

Individual variation in spontaneous problem-solving performance among wild great tits

Ella F. Cole*, Dominic L. Cram¹, John L. Quinn

Edward Grey Institute, Department of Zoology, University of Oxford

ARTICLE INFO

Article history:

Received 13 April 2010

Initial acceptance 8 June 2010

Final acceptance 22 November 2010

Available online 28 December 2010

MS. number: 10-00255R

Keywords:

cognition
consistency
exploration behaviour
great tit
individual variation
innovation
neophobia
Parus major
personality
problem solving

Behavioural traits generally and cognitive traits in particular are relatively understudied in an evolutionary ecological context. One reason for this is that such traits are often difficult to characterize among large numbers of individuals, without the influence of diverse environmental effects swamping intrinsic individual differences. We conducted standardized assays on a natural population of great tits, *Parus major*, to quantify and characterize individual variation in problem-solving performance, a simple cognitive trait often linked to innovative foraging ability. Forty-four per cent of 570 birds solved a food-motivated, lever-pulling problem and this proportion was consistent across three seasons. Individual performance was consistent within and across captivity sessions, across seasons, and between two different problem-solving tasks (lever and string pulling). Problem-solving performance was not explained by differences in latency to approach the empty task, nor latency to feed after human disturbance. Variation was unrelated to body condition, while age and natal origin explained significant but minimal amounts of variation, the importance of which varied between seasons. Problem-solving performance did not covary with exploration behaviour of a novel environment, suggesting that individual differences in problem solving represent an independent source of behavioural variation in our population. Rather than simply reflecting covariance with state or with other behavioural traits, our results suggest that variation in problem-solving performance represents inherent individual differences in the propensity to forage innovatively. We suggest that standardized problem-solving assays may prove ideal for studying the evolutionary ecology of simple cognitive traits.

© 2010 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

The link between cognition and ecology has received growing interest in recent years (see Healy & Braithwaite 2000; Dukas 2004; Biernaskie et al. 2009; Roth et al. 2010). Cognitive processes are likely to have considerable ecological importance because they fundamentally influence how individuals acquire and manage resources, avoid predation, invade new habitats, and maintain social relationships (Cheney et al. 1986; Curio 1988; Pravosudov & Clayton 2002; Sol et al. 2002; Turner et al. 2006; Keagy et al. 2009). Individual variation in cognitive traits is therefore likely to influence survival and reproductive success (Gailey et al. 1985; Morse & Stephens 1996; Egas & Sabelis 2001). Little is known about the underlying causes of such variation in wild populations (Roth et al. 2010). One reason for this is that behavioural traits are influenced by a great variety of environmental factors making it difficult to identify intrinsic causal factors (Merilä & Sheldon 2001). Another is that the large sample

sizes needed for the study of individual variation are often difficult to obtain, especially for cognitive traits, the measurement of which commonly requires extensive subject training or repeated trials (e.g. Weir et al. 2002; Tebbich & Bshary 2004; Heinrich & Bugnyar 2005). Here we attempted to overcome these difficulties by examining a variety of sources of individual variation in a simple cognitive trait, problem-solving performance, using rapid assays among large numbers of wild-caught great tits, *Parus major*, temporarily held in captivity under standardized conditions.

In recent years a growing number of studies have quantified repeatable or consistent behavioural differences between individuals from natural populations (Bell et al. 2009). However, we know of no similar repeatability estimates for cognitive traits. Repeatability is defined as the proportion of phenotypic variation explained by differences between individuals (Lessells & Boag 1987), which can be caused by a variety of factors: genetic, developmental and environmental (Réale et al. 2007; Bell et al. 2009). Selection experiments and common garden studies have provided considerable evidence that genes play an important role in shaping cognitive abilities (McGuire & Hirsch 1977; Lofdahl et al. 1992; Galsworthy et al. 2005; Roth et al. 2010). Developmental factors also affect cognitive performance (Vince 1958, 1960; Laland & Reader 1999; Reader & Laland 2001;

* Correspondence: E. F. Cole, Edward Grey Institute, Department of Zoology, University of Oxford, Oxford OX1 3PS, U.K.

E-mail address: eleonor.cole@zoo.ox.ac.uk (E.F. Cole).

¹ D. L. Cram is now at the Centre for Ecology and Conservation, School of Biosciences, University of Exeter, Cornwall Campus, Penryn TR10 9EZ, U.K.

Botero et al. 2009). For example, juvenile birds direct more attention to novel problems than adults (Vince 1958, 1960; Greenberg 2003) and female guppies, *Poecilia reticulata*, are more likely to forage innovatively than males (Laland & Reader 1999). In addition, environmental conditions can also influence cognitive traits (Cooper & Zubek 1958; Seibt & Wickler 2006; Arnold et al. 2007). For example, the diet of blue tit, *Cyanistes caeruleus*, nestlings has been shown to influence their spatial-learning ability as adults (Arnold et al. 2007). Social environmental factors can also be important. High cognitive performance is often more prevalent among subordinates, which frequently lose out in competition for limited resources (Laland & Reader 1999; Reader & Laland 2000), and drops in social rank can be associated with an immediate improvement in problem-solving performance (Bunnell & Perkins 1980). Empirical data therefore suggest that consistent individual differences in cognitive traits can be caused by a wide range of factors.

Individual variation in cognitive ability may also covary with other behavioural traits. Correlated suites of behaviours, also called behavioural syndromes or personality, have been documented in a wide variety of taxa (reviewed in Gosling 2001; Sih et al. 2004; Réale et al. 2007). For example, in the great tit, in contrast to 'proactive' individuals, 'reactives' are relatively neophobic (Verbeek et al. 1994), less aggressive (Verbeek et al. 1996), explore new environments more slowly and thoroughly (Verbeek et al. 1994), are more responsive to their environment (Verbeek et al. 1994), are less likely to form routines and readily explore new feeding options (Drent & Marchetti 1999). It has been suggested that because 'reactive' individuals explore their environment more thoroughly than 'proactive' individuals, and are better able to adjust their foraging behaviour when faced with environmental change, they may perform better in some cognitive tasks (Benus et al. 1987; Verbeek et al. 1994; see Sih & Bell 2008 for discussion). In contrast, it has also been proposed that bolder, fast-exploring individuals may readily acquire information about novel situations and learn novel tasks more quickly than shy, slow explorers (Sneddon 2003). Few studies have investigated cognitive performance in the context of personality, and those that have report conflicting results, so the generality of this relationship remains unclear (see Carere 2003; Sneddon 2003; Guillelte et al. 2009). Evidence from a variety of taxa suggests that neophobic individuals are less likely to engage in novel feeding situations and consequently are unlikely to solve problems or innovate (Searle 1949; Seferta et al. 2001; Webster & Lefebvre 2001; Greenberg 2003; Bouchard et al. 2007). However, several studies also demonstrate that a significant amount of variation in learning ability remains after controlling for neophobia and exploration behaviour (Galsworthy et al. 2002; Matzel et al. 2003; Bouchard et al. 2007).

