Behavioral Ecology

The official journal of the ISBE International Society for Behavioral Ecology

Behavioral Ecology (2014), 25(1), 124–135. doi:10.1093/beheco/art094

Original Article

Necessity or capacity? Physiological state predicts problem-solving performance in house sparrows

Veronika Bókony,[a] Ádám Z. Lendvai,[b,c] Csongor I. Vágási,[d,e] Laura Pătraş,[d] Péter L. Pap,[d] József Németh,[f] Ernő Vincze,[a] Sándor Papp,[a] Bálint Preiszner,[a] Gábor Seress,[a] and András Liker[a,g]

[a]Department of Limnology, University of Pannonia, Pf. 158, H-8200 Veszprém, Hungary, [b]Department of Biology, College of Nyíregyháza, Sóstói út 31/b, H-4400 Nyíregyháza, Hungary, [c]Department of Biology, Virginia Tech, 4102 Derring Hall, Blacksburg, VA 24060-0406, USA, [d]Evolutionary Ecology Group, Hungarian Department of Biology and Ecology, Babeş-Bolyai University, Clinicilor street 5–7, RO-400006 Cluj-Napoca, Romania, [e]MTA-DE "Lendület" Behavioural Ecology Research Group, Department of Evolutionary Zoology, University of Debrecen, Egyetem tér 1, H-4032 Debrecen, Hungary, [f]Department of Pharmacology and Pharmacotherapy, University of Debrecen, Nagyerdei krt. 98, H-4029 Debrecen, Hungary, and [g]Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

Received 25 February 2013; revised 7 September 2013; accepted 15 September 2013; Advance Access publication 19 October 2013

Innovative behaviors such as exploiting novel food sources can grant significant fitness benefits for animals, yet little is known about the mechanisms driving such phenomena, and the role of physiology is virtually unexplored in wild species. Two hypotheses predict opposing effects of physiological state on innovation success. On one hand, poor physiological condition may promote innovations by forcing individuals with poor competitive abilities to invent alternative solutions. On the other hand, superior physiological condition may ensure greater cognitive capacity and thereby better problem-solving and learning performance. To test these hypotheses, we studied the behavior of wild-caught house sparrows (*Passer domesticus*) in 4 novel tasks of food acquisition, one of which was presented to the birds in repeated trials, and we investigated the relationships of individual performance with relevant physiological traits. We found that problem-solving performance across the 4 tasks was moderately consistent within individuals. Birds with lower integrated levels of corticosterone, the main avian stress hormone, solved the most difficult task faster and were more efficient learners in the repeated task than birds with higher corticosterone levels. Birds with higher concentration of total glutathione, a key antioxidant, solved 2 relatively easy tasks faster, whereas birds with fewer coccidian parasites tended to solve the difficult task more quickly. Our results, thus, indicate that aspects of physiological state influence problem-solving performance in a context-dependent manner, and these effects on problem-solving capacity, probably including cognitive abilities, are more likely to drive individual innovation success than necessity due to poor condition.

Key words: **animal innovation, coccidiosis, learning, oxidative status, stress physiology.**

INTRODUCTION

Innovations are new or modified learned behaviors such as applying established or new behavioral patterns or tools to solve novel challenges or familiar problems in a novel way (Reader and Laland 2003). Such processes may have vast evolutionary significance as they may allow animals to colonize new habitats, exploit new resources, and cope with environmental change (Reader and Laland 2003; Ramsey et al. 2007). These implications recently generated growing interest by behavioral ecologists, resulting in a proliferation of studies on innovative behaviors in various animal taxa (e.g., Liker and Bókony 2009; Keagy et al. 2011; Morand-Ferron et al. 2011; Overington et al. 2011; Benson-Amram and Holekamp 2012; Cole and Quinn 2012; Sol et al. 2012; Thornton

Address correspondence to V. Bókony. E-mail: vbokony@almos.uni-pannon.hu.

and Samson 2012). Despite this recent research effort, we are still only at the start of understanding the factors and mechanisms that lead certain individuals to innovate.

Several hypotheses have been put forward to explain why and how innovations arise in nature. The "necessity drives innovations" hypothesis (Reader and Laland 2003) proposes that individuals with poor competitive abilities are forced to develop novel solutions because they are outcompeted from prevailing ways of resource acquisition. This idea has been supported by a number of empirical studies that found juveniles or subordinates more innovative than adults or dominants, respectively (e.g., Biondi et al. 2010; Morand-Ferron et al. 2011; Cole and Quinn 2012; Thornton and Samson 2012), whereas others found no effect of age and rank (Bouchard et al. 2007; Keagy et al. 2011; Benson-Amram and Holekamp 2012) or reported more innovations by dominants (Boogert et al. 2006, 2008). Other hypotheses emphasize the role of differences in the capacity, rather than the necessity, of individuals for innovative problem solving. For example, innovative abilities may be determined by cognitive skills such as the capacity for learning and reasoning (Reader and Laland 2003). In line with this suggestion, the largest number of field reports (Lefebvre et al. 2004) and perhaps the most compelling cases (Seed et al. 2009) of animal innovations come from large-brained avian and primate species. These hypotheses of necessity and capacity are not mutually exclusive, and each predicts that individuals may differ consistently in their propensity to innovate, as has been found for some species (e.g., Laland and Reader 1999; Cole et al. 2011), although individual consistency may not be present across all contexts (e.g., Morand-Ferron et al. 2011; Sol et al. 2012).

One plausible yet largely overlooked influence on problem-solving performance is the physiological state of individuals, which may affect both necessity and capacity for innovative behaviors in several ways. On one hand, the necessity hypothesis predicts that because competitive ability may be directly linked to actual physiological condition and health, individuals in poor condition may be more motivated to innovate (Laland and Reader 1999). This prediction is scarcely supported by indices of energy reserves (Laland and Reader 1999; Boogert et al. 2008; Cole et al. 2011; Overington et al. 2011; Morand-Ferron et al. 2011; Thornton and Samson 2012), but it is yet to be tested for more direct measures of physiological state and health. On the other hand, the capacity hypothesis predicts that cognitive skills are important for innovations; thus, physiological effects that shape the ontogeny and function of brain regions and thereby cognitive capacity may influence innovation success. First, glucocorticoid hormones released in response to stress have a complex effect on brain function and development depending on life-history stage and the type and intensity of stressor (reviewed by McEwen and Sapolsky 1995; Lupien et al. 2009). Greater stress typically results in reduction of cognitive performance and neuron loss in various brain regions involved in cognitive processes, although moderate short-term stress in adults can enhance hippocampus-mediated learning and memory ("hormetic effect"; Pravosudov 2003; Lupien et al. 2009). Second, the brain is also sensitive to oxidative stress (von Schantz et al. 1999; Barja 2004), which occurs when the accumulation of oxidative agents exceeds the organism's ability to mitigate them by antioxidants. As uncontrolled oxidants can degrade biomolecules and cause cellular oxidative damage, oxidative stress can lead to neurodegeneration and thus impair learning and cognition (e.g., Liu et al. 2003; reviewed by von Schantz et al. 1999). Increased oxidative stress and decreased antioxidant levels are also believed to play an important role in senescence and associated cognitive declines (reviewed by Barja 2004; Dröge and Schipper 2007). Third, both hormonal and oxidative stress can be triggered by parasitic infections (Lindström et al. 2005; Raouf et al. 2006; Sepp et al. 2012a), and both brain development and fighting off infectious diseases are metabolically costly. Therefore, the "parasite-stress hypothesis of intelligence" (Eppig et al. 2010) proposes that cognitive function is affected negatively by parasite load. In line with this idea, cognitive output often declines with greater intensity of infectious diseases (e.g., Kavaliers et al. 1999; Gegear et al. 2005; Eppig et al. 2010).

The above information comes mostly from studies on humans and laboratory rodents, whereas the physiological background of innovation and cognitive performance is almost unexplored in wild species (Pfeffer et al. 2003; Gegear et al. 2005). The present study is a beginning step to unravel the role of physiological variation in innovative behaviors of nonhuman, not captive-bred animals. To this end, we studied wild-caught individuals of the house sparrow (*Passer domesticus*), a passerine species that is known to innovate both in its free-living populations (e.g., Breitwisch and Breitwisch 1991; Suárez-Rodríguez et al. 2013) and in the laboratory (Liker and Bókony 2009). We observed whether and how the birds' performance in 4 novel food-extracting tasks is related to 4 relevant physiological aspects: 1) integrated levels of corticosterone, the main avian stress hormone; 2) markers of antioxidant capacity and oxidative damage; 3) infestation by coccidians that are intestinal parasites that inhibit the uptake of essential dietary components and can have significant negative effects on condition, physiology, and behavior in sparrows (Dolnik and Hoi 2010; Pap et al. 2011, 2013); and 4) body condition index reflecting the amount of energy reserves. We predicted that if innovations are promoted by necessity, individuals in inferior physiological condition such as low levels of antioxidants, high levels of oxidative stress, more parasites, and lower body condition index will perform better. On the other hand, if innovations are promoted by capacity, such as cognitive skills, then individuals in superior physiological state, that is, those with lower levels of stress hormones, parasite load and oxidative stress, and higher levels of antioxidants should be more successful.

