ORIGINAL PAPER

The repeatability of avian egg ejection behaviors across different temporal scales, breeding stages, female ages and experiences

Tomáš Grim · Peter Samaš · Mark E. Hauber

Abstract An evolutionarily fundamental, yet rarely examined, aspect of host–parasite arms races is the repeatability of individual host responses to parasitism. We examined the repeatability of egg ejection, and the latency to eject, across a time-scale of days (within one breeding attempt), weeks and months (between breeding attempts within one breeding season), and years (across different breeding seasons). The ejection of non-mimetic model eggs by European blackbirds, *Turdus merula*, showed overall high repeatability ($r \sim 0.70$) and, similar to empirical patterns on most other behavioral traits already studied, showed decreasing repeatability with time. In contrast, latency to ejection showed negligible repeatability overall ($r \sim 0.20$) and did not change with time. Ejection rates, latencies to ejection and the repeatability of egg ejection did not differ between young and old females. Previous experience with experimentation (number of model eggs the female received before the focal trial) marginally non-significantly covaried with egg ejection (positively) and with latency to egg ejection (negatively). Repeatability estimates for both egg ejection and latency to egg ejection at the within one breeding attempt time-scale did not statistically differ from those reported in a previous study of a different blackbird population

(introduced population in New Zealand). To our knowledge, this is the first comparison of behavioral repeatability between a native vs. introduced population of any animal species.

Introduction

Patterns of individual variation vs. predictability in animal behavior have always been in the focus of classical and modern ethological studies (Bell et al. 2009). The social- and ecological context-specificity of repeatability can inform research into both the developmental and cognitive basis of behavioral consistency and it also provides the substrate for selection to act on decision making algorithms and behavioral repertoires in general. Repeatability in the context of anti-parasitic responses of hosts of avian brood parasites (e.g., egg rejection; Soler 2014), both at the level of individual traits and regarding the potential covariation of the entire suite of different anti-parasitic behaviors have only recently gained prominence (Avilés and Parejo 2011; Samaš et al. 2011).

Any empirical or theoretical study of host egg discrimination implicitly assumes that observed host response (egg accepted or rejected) reflects an intrinsic heritable property of the individual – this is obvious from the terminology of host individuals being described as "acceptors" or "rejecters" (see discussion and references in Samaš et al. 2011). In some hosts, responses to foreign eggs vary according to actual parasitism rate and/or perceived risks of parasitism (e.g., Thorogood and Davies 2013). Such phenotypic plasticity (Welbergen and Davies 2012) is not contradicting the assumption of heritability and repeatability of egg rejection: by definition (Nakagawa and Schielzeth 2010), repeatability refers to host responses to a repeatedly presented identical cue (e.g., plain blue egg

T. Grim (✉) · P. Samaš
Department of Zoology and Laboratory of Ornithology, Palacký University, 17. listopadu 50, 771-46 Olomouc, Czech Republic
e-mail: tomas.grim@upol.cz

M. E. Hauber
Department of Psychology, Hunter College and the Graduate Center of the City University of New York, 695 Park Avenue, New York, NY 10065, USA

model) in the same context (e.g., high parasitism risk) no matter how the same host's responses vary across different cues (e.g., blue vs. spotted eggs; Samaš et al. 2011) or conditions (e.g., high vs. low parasitism risk; Thorogood and Davies 2013). In other words, researchers assume that observed host responses, *other things being equal*, are not random, do not represent inconsistent behavioral noise but, instead, reflect repeatable host decisions to respond to particular egg cue and level of perceived parasitism risk. In any scientific research area it is essential to revisit such basic assumptions that underlie any empirical and theoretical work (see Samaš et al. 2011).

Repeatability estimates in host–parasite interactions are also critical for any study comparing host responses across populations or species. This is because such comparative studies implicitly assume that per population or species estimates of egg rejection rates are repeatable at these respective levels of biological complexity (i.e., do not represent a random noise). For example, many published estimates of egg rejection rates are from studies where host responses were tested only in the host laying period (e.g., Moskát and Hauber 2007), whereas other studies were conducted in both laying and incubation periods (e.g., Grim et al. 2011). If individual hosts change their responses between laying and incubation (e.g., Moksnes et al. 1993; Moskát et al. 2014), then it is not meaningful to directly compare egg rejection rates from studies made at different breeding periods when proportions of experiments done at various breeding stages differ among studies.

Predictions about repeatability depend on exact timing of subsequent parasitism events within or across breeding stages (laying or incubation) and temporal scale (days, months, years; Table 1). Generally, the repeatability of avian breeding behaviors may be studied at three major temporal scales: within one breeding attempt (hereafter WBA), between breeding attempts within one breeding season (hereafter BBA), or between breeding attempts across different breeding seasons (hereafter BBS). Empirically, the repeatability of nearly all behavioral traits ever tested decreases with time passed (Bell et al. 2009). Based on this well-documented pattern, we call our first tested scenario "behavioral decay". This scenario predicts decreasing repeatability in time (WBA>BBA> BBS). This is a null hypothesis where temporal changes in host responses are simply a result of noise (e.g., recognition errors, Reeve 1989) which, other things being equal, increase with increasing time frame of host repeated responses (Reeve 1989).

An alternative "coevolutionary" scenario takes into account costs and benefits of host behaviors in response to parasitism and assumes that the temporal patterns of host responses and, consequently, their repeatability, reflect a specifically evolved host adaptation (Table 1). For example, parasite eggs laid later during the host breeding attempt typically cause less fitness loss than earlier laid eggs, because

later hatched parasite chicks may fail to evict host progeny (Grim et al. 2009) or may not hatch at all, and impose only a slight detrimental effect through reduced host incubation efficiency (Tuero et al. 2007) which may be outweighed by costs of ejection (Antonov et al. 2006). Therefore, the coevolutionary scenario predicts a low repeatability of ejection between laying and incubation WBA caused by a switch in host decisions from adaptive egg ejection in the laying period to adaptive egg acceptance in the incubation period (Moskát et al. 2014). Following the same reasoning, the coevolutionary scenario predicts high repeatability of egg ejection within the laying period (repeated parasite eggs should be ejected) and also high repeatability within incubation period (repeated parasite eggs should be more likely accepted; Table 1). In contrast, BBA and BBS repeatability could either be high and similar to WBA repeatability, when egg ejection is experience-independent (i.e., not affected by learning), or lower, when the egg ejection has a strong learning component, i.e., when a host female switches from acceptance (when she is young and naive) to ejection (when she is old and experienced; Lotem et al. 1995; Stokke et al. 2007). Currently, there are no sufficiently detailed data *across the whole life period* for any hosts of brood parasitic birds on whether and how exactly individual hosts update and refine their egg recognition templates throughout life (for snapshots of effects of host age, see Davies and Brooke 1988; Lotem et al. 1995; Soler et al. 2000; Amundsen et al. 2002; Stokke et al. 2004; Soler et al. 2013). Although this prevents us from formulating more exact predictions for the coevolutionary scenario, this theoretical context still does allow us to predict major patterns and differences between coevolutionary and behavioral decay scenarios (Table 1).

