

Supplementary Notes

GLMM analysis of experimental results

Here, we compare the results of the Host First and Foreign First experiments with generalized linear mixed model (GLMM). We conducted these analyses using the R software¹ package lme4². We constructed a full model using survival as the response variable with binomial error structure, brood identity as a random effect, and hatch order, experiment type (Host First or Foreign First), chick type (host or foreign), and the experiment type x chick type interaction term as fixed effects (Supplementary Table 1). There was a strong effect of the experiment type x chick type interaction ($Z = 5.0$, $n = 30$ broods, 251 chicks, two-tailed $P < 0.0001$), as expected from the reversed pattern of host chick versus foreign chick survival. This confirms that the types of chicks (host or foreign) presented as referents on the first day affects the non-random survival of host and foreign chicks that are introduced on subsequent days.

Supplementary Table1: Fixed-effect parameters for full GLMM model

	Parameter estimate	Standard Error	Z	P
Intercept	2.28	0.52	4.41	<0.0001
Chick Type (host or foreign)	-0.81	0.40	-2.02	0.04
Experiment Type (Host First or Foreign First)	-1.92	0.51	-3.76	<0.001
Hatch Order	-0.58	0.11	-5.19	<0.0001
Chick Type x Experiment Type	3.03	0.61	5.0	<0.0001

We also provide an additional separate analysis of the Host First and Foreign First experiments using a GLMM analysis, as some readers may prefer this statistical

approach. However, this analysis is simply an alternative approach to the matched-pair analyses presented in the text. For each experiment, we constructed a full model with survival as the bivariate response variable, binomial error structure, brood as random effect, and hatch order, chick type and their interaction term as fixed effects. There was no significant interaction effect of hatch order and chick type on survival, so we eliminated this parameter from the model. We then tested the hypothesis that there was non-random mortality between host and foreign chicks for each treatment by removing chick type from the model and conducting a likelihood ratio test. As with the matched-pair comparisons, host chicks survived better than foreign chicks in Host First broods ($\chi^2 = 28.25$, $d.f. = 1$, one tailed $P < 0.001$). In contrast, foreign chicks had higher survival than host chicks in Foreign First broods ($\chi^2 = 3.79$, $d.f. = 1$, one-tailed $P = 0.025$). These results are nearly identical to those presented in the text using Wilcoxon signed-rank tests.

The costs of learning to recognize: a natural case of misimprinting

A remarkable observation in 2007 demonstrates how the costs of learning errors occur in nature. At nest 7322, we observed a pair of birds feeding two chicks from the neighboring territory on a few different days before their own chicks hatched. During a behavioral observation during the hatching stage, BEL witnessed parents vigorously pecking at their own chicks, a type of aggression seen only in experimental Host First and Foreign First broods and towards parasitic chicks in a natural brood (see below). Some chicks disappeared during the course of that observation and the rest disappeared over the course of the next few days. This observation suggested that these particular parents had learned the wrong chicks as their own and were paying the cost of this learning error.

Testing for recognition by discordancy

By design, broods in both the Host First and Foreign First experiments had an unequal total number of host and foreign chicks (Fig. 2a,b), at least during the first few days of the experiment, so an alternative explanation for the observed survival patterns is feasible. The results of these two experiments could simply reflect parental recognition of the majority phenotype, or ‘recognition by discordancy’³, because the chicks predicted and observed to have higher survival by the hatch order hypothesis in each experiment were also in the majority (Fig. 2a,b). However, four lines of evidence indicate that recognition based on the majority type, or discordancy, cannot explain our findings:

(1) *Lack of recognition in experimental broods where birds could have used discordancy but not hatching order to recognize foreign chicks:* By design, hatch order information was not available to birds in the Mixed Synchrony experiment, but in most broods host chicks were the majority type because the foreign chicks came from a few different nests. Thus, if birds use recognition by discordancy, foreign chicks should have suffered lower survival in broods with host chick majority, but they did not (Wilcoxon signed-rank, $W = 15.0$, $n = 21$, one-tailed $P = 0.28$).