We used a wild population of great tits to quantify and characterize individual variation in spontaneous problem-solving performance. Individual variation in novel problem solving is poorly understood but may reflect general cognitive ability (Roth & Dicke 2005) and innovativeness (Laland & Reader 1999; Webster & Lefebvre 2001). The great tit is a generalist species with an extensive geographical range (Gosler 1993). Great tits frequently investigate novel objects in the wild (Gibb 1957) and anecdotal evidence suggests considerable innovative foraging behaviour; examples include the opening of milk bottle tops (Fisher & Hinde 1949; Hinde & Fisher 1951), the tapping of acorns to determine the presence of hidden larvae (Ennion 1962), the use of a pine needle to extract larvae from bark (Duyck & Duyck 1984), and predation of pipistrelle bats, *Pipistrellus pipistrellus* (Estók et al. 2009). Food-related problem solving may therefore be an ecologically relevant trait for this species.

Our aims were (1) to determine whether problem-solving performance was consistent within individuals, across short and long

time periods, and between two different tasks; (2) to establish whether our assay measured goal-oriented problem-solving propensity, rather than quantifying general activity, neophobia or motivational differences; (3) to examine the relationship between a well-studied personality trait, exploration behaviour of a novel environment (henceforth, exploration behaviour), and problem-solving performance; and (4) to determine whether state variables such as sex, age and natal origin could account for observed individual differences in problem-solving performance.

METHODS

Study Site and Catching Procedure

All behavioural assays were carried out on great tits caught from a wild population in Wytham Woods (51°46'N, 1°20'W), Oxfordshire, U.K. We caught 662 individuals using mist nets at temporary feeders erected at sites throughout the wood during the winters of 2006–2007, 2007–2008 and 2008–2009, between November and March. From here onwards these three winter catching periods are referred to as seasons. Additionally, birds were caught in the final week of each season by removing them from their nestboxes after dusk. All birds were caught under ringing licences from the British Trust for Ornithology (BTO). Birds were aged (adult or juvenile if less than 1 year old) and sexed based on plumage (Svensson 1992), and biometrics (wing length, mm, and body mass, g) were taken. All unringed individuals were fitted with a unique BTO metal leg-ring. During the 2007–2008 and 2008–2009 seasons all great tits were also fitted with a unique passive integrated transponder (PIT, EM4102 tag, HID Global GmbH, Walluf, Germany) measuring 12.4 mm × 2.1 mm, which was attached to a celluloid split colour leg-ring. Up to 16 birds per catching session were transported in individual bird bags to the nearby John Krebs field station. Great tits were taken into captivity under Natural England licence.

On arrival at the field station, usually mid-afternoon, birds were housed singly in wire cages (45 × 45 cm and 68 cm high) and were visually isolated from one another. Two housing rooms were used, with a minimum of two and a maximum of eight occupied cages per room. Each cage contained three dowel perches, the unset problem-solving devices and bowls containing fresh water, husked sunflower seeds and mealworms, *Tenebrio molitor*. The birds were kept under a natural light regime using automated timers and the housing rooms were kept at ambient temperature (ranging from 3 to 16 °C) and ventilated using an air conditioning system. All birds were released within 24 h with the exception of a subsample of 80 birds during the 2007–2008 season that were kept in captivity for 4 days to carry out a battery of repeat trials and an object neophobia assay. All birds were released at the catching site.

Problem-solving Assay

All individuals (with the exception of 92 birds from the 2006–2007 season that were used in separate experiments, see below) were presented with a lever-pulling task on the afternoon they were housed ($N = 570$). The device used in this task consisted of a vertical transparent Perspex tube containing a platform, which was supported by a horizontal lever (Fig. 1a, b). The device was baited with four waxworms, *Pyralis farinalis*, placed on the platform. To solve the task, birds had to remove the lever from the device, causing the platform to drop and the waxworms to fall into a feeding dish. When the task was set, a single waxworm was placed in this dish to attract the bird to the device. This waxworm was eaten by the vast majority of birds (99% of birds, $N = 298$, ate the freely available waxworm in winter 2010). All birds were exposed to this task for approximately 3 h without being disturbed: 1 h before the aviary lights went off in

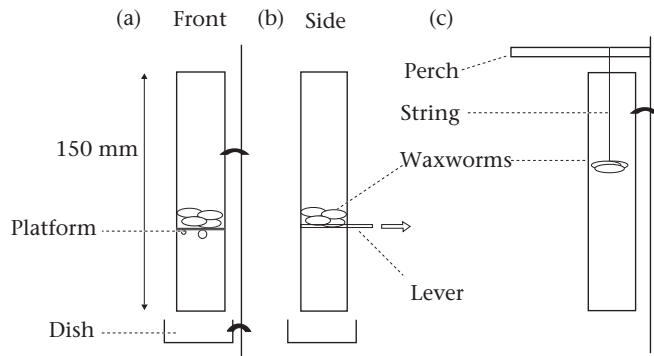


Figure 1. Problem-solving devices: (a) front view of the lever-pulling device as it appears to the bird, (b) side view and (c) string-pulling device. The lever-pulling device consisted of a transparent Perspex tube containing a platform supported by a horizontal lever. To solve the task, the bird had to remove the lever, releasing the platform and food reward. The string-pulling device consisted of a food reward suspended on the end of an 8 cm string, hung from a perch and protected by a transparent Perspex tube. To solve this task, the bird had to lift the string from above.

the evening and 2 h the morning after when the aviary lights had come on. Birds were not food deprived during this trial; however, the food reward in the device (waxworms) was highly preferable to the standard food available in the cage (E. F. Cole, personal observations).

Repeatability and Consistency

In addition to the overnight trial, subsamples of birds from the 2006–2007 and 2007–2008 seasons ($N = 51$ and 80 , respectively) and all birds caught in the 2008–2009 season ($N = 218$) were also presented with the lever-pulling task again on the day after capture to test the individual repeatability of solving performance within a captivity session. Here birds were given 90 min to solve the problem, and were food deprived for 30 min preceding the trial and for the trial duration. In 2007–2008 the birds were moved to new cages just before they were presented with the daytime task for logistical reasons; similarly, they were also given access to their roosting boxes during these trials in 2006–2007 and 2007–2008, which meant they could avoid observing the problem-solving device. These factors probably contributed to the lower solving rates for this assay in these two seasons (see *Results*).