The same predictions described above also hold for the latency to egg ejection, i.e., time delay between introducing the parasitic egg and the egg ejection response (Davies and Brooke 1989; Grim et al. 2011; Table 1). Latency is an important component of anti-parasite defense portfolios because early removal of foreign eggs decreases costs of misdirected incubation effort (Visser and Lessells 2001). Also, low latency to reject foreign eggs avoids multiple parasitism, which would otherwise cause reduced recognition efficiency (Stevens et al. 2013).

We are aware of some empirical data that are apparently inconsistent with our specific predictions derived above (e.g., no co-variation between breeding stage and host response latencies in some species, e.g., Davies and Brooke 1989; Polačiková and Grim 2010; Grim et al. 2011). But those data are from experiments focusing on differences *among* host individuals (i.e., not repeated tests of the same individuals), and so they do not have bearing on repeatability (i.e., *within* individual level).

The repeatability of egg rejection is an issue of fundamental theoretical importance for coevolutionary hypotheses regarding host–parasite arms races in general (details in Samaš et al. 2011). High repeatability is especially critical to persist in any

Table 1 Predictions regarding the repeatability of avian anti-parasite behaviors (egg rejection and its latency) within one breeding attempt (WBA), across breeding attempts within one season (BBA) and across seasons (BBS) based on "behavioral decay" (null hypothesis) vs. "coevolutionary" scenarios (see Introduction for details)

First model egg	Second model egg	Scenario	
		Behavioral decay	Coevolution
WBA			
Laying	Laying	High	High (repeated ejection)
Laying	Incubation	High	Low (ejection→acceptance)
Incubation	Incubation	High	High (repeated acceptance)
BBA			
First nest	Second nest	Lower	High (no learning)
			Lower (learning)
BBS			
First season	Second season	Lowest	High (no learning)
			Lowest (learning)

parasite system where hosts are parasitized multiply — if host ejection response were not very highly repeatable than there would be little difference between fitness of acceptors and rejecters (this is clear even without any sophisticated modelling as shown by the empirical patterns in Stevens et al. 2013). Yet, our literature review shows that repeatability has been rarely studied in hosts of European cuckoos, *Cuculus canorus* (hereafter: cuckoo), brown-headed cowbirds, *Molothrus ater*, or any other inter-specific or conspecific parasites (Table S1). Most importantly, all of these studies also included some methodological limitations (discussed in Samaš et al. 2011), e.g., pooled host responses to different types of foreign eggs for a single study species (Palomino et al. 1998; Croston and Hauber 2014) or for some of the multiple study species (Peer and Rothstein 2010) which violated the assumptions of methods of repeatability estimation (Bell et al. 2009; Nakagawa and Schielzeth 2010). Most studies conducted repeated experiments on the WBA temporal scale (Davies and Brooke 1989; Honza et al. 2007; Peer and Rothstein 2010; Samaš et al. 2011), and only rarely on longer time scales, either at BBA (Alvarez 1996; Lotem et al. 1995) or BBS (Palomino et al. 1998; Soler et al. 2000; Table S1). Just one study included host responses across varying temporal periods but pooled data into one estimate of repeatability, thus preventing the test of how increasing time gaps between successive parasitism events affect egg rejection repeatability (Soler et al. 2000).

Another fundamental yet poorly studied issue concerns the effects of age and/or individual experience on host defenses against brood parasitism (Lotem et al. 1992; Amundsen et al. 2002; Langmore et al. 2009; Soler et al. 2013). A female's egg ejection decision may be affected by two specific components of her individual experience. (1) Accumulated experience with her *own* eggs. As a female becomes older, the more clutches she has laid and viewed. Therefore, old females should be more familiar with their own egg phenotypes than young females and so old females are predicted to reject

foreign eggs more often than young females (Lotem et al. 1992). (2) Parasitism-related experience with *foreign* eggs, including the particular model egg used in our experiments. Individual females in our study population, that were previously experimentally tested by us, should recognize any foreign egg, including the model used in the present study, better than females with no experience with experimental parasitism, *independently of age*. Therefore, the number of previous experiments should positively predict ejection (note that nest desertion is not a specific response to parasitism in our study populations, see Methods).

In summary, no study so far has (1) estimated the repeatability of egg ejection separately for WBA, BBA and BBS, (2) estimated repeatability of the latency to ejection for the three temporal scales, (3) examined repeatability in relation to laying vs. incubation stages, (4) tested how repeatability is affected by female age or (5) long-term previous experience. In the present study we attempted to fill all these research gaps by a detailed study of the European blackbird, *Turdus merula* (hereafter blackbird; see below for the rationale behind studying this particular host). By performing the WBA experiments in the native range of blackbirds, we provide "exact replication" (sensu Kelly 2006) of the study of Samaš et al. (2011) which was done in the introduced, New Zealand range of blackbirds (Samaš et al. 2011; that study did not provide data on experiments at BBA and BBS scales). Such meta-replication, including exact replication, is a crucially important, yet typically neglected, part of behavioral and ecological research (see Johnson 2002; Kelly 2006).

Methods

Study species

We followed methodology recommended for egg ejection repeatability studies (see Samaš et al. 2011 for detailed

rationale behind each criterion). Accordingly, we chose a host species where (1) only one sex rejects parasitic eggs, (2) egg rejection decisions vary between individuals, and (3) egg rejection is by ejection (i.e., nest desertion effectively prevents estimations of WBA repeatability, but not those of BBA and BBS repeatability). Also, (4) we used egg models that are known to elicit intermediate ejection responses, and (5) we used appropriate state of the art statistical tools that can control for covariates, i.e., generalized linear mixed models (GLMM; see below).

Additionally (6), adult blackbirds in our study population show high philopatry (Samaš et al. 2013a), thus increasing the chances of repeated experiments at BBA and BBS time scales. Finally, (7) blackbird egg-rejection behavior does not consistently depend on presence (sympatry) or absence (allopatry) of the cuckoo (Grim et al. 2011). Thus, both sympatric and allopatric populations are similarly heuristically suitable and conclusions from their study can be meaningfully generalized as empirically confirmed by Grim et al. (2011). Future work should also focus on other host defenses, including aggression against adult parasites (Campobello and Sealy 2011; Trnka et al. 2013) or the desertion of parasite chicks (Grim 2007; Langmore et al. 2009) in sympatry vs. allopatry with brood parasites, in this and other hosts that are known to vary some aspects of their anti-parasite behavior with local density of brood parasites (Stokke et al. 2008; Langmore et al. 2009).