(2) *Discordancy does not predict survival of parasitic chicks in naturally parasitized nests.* In naturally parasitized nests, parasitic chicks that hatched on the first hatching day survived as well as host chicks (Fig. 1, Fisher Exact one-tailed $P = 0.59$), whereas later-hatched parasitic chicks suffer higher mortality (pooling all later-hatched chicks, (Fisher Exact one-tailed $P = 0.009$). The hatch order mechanism predicts the absence of rejection of the first-hatched parasite chicks. Conversely, the discordancy hypothesis predicts that

these chicks should have been recognized and rejected because parasitic chicks are always less common than host chicks at naturally parasitized nests.

(3) *Two Mixed Synchrony experimental broods where hatch order but not discordancy was possible.* By chance, only host chicks were introduced on the first hatching day in these two broods (they are excluded from analyses presented in the text), and a roughly equal number of foreign chicks were returned on the second day (which yielded total brood sizes of 3 host & 3 foreign chicks in one nest, 4 host & 5 foreign chicks in the other). All host chicks in these broods survived (7 alive, 0 dead), but all but one foreign chick died (1 alive, 7 dead)—a non-random pattern (two-tailed Fisher's exact test: $P = 0.001$). Thus, recognition occurred in broods where discordancy was either not possible (first brood) or predicted an opposite pattern to what was observed (second brood).

(4) *Case of natural misimprinting.* The example of natural misimprinting at nest 7322, as described above, cannot be explained by discordancy because these birds rejected their own chicks despite being the majority. In fact there were no longer any of the neighboring chicks present when the parents killed off their own chicks. Instead, it seems clear that the birds misimprinted on the neighbor chicks and then rejected their own chicks once they hatched.

Evidence that learning occurs annually

The issue of when in their lifetime birds learn recognition cues is important for understanding the evolution of chick recognition as well as for interpreting the results of experiments. If birds learn through a single time imprinting^{4,5}, then experimentally altering the referents for recognition, either with foreign chicks (Foreign First) or a mix

of host and foreign chicks (Mixed Synchrony), would not affect the responses of older birds that would have already formed a recognition template from previous breeding attempts. With single time imprinting most birds that had bred previously would have formed correct templates in the past and would not have been fooled by our experiments. Thus, with single-time imprinting, the proportion of experienced versus first time breeders will affect the outcome of Foreign First and Mixed Synchrony experiments. We first discuss the demographic makeup of our population, and then the specific implications for our two experiments.

We determined the approximate age of a subset of adults in our study population by carefully quantifying rank leg coloration following a system that Crawford⁶ developed using a group of known-age coots. Though this method is not perfectly accurate (i.e., leg color is a strong but not perfect predictor of age⁶), it does enable us to determine whether the population was highly skewed towards yearlings or older adults.

We determined the age distribution for the two different phases of the study using slightly different methods. From 1988-1990 birds were captured at their nests and leg colors were described in detail from birds in the hand. Adult philopatry rates were so low⁷ that we were unable to collect demographic information on the same individuals across years and connect leg color information directly to demographic data. In 2008 we observed and categorized leg colors from free-ranging birds observed at close range from floating blinds. In all cases, we compared leg colors to reference color chips in the field; Crawford used Smithe's⁸ color system in his ageing system so we were able to match color chips to the Smithe colors described by Crawford.

In the 1988-1990 sample, 54% (34 of 79) of females were estimated to be at least two years old based on leg color; for males the estimate was 71 % (55 of 78). In the latter phase of the study we quantified leg coloration only in 2008, the year we focused on the Host First and Foreign First experiments. In that year, we estimated that 86% of males (24 of 28) and 66% of females (21 of 32) were older than one year. Overall, these results suggest that at least half of the breeding population were birds older than one year.

Given that more than half of our study population consists of adults older than one year, we would not expect to observe the patterns we observed in our experimental study if cues were learned only in the first breeding attempt. For example, if experienced breeders had learned recognition cues from their previous breeding attempt, we would have detected some discrimination in the Mixed Synchrony broods. No discrimination was observed. With similar logic, older birds in the Foreign First experiment should not have been fooled by the foreign chicks on the first hatching day, and their correct recognition should have canceled out the incorrect recognition we induced in young, first time breeders, effectively predicting no evidence for recognition. Contrary to this prediction, we observed clear evidence of recognition.