METHODS

Protocol

Between 2 January and 6 March 2012, we captured 10–14 house sparrows each week at various localities in Hungary (n = 104 birds in total; 50 males and 54 females) using mist-nets (Ecotone, Gdynia, Poland). On capture, we measured body mass (±0.1 g) and tarsus length (±0.1 mm). Birds were brought into captivity and housed for 2 weeks in individual cages (53 × 27 × 41 cm) each containing 2 perches and a vertical plastic sheet hanging from the top of the cage as shelter. We provided ad libitum food (a mixture of millet, wheat, oat, and sunflower seeds) and tap water with multivitamin droplets throughout the study, except for the duration of tests and the preceding fasting periods. During all fasting periods and problem-solving tests, birds were visually isolated from each other by opaque plastic sheets.

After 1–2 days of acclimation, birds participated in 4 problem-solving tests as follows. On days 3–5, birds were food deprived each morning between 8:00 and 9:30; then they were presented with a food-extracting task (i.e., closed feeder) and their behavior was observed by a single person through a 1-way window between 9:30 and 11:00. Each task involved a different feeder; each feeder was placed open into the cage as the sole source of food during the

day preceding the respective test to familiarize the birds with each feeder (i.e., from 11:00 till next morning 8:00). After the third test, birds were weighed and transferred into another room where they were allowed to acclimate for 2 days. On days 8–12, we presented a subset of birds (n = 72) with a fourth task repeatedly in a total of 19 trials (the rest of the birds were untested because they participated in another experiment). There were four 30-min trials each day (excepting the last day) between 9:00 and 16:00, each preceded by 60-min fasting and followed by 15 min when feeder 4 was fixed in opened position, so the birds were allowed to feed from it. The behavior of the birds was observed during each trial by another person. In each test, observers recorded for each bird the latency to first attempt of problem solving (i.e., time elapsed from the start of test until the bird first manipulated the feeder with its beak), and latency to solve (i.e., time elapsed from the start of test until the bird started to feed).

Sample size differs between tests because 7 birds (6.7%) died during the study; this rate of mortality is similar to that found in other studies of captive sparrows (e.g., Liker and Bókony 2009; Bókony et al. 2012a; Pap et al. 2013). Although we did not see any sign of disease, birds that died might have been physiologically stressed already at capture (as suggested by their high malondialdehyde [MDA] levels; see Results). Food deprivation never exceeded 3 consecutive hours per day, and birds always were allowed to feed before and after the lights-off period for at least half an hour. All captures and keeping of birds were in accordance with Hungarian laws and licensed by the Middle Transdanubian Inspectorate for Environmental Protection, Natural Protection and Water Management (permission number: 31559/2011).

Problem-solving tests

All birds were presented with the 4 food-extracting tasks in the same order so that performance can be directly compared among individuals for each task (see Sol et al. 2011; Bókony et al. 2012a). The 4 feeders used in the tasks are shown in Figure 1. Feeder 1 was an 8.5- × 8.5- × 2.5-cm transparent plastic box with a 2.5-cm hole on the top; this hole was uncovered on the day before test 1 but covered by a transparent plastic card, fixed by 2 wooden sticks, during test 1. To reach the seeds, birds had to pull out one or both sticks and toss the card away, or pull the card upwards until it came off the sticks. Feeder 2 was a 7.5-cm diameter, 3.5-cm high transparent plastic dish that was covered by white bakery paper on the top, fixed by sticky tape on the sides, during test 2. Birds' only way of accessing the food was piercing the paper with the beak. Feeder 3 was an 11-cm high commercial bird feeder with a slot cut into it at about 8 cm height; during test 3, a small transparent plastic card was placed into this slot to keep the seeds from falling down. To have the seeds flow out, birds had to remove the card by pulling it out with their beak; some birds achieved this by heavily shaking the feeder. Feeder 4 was an 8.5- × 8.5- × 2.5-cm white plastic box with a transparent side and a lid on the top. Before and between the trials of task 4, the lid was held open by placing a small transparent cup into the feeder. During the trials, the lid was closed, and birds had to insert their beak and head under it and push it up to reach the food. Thus, in contrast to the first 3 tasks, feeder 4 did not remain open after the bird first fed from it; instead, it had to be opened every time to peck a seed.

Markers of oxidative physiology

At capture, we took a blood sample into heparinized capillaries (~100 µL) by brachial venipuncture from each bird that we could free from the net and bleed within 20 min (handling time: mean ± standard error [SE] = 9.42 ± 0.66 min; n = 33). Samples were kept on ice in a dark cooler box for a few hours, and then centrifuged at 8000 rpm for 5 min in the laboratory to separate plasma and packed cell fractions. Both fractions were stored at –20 °C until analyses. We measured the following 4 biochemical markers

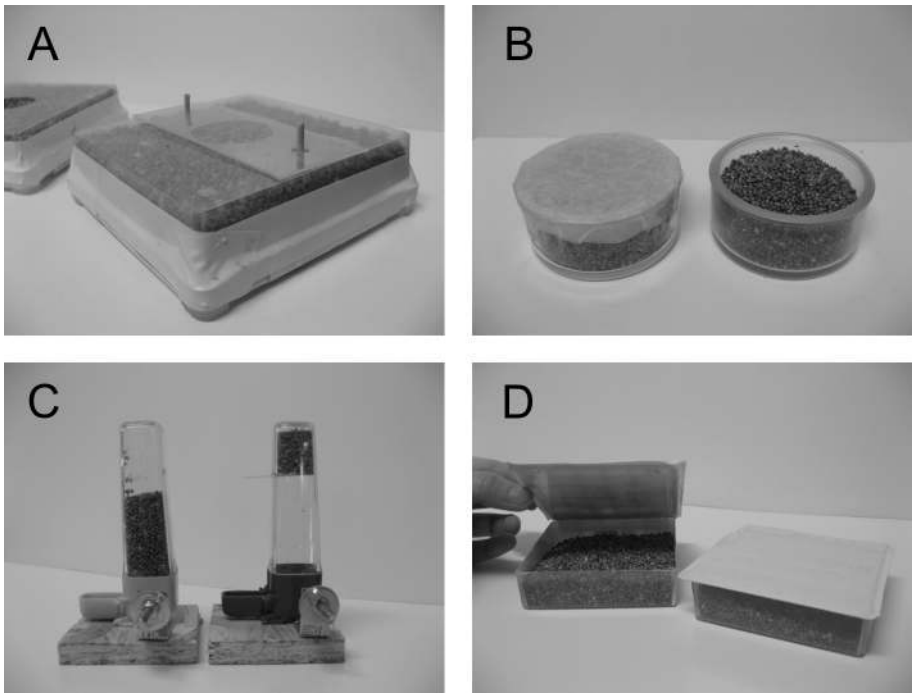

Figure 1
Feeders used in the problem-solving tasks. (A) Feeder 1 closed in the front, open in the back. (B) Feeder 2 closed on the left, open on the right. (C) Feeder 3 open on the left, closed on the right. (D) Feeder 4 open on the left, closed on the right.

of oxidative homeostasis: total antioxidant status (TAS), uric acid (UA), and MDA concentrations from plasma, and total glutathione (tGSH) concentration from erythrocytes. Neither measurement varied significantly with handling time ($P > 0.16$). Due to the small amount of blood samples available from sparrows, we did not run duplicate assays, but a subset of samples was tested twice for repeatability (see below). Absorbance in TAS, UA, and tGSH assays was measured with an automated plate reader (FLUOstar Omega, BMG Labtech, Germany). Note that markers of oxidative physiology are usually repeatable in birds over time intervals comparable with the 1–10 days between blood sampling and various behavioral tests in our study (Norte et al. 2008; Galván and Alonso-Alvarez 2009; Sepp et al. 2010, 2012b).

TAS is a composite measure of antioxidant capacity, expressing the cumulative ability of all nonenzymatic antioxidants found in plasma, such as vitamins, sulfhydryl groups of proteins, and uric acid, to combat a simulated free radical insult. The TAS assay was based on a commercial kit (Cayman Chemical, Ann Arbor, MI). Plasma (5 µL) was diluted 1:10 (v/v) in assay buffer and mixed with chromogen ABTS (2,2′-azinobis-[3-ethylbenzthiazoline-6-sulfonate]), metmyoglobin, and hydrogen peroxide (H_2O_2), and was incubated for 5 min at room temperature. On H_2O_2 addition, the oxidation of ABTS by metmyoglobin leads to the production of the radical cation ABTS+, which generates a blue-green color. Antioxidants in the plasma samples inhibit the oxidation of ABTS, causing the suppression of absorbance at 750 nm to a degree proportional to their concentration (Rice-Evans and Miller 1994). This cumulative antioxidant activity is compared with that of Trolox, a water-soluble tocopherol analogue (standard calibration curve, $R^2 = 0.98$), and is expressed as millimoles per liter Trolox equivalents. Repeatability of a subsample measured twice was moderate but significant (intraclass correlation coefficient: ICC = 0.54, 95% confidence interval [CI] = 0.03–0.83, $F_{12,13} = 3.37$, $P = 0.019$). Because UA is a component of TAS and a product of amino acid catabolism, we controlled TAS for UA levels by ordinary least squares regression and calculated residual TAS as suggested by Cohen et al. (2007).

