Can experienced birds select for Müllerian mimicry?

Eira Ihalainen, Leena Lindström, Johanna Mappes, and Sari Puolakkainen
Department of Biological and Environmental Science, PO Box 35, FI-40014,
University of Jyväskylä, Finland

Field experiments have shown that avian predators in the wild can select for similarity of warning signals in aposematic prey (Müllerian mimicry) because a common signal is better protected than a signal that is novel and rare. The original theory of Müllerian mimicry assumes that the mechanism promoting mimicry is predator learning; by sharing a signal, the comimic species share the mortality that is due to sampling by inexperienced predators. Predation events have not been observed in the wild, and learning experiments with naive bird predators in a laboratory have not unambiguously shown a benefit of a uniform signal compared with different signals. As predators in the field experiments are likely to be more experienced compared with previous laboratory experiments, we studied selection by experienced predators on a novel imperfect mimic. We trained great tits *Parus major* to avoid artificial aposematic models and subsequently introduced perfect and imperfect mimics at different frequencies. Birds with prior experience on the models selected against the imperfect mimics that were at a disadvantage also in a memory test conducted a week after their introduction. Selection against the imperfect mimics was antiapostatic. However, the imperfect mimics also benefited from some signal generalization to the models and possibly gained protection because the birds were familiar with the alternative cryptic prey that was also present. Our results suggest that experienced predators might be more important to the evolution of mimicry than the learning-based theory assumes. *Key words:* antiapostatic selection, aposematism, generalization, learning, memory. *[Behav Ecol]*

Predators can learn to recognize unprofitable prey and reject them by sight. Aposematic species, which are unprofitable due to, for example, toxicity, have typically conspicuous appearance (Poulton 1890). Conspicuous coloration can act as a warning signal that identifies the prey as unprofitable (Wallace 1867; Sherratt and Beatty 2003; Ham et al. 2006; see also Jansson and Enquist 2003) and facilitates the avoidance learning of predators (Gittleman and Harvey 1980; Gittleman et al. 1980; Sillén-Tullberg 1985; Roper and Wistow 1986; Alatalo and Mappes 1996). Aposematic prey are typically expected to suffer highest mortality when rare, and such antiapostatic selection is considered a barrier for the evolution of aposematic defense strategy (Lindström et al. 2001b; see also Endler 1988 and Ruxton et al. 2004 for a review of the "rarity problem").

Müllerian comimics are aposematic species that possess a similar warning signal; the best-known examples are insects with similar color patterns (see, e.g., Müller 1879; Plowright and Owen 1980; see also Ruxton et al. 2004). In his original theory of Müllerian mimicry, Müller (1879) proposed that a similar warning signal is beneficial for the comimics because species that mimic each other share the mortality that is due to "predator education." This theory is in line with the observation of antiapostatic selection on aposematic prey (Greenwood et al. 1989; Lindström et al. 2001b). If naive predators have to sample a certain number of individuals with a given signal to learn to reject prey with that particular signal by sight, comimics that use the same warning signal gain by reducing their per capita risk of getting sampled; the same amount of individuals that is needed to educate the predators is divided between the mimetic species. A common appearance is thus mutually (although not necessarily equally) ben-

eficial to all individuals using the signal. In his original formulation, Müller (1879) used a numerical example of 2 species of different population sizes to show how the benefit of mimicry is relatively greater to a rarer species if predators sample a fixed number of prey individuals from across 2 species.

