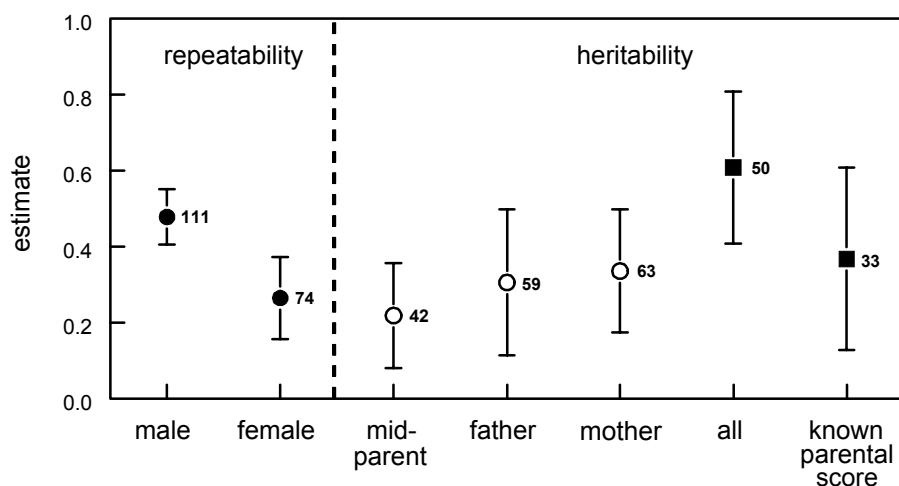


Figure 5 Repeatability (●) and heritability (○,■) estimates $\pm SE$ of exploration score in wild-caught great tits in the Westerheide study area. Heritability estimates are for parent–offspring regressions (○) and sibling analyses (■). Numbers represent sample sizes for each analysis.

selection for a composite measure of exploratory behaviour (broad-sense heritability = 0.54; Drent et al. 2003). These hand-reared birds were originally derived from natural populations, thus supporting our evidence for the existence of heritable variation in this trait in wild-caught birds.

The heritability estimates derived from parent–offspring regressions were not significantly different from each other, although regressions that included the father’s exploration score showed nonsignificant trends (table 3, figure 5). However, in these analyses the heritability estimate might have been underestimated and its standard error and *P* value inflated, because paternity of the offspring was uncertain owing to the occurrence of extrapair paternity in this species (3–18%; Gullberg et al. 1992; Lubjuhn et al. 1993; Blakey 1994; Verboven & Mateman 1997). Furthermore, in great tits, maternity of the offspring is certain, because brood parasitism has not been reported and is likely to be very rare (Kempnaers et al. 1995; references cited above).

We showed consistent individual differences in exploratory behaviour in a novel environment in males and females in two populations of wild-caught great tits. Since repeatability normally sets an upper limit to the heritability (Lessells & Boag 1987; Falconer & Mackay 1996), our repeatability estimates ($r=0.27$ – 0.66) confirm the moderate field heritabilities for this behaviour. Significant repeatabilities could be related to other differences between individuals, for instance in body size, age or body condition (Boake 1989; Falconer & Mackay 1996). However, within-individual variation in exploratory behaviour was not related to age or changes in body weight (table 1). Furthermore, there was significant between-individual variation in exploration score when controlling for all other explanatory variables in table 1. Thus, individual variation in exploratory behaviour cuts across categories of sex, size, age and state, confirming results from earlier laboratory studies (Verbeek et al. 1994; Drent & Marchetti 1999).

We have no evidence that exploratory behaviour was influenced by rearing condition. First, the heritability estimate based on sibling analysis was not inflated compared to the estimate derived from parent–offspring regressions (figure 5), which suggests that common environment effects were relatively unimportant (Falconer & Mackay 1996). Second, rearing condition (i.e. fledging weight) did not affect within- or between-brood variation in exploratory behaviour, in agreement with earlier results (Verbeek et al. 1994, 1996). However, we might have been unable to detect effects of rearing conditions if exploratory behaviour is affected by fledging weight only in individuals of extremely poor quality: the offspring included in our analyses were most probably individuals of higher quality, because survival is positively related to fledging weight in great tits (Tinbergen & Boerlijst 1990; Both et al. 1999).