Plasma UA concentration was determined spectrophotometrically by an uricase/peroxidase method using a commercial kit (Uric Acid liquicolor, Human, Wiesbaden, Germany). This test provides an enzymatic colorimetric determination of UA and eliminates the falsely elevated results generated by the turbidity of lipemic specimens through the lipid-clearing factor. Briefly, we added 250 µL of reagent mix to 5 µL of plasma and incubated this solution at 37 °C for 5 min. The absorbance was read at 520 nm. Results are given as mg/dL plasma. Repeatability of duplicate measures was very high (ICC = 0.99, 95% CI = 0.97–0.99, $F_{16,17} = 153$, $P < 0.001$).

MDA is a carbonyl compound that results from the peroxidative degeneration of membrane polyunsaturated fatty acids by reactive oxygen species, and thus, it is a widely used marker of oxidative stress (Del Rio et al. 2005). Plasma MDA concentration was assessed by high-performance liquid chromatography (HPLC), instead of TBARS (thiobarbituric acid reactive substances) assay, due to its higher precision and reliability (Del Rio et al. 2005). We followed the fast and sensitive method employed by Karatas et al. (2002) with adaptation to small plasma volume of birds as per Noguera et al. (2011). Briefly, 10 µL of plasma was mixed with 50 µL of 0.1 M perchloric acid and 90 µL dH_2O (i.e., 1:15 v/v dilution). This solution was spun at 4500 × g for 5 min, and 100 µL from the supernatant was injected in the HPLC (SUPELCOSIL™ LC-18 column, 5 µm particle size; Sigma-Aldrich) with UV detection at 254 nm (Jasco, UV-2075 Plus, Japan). The mobile phase was 30 mM monopotassium phosphate–methanol (65:35, v/v %), and the flow rate was 0.5 mL/min. The retention time of MDA recorded was around 6 min. MDA concentration in the sample was determined using a calibration curve ($R^2 = 0.99$) of a series of standards generated by acidic hydrolysis of 1,1,3,3-tetraethoxypropane (Sigma-Aldrich). Results are given as µg/mL plasma and are not corrected for the above 1:15 dilution factor. The repeatability of the subsample measured twice was very high (ICC = 0.97, 95% CI = 0.91–0.99, $F_{14,15} = 62.2$, $P < 0.001$).

Glutathione (GSH) is the most significant intracellular, endogenous, nonenzymatic antioxidant (Galván and Alonso-Álvarez 2008). Total GSH concentration was assayed by means of a commercial kit (Sigma-Aldrich, St Louis, MO) and according to Galván and Alonso-Álvarez (2008) and Hõrak et al. (2010) with modifications. After thawing on ice, the erythrocyte pellet was washed 3 times with phosphate-buffered saline and spun at 600 × g for 10 min at 4 °C. After that, because the pellet had different volumes between individuals, we weighed it (±0.001 mg) and deproteinized with 5% 5-sulfosalicylic acid (SSA; 1:1 w/v), so for example, a 100-mg pellet was diluted in 100 µL of 5% SSA. This solution was vigorously vortexed, kept on ice for 10 min, and then centrifuged at 10 000 g for 10 min and at 4 °C to remove the precipitated proteins. The supernatant (5 µL) was pipetted to another test tube, diluted 10-fold and used subsequently for tGSH detection, according to the manufacturer's instructions. Total GSH determination is based on a kinetic assay in which nanomoles of GSH cause a continuous reduction of 5,5′-dithiobis(2-nitrobenzoic acid) to 5-thio-2-nitrobenzoic acid (TNB) and the GSH oxidized to GSSG is subsequently recycled by glutathione reductase and NADPH. The yellow color of the TNB product is proportional to the GSH concentration, and its absorbance was measured spectrophotometrically at 412 nm at 1 min intervals for 5 min. The change in absorbance was compared with that of a standard curve ($R^2 = 0.99$) generated by serial dilution of reduced GSH. Results are given in nanomoles per milligram of pellet. The repeatability of the subsample measured twice was high (ICC = 0.82, 95% CI = 0.53–0.94, $F_{13,14} = 9.88$, $P < 0.001$).

Parasite load

We collected fecal samples during 2 consecutive days, after test 1 and test 2, from $n = 66$ birds to estimate coccidian infestation (Pap et al. 2009). Samples were weighed (±0.001 g), diluted in 1 mL distilled water, vortexed gently to homogenize, and stored at 4 °C until analysis, which took place within 1 week after collection. The number of coccidian oocysts was counted as described by Pap et al. (2009), and concentration was expressed as number of oocysts per gram feces. Oocyst concentration was highly repeatable (recounts of $n = 22$ samples: ICC = 0.994, 95% CI = 0.987–0.998, $F_{21,22} = 358$, $P < 0.001$) and correlated strongly between the 2 days (Spearman rank correlation: $r_s = 0.74$, $P < 0.001$, $n = 64$). We used the log-transformed mean values of the oocyst concentrations of 2 days for each individual in further analyses.

Corticosterone assay

At the end of test 3, from each bird we plucked the 2 outermost tail feathers and stored them at room temperature in paper envelopes until analysis (only 1 feather was analyzed per bird). Feather corticosterone concentrations were measured by radioimmunoassay (RIA) following a methanol-based extraction using the method of Bortolotti et al. (2008). Briefly, the calamus of the feathers was cut, and the remaining part of the feather was minced into pieces less than 2 mm^2 into a test tube. The total mass of the cut feather

fragments was weighed to the nearest 0.1 mg. We then added 5 mL methanol to the feather fragments, the solution was placed in a sonicating water bath for 30 min, and then incubated for at least 20 h in a heated shaker (at 50 °C). The methanol was removed from the feather particles by filtration through a syringe filter (PTFE filter with 0.2 µm pore size, VWR). The tube with the remaining feather particles was washed and filtered twice again with additional 2 × 2.5 mL volumes of methanol. The methanol was then evaporated under a fume hood at room temperature under a current of air. The extraction was reconstituted with PBS buffer used in the RIA. We used a commercial antiserum, raised in rabbits against corticosterone-3-(O-carboxymethyl) oxime bovine serum albumin conjugate (Sigma-Aldrich, St Louis, MO; product number: C8784). The reconstituted extracts were incubated for 48 h at 4 °C with 100 µL of [3H]corticosterone (Perkin Elmer, product number: NET399250UC) and antiserum. The total volume of the assay was 1 mL. The radioactively labeled corticosterone had an activity of cca. 10 000 dpm. Bound and free corticosterone were separated by adding 100 µL dextran-coated charcoal. After centrifugation, the 800 µL of the bound fraction was added to 6 mL of scintillation cocktail (Optima Gold, Perkin Elmer) and counted in a liquid scintillation counter (Tri-carb 2800TR, Perkin Elmer). The minimum detectable level of corticosterone was 3.90 pg/tube (our lowest measurement was 101.7 pg/tube, i.e., much higher than the detection threshold). Corticosterone levels were corrected for sample mass by dividing the measured corticosterone concentration by feather mass; we refer to this variable as CORT henceforth. CORT showed a positive correlation with feather mass ($r = 0.30$, $P = 0.003$, $n = 97$), which disappeared after excluding 5 samples with feather mass below the 5% percentile (<10 mg) or above the 95% percentile (>16 mg; $r = 0.13$, $P = 0.203$, $n = 92$). Therefore, we repeated all analyses of CORT and problem solving by excluding these 5 outliers. CORT was log-transformed in all analyses.

Body condition index

After the 19th trial of test 4, we measured the birds' body mass for the third time. To quantify body condition as body mass relative to body size, we calculated the scaled mass index following Peig and Green (2009). This index adjusts the mass of all individuals to that which they would have if they had the same body size, using the equation of the linear regression of log mass on log length estimated by type-2 SMA regression. We used the equation of Bókony et al. (2012b) that was derived from the data of >2000 house sparrows: body mass index = body mass × $(19/\text{tarsus length})^{1.71}$, where 19 is mean tarsus length, and 1.71 is the slope of the SMA regression between log-mass and log-tarsus length.

Statistical analyses

We used chi-square (χ^2) tests to compare the proportion of successful individuals between various tasks. We tested whether individual performance was repeatable across different tasks by calculating the ICC; we rank-transformed solving latencies to meet the statistical requirements of the test (Nakagawa and Schielzeth 2010). We applied the same approach to test the repeatability of attempt latencies.

We analyzed the relationships between physiological variables and problem-solving latency in each task using Cox proportional hazards models, a nonparametric survival-analysis method (Sol et al. 2011). We used the latency to solve as dependent variable in each model, expressed in tasks 1–3 as the time (in minutes) elapsed from the start of the test until first feeding and in task 4 as the number of trials needed to first open the feeder. Birds that did not solve a task were assigned a value one unit higher than the maximal latency (i.e., 93 min in tasks 1–3, and 20 trials in task 4) for the respective task and were treated as censored observations in the analyses. We included the latency to first attempt as covariate in each model to control for any individual difference in motivation or emotionality (sensu Sol et al. 2012). For task 4, average attempt latency was calculated as the mean of attempt latencies across trials until the first successful trial. Because Cox analyses model the probability of not solving as a function of test time, positive parameter estimates mean shorter latencies (i.e., faster decrease of probability of not solving during the test), whereas negative parameter estimates mean longer latencies.