Despite examples that show how signal similarity must be beneficial to aposematic prey in nature (Benson 1972; Mallet and Barton 1989; Dumbacher and Fleischer 2001; Kapan 2001; Symula et al. 2001), the mechanism promoting mimicry remains unclear. The learning phase of predators has been studied with laboratory experiments. Studies where great tits (*Parus major*) were used as inexperienced predators of artificial prey have not shown constant benefits of a shared warning signal versus different warning signals (Rowe et al. 2004; Ham et al. 2006; Lindström et al. 2006; Ihalainen et al. 2007; but see Rowland et al. 2007 and also Beatty et al. 2004). Instead, in field studies where live aposematic prey was used, transferred butterflies with a wing pattern that matched that of the locally abundant morph survived better than individuals with a different wing pattern (Benson 1972; Mallet and Barton 1989; Kapan 2001), and this difference in survival was likely due to predation by birds. In the wild, even simple predator communities can be heterogeneous in the level of experience the individuals have (see Endler and Mappes 2004). Although predation was never directly observed in the experiments where butterflies were transferred between locations, it is possible that the butterflies were attacked not solely by naive predators that were learning to avoid aposematic prey but by birds that were already familiar with the local butterfly morphs and did not generalize their learned avoidances to the transferred morphs. Experienced jacamars, which were potential predators of the transferred butterflies (see Mallet and Barton 1989), have indeed been shown to attack novel aposematic prey (Langham 2006). Kapan (2001) also found some evidence that selection against the introduced morphs was relaxed when they were released in a higher density, which also emphasizes the principle that aposematic species and
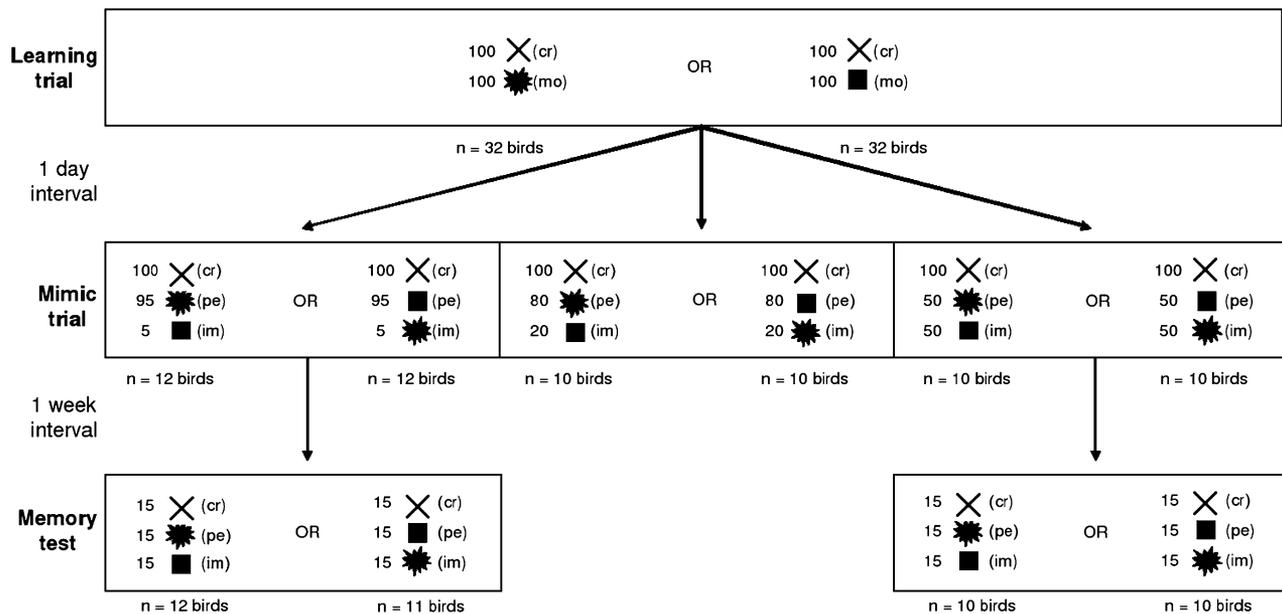
Figure 1
The experimental setup, number of birds, and the types and numbers of prey items presented in the learning trial, mimic trial, and memory test. Edible cryptic prey (cr) was marked with the cross signal. The star and square signals alternated as models (mo), perfect mimics (pe), and imperfect mimics (im).

comimics have "strength in numbers" (see also Rowland et al. 2007).

The field experiments differ from the learning experiments in their starting point: in field experiments, the predator community is presented with a new morph (in addition to naturally occurring species that they have previous experience of), whereas learning experiments present naive predators with different morphs simultaneously. Thus, the predators have had different level of experience in these studies; the field experiments focus more on generalization, and the laboratory experiments focus on the learning process.

Here, we present data from a laboratory experiment where bird predators had previous experience on aposematic prey. Great tits were first trained to forage on artificial prey and to discriminate between cryptic, edible prey, and highly unpalatable aposematic prey (models). Subsequently, we presented the birds with a perfect mimic (identical to the model) and an imperfect mimic that was not only visually different to the model (and thus unfamiliar to the birds) but also less aversive in taste. We presented the mimics at different frequencies. A week later, we tested the level of avoidance the birds had for the signaling prey in a "memory test." The results show that when studying predation on mimetic prey the context of presentation is important and that experienced predators could readily select for accurate Müllerian mimicry.

METHODS

The experiments were conducted at the Konnevesi Research Station in Central Finland from September to December 2001. Capturing and using the birds in the study was licensed by the Central Finland Regional Environment Center (permission number LS-46/01, 0901L0448/254) and the Experimental Animal Committee of the University of Jyväskylä (permission number 19/5.6.2001).

Birds

Great tits were caught from feeding sites with traps and ringed for identification. After the experiments, the birds were re-

leased back into the wild at the location of their capture. They were housed individually in illuminated plywood cages, sized $64 \times 46 \times 77$ cm with a daily light period of 11 h. Sunflower seeds, peanuts, tallow, and fresh water were available ad libitum, and occasionally, the birds were offered mealworms (larvae of *Tenebrio molitor*). Prior to the experiment, the birds were food deprived (ca. 2 h) to ensure motivation to search for the artificial prey but water was always available.