To assess whether performance in the lever-pulling task represented a more general propensity to problem solve, all birds assayed during the 2007–2008 and 2008–2009 seasons were also presented with a string-pulling task during the afternoon of the day after capture (Fig. 1c). Here birds obtained their food reward (three waxworms) by lifting an 8 cm string from the top of a Perspex tube. Birds were food deprived for 30 min preceding the trial and during the 60 min trial itself. During the 2007–2008 season, a subsample of birds ($N = 80$) was presented with this task a second time during the second evening in captivity to assess individual repeatability of string-pulling performance.

Over the 3-year study 67 individuals were taken into captivity twice, with a minimum of 2 weeks between captivity sessions, allowing us to estimate the repeatability of performance, both within and between the two tasks over longer time periods.

Characterizing Problem-solving Performance

Activity levels

To confirm that success in our lever-pulling task was a measure of goal-oriented problem-solving propensity rather than random activity in the cage, during the 2006–2007 season a subsample ($N = 45$) was presented with the daytime lever-pulling task

(as described above) without the device being loaded with a food reward. A further subsample ($N = 47$) was presented with the normal loaded task. Both groups were experiencing the task for the first time in this trial and did not undergo any other problem-solving tests. If problem solving was motivated by food and not merely a reflection of individual activity levels, we predicted birds would primarily solve the device in the presence of food.

Neophobia

To determine whether individual variation in problem-solving performance could be explained by variation in neophobic response, we measured object neophobia among a subsample of individuals ($N = 61$) taken into captivity during the 2007–2008 season. Neophobia was assayed by comparing an individual's latency to feed with and without a novel object placed next to a food dish (modified methods from Greenberg 1984). This test was carried out on the morning before birds were released after all other assays had been completed. Normal food bowls were first removed from a bird's cage and three waxworms were placed in a dish permanently attached to the inside of the cage. The bird was then shut out of its roosting box and given 30 min to feed from the dish. Six birds did not feed during this period of habituation to the new feeding regime, and were therefore excluded from the experiment. Immediately after this initial 30 min period, the dish was refilled with three waxworms and the bird was randomly presented with either a novel object (white plastic disc, 10 cm in diameter, with six protruding, 2 cm long, coloured pipe cleaners) next to their food dish or the control treatment where only the food was added; the bird was then given 30 min to feed. Immediately after this, the final phase of the test was carried out, which involved the opposite treatment; either experimental or control. Latency to feed from the dish in the experimental and control treatments was extracted from video footage (Sony Handycam DCR-SR33E camcorder).

During the 2008–2009 season, 10 of the 16 housing cages were fitted with PIT detectors (Francis Scientific Instruments Ltd, Cambridgeshire, U.K.), positioned so that they acted as a perch in front of the lever-pulling device. All birds were fitted with PIT tags and therefore we could quantify the time it initially took individuals to approach the device when housed, a measure that we assumed reflected neophobia directed towards the device.

Exploration Behaviour

All birds underwent exploration behaviour assays on the morning after capture. Birds were released into a novel room (Verbeek et al. 1994) in which their movements were recorded for 8 min using an event recorder (Psion Workabout, with Observer mobile software, Noldus Information Technology, Wageningen, The Netherlands). Information on the number of hops, the number and duration of flights, and the areas of the room utilized were then entered into a principal components analysis, and the first component (all major loadings referred to increased activity) used in a general linear model (GLM) to produce an exploration behaviour score for each bird from the parameter estimates derived from using individual as a fixed effect, controlling for time of year (see Quinn et al. 2009 for further details).

Ethical Note

The protocols outlined were all subject to ethical review by the Department of Zoology (Oxford) ethical committee. Behaviour during captivity was closely monitored to ensure that any deterioration in health could be noticed quickly. Birds invariably began feeding shortly after they were transferred into their housing cages. One of the 662 individuals taken into captivity stopped eating but

recovered after being transferred to a heated cage and fed sugar water with a pipette. Changes in body mass between capture and release (mean \pm SE = -0.73 ± 0.033 g) were well within the great tit's natural range of diurnal mass fluctuations (Macleod et al. 2005). All birds were released at the site of capture, where there was a supply of supplemented food to allow birds to feed before dusk. The PIT tags weighed approximately 0.13 g, small compared to the normal diurnal fluctuations of great tits (1.1 ± 0.02 g; Macleod et al. 2005). This added weight results in no adverse effects in terms of individual survival or body condition (Nicolaus et al. 2008).

Statistical Analysis

At the population level, solving-rate consistency between seasons was calculated using binary two-sample tests (Z statistic estimated with a normal approximation). A binary two-sample test was also used to test whether solving performance was goal oriented. Individual repeatability in solving performance between trials and consistency across tasks was calculated using GLMs, assuming a binomial error distribution, and using individual identity as the explanatory factor, adapted from Lessells & Boag's (1987) measure of repeatability. A logistic binary regression was used to determine which effects influenced solving performance. These included fixed factors age (juvenile/adult), sex, natal origin (immigrant/resident-born), cage (1–16) and room (1/2), and the continuous variables body condition and exploration behaviour. 'Days since 1 September' (September date) was also included as a continuous variable to control for a variety of possible time-of-year effects, such as hunger-related motivation and experience. Body condition was estimated for each bird by dividing the residual from a regression of body mass against the time of day birds were caught, by its wing length. Early environmental factors such as diet (Perrins 1991) and habitat types (Blondel et al. 1993) are likely to vary between years, and therefore we included interactions between age and season, and natal origin and season. Only first tests were included for birds that were taken into

captivity multiple times. Main effect statistics were calculated including all main effects without interactions, dropping each one from the full model to test its significance; statistics for the interaction terms were estimated from full models containing all interactions. Nonparametric tests were used to test the relationships between neophobia measures and problem-solving performance. All analyses were conducted in either SPSS (SPSS Inc., Chicago, IL, U.S.A.) or Genstat (VSN International, Hemel Hempstead, U.K.).

RESULTS

Rate of Problem Solving

Forty-four per cent of birds solved the overnight lever-pulling task and this was remarkably consistent over the three seasons (Fig. 2). Fewer birds (26%) solved the daytime lever-pulling task but this was not consistent across seasons (Fig. 2). This inconsistency can be attributed to the lower daytime solving rates in 2006–2007 and 2007–2008, which we attribute to differences in methodology (see Methods). The string-pulling task was solved by 25% and this rate was consistent across seasons (Fig. 2). Therefore fewer than half of all birds solved our main problem-solving tasks, demonstrating individual variation in problem-solving propensity in our population.