To investigate the efficiency of learning, we compared the time needed to solve task 4 for the first and second time (i.e., in the first 2 successful trials), following Thornton and Samson (2012). Because attempt latency decreased significantly from the first to the second successful trial (on average by 4.35 ± 1.25 min, paired t-test: $t_{59}=3.48$, $P < 0.001$), to control for this effect, in each trial we calculated solving time as the time between the first attempt to open the feeder and the subsequent successful solution of the task (i.e., eating seed from the feeder). The change in solving time from the first to the second successful trial was used as a proxy for learning efficiency, that is, individuals that reduced their solving time to a greater extent were considered more effective in recalling and processing the information obtained during their first successful trial. This variable (i.e., difference in solving time between first and second successful trials) was used to test the correlation of learning efficiency with problem-solving latencies. To analyze the relationships of learning efficiency with physiological variables, solving time was used as dependent variable in a generalized linear mixed-effect model with quasi-Poisson error distribution (log-link function), containing individual as random factor, the rank of successful trial (first or second) as fixed factor, and the interactions of trial rank with physiological variables. Note that the latter analyses concern only those individuals that solved task 4 at least twice ($n = 60$); we did not include more than 2 successful trials because solving time showed little variation among individuals after the second successful trial (see Results).

In all analyses of problem-solving latencies and learning efficiency, the initial models also included the potentially confounding effects of sex, date, tarsus length (as measure of body size), acclimation time (i.e., time spent in captivity before test 1, in days), and their interactions with the physiological variables. Then we omitted nonsignificant explanatory variables from the models stepwise, and we report the final models that contain significant effects only ($P < 0.05$; note that we never omitted attempt latency). Models containing tGSH or residual TAS were re-ran including MDA levels to control for the extent of oxidative damage to tease out if high antioxidant levels can be interpreted as measures of superior condition or are associated with high oxidative damage (i.e., poor condition). All analyses were done in R 2.15.1 (R Core Team 2012). All statistical tests are 2-tailed with a 95% CI. Mean values are reported with ±SE.

RESULTS

Problem-solving and learning performance

Out of the 4 tasks, 3 were solved by the majority of birds, whereas a significantly lower proportion of individuals was successful in task 2 (Tables 1 and 2). There was considerably longer time available for solving task 4 (30 min × 19 trials; 570 min in total) than the rest (90 min per task); during the first 30-min trial, task 4 appeared similarly difficult to the birds (19.4%, 14 out of 72 were successful) as task 2 (Table 2), whereas by the third trial (i.e., after 90 min in total),

Table 1
Summary statistics of performance variables

Variable (unit)	Minimum	Median	Maximum	*n*	% Solved
Task 1 solving latency (min)	3	24	93	104	79.8
Task 2 solving latency (min)	12	93	93	101	22.8
Task 3 solving latency (min)	3	36	93	98	72.4
Task 4 solving latency (number of trials)	1	4	20	72	83.3
Task 4 learning efficiency (min)	−15	3	21	60	n/a

For solving latencies, 93 min and 20 trials designate birds that did not solve the respective task. Learning efficiency was calculated as the difference in solving time (i.e., time elapsed from first attempt to first feeding) between the first and second successful trial in task 4. n/a, not applicable.

Table 2
Comparisons of performance across problem-solving tasks

Variable	Task 1 solving latency	Task 2 solving latency	Task 3 solving latency	Task 4 solving latency	Task 4, 1st trial	Task 4, 3rd trial
Task 1 solving latency	—	**64.49 (<0.001)**	1.13 (0.288)	0.15 (0.695)	**60.25 (<0.001)**	**21.96 (<0.001)**
Task 2 solving latency	0.15 (0.143; 101)	—	**47.28 (<0.001)**	**59.37 (<0.001)**	0.11 (0.735)	**8.13 (0.004)**
Task 3 solving latency	**0.21 (0.041; 98)**	0.10 (0.314; 98)	—	2.20 (0.138)	**44.55 (<0.001)**	**12.48 (<0.001)**
Task 4 solving latency	0.18 (0.137; 72)	0.17 (0.155; 72)	**0.28 (0.017; 72)**	—	—	—
Task 4 learning efficiency	0.004 (0.978; 60)	−0.02 (0.887; 60)	0.20 (0.130; 60)	−0.04 (0.776; 60)	—	—

Values above the diagonal are χ^2 statistics (df = 1), with *P* value in brackets, from χ^2 tests comparing the proportion of successful individuals between tests. Values below the diagonal are coefficients of Spearman rank correlations between performance variables, with *P* value and sample size in brackets. Significant results are highlighted in bold. See Table 1 for units of measurement.

32 out of 72 birds (44.4%) solved task 4, which is higher than the success rate in task 2 but lower than in tasks 1 and 3 (Table 2).

Individuals that solved task 3 sooner were also faster in task 1 and took less trials to first solve task 4 (Table 2). Although the latency to solve task 2 was not significantly consistent with the speed of solving other tasks (Table 2), those few birds that solved task 2 ($n = 23$) were significantly faster than unsuccessful birds in task 1 (independent-sample *t*-test: $t_{47.28} = 2.23$, $P = 0.030$) and task 4 ($t_{56.78} = 3.73$, $P < 0.001$), but there was no such difference in task 3 ($t_{30.98} = 0.95$, $P = 0.351$). Thus, repeatability of performance across the 4 tasks was weak but significant (ICC = 0.17, 95% CI = 0.06–0.30, $F_{71,216} = 1.83$, $P < 0.001$).

Out of the 60 birds that solved task 4 over the 19 trials, 54 opened the feeder in every trial following their first solution, and 5 birds were unsuccessful in 1 or 2 consecutive trials after their first or second successful trial but performed consistently afterwards (1 bird died after its second successful trial). Solving time decreased significantly from the first successful trial to the second one (on average by 4.5 ± 0.94 min, paired *t*-test: $t_{59} = 4.78$, $P < 0.001$); in the following trials, the majority of birds started to feed immediately (Figure 2). Learning efficiency in task 4 was not correlated with problem-solving speed in any of the 4 tasks (Table 2).

Physiological traits and attempt latency

Body mass index was highly repeatable during the study (ICC = 0.78, 95% CI = 0.71–0.84, $F_{96,192} = 11.8$, $P < 0.001$), whereas it was unrelated to most physiological measurements, except that birds with higher CORT lost more weight after capture (Pearson correlation: $r = 0.28$, $P = 0.007$, $n = 96$) and therefore had lower body mass index in captivity than birds with lower CORT (Table 3). CORT levels and parasite load were not correlated with each other or with any measure of oxidative status, except for tGSH, which showed a positive trend with CORT (Table 3). Residual TAS and MDA correlated positively but neither showed significant relationship with tGSH (Table 3). Out of the 7 birds that died during the study, we had blood samples for only 2; both of them had higher MDA concentration (4.09 and 4.59 μg/mL, respectively) than the upper quartile of MDA levels for birds that did not die (3.98 μg/mL).

Attempt latency was weakly but significantly repeatable over the 4 tasks (ICC = 0.15, 95% CI = 0.04–0.28, $F_{71,216} = 1.69$, $P = 0.002$) and showed no significant relationship with the physiological variables in all but one case (Table 3).

Problem solving and physiology

With the exception of task 3, birds that made their first attempt later took significantly more time to access food (Table 4). CORT had opposing significant effects in 2 tasks (Table 4): birds with higher CORT solved task 2 later (Figure 3A), but they needed less trials to first solve task 4 than birds with lower CORT (Figure 3B). When we excluded 5 samples with very low or high feather mass, the former effect remained significant and the latter did not (Table 4). Total GSH concentration had consistent significant effect in the other 2 tasks: birds with higher tGSH were faster in tasks 1 and 3 (Table 4, Figure 3C,D); these relationships remained qualitatively unchanged when MDA level was included in the models (Table 4). Birds that solved task 2 faster had significantly fewer coccidian parasites (Cox model: $b \pm SE = -0.48 \pm 0.21$, e^b [95% CI] = 0.62 [0.41–0.94], $P = 0.024$; Figure 3E); however, when parasite load was included in the final model of task 2, its effect on solving latency was not significant ($P = 0.235$). Body mass index was unrelated to performance in most tasks; although leaner birds solved task 3 sooner (Table 4), this relationship became marginally nonsignificant when MDA level was included in the model (Table 4).