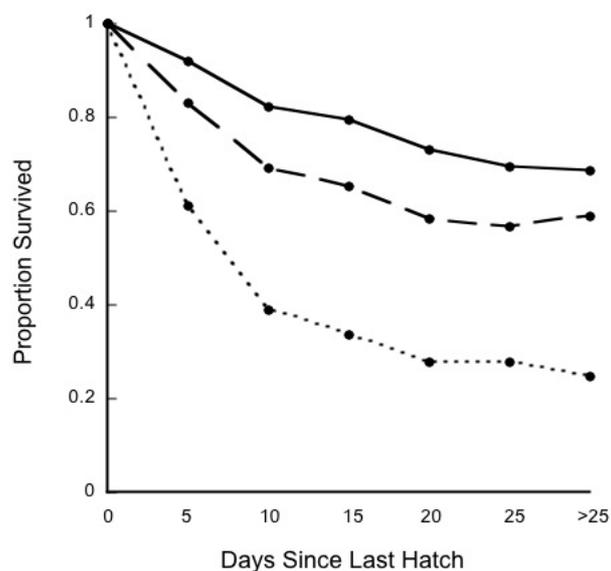
On the other hand, there is evidence that some birds may use previously acquired cues for recognition. If templates were only learned every year, we would expect no differences between the Host First and Foreign First experiments in the survival of reference chicks. However, we found that reference chicks in Foreign First broods suffered higher mortality than those in Host First broods (Restricted Maximum Likelihood, with nest as random factor: $F = 6.03$, $d.f. = 1$, $P = 0.02$), suggesting that parents were less likely to accept foreign chicks as referents. This suggests that some parents had prior information

on cues that reliably identify their own chicks and were less likely to accept foreign chicks as referents compared to their own chicks. The difference between the two experiments is expected because in the Foreign First experiment predictions for previous breeders and first time breeders are in opposite directions, whereas in the Host First experiment, predictions for previous and naïve breeders are in the same direction—both will favor host chicks over foreign chicks. Together, these observations suggest that recognition is largely based on templates that are completely relearned annually, but at least some birds appear to use previously acquired templates for recognizing chicks.

Evidence that rejection occurs early in parental care stage

The ease of behavioral observations at coot nests can vary across individuals, vegetation density, and time. American coot chicks are able to leave the nest for periods of time within a day of hatching, but rarely do so for the first several days. Throughout the 3 to 10 day hatching period, and for several days after, parents often feed the young in the nest, or very close to it, and most nests are hidden back in vegetation, making observations difficult. Therefore, our ability to conduct accurate censuses and feeding observations were often very limited during the first days of parental provisioning. In Host First broods, the timing of first census ranged from 1-13 days after the last hatching day (median 4 days). In Foreign First broods, first censuses were conducted 1-10 days after last hatching day (median 3 days). A significant proportion of the total mortality observed occurred before our first census (74% for Host First broods, and 43% for Foreign First broods). As a result, biased mortality of foreign chicks was already apparent in Host First broods at the first census (Wilcoxon signed-rank, $W = 38.5$, $n = 15$, one-tailed $P = 0.002$) and there was a trend for more host chicks to have died compared to

foreign chicks by the first census in Foreign First broods (Wilcoxon signed-rank, $W = -11.5$, $n = 15$, one-tailed $P = 0.059$). It is also clear that beyond 10 days after the last hatching day, mortality is less biased, as template, host and foreign chicks all suffer similarly low level of mortality (Supplementary Fig. 1). Overall, the combination of rapid mortality and difficulty in observing broods early in the chick stage meant that we had detailed early behavioral observations of the process of chick rejection for very few nests.



Supplementary Figure 1: Comparison of survival in Host First and Foreign First experiments. Lines show the proportions of “template” chicks (solid line), “in” chicks (i.e. host chicks in Host First broods and foreign chicks in Foreign First broods, dashed line), and “out” chicks (i.e. foreign chicks in Host First broods and host chicks in Foreign First broods, dotted line) across census periods during parental care.

Additional anecdotal evidence for chick recognition and rejection

Two observations not mentioned in the text add anecdotal evidence that chick recognition and rejection occurs in a natural context: a nest where a pair adopted back their own

chicks that had been laid as parasitic eggs in a neighbor's nest and a nest where parents killed a brood parasitic chick.