Behavioural Repeatability and Consistency

Using scores from all individuals that were taken into captivity twice over the 3-year period, we found that both lever-pulling and string-pulling performances were repeatable within individuals over time (Table 1, Fig. 3a, b). Solving success in both tasks was also repeatable for individuals repeatedly assayed a year or more apart (Table 1). Again lever-pulling and string-pulling performances were consistent when we compared performance within a single captivity session (Table 1). Comparing performance between the two different

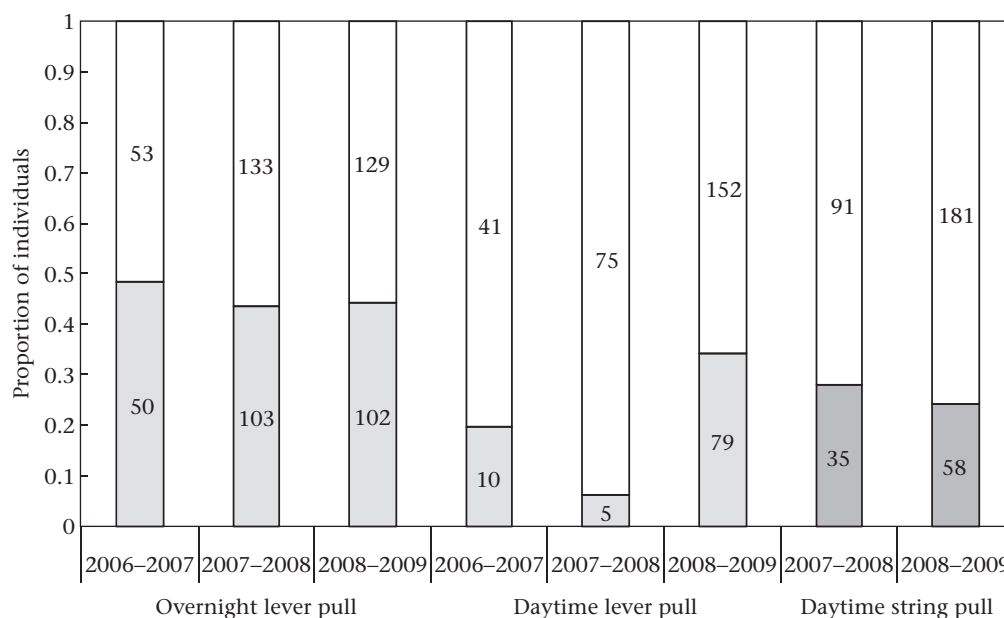


Figure 2. Proportion of individuals that solved the different tasks. Light grey represents solvers of the lever-pulling tasks, dark grey solvers of the string-pulling task and white nonsolvers. The numbers inside the bars correspond to the number of individual solvers and nonsolvers in each season. The proportion that solved the overnight task did not differ significantly between seasons: comparison of seasons 1 and 2: two-sample binomial test: $Z = 0.736$, $P = 0.462$; comparison of seasons 1 and 3: $Z = 0.744$, $P = 0.457$; comparison of seasons 2 and 3: $Z = 0.052$, $P = 0.958$. The proportion that solved the daytime lever-pulling task differed significantly between seasons: comparison of seasons 1 and 2: $Z = 2.341$, $P = 0.019$; comparison of seasons 1 and 3: $Z = -2.029$, $P = 0.042$; comparison of seasons 2 and 3: $Z = -4.623$, $P < 0.001$. The proportion that solved the string-pulling task did not differ significantly between seasons: comparison of seasons 2 and 3: $Z = 0.732$, $P = 0.464$.

Table 1
Repeatability of individual lever-pulling and string-pulling performance within and between captivity sessions

		Year	N	Repeatability	SE	P
Within captivity session	Lever-pulling performance between the overnight and daytime trials	All	347	0.34	0.05	<0.001
	String-pulling performance between the overnight and daytime trials	2007–2008	80	0.37	0.10	<0.001
Between captivity sessions	Lever-pulling performance between years	All	35	0.27	0.16	0.046
	Lever-pulling performance between captivity sessions over 3 years	All	67	0.40	0.10	<0.001
	String-pulling performance between years	All	23	0.54	0.15	<0.001
	String-pulling performance between captivity sessions over 3 years	All	47	0.37	0.13	<0.01

Performance between captivity sessions includes all repeated trials in which birds were taken into captivity twice with at least 2 weeks between sessions, whereas performance between years only includes birds taken into captivity twice with a year or more between sessions. Repeatability was estimated using methods from Lessells & Boag (1987).

tasks also showed a marked consistency in individuals' solving propensity both within and between captivity sessions (Table 2).

Characterizing Problem-solving Performance

None of the individuals presented with an empty problem-solving device solved the task, and this was significantly fewer than the number that solved the loaded task (binomial two-sample test: $Z = -3.635$, $N = 92$, $P < 0.001$).

Birds demonstrated marked consistency in the time they took to feed in the control and neophobia trials in captivity (Spearman rank correlation: $r_s = 0.360$, $N = 55$, $P = 0.007$). However, latency to feed was similar in the neophobia and control trials (Wilcoxon signed-ranks test: $T = -1.593$, $N = 55$, $P = 0.111$; neophobia: mean \pm SE = 358 ± 63.1 s; control: mean \pm SE = 467.2 ± 82.2 s). Rather than quantifying neophobia, therefore, this test provided a repeatable measure of motivation to feed after human disturbance. This measure did not significantly explain any of the observed individual variation in problem-solving performance (Mann–Whitney U test: $U = -0.726$, $P = 0.726$; solvers: mean \pm SE = 473 ± 152 s, $N = 18$; nonsolvers: mean \pm SE = 464 ± 99 s, $N = 37$). Furthermore body condition also did not explain variation in problem-solving performance (Table 3).

The time taken to approach the lever-pulling device when initially housed, which was measured in a subset of individuals during the 2008–2009 season, did not differ between solvers and nonsolvers (Mann–Whitney U test: $U = -0.624$, $P = 0.533$; solvers: mean \pm SE = 493 ± 66 s, $N = 69$; nonsolvers: mean \pm SE = 561 ± 73 s, $N = 70$). This suggests that variation in solving performance was not explained by device neophobia.

Variation in exploration behaviour was unrelated to variation in solving performance (Table 3).

State and Environmental Sources of Variation

There was no difference in solving performance between males and females (Table 3). Juveniles tended to solve the lever-pulling task more often than adults (Table 3), an effect that was present in only one of the three seasons (Table 3, Fig. 4a). Residents were also more likely to solve the task than immigrants and again this effect was present in only one of three seasons (Table 3, Fig. 4b). Birds solved the lever-pulling task more often with increasing September date (Table 3, Fig. 4c). Overall, the model explained just 10.7% of the observed variation in problem-solving success, which is small relative to the individual variation reported above.

DISCUSSION

Spontaneous problem-solving performance in our assays was goal oriented and consistent within individuals over time and between two distinct tasks. Individual variation was independent of differences in body condition, latency to approach the empty device, latency to feed after human disturbance and exploration behaviour. Males and females did not differ in their problem-solving performance and only a minimal amount of individual variation was accounted for by differences in age and natal origin. Together, these results suggest that, rather than reflecting motivational differences or covariance with other behavioural traits such as activity and exploration, the variation we characterize may largely represent inherent individual differences in the propensity or ability to solve problems related to innovative foraging.