Learning efficiency and physiology

Birds with higher CORT reduced their solving time (i.e., time elapsed from first attempt to opening the feeder) between their first

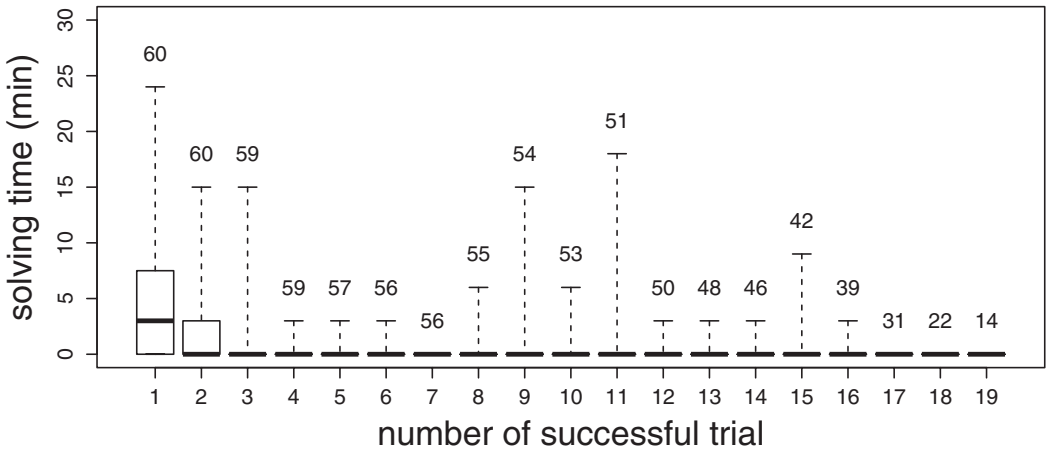

Figure 2
Solving time (i.e., time elapsed from first attempt to first feeding) in successful trials in task 4. Boxplots show the median (thick line), interquartile range (box), and data range (whiskers). Numbers above the whiskers show sample sizes (i.e., number of birds that solved task 4 in respective number of trials; for example, only 14 birds were successful in all 19 trials). Note that the difference in solving time between the first 2 successful trials expresses learning efficiency.

2 successful trials to a lesser extent, that is, they learned more slowly than birds with lower CORT (Table 4, Figure 3F). Parameter estimates indicate that this difference was driven mainly by the second trial, that is, CORT had no significant effect on solving time in the first successful trial but birds with higher CORT had longer solving time in the second successful trial (Table 4). Excluding 5 birds with extreme feather mass did not change this result qualitatively (Table 4). Despite their lower CORT levels, birds with higher body mass index did not learn faster than leaner birds ($P = 0.102$). Parasite load and oxidative status showed no relationship with learning efficiency ($P > 0.18$).

DISCUSSION

Our study showed that problem-solving and learning performance is related to several ecologically relevant physiological traits in a wild species. Propensity for problem solving exhibited individual consistency across some but not all tasks, and the efficiency of learning was not correlated with solving success in any task. In parallel with this, the effect of physiological traits was also context specific, as performance in different contexts was predicted by different measures of individual condition. In most cases, however, birds in better physiological state were more successful, supporting the capacity hypothesis that innovations are likely driven by the ability of individuals to come up with and learn novel solutions.

From our results, we may infer that the 4 tasks can be divided in 2 groups: task 1 and task 3 were relatively easy for the birds (72–80% successful), whereas task 2 and initially task 4 were more difficult (19–23% successful). This is so probably because tasks 1 and 3 both could be solved by several techniques, and birds could remain in visual contact with the food during attempts that led to the solution (e.g., pecking at the card in feeder 3). In contrast, the food in feeder 2 was visible only from side view, but pecking at the side of the feeder did not result in problem solving as birds had to peck at the nontransparent paper at the top to reach the seeds. Thus, in order to solve task 2, birds probably had to inhibit ineffective behaviors stimulated by the sight of food. Similarly in task 4, birds had to realize that lifting the lid by their beak is not effective because they had to insert their head below the lid to be able to peck seeds. This notion that tasks 2 and 4 were cognitively more challenging than tasks 1 and 3 is paralleled by our finding that performance in these 2 groups of tasks was differentially related to physiological variables.

In the easier tasks, problem-solving performance was consistently positively correlated with levels of tGSH, a key antioxidant. These relationships remained significant when we controlled for MDA levels, suggesting that they were due to superior antioxidant capacity and not to increased oxidative damage. We speculate that tasks 1 and 3 could be solved by mere perseverance (sensu Thornton and Samson 2012) and/or simple cognitive functions such as responding to obstacle movement cues (Overington et al. 2011). These processes might be specifically related to tGSH levels or some associated trait, for example, activity or age (Metcalfe and Alonso-Alvarez 2010). Alternatively, tGSH levels may have a more general relationship with problem-solving performance, which we might have failed to detect in tasks 2 and 4 due to power limitations (e.g., we had tGSH data for only 5 birds that solved task 2). In either case, the positive relationship between tGSH levels and innovation success is in accordance with the findings that oxidative stress induces neural damage and cognitive dysfunction in young rats (Song et al. 2009; Boksa 2010) and age-related decrease in neural performance parallels that of GSH decrease in humans (Barja 2004; Dröge and Schipper 2007). It remains unclear, however, why plasma TAS and the level of MDA showed no association with problem solving. Although markers of oxidative physiology are often repeatable within individuals both in the wild and in captivity (Norte et al. 2008; Galván and Alonso-Alvarez 2009; Sepp et al. 2010, 2012b), some studies suggest that GSH levels are more consistent over time than TAS (Galván and Alonso-Alvarez 2009; Sepp et al. 2010; but see Sepp et al. 2012b). Because we took blood samples a few days before the tests, the latter might explain why we detected an effect of tGSH but none of TAS. Our results, thus, raise the intriguing possibility that intracellular antioxidants such as tGSH might be better predictors of innovation success than plasma antioxidants and plasma measures of oxidative damage, which is to be tested by future experiments.

In the more difficult tasks, the effects of physiological traits were more complex. Two results indicated that birds with higher CORT levels were less successful, as they solved task 2 later or not at all, and they were less efficient learners in task 4, that is, after their

Table 3
Correlations of physiological variables and attempt latencies

Variable	CORT level	Parasite load	Residual TAS	MDA level	tGSH level	Body mass index prior to the test
Parasite load[a]	0.15 (0.259; 62)	—	—	—	—	—
Residual TAS[a]	0.28 (0.160; 27)	0.23 (0.308; 21)	—	—	—	—
MDA level[a]	−0.02 (0.934; 29)	−0.13 (0.569; 21)	**0.37 (0.044; 30)**	—	—	—
tGSH level[a]	0.37 (0.051; 28)	−0.15 (0.514; 22)	0.12 (0.524; 29)	0.10 (0.614; 29)	—	—
Body mass index at capture[a]	0.01 (0.952; 97)	−0.11 (0.368; 64)	−0.07 (0.732; 30)	0.32 (0.074; 32)	0.06 (0.740; 31)	—
Body mass index after test 3[a]	**−0.21 (0.041; 97)**	−0.03 (0.809; 62)	−0.02 (0.931; 28)	0.17 (0.382; 30)	<0.01 (0.985; 29)	—
Body mass index after test 4[a]	**−0.20 (0.048; 96)**	−0.09 (0.476; 62)	−0.12 (0.527; 28)	0.07 (0.723; 30)	−0.07 (0.708; 28)	—
Task 1 attempt latency[b]	−0.08 (0.416; 97)	−0.10 (0.410; 64)	−0.02 (0.905; 30)	−0.11 (0.555; 32)	−0.16 (0.395; 31)	0.12 (0.238; 104)
Task 2 attempt latency[b]	−0.13 (0.207; 97)	0.23 (0.062; 64)	−0.27 (0.156; 29)	0.23 (0.205; 31)	0.04 (0.814; 30)	0.08 (0.435; 101)
Task 3 attempt latency[b]	−0.07 (0.482; 97)	−0.02 (0.847; 62)	**−0.44 (0.019; 28)**	−0.29 (0.116; 30)	−0.31 (0.102; 29)	−0.12 (0.257; 98)
Task 4 mean attempt latency[b]	−0.18 (0.134; 72)	−0.15 (0.328; 45)	−0.14 (0.518; 23)	0.22 (0.291; 24)	−0.38 (0.069; 24)	0.20 (0.100; 72)

Correlation coefficients are shown with *P* value and sample size in brackets. Significant results are highlighted in bold.
[a]Pearson correlations.
[b]Spearman rank correlations.

Table 4
Problem-solving speed and learning efficiency in relation to various physiological variables and the latency to first attempt