Adoption of own chicks laid as parasitic eggs in neighboring nest: On June 16, 2005, D.S. observed a rare case of adoption at Nest 5009. Three pairs of chicks with the same band combinations were observed (orange-orange, white-white and blue-blue), despite our records showing only one chick of each of these combinations were returned to the nest. Our records also indicated that between June 11 and June 13, three chicks with these same tag combinations disappeared from Nest 5026 on the adjacent territory. Therefore, we can be confident that these three chicks from Nest 5026 were adopted by parents at Nest 5009. Retrospective comparisons of eggs and hatching data indicated that these three adopted chicks hatched from eggs that were among the four eggs that female 5009 had laid parasitically in Nest 5026, prior to laying eggs in her own nest. Therefore, these three chicks were parasitic chicks that were simply 'adopted' back by their biological parents. Given that adoption is very rare in our population, the adoption of three chicks back by biological parents suggests that chick recognition was involved—even though these parents had no previous experience with the adopted chicks, the chicks presumably had the same recognition cues as the adopting parents' other chicks. The alternative explanation, that the adoption of three biological offspring was a chance event, can be ruled out. Given the fraction of host ($n = 10$) and parasitic chicks ($n = 4$) in the donor nest, the probability that all three adopted chicks were the genetic offspring of the adopting female was non-random (randomization test, $P = 0.01$, 10000 runs).

Infanticide of natural brood parasitic chick: On June 12, 2005, B.E.L. observed the first definitive instance of infanticide of an offspring in this population (documented in

Supplemental Movie 2) at Nest 5012 at Jaimeson Meadow. The chick was one of two parasitic chicks present in the brood of eight chicks, and was 4 days old at the time of observation. All six of the host chicks hatched before the two parasitic chicks. Both parents were involved in extreme parental aggression toward the parasitic chick, often chasing it even when it was well separated from the rest of the brood. The parents appeared to respond particularly strongly in response to distress calls of the chick. The chick was last seen swimming towards shore, outside of territory borders and D.S. found the chick dead on shore a few hours later. The identity of the chick was confirmed by both the nape tag color and the brood number written in permanent marker on the tag (we write brood numbers on tags to keep chick identities clear when setting up experiments but these numbers typically disappear over time).

Parental behaviors involved in the infanticide of this parasitic chick in 2005 share many common features with observations of chick rejection at one control nest in 2007, where parents appeared to mistakenly learn the neighbor's chicks as their own (see S5 above), and at three experimental broods in 2008 (documented in Supplemental Movies 2-4). In contrast to "normal" tousling behavior seen in both European coots, *Fulica atra*⁹ and American coots (Shizuka and Lyon unpublished manuscript, Supplemental Movie 1) these instances of infanticidal behavior included intense pecking, holding the chicks under water, and dragging the chicks by the head. In addition, parents were observed to actively seek out targeted chicks from up to 10m away. The fact that we observed similar behaviors in both natural instances as well as experimental broods suggest that active aggression is likely to be involved in the process of parasitic chick rejection. However, infanticide may not be unique to rejection of parasitic chicks, as a previous study on

European coots suggested that parental infanticide may also be used in the context of brood reduction, where parents reduce their brood size by killing off some of their own chicks⁹.

References

1. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>. (2009).
2. Bates, D., Maechler, M. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31. <http://CRAN.R-project.org/package=lme4> (2009).
3. Rothstein, S. I. Experimental and teleonomic investigation of avian brood parasitism. *Condor* **77**, 250-271 (1975).
- 4 Lotem, A. Learning to recognize nestlings is maladaptive for cuckoo *Cuculus-canorus* hosts. *Nature* **362**, 743-745 (1993).
5. Lotem, A., Nakamura, H. & Zahavi, A. Constraints on egg discrimination and cuckoo-host co-evolution. *Anim. Behav.* **49**, 1185-1209 (1995).
6. Crawford, R. D. Tarsal color of American coots in relation to age. *Wilson Bull.* **90**, 536-543 (1978).
7. Lyon, B. E. *The Ecology and Evolution of Conspecific Brood Parasitism in American Coots (Fulica americana)*. Thesis, Princeton University, Princeton N.J. (1992).
8. Smithe, F. B. *Naturalist's color guide*. American Museum of Natural History, New York (1975).
9. Horsfall, J. A. Brood reduction and brood division in coots. *Anim. Behav.* **32**, 216-225 (1984).