Both lever- and string-pulling performances were relatively consistent within individuals between trials, both within captivity sessions and over long time periods, suggesting that consistent individual differences in performance exist within the population. Our levels of behavioural consistency are comparable with those of

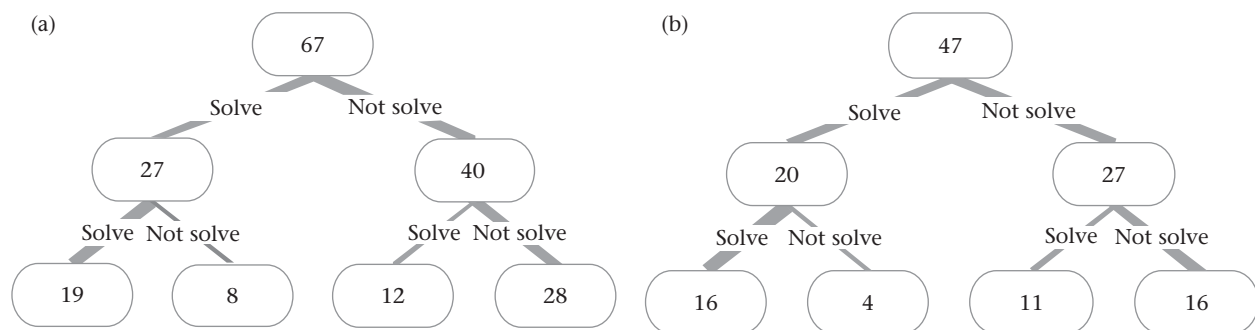


Figure 3. Decision trees showing the repeatability of problem-solving performance over time. Primary branches represent solvers and nonsolvers of the first trial and secondary branches represent solvers and nonsolvers of the same task when taken into captivity again (minimum gap between tests = 2 weeks) for (a) lever pulling and (b) string pulling. Branch width represents the proportion of individuals that solved/did not solve the task, and numbers in the boxes indicate the numbers of individuals in each category.

Table 2
Consistency of individual performance between the lever-pulling and string-pulling tasks

		Year	N	Repeatability	SE	P
Within captivity session	Daytime lever-pulling and daytime string-pulling performance	2007–2008 & 2008–2009	297	0.24	0.05	<0.001
Between captivity sessions	Overnight lever-pulling and daytime string-pulling performance, between years	All	23	0.31	0.15	0.035
	Overnight lever-pulling and daytime string-pulling performance, between captivity sessions over 3 years	All	46	0.35	0.13	0.017

Consistency was estimated using a measure of individual repeatability (adapted from Lessells & Boag 1987).

other behavioural traits (reviewed in Bell et al. 2009). Repeatability estimates for behavioural traits are generally higher when the interassay interval is short (Bell et al. 2009), but our assay was repeatable for assays conducted a year or more apart. Taken together with the fact that solving performance was not influenced by body condition, these results suggest that performance is largely independent of state factors.

By definition, goal-oriented behaviour is expected to be repeatable because successful behaviour is reinforced by a reward (Skinner 1938). However, if the individual consistency in solving performance we observed was solely the result of reinforcement we would expect consistency to reduce as the period of time elapsed between trials increased (Shettleworth 1998). We did not find support for this, as our estimates for within and between captivity sessions are comparable. In addition, we found solving performance to be relatively consistent between the two distinctly different problem-solving tasks, even when the gap between the two tasks was 2 weeks or longer. Together, these results suggest that our assay captures consistent individual variation in the ability to solve foraging-related tasks.

Birds did not solve the task in the absence of a food reward, demonstrating that solving the task was motivated by the food, rather than a result of random activity patterns. Furthermore, although our trials were preceded by human disturbance and used relatively novel devices, both of which could have influenced motivation to solve tasks, several lines of evidence suggest this is unlikely. First, we failed to detect any evidence of object neophobia using Greenberg's (1984) approach, because there was no increase in latency to feed with novel objects compared with control trials. This may be because novel objects in a novel environment can elicit a less intense response than in a familiar one (Greenberg & Mettke-Hofmann 2001). We also found no relationship between solving performance and latency to approach the problem-solving device

Table 3
Binomial logistic regression model to establish whether variation in problem-solving performance can be explained by both state and environmental factors

	Overnight lever-pulling task		
	df	W	P
Season	2	0.597	0.742
September date	1	5.265	0.022
Age	1	3.281	0.070
Sex	1	0.014	0.905
Natal origin	1	2.043	0.153
Exploration behaviour	1	0.341	0.559
Body condition	1	0.775	0.379
Test time	1	5.033	0.754
Room	1	0.240	0.624
Cage	15	21.020	0.136
Age*Season	2	6.124	0.047
Natal origin*Season	2	9.402	0.009

The model was run first without the interactions to give all the main-term statistics and then with the interactions. September date refers to days since 1 September; age is classed as either juvenile or adult, and natal origin classed as Wytham born or immigrant. *P* values significant at the 0.05 level are shown in bold. *N* = 509 individuals.