Locomotor exploration

The assumption behind novel environment studies is that the amount of movement is an index of exploration, since animals have to move around to explore the environment (Russell 1983). The use of novel environment tests has been criticized, because locomotion in forced tests may also reflect behaviour that is unrelated to exploration (Barnett & Cowan 1976; Renner 1990; Hughes 1997). For instance, the effect of sequence and interval (figure 3) may indicate that the birds experienced less fear (Maier et al. 1988; Budaev 1997) when recaptured shortly after the first capture, and that stress affected the behaviour during the first measurements. Despite these limitations, the behaviour in novel environments is likely to result in

information acquisition and therefore reflects some aspects of exploration (Russell 1983).

Earlier studies on hand-reared great tits showed that activity in the home cage was unrelated to measures of exploration in novel object and novel environment tests (Verbeek et al. 1994, 1998). In concordance with our study (table 1), hand-reared birds also explored novel situations more quickly when confronted with the same challenge a second time (Verbeek et al. 1994). These results imply that the birds used information acquired during the first tests (Welker 1961; Renner 1990), and strongly suggest that the individual differences in behaviour of great tits reflect differences in aspects of exploration. Furthermore, the behaviour in the novel environment of hand-reared offspring and wild-caught parents is significantly correlated (Drent et al. 2003), providing direct evidence for a link between the behaviours measured in hand-reared and wild-caught great tits. Exploration tests in the field would provide an interesting new step to evaluate whether the behaviour in wild-caught great tits in the laboratory also extends to the wild.

Adaptive variation in behavioural traits

We now have three independent samples showing that heritable variation in exploratory behaviour exists in great tits (Verbeek et al. 1994; Drent et al. 2003; this study). Apparently these genotypes coexist. From an evolutionary perspective, genetic variation in behavioural traits may be adaptive and be maintained by natural selection (Wilson 1998; Foster & Endler 1999; Mousseau et al. 2000). The mechanisms responsible for the maintenance of variation in psychological traits that are related to how animals cope with novel situations are largely unstudied (Wilson et al. 1994; Wilson 1994), partly because of the difficulty in determining both fitness and behavioural phenotypes in the same species (but see Armitage 1986; Eaves et al. 1990; Mealey & Segal 1993; Réale et al. 2000). The few studies that have measured the consequences of these behavioural traits for survival, dispersal or reproduction either lack evidence for a genetic basis (Wilson et al. 1993; Réale et al. 2000; Fraser et al. 2001) or suffer from insufficient statistical power (Armitage 1986). Only studies on personality traits in humans have aimed to integrate both the genetic background (e.g. Kagan et al. 1988; Bouchard 1994) and fitness consequences (Eaves et al. 1990; Mealey & Segal 1993) of traits related to how individuals respond to novelty. The great tit is one of the few wild-living species for which we now have substantial evidence for a heritable basis of a behavioural trait both from a laboratory (Verbeek et al. 1994; Drent et al. 2003) and a wild population (this study). Because individual life histories are relatively easily obtained in wild populations of great tits, these findings provide a unique opportunity to study the consequences of individual behavioural strategies for major life history traits, including survival, dispersal and reproduction. In addition, such field studies are necessary to identify the processes responsible for the maintenance of heritable variation in temperament traits in natural populations (Wilson 1994, 1998).

ACKNOWLEDGEMENTS

We thank Piet de Goede for general assistance. Henri Bouwmeester, Leontien Witjes, Anna den Held, Kathelijne de Meijer, Maarten Hageman, Michel Geven and Dirk Zoetebier for help in the field and laboratory, Kate Lessells and Christa Mateman for sex determination using molecular markers, Bart van IJmeren, Mary-Lou Aaldering and Tanja Thomas for taking care of the birds in the laboratory and Jan Visser for managing the great tit databases. We are grateful to 'Het Gelders Landschap' and Mrs van Boetzelaer for permission to work in their properties. Helpful suggestions on the manuscript were made by Ton Groothuis, Jaap Koolhaas, Kate Lessells, David Thomson, Joost Tinbergen, Marcel Visser and two anonymous referees. N.D., C.B. and K.v.O. are supported by the Life Sciences Foundation (ALW grants 805-33-321, -324, and -323), which is subsidized by the Netherlands Organisation for Scientific Research (NWO).