General field procedures

We followed well-established protocols for field work (mist-netting, nest searching and checking) and data analyses (see below for specific details) that are used as a standard in similar studies. We collected empirical data in the city of Olomouc, Czech Republic (CZ; 49°35′N, 17°15′E) in 2009–2012 (for details, see Grim et al. 2011).

The birds were captured using 19-mm mesh sized, 3- to 18-m-long mist nets. We caught the female always solely after the first WBA experiment was finished. Therefore, the stress of being captured and handled could not influence rejection behavior in WBA experiments, although it might still influence egg rejection response between breeding attempts within or between seasons. However, it is impossible to avoid this putative effect in any animal study where individuals are captured and handled. Individual birds were ringed with both the standard metal ring and a unique combination of color bands to enable individual recognition without the need to re-trap and disturb them. We determined the age of captured birds as young (i.e., yearlings) or old (any older birds) according to Svensson (1992). The age estimates were missing for females that did not show clear age-related traits (therefore, sample sizes differ across analyses).

We recorded the final clutch size (which potentially affects egg ejection; Hauber 2003; Hauber et al. 2014) in all nests except those predated before clutch completion. Repeated nesting attempts were then searched both within the same

breeding season (within several days after the end of the previous breeding attempt; Cramp 1988) and in the subsequent years. We focused on color-ringed females because extensive video-recordings confirmed that solely females eject foreign eggs in our CZ study population (J. Weiszensteinová, PS, TG, unpublished data). Although we ringed a large number of blackbirds, only 19 females (out of 143 mist-netted adult females) were successfully relocated and tested in subsequent years. None of the 267 ringed chicks was later found as a breeding female. Thus, sample sizes for BBS repeatability estimates were limited (but still statistically robust, see below) due to large nest, post-fledging, and adult mortality (PS, TG, unpublished data), dispersal of some individuals out of the study site (Samaš et al. 2013a) and some repeat nests being located at inaccessible places (e.g., private gardens).

Experimental procedures

We used the plain light blue artificial egg which is the most commonly used model in studies of host egg rejection across Europe (Davies 2000; Polačiková and Grim 2010; Grim et al. 2011). This decision facilitates comparisons of host behavior with other populations and species. Non-mimetic models were made from polysynthetic material and painted with acrylic paints to resemble eggs laid by the cuckoo into the nests of the common redstart, *Phoenicurus phoenicurus*. The size ($x\pm$ SD$=22.7\pm0.54\times17.4\pm0.48$ mm, $n=10$), mass (3.7 ± 0.45 g, $n=10$) and the shape of these non-mimetic blue egg models were similar to real, cuckoo eggs (size range: 20–26×15–19 mm, mass range: 2.9–3.8 g, Cramp 1985). For reflectance spectra, see Samaš et al. (2011).

For each individual trial (i.e., the introduction of the model egg into a host nest) we followed standard procedures established in previous studies (Davies and Brooke 1989; Grim et al. 2011). First experimental trial at each nest was done during the laying stage or during the first 5 days of incubation (nests were visited daily, thus, clutch ages were not estimated but known exactly). Some previous studies showed that blackbird egg rejection responses (ejection and desertion pooled) do not depend on nest age (Davies and Brooke 1989; Polačiková and Grim 2010; Grim et al. 2011), whereas others detected slight differences between laying and incubation stages (Samaš et al. 2011). However, in the present study we specifically focused on experimental parasitism in both laying and incubation stages to test predictions from behavioral decay vs. coevolutionary scenarios (Table 1).

We added the egg model during laying stage after the second own egg was laid. We did it intentionally to reduce the possible effect of learning by the host through inspecting only the experimental egg in the nest without any host eggs being present (Strausberger and Rothstein 2009). We did not remove any host's egg(s) because egg removal does not affect rejection probability of our type of model egg in blackbirds

(Davies and Brooke 1989; Grim et al. 2011). We monitored the nest contents daily until ejection, desertion or final acceptance after the standard 6-day exposure period at active nests (Grim et al. 2011). The accepted egg models were removed on the sixth day. Previous studies included nest desertions as a rejection response to parasite eggs (Polačiková and Grim 2010). However, we excluded deserted nests because work in our study population showed that desertion rates did not differ between experimentally parasitized and non-manipulated control nests, confirming that in our study population desertion was not a specific response to parasitism when using the particular egg types employed here (TG et al. unpublished data; see also Samaš et al. 2011; cf. Hauber et al. 2014).

In the WBA treatment, we introduced another egg model into the host nest 2 days after the first trial was completed (resulting in either ejection or acceptance). We again monitored the nest daily until ejection, desertion, or acceptance up to 6 days. Egg laying and incubation periods in European populations of blackbirds last approximately 18 days (5 days of laying and 13 days of incubation; Cramp 1988), providing enough time to test pure acceptors repeatedly (i.e., individuals that accepted both the first and second experimental eggs; 6+1+6=13 days). We excluded from our analyses those nests that were depredated or deserted before the end of the first or second trial.

Nests that were depredated after the first trial successfully ended could not be included in WBA treatment; however, they were useful for estimates of BBA and BBS repeatability. Specifically, out of total 23 BBA females 17 received one model and six received two models before the second BBA trial (i.e., in their first nest in the same season). Out of total 19 BBS females, three received one model, 11 received two models and five received three models before the second BBS trial (i.e., in their first nest(s) in the previous breeding season). This variation in number of previous experiments was useful for testing the effects of previous experience with the specific model (see above) independently of female age (see the next section).

Statistical analyses

In all analyses, we followed the recommendations of Martin and Bateson (2008), Nakagawa and Schielzeth (2010) and Dingemanse and Dochtermann (2013). We present three types of repeatability estimates: (1) Spearman correlations, (2) simple "repeatability" (synonymous with "agreement repeatability"; Nakagawa and Schielzeth 2010) estimated by GLMM with no covariates, and (3) "adjusted repeatability" estimated by GLMM with same covariates forced to all models to make estimates meaningfully comparable (S. Nakagawa, personal communication).