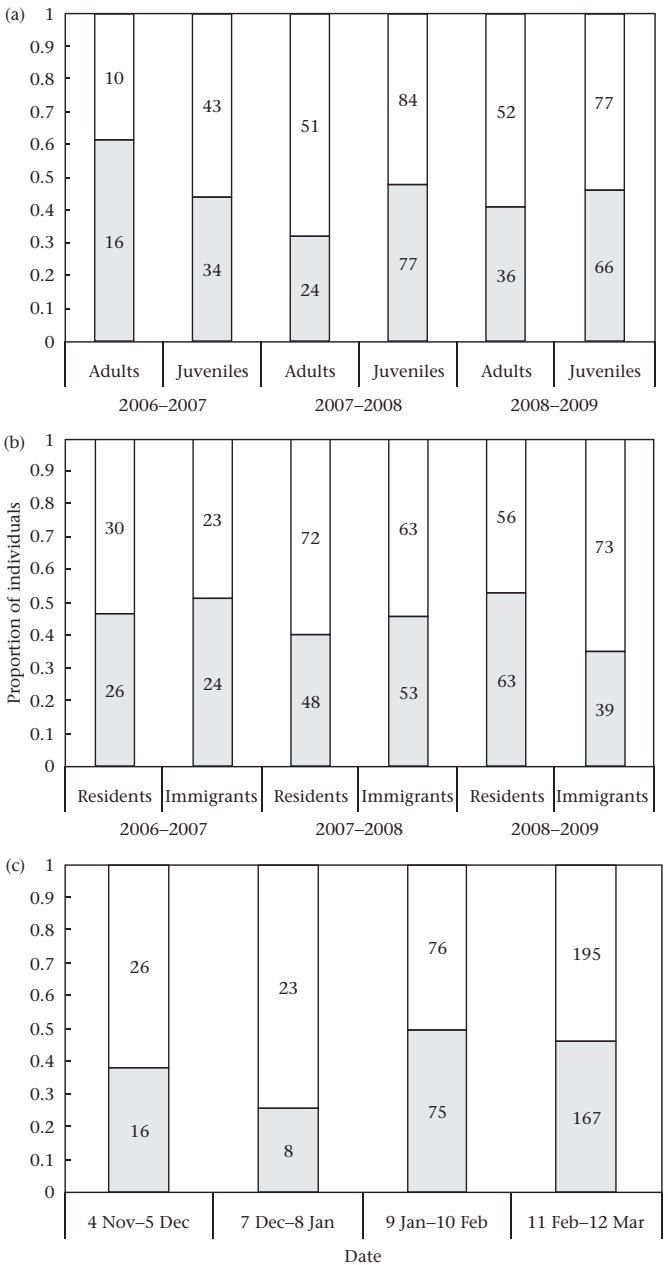


Figure 4. Factors contributing to individual variation in solving success in the overnight lever-pulling task: (a) age, (b) natal origin and (c) time of year. Light grey represents solvers of the lever-pulling tasks and white nonsolvers. See Table 3 for statistics. The numbers inside the bars correspond to the number of individual solvers and nonsolvers in each category.

when initially housed, suggesting that neophobia directed towards the task itself does not account for the variation we observed in problem solving. Neophobia, therefore, seems not to have been a factor in our system. In addition, both object neophobia (Verbeek et al. 1994) and response to stress (Carere et al. 2003; Carere & van Oers 2004) have been shown to covary with exploration behaviour in great tits but we found no link between exploration behaviour and problem-solving performance in our population. Finally, we found no relationship between solving performance and time to feed in the control trial after freely available food had been placed in the food dish. As both of these assays were preceded by human disturbance, the absence of a relationship suggests that response to disturbance did not account for the observed variation in problem solving. Even though we cannot rule out the possibility that our results might have been influenced by these possible confounding effects, our additional analyses suggest these factors are unlikely to be a major driver in problem-solving variation.

We found no support for either a positive or negative relationship between exploration behaviour and problem-solving performance. This supports previous work on the great tit that found no difference between fast and slow explorers in performance in foraging tasks requiring associative learning and spatial memory (Carere 2003). In contrast, a recent study on a related species, the black-capped chickadee, *Poecile atricapilla*, found a positive relationship between readiness to enter a novel environment and learning speed in an acoustic discrimination task (Guillette et al. 2009). It seems, therefore, that links between exploration behaviour and problem-solving performance are likely to be task and species dependent, which is unsurprising given that behavioural correlations can differ between populations of the same species (Dingemanse et al. 2007).

Great tits were more likely to solve tasks as the winter progressed. A similar time-of-year effect is observed in exploration behaviour, where individuals become 'faster' closer to the breeding season (Dingemanse et al. 2002; Quinn et al. 2009). Avian temperature decreased with September date, but we suspect it is unlikely that the change in solving rate was a result of temperature change because, throughout the winter, there was a consistent difference in temperature between the two housing aviary rooms (approximately 4 °C), which did not account for any variation in performance. Post hoc analysis shows that the effect was observed in both adults and juveniles (binary logistic regression: age*time of year: $W_2 = 0.37$, $P = 0.847$), suggesting that cumulative experience does not explain this seasonal trend. Therefore the change in solving performance observed may reflect changes in foraging behaviour or intrinsic motivational levels. Alternatively, differential overwinter survival rates of solvers and nonsolvers could be important, or variation in their catchability, resulting in solvers being more likely to be caught later in the season when conditions are harsher.

Little variation in solving performance was accounted for by differences in age and natal origin. We found a tendency for juveniles to be more likely to solve the task than adult birds, but only in some seasons. Therefore variation cannot be explained by greater curiosity among juveniles alone, a phenomenon commonly observed in birds (Vince 1958, 1960; Greenberg 2003). We also found that immigrants were more likely to solve the task in one of the three seasons. These season effects suggest that problem-solving performance may be influenced by early environmental factors that vary between years and habitat types. Several studies have found that variation in factors, such as natal diet and foraging experience, result in individual differences in cognitive abilities later in life (e.g. Cooper & Zubek 1958; Seibt & Wickler 2006; Arnold et al. 2007). Furthermore, a recent interspecific cross-fostering experiment demonstrated that early learning has a strong and lasting effect on the foraging behaviour of wild great tits (Slagsvold & Wiebe 2007). A multitude of biotic and abiotic factors influence

relative prey abundance and foraging conditions between years (Gibb 1950; Perrins 1979). In addition these factors are likely to affect different populations in different ways. For example, whereas great tits breeding in oak-rich habitats, such as Wytham, rely mainly on winter moth larvae, *Operophtera brumata*, to feed their young, birds breeding in coniferous or garden habitats are dependent on different prey species (van Balen 1973; Perrins 1979). Therefore a 'good food year' in Wytham may not be a good year for external populations, resulting in the relative behavioural phenotypes of residents and immigrants varying between years. Clearly more detailed analysis using data on natal habitat quality and fledging conditions is required to establish whether these sorts of cohort effects are occurring in our population.

The functional significance of problem-solving performance remains poorly understood. Interspecific comparative analyses have shown positive relationships between innovative foraging ability and success in both invading new habitats and dealing with seasonal environmental change (Sol et al. 2002; Lefebvre et al. 2004). At the individual level, recent work on the black-capped chickadee demonstrated that nestlings taken from a harsh environment, where food is difficult to locate, are better problem solvers than those from a favourable one (Roth et al. 2010). This suggests that the ability to problem solve may allow individuals to survive in conditions where foraging is particularly difficult. In our population, individual variation in problem solving may largely represent inherent differences in an independent behaviour trait. The relative importance of different proximate sources of variation and their functional significance remain to be determined in our population, but we suggest that the use of simple assays such as ours may provide new insight into the evolutionary ecology of cognition.

Acknowledgments

This work was supported by a Biotechnological and Biological Sciences Research Council (BBSRC) studentship to E.F.C. and a Royal Society research grant to J.L.Q. We thank J. Morand-Ferron and B. Sheldon for their comments on the manuscript, and for support in a variety of ways. The ideas in this paper were influenced by useful discussion with A. Kacelnik, who also commented on the manuscript. We also thank D. Wilson for bird husbandry and members of the EGI for help catching and assaying birds, especially S. Bouwhuis, S. Patrick, M. Wood and A. Gosler.