	Model set A			Model set B[a]		
Model	$b \pm$ SE	e^b [95% CI]	*P*	$b \pm$ SE	e^b [95% CI]	*P*
Task 1, solving latency (30; 28)[b]						
Attempt latency	−0.03 ± 0.01	0.96 [0.94, 0.99]	0.007	−0.03 ± 0.01	0.97 [0.94, 0.99]	0.020
tGSH	0.42 ± 0.16	1.53 [1.11, 2.12]	0.010	0.45 ± 0.16	1.57 [1.13, 2.19]	0.007
MDA	—	—	—	0.48 ± 0.28	1.62 [0.93, 2.82]	0.086
Task 2, solving latency (97; 92)[b]						
Attempt latency	−0.05 ± 0.02	0.95 [0.92, 0.99]	0.006	−0.05 ± 0.02	0.95 [0.92, 0.99]	0.007
Acclimation time	−0.91 ± 0.40	0.40 [0.16, 0.98]	0.046	−0.84 ± 0.43	0.43 [0.18, 1.05]	0.065
log(CORT)	−1.33 ± 0.26	0.26 [0.10, 0.71]	0.008	−1.30 ± 0.27	0.27 [0.10, 0.73]	0.009
Task 3, solving latency (28; 26)[b]						
Attempt latency	−0.01 ± 0.01	0.99 [0.96, 1.01]	0.341	−0.01 ± 0.01	0.99 [0.96, 1.01]	0.339
Body mass index	−0.25 ± 0.12	0.78 [0.61, 0.99]	0.044	−0.22 ± 0.13	0.80 [0.62, 1.04]	0.092
tGSH	0.55 ± 0.20	1.75 [1.17, 2.61]	0.006	0.53 ± 0.21	1.71 [1.13, 2.58]	0.011
MDA	—	—	—	0.12 ± 0.22	1.13 [0.72, 1.76]	0.592
Task 4, number of trials (72; 67)[b]						
Mean attempt latency	−0.09 ± 0.01	0.91 [0.88, 0.94]	0.000	−0.09 ± 0.01	0.91 [0.88, 0.94]	<0.001
log(CORT)	0.53 ± 0.25	1.71 [1.04, 2.79]	0.033	0.44 ± 0.27	1.56 [0.91, 2.67]	0.107
Learning efficiency (60; 56)[c]						
Intercept (1st trial)	2.72 ± 0.94	15.29 [2.40, 97.47]	0.006	3.62 ± 1.03	37.61 [5.65, 284.27]	0.001
Trial (2nd)	−5.54 ± 1.51	0.004 [0.0002, 0.08]	0.001	−5.98 ± 1.71	0.003 [0.0001, 0.07]	0.001
log(CORT)	−0.31 ± 0.27	0.73 [0.43, 1.25]	0.253	−0.60 ± 0.30	0.55 [0.30, 0.99]	0.053
Trial (2nd) × log(CORT)	1.18 ± 0.41	3.26 [1.45, 7.34]	0.006	1.28 ± 0.48	3.63 [1.41, 9.35]	0.010

[a]Final models of Model set A were repeated with including MDA levels for tasks 1 and 3, and excluding CORT samples with extreme feather mass for tasks 2 and 4 (see Methods).
[b]Cox's proportional hazards models. Exponentially transformed parameter estimates (e^b) show the proportional change of hazard ratio, that is, the probability of solving the task, in response to unit change of predictors. Numbers in brackets are sample sizes in model sets A and B.
[c]Generalized linear mixed-effect model with quasi-Poisson error distribution. Exponentially transformed parameter estimates (e^b) show the proportional change of solving time between the first 2 successful trials in response to unit change of predictors. Numbers in brackets are sample sizes in model sets A and B.

first successful solution they needed similar or even longer time to open the feeder than for the first time. In contrast, birds with higher CORT made their first solution sooner during repeated exposure to task 4, although the latter relationship was less robust. Because we measured the amount of stress hormones that had been deposited into tail feathers over the prolonged period of molt several months before the problem-solving tests, any relationship between CORT and performance indicates that our hormonal

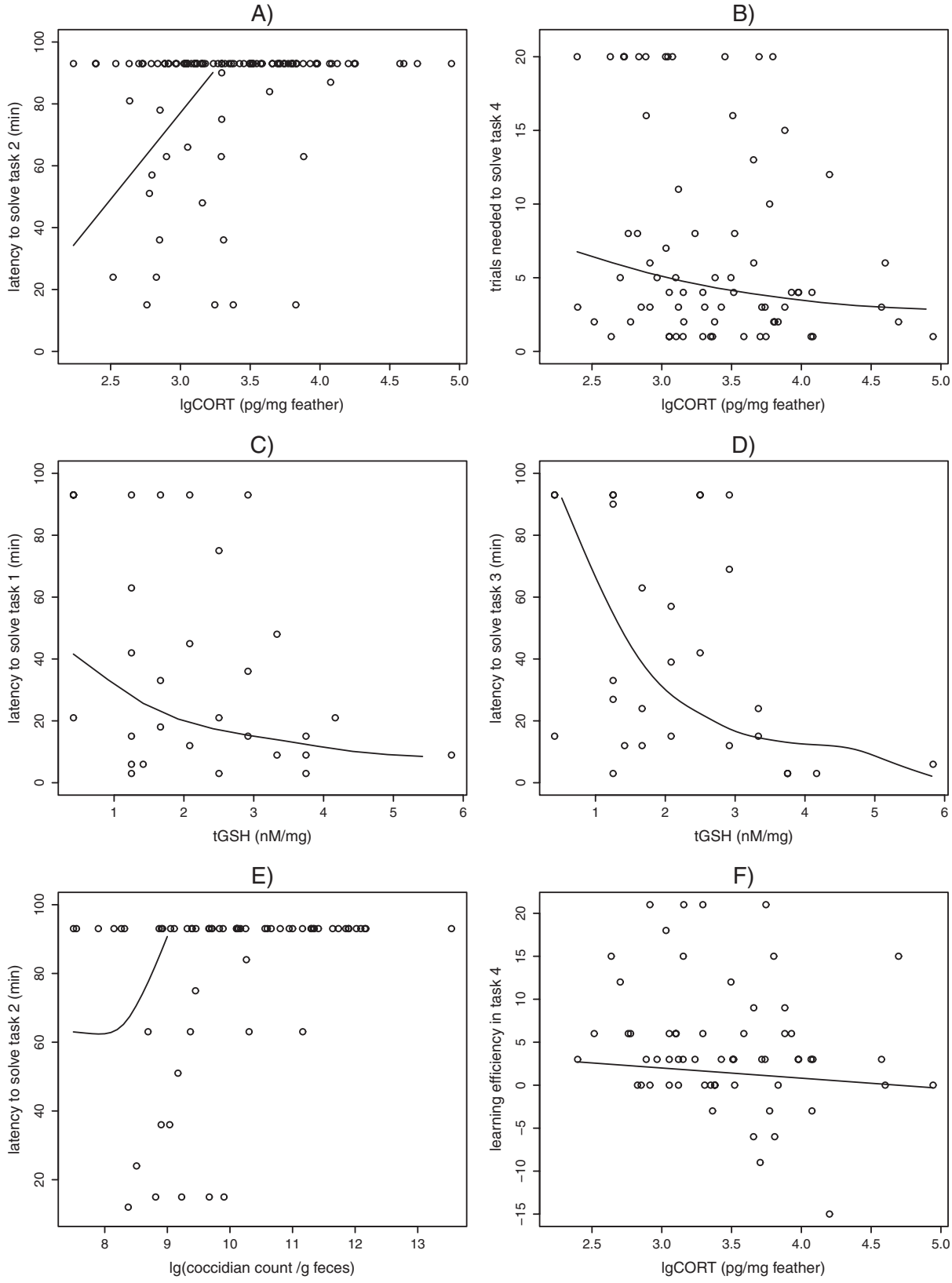

Figure 3
Problem-solving performance and learning efficiency in relation to corticosterone concentration in tail feathers (CORT), total glutathione concentration (tGSH), and abundance of coccidian oocysts in feces. Learning efficiency is illustrated as the change in solving time (i.e., time elapsed from first attempt to first feeding) from the first to the second successful trial; positive values mean that solving time decreased from first to second successful trial. Curves show predicted values from Cox's proportional hazards models (A–E) and mixed-effect model with quasi-Poisson error (F).

measurement captured an aspect of stress physiology that exhibits long-term individual consistency (Cockrem et al. 2009; Rensel and Schoech 2011). Birds with higher CORT in their feathers may have been exposed to stressors more often if they lived in a more challenging environment or had an increased metabolic rate during the molt. Because feather hormone concentration is an integrated measure of hormone levels over an extended period, it is not possible to differentiate between short-term spikes in hormone levels and chronic moderate elevation of baseline levels (except for large feathers allowing sectional analysis; Bortolotti et al. 2009). Alternatively, high CORT may have arisen by individuals producing stronger and/or more frequent responses to stressors during molt, indicating that they were more sensitive to stress. This is supported by our result that individuals with higher CORT lost more weight in captivity, which is likely a stressful condition for wild birds (Martin et al. 2012). By this logic, our findings mean that less stress-resistant individuals were less successful in task 2, which was probably our cognitively most demanding task requiring enhanced inhibitory control of inappropriate attempts and providing no obstacle movement cues. The contrasting tendency in one aspect of task 4, that is, that birds with higher CORT needed less trials to first open the feeder, might have been spurious due to a few outlier samples; although it is similar to the previous finding that greylag geese (*Anser anser*) that became innovators had higher fecal CORT concentrations (Pfeffer et al. 2003). That finding was interpreted as a result of the enhancing effect of mild acute stress on memory consolidation processes (Pfeffer et al. 2003; Lupien et al. 2009). However, this scenario is not compatible with our result that birds with higher CORT showed weaker performance in the other aspect of task 4, namely learning efficiency. On average, 4.45 trials were needed to invent the technique of opening feeder 4; after that the birds quickly mastered this technique; and by their third successful trial, they used it consistently and rapidly at the start of each trial. The efficiency of learning, measured as the decrease in solving time between the first 2 successful trials, was greater in individuals with lower CORT levels. This result fits well with the findings on humans that high levels of glucocorticoid hormones impair working memory and suppress the ability to filter out irrelevant information (reviewed by McEwen and Sapolsky 1995). Thus, our results for task 2, as well as for learning efficiency, agree with the general pattern that stress reduces cognitive capacity (Lupien et al. 2009).