First, we calculated Spearman's correlation coefficients (r_s) and their confidence intervals (CIs). Although this is a statistically correct method (Martin and Bateson 2008, pp. 74–78),

its univariate non-parametric approach does not allow accounting for covariates and some authors recommend to estimate repeatability using only Generalized Linear Mixed Model (GLMM; Nakagawa and Schielzeth 2010). We built such models separately for WBA, BBA and BBS treatments and present results from both simple and covariate-adjusted GLMMs, i.e., both raw phenotypic repeatability and adjusted repeatability (Table 2). Reporting of the raw metric of phenotypic repeatability is crucial as it can be included in meta-analyses (Bell et al. 2009). The adjusted repeatability enabled us to answer the question whether conclusions based on agreement repeatability are not affected by following covariates: "nest age" (continuous; days; starting from the first laid egg), "first egg-laying date" (continuous; including its quadratic term to test for possible non-linear seasonal trends; Samaš et al. 2013b), and "final clutch size" (continuous; number of eggs). "Female identity" was entered as a categorical random effect. Originally, we included additional random effect of "year" but it did not explain any significant variation in the data; when removed, the resulting simpler models with same fixed effects had a much better fit (much lower AIC$_c$) and very similar parameter estimates. Therefore, we present the results of the models without "year" random effects (following recommendations of Bolker et al. 2009). We also calculated alternative models which included actual clutch size at the start of experiment (instead of final clutch size). These models yielded statistically same estimates of repeatability (results not shown).

In additional analyses, we included "female experience" (continuous; number of model eggs the female experienced before the focal trial). This variable was applicable only in BBA and BBS GLMM models (females in WBA treatment were not tested previously by definition).

We tested the effect of female age per se on the subset of first egg trials, i.e., we included only the first model egg trial in her life per each female. This excludes any possible effect of parasitism-related experience with the particular model egg (although we naturally cannot exclude a possibility that some of these females were previously parasitized by conspecifics). We then tested the effect of previous experience with the egg model in the BBA and BBS data sets (see above).

Further, we tested whether the repeatability of egg ejection differed between the following groups of WBA nests: (1) both first and second model eggs introduced within the laying period, (2) both first and second model eggs introduced within incubation, or (3) the first egg model introduced within the laying period and the second egg model introduced within the incubation stage (Table 1). We pooled data from (1) and (2) because there was no variation within the laying stage data (n=4 paired trials, all ejections), which prevented a meaningful estimation of repeatability for this subgroup (see Samaš et al. 2011).

The repeatability of latency to egg ejection (continuous; days) was modelled using Linear Mixed Models (LMM).

Table 2 Repeatability (estimates with 95 % CIs) of blackbirds' anti-parasite behaviors during repeated trials on the same females

Host behavior	Ejection (%)			Latency to ejection (days±SE)		
Time scale	WBA	BBA	BBS	WBA	BBA	BBS
Repeatability						
Spearman correlation	0.91 (0.79,1.00)	0.57 (0.20,0.94)	0.57 (0.23,0.92)	0.13 (−0.19,0.45)	0.44 (0.02,0.86)	−0.26 (−0.46,−0.05)
Repeatability	0.92 (0.88,0.95)	0.57 (0.21,0.80)	0.70 (0.36,0.88)	0.00[a]	0.07 (0.00,0.51)	[b]
Adjusted repeatability	0.92 (0.88,0.95)	0.74 (0.47,0.88)	0.83 (0.60,0.93)	0.07 (0.00,0.34)	0.31 (0.00,0.71)	[b]
Covariates						
Nest age	0.19 (0.13)	0.09 (0.21)	0.50 (0.46)	−0.06 (0.02)	−0.09 (0.04)	−0.11 (0.04)
Clutch	0.90 (2.34)	−0.58 (0.98)	−0.33 (3.57)	–	–	–
FEG	0.06 (0.20)	−0.12 (0.12)	−0.12 (0.23)	–	–	–
FEG^2	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	–	–	–
Variances – r						
Between-individual	39.1	4.38	7.57	0.00	0.08	0.00
Within-individual	$\pi^2/3$	$\pi^2/3$	$\pi^2/3$	0.72	1.00	0.99
Variances – adjusted r						
Between-individual	37.66	9.18	16.28	0.05	0.33	0.00
Within-individual	$\pi^2/3$	$\pi^2/3$	$\pi^2/3$	0.59	0.74	0.79

Parameter estimates [mean (SE)] for covariates are from GLMM that was used to calculate the adjusted repeatability. Following published recommendations (see Methods), we present three estimates: "Spearman correlation" coefficient (r_s), simple "repeatability" synonymous with "agreement repeatability" (GLMM without covariates), and "adjusted repeatability" (GLMM with covariates). To make the estimates meaningfully comparable with each other, all adjusted repeatability models included all the same covariates (S. Nakagawa, personal communication). Note that the removal of non-significant covariates (across all models) did not affect our conclusions (results not shown). *Nest age* age of nest at the start of experiment, *Clutch* final clutch size, *FEG* first egg laying date, FEG^2 squared first egg laying date (see Methods)

[a] Impossible to calculate CIs. For sample sizes, see Fig. 1

[b] Impossible to estimate due to negative variances

Models included response variable "latency to egg ejection", "female identity" as categorical random variable and "nest age" (continuous) as fixed covariate. Only individuals that ejected the model egg in both first and second trials were included in this analysis. Consequently, sample sizes were smaller for latency analyses than for egg ejection analyses, especially for BBA and BBS. Therefore, latency models did not include other covariates (i.e., including other variables resulted in biased estimates of variances or models did not converge in some BBA and BBS analyses; for consistency across analyses we present the simpler latency models for all temporal periods). When the latency to ejection was modelled assuming negative binomial or Poisson distributions, respectively, the results remained the same (results not shown).

We calculated both raw phenotypic and adjusted repeatability (i.e., repeatability calculated after controlling for covariates listed above) using GLMM with binary response by formula $r=V_B/(V_B+V_E+\pi^2/3)$, where V_B denotes between-individual variance, V_E is the residual variance always fixed to 0 for binary response variable, and $\pi^2/3$ is the inherent distribution-specific variance (Nakagawa and Schielzeth 2010). The adjusted repeatability of latency to egg ejection was calculated as $r=V_B/(V_B+V_W)$, where V_B denotes between-individual variance and V_W denotes within-individual variance (Nakagawa and Schielzeth 2010).

Spearman's correlation coefficient with exact 95 % CIs was calculated using StatXact 7 (Cytel Inc 2005). GLMM and LMM were calculated in R 2.15.2 (R Core Team 2012; package "lme4" v. 0.999999-2; Bates et al. 2012). We used adaptive Gauss–Hermite approximation in our GLMMs. We estimated asymptotic 95 % CIs for repeatability estimates from GLMM model and tested differences between correlations using R package "psych" v. 1.2.8 (calculation based on the Fisher *r*-to-*z* transformation; Cohen et al. 2003; note that R package "rptR" cannot do the calculation of adjusted *r* and its CIs based on bootstrapping for binary responses). To estimate 95 % CIs for repeatability of latency to egg ejection, we used parametric bootstrapping in R package "rptR" v. 0.6.404 (Nakagawa and Schielzeth 2010). All estimates are mean±SE unless stated otherwise.