References

- Arnold, K. E., Ramsay, S. L., Donaldson, C. & Adam, A. 2007. Parental prey selection affects risk-taking behaviour and spatial learning in avian offspring. *Proceedings of the Royal Society B*, **274**, 2563–2569.
- van Balen, J. H. 1973. A comparative study of the breeding ecology of the great tit *Parus major* in different habitats. *Ardea*, **61**, 1–93.
- Bell, A. M., Hankison, S. J. & Laskowski, K. L. 2009. The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, **77**, 771–783.
- Benus, R. F., Koolhaas, J. M. & van Oortmerssen, G. A. 1987. Individual differences in behavioural reaction to a changing environment in mice and rats. *Behaviour*, **100**, 105–122.
- Biernaskie, J. M., Walker, S. C. & Gegeer, R. J. 2009. Bumblebees learn to forage like Bayesians. *American Naturalist*, **174**, 413–423.
- Blondel, J., Dias, P. C., Maistre, M. & Perret, P. 1993. Habitat heterogeneity and life-history variation of Mediterranean blue tits (*Parus caeruleus*). *Auk*, **110**, 511–520.
- Botero, C. A., Rossman, R. J., Caro, L. M., Stenzler, L. M., Lovette, I. J., de Kort, S. R. & Vehrencamp, S. L. 2009. Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. *Animal Behaviour*, **77**, 701–706.
- Bouchard, J., Goodyer, W. & Lefebvre, L. 2007. Social learning and innovation are positively correlated in pigeons (*Columba livia*). *Animal Cognition*, **10**, 259–266.
- Bunnell, B. & Perkins, M. 1980. Performance correlates of social behavior and organization: social rank and complex problem solving in crab-eating macaques (*M. fascicularis*). *Primates*, **21**, 515–523.
- Carere, C. 2003. Personalities as epigenetic suites of traits: a study on a passerine bird. Ph.D. thesis, University of Groningen.