Birds with lower coccidian infestation tended to solve task 2 faster, which is in line with the prediction of the parasite-stress hypothesis of intelligence (Eppig et al. 2010) that parasite load reduces cognitive performance. Although the relationship between parasite load and solving latency was not significant when controlled for attempt latency and CORT, this might have been due to the low number of successful birds in task 2 and the restricted sample size for coccidian counts. We can think of 2 possible explanations as for why parasite load showed an effect exclusively in task 2. First, because fecal samples were collected before and after task 2, it is possible that our parasite counts were representative only for this period; however, the strong correlation between the 2 sampling days agrees with the earlier finding that coccidian infestation is highly consistent within individuals in captive house sparrows (Pap et al. 2013). Second, because task 2 appeared the most difficult for our birds, we might speculate that coccidian infestation influences performance only when the latter requires demanding cognitive effort. For example, mice subclinically infected with coccidian parasites display reduced spatial learning, which is believed to be a side effect of the host's immunological and neuromodulatory responses to infection such as opioid neuropeptides and cytokines that decrease cognitive performance (Kavaliers et al. 1995).

Our study provided little support for the necessity hypothesis, because performance was not enhanced by poor physiological condition as measured by low antioxidant levels, high oxidative damage, and high parasite load. Also, performance was not related to body mass index in any situation excepting task 3, suggesting that actual energetic state is a weak predictor of problem solving and learning, in accordance with several recent studies on other species (Pfeffer et al. 2003; Boogert et al. 2008; Cole et al. 2011; Morand-Ferron et al. 2011; Overington et al. 2011; Thornton and Samson 2012). Although we tried to standardize the amount of energy reserves across birds by providing ad libitum food between tests and standardized fasting periods before tests, individuals still varied almost as greatly in body mass index during the tests (range: 9.99 g) as at capture (range: 11.37 g). This variation was likely due to individual differences in stress sensitivity, as birds with more CORT lost more weight and became leaner in captivity than birds with less stress hormones. Although this correlation would predict a positive effect of CORT on problem-solving performance via hunger levels, we did not find such an effect in any but one case (task 4, and this relationship was sensitive to outliers).

It is important to note that there may be other individual traits besides cognitive skills that influence the capacity for innovations. First, good physiological condition may enhance problem solving through increased endurance or physical force, which may affect the amount and/or effectiveness of attempts. Second, consistent between-individual differences in behavioral tendencies, that is, personality or temperament traits (Réale et al. 2007), may predestine explorative individuals to encounter and overcome novel challenges more often compared with less bold conspecifics (Reader and Laland 2003). This suggestion is supported by some studies (Webster and Lefebvre 2001; Bouchard et al. 2007; Overington et al. 2011; Sol et al. 2011, 2012; Benson-Amram and Holekamp 2012) and conflicted by others (Boogert et al. 2008; Liker and Bókony 2009; Biondi et al. 2010; Cole et al. 2011), and experimental results caution that the relationship between exploration and learning may be more complex (Matzel et al. 2011). Furthermore, individuals can vary consistently in physiological traits such as hormonal reactivity to stress (Cockrem et al. 2009; Rensel and Schoech 2011), antioxidant capacity (Norte et al. 2008; Galván and Alonso-Alvarez 2009), or susceptibility to infection (Hõrak et al. 2006), and variation in physiological coping styles often covaries with behavioral tendencies such as exploration or neophobia (Groothuis and Carere 2004; Lendvai et al. 2011), although the presence and direction of causality in the latter relationships is unclear (Koolhaas 2008; Koolhas et al. 2010). In this study, we did not quantify neophobia as our previous work showed that this trait is not correlated with problem-solving success in house sparrows (Liker and Bókony 2009); instead we tried to minimize any effect of novelty by familiarizing the birds with the feeders before tests. Nevertheless, the repeatability of attempt latency and its positive correlation with solving latency in our study may reflect the effect of some personality trait such as neophobia or stress sensitivity. However, this trait is less likely to have driven the relationship between physiology and innovative performance than cognitive and/or physical capacity, because attempt latency was largely unrelated to the physiological variables. Alternatively, superior performance might have put innovative individuals in good condition, leading to a correlation between performance and physiology.

To ascertain the direction of causality in this relationship, experimental studies are needed.

In sum, our study revealed several correlations in a wild species between problem-solving performance and aspects of physiology that do not support the hypothesis that inventing and learning novel behaviors are primarily driven by the need of individuals. Instead our results suggest that the capacity for innovative behaviors, probably including cognitive abilities, is an important predictor of performance, being promoted by superior physiological state. We hope that these explorative findings will encourage more detailed and experimental studies on the proximate mechanisms of animal innovations. Such integration of physiology, cognition, and ecology will be crucial for understanding the evolutionary significance of the ways by which animals respond to novel challenges.

FUNDING

We were supported by the Hungarian Scientific Research Fund (OTKA, K84132 to A.L., PD76862 to Á.Z.L., and K75965 to J.N.), TÁMOP-4.2.2.A-11/1/KONV-2012-0064 to A.L., the Romanian Ministry of Education and Research (PN II RU TE 291/2010 to P.L.P., C.I.V., and L.P.); Á.Z.L was supported by the National Science Foundation (1145625) and A.L. by a Marie Curie Intra-European Fellowship.

We thank Katalin Pap, Viktória Bakó, Mónika Szathmáry, and Henriett N. Uhrin for assistance with corticosterone assays. We are highly indebted to Alina Sesarman and Manuela Banciu without whom the analyses of oxidative stress would not have been possible. Adela Pintea, Elin Sild, José Carlos Noguera, and Jonathan Blount kindly provided further help with MDA and tGSH assays.

Handling editor: Alison Bell

REFERENCES

Barja G. 2004. Free radicals and aging. TRENDS Neurosci. 27:595–600.
Benson-Amram S, Holekamp KE. 2012. Innovative problem solving by wild spotted hyenas. Proc R Soc B. 279:4087–4095.
Biondi LM, Bó MS, Vassallo AI. 2010. Inter-individual and age differences in exploration, neophobia and problem–solving ability in a Neotropical raptor (*Milvago chimango*). Anim Cogn. 13:701–710.
Bókony V, Kulcsár A, Tóth Z, Liker A. 2012a. Personality traits and behavioral syndromes in differently urbanized populations of house sparrows (*Passer domesticus*). PLoS One. 7:e36639.
Bókony V, Seress G, Nagy S, Lendvai ÁZ, Liker A. 2012b. Multiple indices of body condition reveal no negative effect of urbanization in adult house sparrows. Landscape Urban Plan. 104:75–84.
Boksa P. 2010. Effects of prenatal infection on brain development and behavior: a review of findings from animal models. Brain Behav Immun. 24:881–897.
Boogert NJ, Reader SM, Hoppitt W, Laland KN. 2008. The origin and spread of innovations in starlings. Anim Behav. 75:1509–1518.
Boogert NJ, Reader SM, Laland KN. 2006. The relation between social rank, neophobia and individual learning in starlings. Anim Behav. 72:1229–1239.
Bortolotti GR, Marchant T, Blas J, Cabezas S. 2009. Tracking stress: localisation, deposition and stability of corticosterone in feathers. J Exp Biol. 212:1477–1482.
Bortolotti GR, Marchant T, Blas J, German T. 2008. Corticosterone in feathers is a long-term, integrated measure of avian stress physiology. Funct Ecol. 22:494–500.
Bouchard J, Goodyer W, Lefebvre L. 2007. Social learning and innovation are positively correlated in pigeons (*Columba livia*). Anim Cogn. 10:259–266.
Breitwisch R, Breitwisch M. 1991. House Sparrows open an automatic door. Wilson Bull. 103:725–726.
Cockrem JF, Barrett DP, Candy EJ, Potter MA. 2009. Corticosterone responses in birds: Individual variation and repeatability in Adelie penguins (*Pygoscelis adeliae*) and other species, and the use of power analysis to determine sample sizes. Gen Comp Endocrinol. 163:158–168.
Cohen A, Klasing K, Ricklefs R. 2007. Measuring circulating antioxidants in wild birds. Comp Biochem Physiol B. 147:110–121.
Cole EF, Cram DL, Quinn JL. 2011. Individual variation in spontaneous problem-solving performance among wild great tits. Anim Behav. 81:491–498.
Cole EF, Quinn JL. 2012. Personality and problem-solving performance explain competitive ability in the wild. Proc R Soc B. 279:1168–1175.
Del Rio D, Stewart AJ, Pellegrini N. 2005. A review of recent studies on malondialdehyde as toxic melocule and biological marker of oxidative stress. Nutr Metab Cardiovasc Dis. 15:316–328.
Dolnik OV, Hoi H. 2010. Honest signalling, dominance hierarchies and body condition in House Sparrows *Passer domesticus* (Aves: Passeriformes) during acute coccidiosis. Biol J Linn Soc. 99:718–726.
Dröge W, Schipper HM. 2007. Oxidative stress and aberrant signaling in aging and cognitive decline. Aging Cell. 6:361–370.
Eppig C, Fincher CL, Thornhill R. 2010. Parasite prevalence and the worldwide distribution of cognitive ability. Proc R Soc B. 277:3801–3808.
Galván I, Alonso-Álvarez C. 2008. An intracellular antioxidant determines the expression of a melanin–based signal in a bird. PLoS One. 3:e3335.
Galván I, Alonso-Alvarez C. 2009. The expression of melanin-based plumage is separately modulated by exogenous oxidative stress and a melanocortin. Proc R Soc B. 276:3089–3097.
Gegear RJ, Otterstatter MC, Thomson JD. 2005. Does parasitic infection impair the ability of bumblebees to learn flower-handling techniques? Anim Behav. 70:209–215.
Groothuis TGG, Carere C. 2004. Avian personalities: characterization and epigenesis. Neurosci Biobehav Rev. 29:137–150.
Hõrak P, Saks L, Karu T, Ots I. 2006. Host resistance and parasite virulence in greenfinch coccidiosis. J Evol Biol 19:277–288.
Hõrak P, Sild E, Soomets U, Sepp T, Kilk K. 2010. Oxidative stress and information content of black and yellow plumage coloration: an experiment with greenfinches. J Exp Biol. 213:2225–2233.
Karatas F, Karatepe M, Baysar A. 2002. Determination of free malondialdehyde in human serum by high performance liquid chromatography. Anal Biochem. 311:76–79.
Kavaliers M, Colwell DD, Choleris E. 1999. Parasites and behavior: an ethopharmacological analysis and biomedical implications. Neurosci Biobehav R. 23:1037–1045.
Kavaliers M, Colwell DD, Galea LAM. 1995. Parasitic infection impairs spatial learning in mice Anim Behav. 50:223–229.
Keagy J, Savard J-F, Borgia G. 2011. Complex relationship between multiple measures of cognitive ability and male mating success in satin bowerbirds, *Ptilonorhynchus violaceus*. Anim Behav. 81:1063–1070.
Koolhaas JM. 2008. Coping style and immunity in animals: Making sense of individual variation. Brain Behav Immun. 22:662–667.
Koolhas JM, de Boer SF, Coppens CM, Buwalda B. 2010. Neuroendocrinology of coping styles: Towards understanding the biology of individual variation. Front Neuroendocrin. 31:307–321.
Laland KN, Reader SM. 1999. Foraging innovation in the guppy. Anim Behav. 57:331–340.
Lefebvre L, Reader SM, Sol D. 2004. Brains, innovations and evolution in birds and primates. Brain Behav Evol. 63:233–246.
Lendvai ÁZ, Bókony V, Chastel O. 2011. Coping with novelty and stress in free-living house sparrows. J Exp Biol. 214:821–828.
Liker A, Bókony V. 2009. Larger groups are more successful in innovative problem solving in house sparrows. Proc Natl Acad Sci USA. 106:7893–7838.
Lindström KM, Hawley DM, Davis AK, Wikelski M. 2005. Stress responses and disease in three wintering house finch (*Carpodacus mexicanus*) populations along a latitudinal gradient. Gen Comp Endocrinol. 143:231–9.
Liu R, Liu IY, Bi X, Thompson RF, Doctrow SR, Malfroy B, Baudry M. 2003. Reversal of age–related learning deficits and brain oxidative stress in mice with superoxide dismutase/catalase mimetics. Proc Natl Acad Sci USA. 100:8526–8531.
Lupien SJ, McEwen BS, Gunnar MR, Heim C. 2009. Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat Rev Neurosci. 10: 434–45.
Martin LB, Brace AJ, Urban A, Coon CAC, Liebl AL. 2012. Does immune suppression during stress occur to promote physical performance? J Exp Biol. 215:4097–4103.