Ethical note

Experiments were done under permission from local authorities (no. SmOl/ZP/55/6181 b/2009/Pr and SMOVZP/55/8542120111Kol), permission to handle animals during biological experiments (no. 065/2002–V2 to TG), ringing license

(Bird-ringing station of the Natural History Museum Prague, no. 1085 to PS) and institutional animal ethics committee permission of Palacký University (no. 45979/2001-1020).

To minimize the disturbance to breeding birds we captured as many birds as possible during winter (close to feeders; this was also part of another research project: Samaš et al. 2013a). When mist-netting birds during the breeding period we did not use playback. We placed mist-nets within several meters from active blackbird nests depending on vegetation structure of particular nests. If a female was not caught within 10 min, we removed the net to allow nest owners to resume their normal behavior. In such cases we made another attempt to catch the female next day. When the female was not caught again, we excluded the nest and female from further analyses. After ringing, each bird was immediately released at the same place where it was caught.

When searching for host nests in suitable vegetation (i.e., bushes, small trees) we carefully minimized any disturbance to vegetation cover around each nest to not change the original nest concealment. When checking each nest repeatedly to determine acceptance or rejection of the model egg (see below) we always took care to minimize the length of our presence nearby the nest, thus, minimizing the risk that predators would get attracted due to our presence. Further, any vegetation moved from its original position by us was arranged back to its original position after each nest check.

Results

Within breeding attempt temporal scale (WBA)

At this shortest time scale, out of 73 females only two (one young, one old) changed their responses, both switching from acceptance to ejection (Fig. 1a). Consequently, WBA egg ejection repeatability was high (Table 2).

We tested whether repeatability differs according to whether both model eggs were introduced within one breeding stage (laying or incubation) or the first egg was introduced in laying and the second in incubation. The repeatability of egg ejection in the subset of nests where first egg model was introduced during laying stage and second egg model during incubation stage was high (Spearman's correlation: $r_s=0.85$, exact 95 % CI=0.66–1.00, $n=24$, $P<0.0001$; GLMM: $r=0.90$, 95 % CI= 0.78–0.96). The combined repeatability of egg ejection of model eggs both introduced during only laying or during only incubation (see Methods why we pooled the data) was also very high (Spearman's correlation: $r_s=1.00$, exact 95 % CI=1.00– 1.00, $n=49$, $P=0.02$; GLMM: $r=0.98$, 95 % CI=0.96–0.99).

The repeatability of the latency to egg ejection between trials WBA was very low and not significantly different from zero (Table 2, Fig. 1b). Latency decreased with nest age in the first trials ($F_{1,56}=8.19$, $P=0.006$), but not in the second trials ($F_{1,56}=0.12$, $P=0.73$).
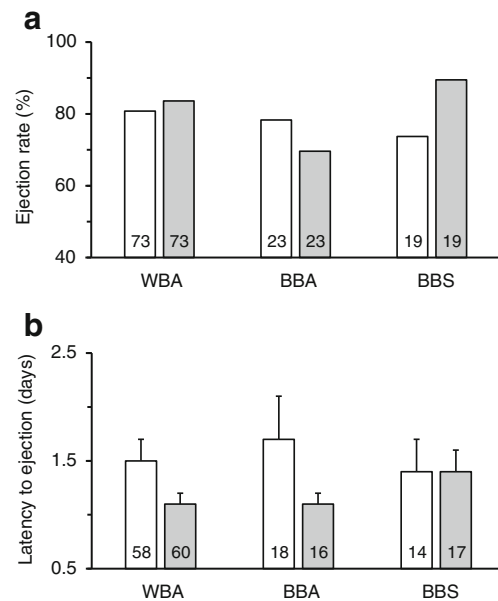

Fig. 1 Blackbird anti-parasite behaviors during repeated trials (first — white bars, second — grey bars) with the same females: **a** egg ejection, **b** latency to ejection (mean±SE). Sample sizes (numbers of tested females; shown within bars) are identical for first and second trial "ejections" (due to the paired nature of the experiment), but not for "latencies" because ejection rates changed between first and second trials. Sample sizes for WBA latencies are lower than expected from ejection rates because we lost latency data for one female

The repeatability of egg ejection did not differ statistically between native Czech ($r_s=0.91$; Table 2) and introduced New Zealand populations of blackbirds ($r_s=0.86$; data from Samaš et al. 2011; Steiger's Z-test following Steiger 1980: $Z=1.16$, $P=0.25$). Also, latency to ejection did not differ statistically between native Czech ($r_s=0.13$; Table 2) and introduced New Zealand populations of blackbirds ($r_s=0.46$; data from Samaš et al. 2011; $Z=1.60$, $P=0.11$).

Between breeding attempt temporal scale (BBA)

At this intermediate temporal scale, out of 23 females only three (all old) changed their responses, one from acceptance to ejection, and two from ejection to acceptance (Fig. 1a). The BBA repeatability of egg ejection was moderate (Table 2) and significantly lower than that for WBA (Table 3).

Repeatability of the latency to egg ejection between trials was low (Table 2, Fig. 1b) and not significantly different from that for WBA (Table 3). Latency to ejection was not predicted by nest age either in the first ($F_{1,13}=0.72$, $P=0.41$) or second breeding attempts within one season ($F_{1,13}=0.41$, $P=0.54$).

Between breeding seasons temporal scale (BBS)

At this longest time scale, out of 19 females only three (two young, one old in the first year of trials) changed their responses, all from acceptance to ejection. The repeatability of

Table 3 Comparison of GLMM egg ejection and latency to ejection repeatabilities across different temporal scales for blackbirds

Temporal scale comparison	Repeatability		Adjusted repeatability	
	Z	P	Z	P
Egg ejection				
WBA>BBA	3.47	0.001	2.52	0.01
WBA≥BBS	2.60	0.009	1.45	0.15
BBA=BBS	0.66	0.51	0.71	0.48
Latency to ejection				
WBA=BBA	–	–	0.80	0.42
WBA vs. BBS	–	–	–	–
BBA vs. BBS	–	–	–	–

Steiger's Z-tests follow Steiger (1980). Note that in most comparisons of latency, it was not possible to apply the Z-test due to non-estimable repeatabilities or their confidence intervals (Table 2)

egg ejection was moderate (Table 2), being statistically similar to BBA but significantly lower than WBA repeatability (Table 3). However, when adjusted for covariates BBS did not marginally differ from WBA repeatability (Table 3).

BBS repeatability of the latency to egg ejection was effectively zero (Table 2, Fig. 1b) and, although the statistical test could not be performed, clearly not different from that for WBA or BBA (Table 3; GLMM cannot calculate negative repeatability because variances are constrained to be non-negative; Nakagawa and Schielzeth 2010). In the first trials, latencies to ejection did not correlate with nest age ($F_{1,12}=$ 3.40, $P=0.09$). In the second trials, latency to ejection decreased with nest age ($F_{1,12}=5.82$, $P=0.03$).