- Carere, C. & van Oers, K. 2004. Shy and bold great tits (*Parus major*): body temperature and breath rate in response to handling stress. *Physiology and Behaviour*, **82**, 905–912.
- Carere, C., Groothuis, T. G. G., Möstl, E., Daan, S. & Koolhaas, J. M. 2003. Fecal corticosteroids in a territorial bird selected for different personalities: daily rhythm and the response to social stress. *Hormones and Behavior*, **43**, 540–548.
- Cheney, D., Seyfarth, R. & Smuts, B. 1986. Social relationships and social cognition in nonhuman primates. *Science*, **234**, 1361–1366.
- Cooper, R. M. & Zubek, J. P. 1958. Effects of enriched and restricted early environments on the learning ability of bright and dull rats. *Canadian Journal of Psychology*, **12**, 159–164.
- Curio, E. 1988. Cultural transmission of enemy recognition in birds. In: *Social Learning: Psychological and Biological Perspectives* (Ed. by T. R. Zentall & B. G. Galef, Jr.), pp. 75–98. Hillsdale, New Jersey: L. Erlbaum.
- 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**, 929–938.
- Dingemanse, N. J., Wright, J., Kazem, A. J. N., Thomas, D. K., Hickling, R. & Dawman, N. 2007. Behavioural syndromes differ predictably between 12 populations of three-spined stickleback. *Journal of Animal Ecology*, **76**, 1128–1138.
- Drent, P. J. & Marchetti, C. 1999. Individuality, exploration and foraging in hand raised juvenile great tits. In: *Proceedings of the 22nd International Ornithological Congress, Durban* (Ed. by N. J. Adams & R. H. Slotow), pp. 896–914. Johannesburg: BirdLife South Africa.
- Dukas, R. 2004. Evolutionary biology of animal cognition. *Annual Review of Ecology, Evolution, and Systematics*, **35**, 347–374.
- Duyck, I. & Duyck, J. 1984. Koolmees, *Parus major*, gebruikt instrument bij het voedselzoeken. *Wielewaal*, **50**, 416.
- Egas, M. & Sabelis, M. W. 2001. Adaptive learning of host preference in a herbivorous arthropod. *Ecology Letters*, **4**, 190–195.
- Ennion, E. H. 1962. Notes on birds seen in Aden and the western Aden Protectorate. *Ibis*, **104**, 560–562.
- Estók, P., Zsebok, S. & Siemers, B. M. 2009. Great tits search for, capture, kill and eat hibernating bats. *Biological Letters*, **6**, 59–62.
- Fisher, J. & Hinde, R. A. 1949. The opening of milk bottles in birds. *British Birds*, **42**, 347–357.
- Gailey, D. A., Hall, J. C. & Siegel, R. W. 1985. Reduced reproductive success for a conditioned mutant in experimental populations of *Drosophila melanogaster*. *Genetics*, **111**, 795–804.
- Galsworthy, M. J., Paya-Cano, J. L., Monleón, S. & Plomin, R. 2002. Evidence for general cognitive ability in heterogeneous stock mice and an analysis of potential confounds. *Genes, Brain and Behaviour*, **1**, 88–95.
- Galsworthy, M. J., Paya-Cano, J., Liu, L., Monleón, S., Gregory, G., Fernandes, C., Schalkwyk, L. & Plomin, R. 2005. Assessing reliability, heritability and general cognitive ability in a battery of cognitive tasks for laboratory mice. *Behavior Genetics*, **35**, 675–692.
- Gibb, J. A. 1950. The breeding biology of the great and blue titmice. *Ibis*, **92**, 507–539.
- Gibb, J. A. 1957. Food requirements and other observations on captive tits. *Bird Study*, **4**, 207.
- Gosler, A. 1993. *The Great Tit*. London: Hamlyn Press.
- Gosling, S. D. 2001. From mice to men: what can we learn about personality from animal research? *Psychological Bulletin*, **127**, 45–86.
- Greenberg, R. 1984. Differences in feeding neophobia in the tropical migrant wood warblers *Dendroica castanea* and *D. pensylvanica*. *Journal of Comparative Psychology*, **98**, 131–136.
- Greenberg, R. 2003. The role of neophobia and neophilia in the development of innovative behaviour. In: *Animal Innovation* (Ed. by S. M. Reader & K. N. Laland), pp. 175–196. Oxford: Oxford University Press.
- Greenberg, R. & Mettke-Hofmann, C. 2001. Ecological aspects of neophobia and neophilia in birds. In: *Current Ornithology*, Vol. 16 (Ed. by V. Nolan & C. F. Thompson), pp. 119–178. New York: Plenum.
- Guillette, L. M., Reddon, A. R., Hurd, P. L. & Sturdy, C. B. 2009. Exploration of a novel space is associated with individual differences in learning speed in black-capped chickadees, *Poecile atricapillus*. *Behavioural Processes*, **82**, 265–270.
- Healy, S. & Braithwaite, V. 2000. Cognitive ecology: a field of substance? *Trends in Ecology & Evolution*, **15**, 22–26.
- Heinrich, B. & Bugnyar, T. 2005. Testing problem solving in ravens: string-pulling to reach food. *Ethology*, **111**, 962–976.
- Hinde, R. A. & Fisher, J. 1951. Further observations on the opening of milk bottles by birds. *British Birds*, **44**, 393–396.
- Keagy, J., Savard, J. & Borgia, G. 2009. Male satin bowerbird problem-solving ability predicts mating success. *Animal Behaviour*, **78**, 809–817.
- Laland, K. N. & Reader, S. M. 1999. Foraging innovation in the guppy. *Animal Behaviour*, **57**, 331–340.
- Lefebvre, L., Reader, S. M. & Sol, D. 2004. Brains, innovations and evolution in birds and primates. *Brain, Behavior and Evolution*, **63**, 233–246.
- Lessells, C. M. & Boag, P. T. 1987. Unrepeatable repeatabilities: a common mistake. *Auk*, **104**, 116–121.
- Lofdahl, K. L., Holliday, M. & Hirsch, J. 1992. Selection for conditionability in *Drosophila melanogaster*. *Journal of Comparative Psychology*, **106**, 172–183.
- McGuire, T. R. & Hirsch, J. 1977. Behavior-genetic analysis of *Phormia regina*: conditioning, reliable individual differences, and selection. *Proceedings of the National Academy of Sciences, U.S.A.*, **74**, 5193–5197.
- Macleod, R., Gosler, A. G. & Cresswell, W. 2005. Diurnal mass gain strategies and perceived predation risk in the great tit *Parus major*. *Journal of Animal Ecology*, **74**, 956–964.
- Matzel, L. D., Han, Y. R., Grossman, H., Karnik, M. S., Patel, D., Scott, N., Specht, S. M. & Gandhi, C. C. 2003. Individual differences in the expression of a 'general' learning ability in mice. *Journal of Neuroscience*, **23**, 6423–6433.
- Merilä, J. & Sheldon, B. C. 2001. Avian quantitative genetics. *Current Ornithology*, **16**, 179–255.
- Morse, D. H. & Stephens, E. G. 1996. The consequences of adult foraging success on the components of lifetime fitness in a semelparous, sit and wait predator. *Evolutionary Ecology*, **10**, 361–373.
- Nicolaus, M., Bouwman, K. M. & Dingemanse, N. J. 2008. Effect of PIT tags on the survival and recruitment of great tits *Parus major*. *Ardea*, **96**, 286–292.
- Perrins, C. M. 1979. *British Tits*. London: Collins.
- Perrins, C. M. 1991. Tits and their caterpillar food supply. *Ibis*, **133**, 49–54.
- Pravosudov, V. V. & Clayton, N. S. 2002. A test of the adaptive specialization hypothesis: population differences in caching, memory, and the hippocampus in black-capped chickadees (*Poecile atricapilla*). *Behavioural Neuroscience*, **116**, 515–522.
- 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.
- Reader, S. M. & Laland, K. N. 2000. Diffusion of foraging innovations in the guppy. *Animal Behaviour*, **60**, 175–180.
- Reader, S. M. & Laland, K. N. 2001. Primate innovation: sex, age and social rank differences. *International Journal of Primatology*, **22**, 787–805.
- Réale, D., Reader, S. M., Sol, D., McDougall, P. T. & Dingemanse, N. J. 2007. Integrating animal temperament within ecology and evolution. *Biological Reviews*, **82**, 291–318.
- Roth, G. & Dicke, U. 2005. Evolution of the brain and intelligence. *Trends in Cognitive Sciences*, **9**, 250–257.
- Roth, T. C., LaDage, L. D. & Pravosudov, V. V. 2010. Learning capabilities enhanced in harsh environments: a common garden approach. *Proceedings of the Royal Society B*, **277**, 3187–3193.
- Searle, L. V. 1949. The organization of hereditary maze-brightness and maze-dullness. *Genetic Psychology Monographs*, **39**, 279–325.
- Seferta, A., Guay, P., Marzinotto, E. & Lefebvre, L. 2001. Learning differences between feral pigeons and zenaïda doves: the role of neophobia and human proximity. *Ethology*, **107**, 281–293.
- Seibt, U. & Wickler, W. 2006. Individuality in problem solving: string pulling in two carduelis species (Aves: Passeriformes). *Ethology*, **112**, 493–502.
- Shettleworth, S. 1998. *Cognition, Evolution and Behavior*. New York: Oxford University Press.
- Sih, A. & Bell, A. M. 2008. Insights for behavioral ecology from behavioral syndromes. *Advances in the Study of Behavior*, **38**, 227–281.
- Sih, A., Bell, A. M. & Johnson, J. C. 2004. Behavioural syndromes: an integrative overview. *Quarterly Review of Biology*, **79**, 241–277.
- Skinner, B. F. 1938. *The Behavior of Organisms*. New York: Appleton Century Crofts.
- Slagsvold, T. & Wiebe, K. L. 2007. Learning the ecological niche. *Proceedings of the Royal Society B*, **274**, 19–23.
- Sneddon, L. U. 2003. The bold and the shy: individual differences in rainbow trout. *Journal of Fish Biology*, **62**, 971–975.
- Sol, D., Timmermans, S. & Lefebvre, L. 2002. Behavioural flexibility and invasion success in birds. *Animal Behaviour*, **63**, 495–502.
- Svensson, L. 1992. *Identification Guide to European Passerines*. Thetford: British Trust for Ornithology.
- Tebbbich, S. & Bshary, R. 2004. Cognitive abilities related to tool use in the woodpecker finch, *Cactospiza pallida*. *Animal Behaviour*, **67**, 689–697.
- Turner, A. M., Turner, S. E. & Lappi, H. M. 2006. Learning, memory and predator avoidance by freshwater snails: effects of experience on predator recognition and defensive strategy. *Animal Behaviour*, **72**, 1443–1450.
- 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.
- Verbeek, M. E. M., Boon, A. & Drent, P. J. 1996. Exploration, aggressive behaviour and dominance in pair-wise confrontations of juvenile male great tits. *Behaviour*, **133**, 945–963.
- Vince, M. A. 1958. 'String-pulling' in birds. (2). Differences related to age in greenfinches, chaffinches and canaries. *Animal Behaviour*, **6**, 53–59.
- Vince, M. A. 1960. Developmental changes in responsiveness in the great tit (*Parus major*). *Behaviour*, **15**, 219–243.
- Webster, S. J. & Lefebvre, L. 2001. Problem solving and neophobia in a columbiform–passeriform assemblage in Barbados. *Animal Behaviour*, **62**, 23–32.
- Weir, A. A. S., Chappell, J. & Kacelnik, A. 2002. Shaping of hooks in New Caledonian crows. *Science*, **297**, 981.