Matzel LD, Wass C, Kolata S. 2011. Individual differences in animal intelligence: learning, reasoning, selective attention and inter-species conservation of a cognitive trait. Int J Comp Psychol 24:36–59.
McEwen BS, Sapolsky RM. 1995. Stress and cognitive function. Cur Opin Neurobiol. 5:205–216.
Metcalfe NB, Alonso-Alvarez C. 2010. Oxidative stress as a life-history constraint: the role of reactive oxygen species in shaping phenotypes from conception to death. Funct Ecol. 24:984–996.
Morand-Ferron J, Cole EF, Rawles JEC, Quinn JL. 2011. Who are the innovators? A field experiment with 2 passerine species. Behav Ecol. 22:1241–1248.
Nakagawa S, Schielzeth H. 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol. Rev. 85:935–956.
Noguera JC, Lores M, Alonso-Álvarez C, Velando A. 2011. Thrifty development: early life diet restriction reduces oxidative damage during later growth. Funct Ecol. 25:1144–1153.
Norte AC, Sheldon B, Sousa JP, Ramos JA. 2008. Repeatability and method-dependent variation of blood parameters in wild-caught Great Tits Parus major. Acta Ornithol. 43:65–75.
Overington SE, Cauchard L, Côté K-A, Lefebvre L. 2011. Innovative foraging behaviour in birds: what characterizes an innovator? Behav Proc. 87:274–285.
Pap PL, Vágási CI, Bărbos L, Marton A. 2013. Chronic coccidian infestation compromises flight feather quality in house sparrows *Passer domesticus*. Biol J Linn Soc. 108:414–428.
Pap PL, Vágási CI, Czirják GÁ, Titilincu A, Pintea A, Barta Z. 2009. Carotenoids modulate the effect of coccidian infection on condition and immune response in moulting House Sparrows. J Exp Biol. 212:3228–3235.
Pap PL, Vágási CI, Czirják GÁ, Titilincu A, Pintea A, Osváth G, Fülöp A, Barta Z. 2011. The effect of coccidians on the condition and immune profile of molting house sparrows (*Passer domesticus*). Auk. 128:330−339.
Peig J, Green AJ. 2009. New perspectives for estimating body condition from mass/length data: the scaled mass index as an alternative method. Oikos. 118:1883–1891.
Pfeffer K, Fritz J, Kotrschal K. 2003. Hormonal correlates of being an innovative greylag goose, *Anser anser*. Anim Behav. 63:687–695.
Pravosudov VV. 2003. Long–term moderate elevation of corticosterone facilitates avian food-caching behaviour and enhances spatial memory. Proc R Soc Lond B. 270:2599–2604.
R Core Team. 2012. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: http://www.R–project.org/
Ramsey G, Bastian ML, van Schaik C. 2007. Animal innovation defined and operationalized. Behav Brain Sci. 30:393–407.
Raouf SA, Smith LC, Brown MB, Wingfield JC, Brown CR. 2006. Glucocorticoid hormone levels increase with group size and parasite load in cliff swallows. Anim Behav. 71: 39–48.
Reader SM, Laland KN, editors. 2003. Animal innovation. Oxford: Oxford University Press.
Réale D, Reader SM, Sol D, McDougall PT, Dingemanse NJ. 2007. Integrating animal temperament within ecology and evolution. Biol Rev Camb Philos Soc. 82:291–318.
Rensel MA, Schoech SJ. 2011. Repeatability of baseline and stress-induced corticosterone levels across early life stages in the Florida scrub-jay (*Aphelocoma coerulescens*). Horm Behav. 59:497–502.
Rice-Evans C, Miller NJ. 1994. Total antioxidant status in plasma and body fluids. Methods Enzymol. 234:279–293.
Seed A, Emery N, Clayton N. 2009. Intelligence in corvids and apes: a case of convergent evolution? Ethology. 115:401–420.
Sepp T, Karu U, Blount JD, Sild E, Hõrak P. 2012a. Coccidian infection causes oxidative damage in greenfinches. PLoS One. 7:e36495.
Sepp T, Sild E, Blount JD, Männiste M, Karu U, Hõrak P. 2012b. Individual consistency and covariation of measures of oxidative status in greenfinches. Physiol Biochem Zool 85:299–307.
Sepp T, Sild E, Hõrak P. 2010. Hematological condition indexes in greenfinches: effects of captivity and diurnal variation. Physiol Biochem Zool. 83:276–282.
Sol D, Griffin AS, Bartomeus I. 2012. Consumer and motor innovation in the common myna: the role of motivation and emotional responses. Anim Behav. 83:179–188.
Sol D, Griffin AS, Bartomeus I, Boyce H. 2011. Exploring or avoiding novel food resources? The novelty conflict in an invasive bird. PLoS One. 6:e19535.
Song L, Zheng J, Li H, Jia N, Suo Z, Cai Q, Bai Z, Cheng D, Zhu Z. 2009. Prenatal stress causes oxidative damage to mitochondrial DNA in hippocampus of offspring rats. Neurochem Res. 34:739–745.
Suárez-Rodríguez M, López-Rull I, Macías Garcia C. Forthcoming 2013. Incorporation of cigarette butts into nests reduces nest ectoparasite load in urban birds: new ingredients for an old recipe? Biol Lett. 9:20120931. doi: 101098/rsbl20120931.
Thornton A, Samson J. 2012. Innovative problem solving in wild meerkats. Anim Behav. 83:1459–1468.
von Schantz T, Bensch S, Grahn M, Hasselquist D, Wittzell H. 1999. Good genes, oxidative stress and condition-dependent sexual signals. Proc R Soc Lond B. 266:1–12.
Webster SJ, Lefebvre L. 2001. Problem solving and neophobia in a columbiform–passeriform assemblage in Barbados. Anim Behav. 62:23–32.