Effects of female age and experience

The ejection rate of blue egg models by experimentally naive blackbirds was 81.3 % ($n=96$ females). Within this subset of data, young females ejected the plain blue model egg (first trials only) at statistically the same rates (69.2 %, $n=13$ females) as old females (81.8 %, $n=22$ females; $\chi^2=0.73$, $df=1$, $P=0.39$). Young females also ejected the plain blue model egg (first trials only) with similar latencies (1.4± 0.4 days, $n=9$ females) as old females (1.8±0.4 days, $n=18$ females; Welch's t-test: $t_{19.4}=-0.61$, $P=0.55$).

Some of these females were then tested repeatedly. Repeatability of WBA egg ejection was identical between young ($r_s=$ 0.77, $n=11$ females) and old females ($r_s=0.77$, $n=11$ females). We did not estimate the repeatability of latencies to ejection separately for young and old females due to small samples. Because (1) ejection rates, latencies to ejection and repeatability of egg ejection were not different between young and old females, (2) we were unable to age all females, and (3) young/old female ratio was unbalanced at BBA and BBS

temporal scales we did not include female age into other analyses.

Previous experience with experimentation was marginally non-significant ($F_{1,129}=3.52$; $P=0.06$) when kept as the major predictor of interest in a model explaining variation in egg ejection. More often tested females tended to eject model eggs more often; logit[egg ejection probability]=9.12(3.75)+ 1.63(0.78) × previous experience. All other covariates (nest age, clutch, FEG, FEG^2) were non-significant and sequentially backward eliminated from the final model (results not shown). Therefore, we did not consider previous experience as a predictor in BBA and BBS models; including this variable (1) would not change our conclusions for BBA–BBS comparisons, and (2) prevent us from comparing adjusted repeatability between WBA and these longer temporal scales because previous experience cannot be applied to WBA treatment (see above).

In a model with the same predictors and backward elimination, the latency to ejection was not significantly predicted by previous experience ($F_{1,153}=3.26$; $P=0.07$), when controlling for a highly significant statistical effect of nest age ($F_{1,151}=$ 16.82; $P=0.0001$; latency to ejection=1.59(0.10)−0.14(0.08) × previous experience −0.05(0.01) × nest age). More experienced females non-significantly tended to eject faster.

Discussion

Repeatability estimates for blackbird egg ejection were highest for the within one breeding attempt period (WBA), lower for between breeding attempts within one breeding season (BBA), and similarly lower for between breeding attempts in different breeding seasons (BBS). Specifically, both Spearman correlations and simple GLMM suggested that WBA *repeatability* was significantly higher than BBA or BBS repeatability, with the latter two statistically similar. In a partial contrast, GLMM estimates of *adjusted repeatability* (GLMM with covariates) suggested that WBA repeatability was still higher than BBA, but statistically similar to BBS, with BBA and BBS also statistically similar. Therefore, all three approaches concurred that repeatability did not differ between the two longer time scales: BBA and BBS, while both were still significantly above zero. The different estimates of WBA repeatability of egg ejection were also consistently high for experiments where first egg model was introduced during the laying and the second egg was introduced during incubation stage, as well as for experiments where both egg models were introduced only in the laying or only in the incubation stages. In contrast to patterns of egg ejection, the repeatabilities of the latencies to ejection were overall much lower, and statistically non-different from zero in all cases. Results for the WBA time-scale, both egg rejection and latency to egg rejection repeatabilities, were quantitatively similar to those obtained from the previous study of a different blackbird population (introduced population in New

Zealand; Samaš et al. 2011). To our knowledge, this is the first comparison of behavioral repeatability in native vs. introduced populations of any animal species.

Taken together, these results are more in line with predictions of the behavioral decay scenario, and more in contrast with the predictions of the coevolutionary scenario (Table 1). Interestingly, most host species that are known to be suitable cuckoo hosts and that are currently frequently parasitized by cuckoos do not *empirically* behave according to coevolutionary theory regarding their responses in laying vs. incubation (see Introduction and Supplementary Table 1). This gives an impetus for future work to test whether the costs of late parasitism (i.e., parasite eggs laid during host incubation stage) are not strong enough to select for host egg ejection after all, e.g., due to increased incubation costs (Visser and Lessells 2001; Tuero et al. 2007).

Nonetheless, high repeatability in response to artificial parasitism is consistent with a genetic basis of egg ejection (Bell et al. 2009; but see Dohm 2002), upon which selection may act to increase, maintain, or reduce the frequency of anti-parasitic behaviors (Shaw and Hauber 2009, 2012). This is further supported by the ability of young blackbird females to eject foreign eggs at the same frequencies and with the same high repeatability as old females. Thus, similarly to the majority of previous studies, we also did not find any strong or consistent age effects on the probability of egg ejection (Davies and Brooke 1988; Soler et al. 2000; Amundsen et al. 2002; Stokke et al. 2004; Soler et al. 2013; but see Lotem et al. 1995). Relatively lower BBA and BBS repeatability of egg ejection, and negligible repeatability of latency to ejection, can be, at least partly, explained by the fine-tuning of individual experiences, which marginally non-significantly positively co-varied with egg ejection in blackbirds in our study, and, additionally, by low repeatabilities of same females clutch coloration both within and between seasons (Honza et al. 2012).

Potential confounding factors

Studies of egg rejection repeatability can be confounded by multiple factors. As Samaš et al. (2011) have already discussed most of those factors at length, we will here only add a short commentary on one critical confound that may become more important when repeatability is estimated over longer temporal periods (months, years).

The quantitative similarity between model and host eggs varies both between individual clutches of the same female and the clutches of different females (Honza et al. 2012). Consequently, experimental parasitism using invariant, artificial model eggs introduced into different clutches with variable appearances may, at first sight, represent a problem for the study of repeatability. Changing the appearance of host eggs across breeding attempts or years could explain why raw repeatability slightly decreased over longer time-scales (Table 2). However,

there is no reason to expect that changes in egg phenotypes should consistently move in a specific direction, for example, consistently increase host–parasite (or model) egg dissimilarity. Indeed there is no evidence for such an effect (Honza et al. 2012 and references therein). If, on the other hand, egg phenotypes change randomly across time, then there is no consistent directional change of the host–parasite egg dissimilarity in time and host–parasite egg similarity cannot confound repeatability estimation. Also note that only raw repeatability was lower at BBA and BBS scales, while adjusted repeatability (which is biologically more meaningful given significant effects of some covariates on host responses) did not differ between WBA and BBS scales, contradicting the hypothesis that changing appearance of host eggs explains the patterns we found.

Most importantly, these considerations may be interesting in theory but have no bearing on biological reality. The changing appearances of individual phenotypes (e.g., host eggs) are natural and inevitable features of any biological system (Honza et al. 2012), are faced by all brood parasites (either conspecific or interspecific), and cannot be avoided in experimental studies, either (see Samaš et al. 2011 for a detailed explanation). Since we did not include the instantaneous manufacture of artificial eggs in our experiments based on the appearance of each new clutch encountered in our search, to generate a constant distance of phenotypic differences between each of the hosts' eggs and our artificial models (which was also not done by any of the many published studies of brood parasitism, but see Martín-Vivaldi et al. 2013), we must simply face this biological reality and potential confound.

In fact, it is questionable whether it would be possible to perform such a study because we would have to know *individual* discrimination thresholds *before* making the repeatability experiments; only experiments would enable us to determine such thresholds leading to inherent circularity. Further, overall variation of host eggs within a clutch cannot be known for any experiments done during laying period (i.e., experimental egg is introduced *before* the clutch is finished making it impossible to know the overall clutch variation at the time of experiment). Finally, eggs vary within each clutch, thus, it is in principle impossible to keep the distance between the model vs. each host egg standardized across *all* eggs within even a single clutch.

To sum up, a "constant distance of phenotypic differences" approach was never applied in any study of brood parasitism, it is logistically impossible to perform in the critical time of host laying period and, most importantly, it would represent a biologically unrealistic context as neither host and parasite eggs ever remain identical in appearance across repeated acts of parasitism. It also contradicts the meaning of repeatability. Repeatability is defined as a temporal constancy of response to a constant *cue* (Nakagawa and Schielzeth 2010) and, therefore, it inevitably involves a change in individual's phenotype.

For example, individual's age *will* always change between two (or more) successive measurements that are the source of primary data for estimation of repeatability. Age impact is correlated with experience, memory capacity and other cognitive phenomena (Bell et al. 2009). Therefore, because the cognitive system itself changes with time, this renders "constant distance" experimental approach a biologically inappropriate concept (see also Antonov et al. 2012; Spottiswoode and Stevens 2010).

Conclusions

In this study, we examined for the first time whether the repeatability of host anti-parasite defenses varies at both short and longer time-scales. We documented relatively high to moderate repeatability of egg ejection decisions by individual female birds across days, months and even years and negligible repeatability of latencies to egg ejection across all time-scales. Overall, our results are in line with known empirical patterns of "behavioral decay", meaning that the repeatability of behavioral traits decreases with increasing time between successive measurements (Bell et al. 2009). In contrast, our results do not support repeatability predictions from a coevolutionary scenario of egg ejection behaviors in blackbirds. The present study also allowed for the first time the comparison of repeatability of a behavior in native vs. introduced populations of the same species (which was, to our knowledge, not done for any animal species yet), showing no statistical differences in repeatability between blackbird populations isolated from each other for at least a century and half. Future studies should focus on other anti-parasite traits (e.g., aggression against adult parasites, specific enemy recognition, chick discrimination; Trnka et al. 2013), their repeatabilities, and the presence and direction of possible covariation between anti-parasitic behaviors with other behavioral and life-history host traits. This will allow to integrate studies of brood parasite–host arms races into the broader framework of ongoing research on behavioral syndromes and animal personalities (Sih et al. 2004; Avilés and Parejo 2011).

References

Alvarez F (1996) Model cuckoo *Cuculus canorus* eggs accepted by rufous bush chats *Cercotrichas galactotes* during the parasite's absence from the breeding area. Ibis 138:340–342

Amundsen T, Brobakken PT, Moksnes A, Røskaft E (2002) Rejection of common cuckoo *Cuculus canorus* eggs in relation to female age in the bluethroat *Luscinia svecica*. J Avian Biol 33:366–370

Antonov A, Stokke BG, Moksnes A, Kleven O, Honza M, Røskaft E (2006) Eggshell strength of an obligate brood parasite: a test of the puncture resistance hypothesis. Behav Ecol Sociobiol 60:11–18

Antonov A, Stokke BG, Fossøy F, Ranke PS, Liang W, Yang C, Moksnes A, Shykoff J, Røskaft E (2012) Are cuckoos maximizing egg mimicry by selecting host individuals with better matching egg phenotypes? PLoS ONE 7:e31704

Avilés JM, Parejo D (2011) Host personalities and the evolution of behavioural adaptations in brood parasite–host systems. Anim Behav 82:613–618

Bates D, Maechler M, Bolker B (2012) lme4:Linear mixed-effects models using S4 classes. R package version 0999999–0 [Online] Available at http://lme4.r-forge.r-project.org/

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White JSS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol 24:127–135

Campobello D, Sealy SG (2011) Nest defence against avian brood parasites is promoted by egg-removal events in a cowbird-host system. Anim Behav 82:885–891

Cohen J, Cohen P, West SG, Aiken LS (2003) Applied multiple regression/correlation analysis for the behavioral sciences, 3rd edn. Lawrence Erlbaum, Mahwah, NJ

R Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Cramp S (1988) The birds of the Western Palearctic, volume V (tyrant flycatchers to thrushes). Oxford University Press, Oxford

Croston R, Hauber ME (2014) High repeatability of egg rejection in response to experimental brood parasitism in the American robin (*Turdus migratorius*). Behaviour (published online, doi:10.1163/1568539X-00003164)

Davies NB (2000) Cuckoos, cowbirds and other cheats. T and A D Poyser, London

Davies NB, Brooke ML (1988) Cuckoos versus reed warblers: adaptations and counteradaptations. Anim Behav 36:262–284

Davies NB, Brooke ML (1989) An experimental study of co-evolution between the cuckoo, *Cuculus canorus*, and its hosts: I. Host egg discrimination. J Anim Ecol 58:207–224

Dingemanse NJ, Dochtermann NA (2013) Quantifying individual variation in behaviour: mixed-effect modelling approaches. J Anim Ecol 82:39–54

Dohm MR (2002) Repeatability estimates do not always set an upper limit to heritability. Funct Ecol 16:273–280

Grim T (2007) Experimental evidence for chick discrimination without recognition in a brood parasite host. Proc R Soc Lond B 274:373–381

Grim T, Rutila J, Cassey P, Hauber ME (2009) Experimentally constrained virulence is costly for common cuckoo chicks. Ethology 115:14–22

Grim T, Samaš P, Moskát C, Kleven O, Honza M, Moksnes A, Røskaft E, Stokke BG (2011) Constraints on host choice: why do parasitic birds rarely exploit some common potential hosts? J Anim Ecol 80:508–518

Hauber ME (2003) Interspecific brood parasitism and the evolution of host clutch sizes. Evol Ecol Res 5:559–570

Hauber ME, Samaš P, Anderson MG, Rutila J, Low J, Cassey P, Grim T (2014) Life-history theory predicts host behavioural responses to experimental brood parasitism. Ethol Ecol Evol. doi:10.1080/03949370.2013.851121

Honza M, Požgayová M, Procházka P, Tkadlec E (2007) Consistency in egg rejection behaviour: responses to repeated brood parasitism in the blackcap (*Sylvia atricapilla*). Ethology 113:344–351

Honza M, Procházka P, Požgayová M (2012) Within- and between-season repeatability of eggshell colouration in the great reed warbler *Acrocephalus arundinaceus*. J Avian Biol 43:91–96

Inc C (2005) StatXact 7 Statistical software for exact nonparametric inference. Cytel Inc., Cambridge, USA

Johnson DH (2002) The importance of replication in wildlife research. J Wildlife Manage 66:919–932

Kelly CD (2006) Replicating empirical research in behavioral ecology: how and why it should be done but rarely ever is. Q Rev Biol 81:221–236

Langmore NE, Cockburn A, Russell AF, Kilner RM (2009) Flexible cuckoo chick-rejection rules in the superb fairy-wren. Behav Ecol 20:978–984

Lotem A, Nakamura H, Zahavi A (1995) Constraints on egg discrimination and cuckoo-host co-evolution. Anim Behav 49:1185–1209

Martin P, Bateson P (2008) Measuring behaviour. Cambridge University Press, Cambridge

Martín-Vivaldi M, Soler JJ, Møller AP, Pérez-Contreras T, Soler M (2013) The importance of nest-site and habitat in egg recognition ability of potential hosts of the common cuckoo *Cuculus canorus*. Ibis 155:140–155

Moksnes A, Røskaft E, Korsnes L (1993) Rejection of cuckoo (*Cuculus canorus*) eggs by meadow pipits (*Anthus pratensis*). Behav Ecol 4:120–127

Moskát C, Hauber ME (2007) Conflict between egg recognition and egg rejection decisions in common cuckoo (*Cuculus canorus*) hosts. Anim Cogn 10:377–386

Moskát C, Hauber ME, Elek Z, Gommers M, Bán M, Groenewoud F, Versluijs TSL, Hoetz CWA, Komdeur J (2014) Foreign egg retention by avian hosts in repeated brood parasitism: why do rejecters accept? Behav Ecol Sociobiol. doi:10.1007/s00265-013-1654-y

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

Palomino JJ, Martin-Vivaldi M, Soler M, Soler JJ (1998) Females are responsible for ejection of cuckoo eggs in the rufous bush robin. Anim Behav 56:131–136

Peer BD, Rothstein SI (2010) Phenotypic plasticity in common grackles (*Quiscalus quiscula*) in response to repeated brood parasitism. Auk 127:293–299

Polačiková L, Grim T (2010) Blunt egg pole holds cues for alien egg discrimination: experimental evidence. J Avian Biol 41:111–116

Samaš P, Cassey P, Hauber ME, Grim T (2011) Repeatability of foreign egg rejection: testing the assumptions of co-evolutionary theory. Ethology 117:606–619

Samaš P, Heryán J, Grim T (2013a) How does urbanization affect dispersal in blackbirds (*Turdus merula*)? Sylvia 49:21–38 (in Czech, summary in English)

Samaš P, Grim T, Hauber ME, Cassey P, Weidinger K, Evans KL (2013b) Ecological predictors of reduced avian reproductive investment in the southern hemisphere. Ecography 36:809–818

Shaw RC, Hauber ME (2009) Experimental support for the role of nest predation in the evolution of brood parasitism. J Evol Biol 22:1354–1358

Shaw RC, Hauber ME (2012) Linking nest predation with brood parasitism in captive zebra finches: a multi-pair study. J Ethol 30:255–262

Sih A, Bell A, Johnson JC (2004) Behavioral syndromes: an ecological and evolutionary overview. Trends Ecol Evol 19:372–378

Soler M (2014) Long-term coevolution between avian brood parasites and their hosts. Biol Rev. doi:10.1111/brv.12075

Soler M, Palomino JJ, Martin-Vivaldi M, Soler JJ (2000) Lack of consistency in the response of rufous-tailed scrub robins *Cercotrichas galactotes*. Ibis 142:151–154

Soler M, Ruiz-Castellano C, Carra LG, Ontanilla J, Martín-Galvez D (2013) Do first-time breeding females imprint on their own eggs? Proc R Soc Lond B 280:20122518

Spottiswoode CN, Stevens M (2010) Visual modeling shows that avian host parents use multiple visual cues in rejecting parasitic eggs. Proc Natl Acad Sci U S A 107:8672–8676

Steiger JH (1980) Tests for comparing elements of a correlation matrix. Psychol Bull 87:245–251

Stevens M, Troscianko J, Spottiswoode CN (2013) Repeated targeting of the same hosts by a brood parasite compromises host egg rejection. Nat Commun 4:2475

Stokke BG, Rudolfsen G, Moksnes A, Røskaft E (2004) Rejection of conspecific eggs in chaffinches: the effect of age and clutch characteristics. Ethology 110:459–470

Stokke BG, Takasu F, Moksnes A, Røskaft E (2007) The importance of clutch characteristics and learning for antiparasite adaptations in hosts of avian brood parasites. Evolution 61:2212–2228

Stokke BG, Hafstad I, Rudolfsen G, Moksnes A, Møller AP, Røskaft E, Soler M (2008) Predictors of resistance to brood parasitism within and among reed warbler populations. Behav Ecol 19:612–620

Strausberger BM, Rothstein SI (2009) Parasitic cowbirds may defeat host defense by causing rejecters to misimprint on cowbird eggs. Behav Ecol 20:691–699

Svensson L (1992) Identification guide to European passerines. British Trust for Ornithology, Thetford

Thorogood R, Davies NB (2013) Reed warbler hosts fine-tune their defences to track three decades of cuckoo decline. Evolution 67:3545–3555

Trnka A, Požgayová M, Samaš P, Honza M (2013) Repeatability of host female and male aggression towards a brood parasite. Ethology 119:907–917

Tuero DT, Fiorini VD, Reboreda JC (2007) Effects of shiny cowbird *Molothrus bonariensis* parasitism on different components of house wren *Troglodytes aedon* reproductive success. Ibis 149:521–529

Visser ME, Lessells CM (2001) The costs of egg production and incubation in great tits (*Parus major*). Proc R Soc Lond B 268:1271–1277

Welbergen JA, Davies NB (2012) Direct and indirect assessment of parasitism risk by a cuckoo host. Behav Ecol 23:783–789