Artificial prey and experimental aviaries

The prey items were 8×8 mm paper shells with small pieces (ca. 0.1 g in weight) of almond glued inside. We created 2 aposematic prey types: models/perfect mimics and imperfect mimics. These prey items were made unpalatable by soaking the almond for an hour in a solution of 30 mL of water and 2 g of chloroquine phosphate (models and perfect mimics) or 0.25 g of chloroquine phosphate (imperfect mimics). These concentrations translate to a difference in taste to the birds (Lindström et al. 2006; Ihalainen et al. 2007). The birds typically react to the bitter taste of chloroquine phosphate by head shaking, beak wiping, and drinking. The difference in unpalatability level was included because it is assumed that a species can be a "mimic" that evolves to resemble a "model" in a Müllerian mimicry complex if it is less defended (see Mallet 1999). It has been previously shown that compared with highly unpalatable prey, learning rate is slower (and thus mortality is higher) for moderately unpalatable prey in this system (Lindström et al. 2006; Ihalainen et al. 2007). Therefore, we expected the selection pressure on the imperfect mimic to be a conservative upper limit estimate. The cryptic edible prey contained a piece of untreated almond.

A black-and-white signal was printed on both sides of all types of prey items. The cryptic pattern was a cross symbol that was also printed on the background of the aviary floor (see below and Figure 1). The 2 different warning signals were a square and an asymmetrical star (see Figure 1). These signals were equally visible, and great tits did not show preference for either of them when palatable (Ihalainen 2006).

The aviary background consisted of white paper with black crosses printed on it. The paper was covered with adhesive plastic, and "fake prey items" (printed crosses cut out of white cardboard) were glued onto it to make the background 3 dimensional. The background formed a grid of 22 columns and 15 rows, that is, 330 cells. Wooden boards were placed between the rows to ease prey handling and movement of the birds. The learning and mimic trials (see below) were run in an aviary with a floor area of 57 m^2 (the black-and-white background covered 41 m^2 of this). There were 8 perches at the height of 0.5 m for prey handling. For the memory test and training of the birds (see below), we used 2 smaller aviaries where the paper background formed a grid of 8 rows and 10 columns and covered an area of 10 m^2. These aviaries had 2 perches. All aviaries were illuminated with energy-saving light bulbs, and fresh water was available to the birds.

The birds were observed through a 1-way window during the trials. We used the coordinates of the grid to identify the prey type that was attacked by the bird as individual cells of the grid were occupied by 1 prey item only (see below).

Pretraining and learning trial

All birds ($n = 64$) were first trained stepwise to handle the artificial prey in their home cages. Blank white paper shells were used for the training prey. The birds were also familiarized with the aviaries and trained to forage from the black-and-white background to ensure they would start foraging in the actual trial and could utilize the perches. To familiarize the birds with the large aviary, several individuals at a time were allowed to feed and overnight there. The background grid was replaced with transparent plastic, and sunflower seeds and peanuts were spread on it. The small aviaries were used for background training during which the birds had to find and eat 12 plain white and 3 cryptic prey items (crosses). These were included to give the birds some experience with the cryptic prey before first encounter with the defended prey (for details of training, see Ihalainen et al. 2007).

In the learning trial, 100 highly unpalatable model prey items and 100 cryptic prey items were randomly distributed on the background grid. For half of the birds the square was the signal of the model, and for the other half it was the star. To gain experience on aposematic prey and to learn to avoid the model, the birds were allowed to kill 30 prey items. On average, it took 1h18min for the birds to complete a trial. A prey item was considered to be killed when the bird opened the paper shell and tasted or ate the almond inside. We recorded the number of each prey type killed.

Mimic trial

The mimic trial was run on the day after the learning trial. Hundred unpalatable prey items were presented: the unfamiliar mildly unpalatable imperfect mimic (with the signal that had not been the model in the learning trial) was introduced alongside the perfect mimics that were identical to the familiar models. The frequencies of the signaling prey were 95/5 or 80/20 or 50/50 for the perfect and the imperfect mimics, respectively. There were also a 100 cryptic prey items available. When dividing the birds to the frequency treatments, we took into account their "strength of avoidance," that is, the number of models the birds killed in the end of the learning trial (within the last 10 prey items attacked); there was no difference in avoidance between the frequency treatments (analysis of variance [ANOVA] $F_{2,58} = 0.326$, $P = 0.723$; also no effect of signal treatment $F_{1,58} = 2.502$, $P = 0.119$ and no interaction between the treatments $F_{2,58} = 0.329$, $P = 0.721$). Furthermore, the total number of models killed in the whole learning trial did not differ between the frequency treatments (ANOVA $F_{2,58} = 0.067$, $P = 0.935$), which ensured that initial differences in avoidance or the amount of negative experience would not confound the possible effect of mimic frequencies on prey mortalities (for further analysis of the whole learning trial, see Results). The birds were again allowed to kill 30 prey items (trials lasted on average ca. 53 min), and the total number of each prey type killed was recorded.

The experiment was therefore a 2 × 3 design (signal of the model × frequencies of the mimics), where there were 12 birds in the 95/5 frequency treatments and 10 birds in the other 4 treatments, so 64 birds altogether (Figure 1).

Memory test

The memory test was carried out in the small aviaries (see Methods) a week after the mimic trial. We included 43 of the 44 birds from the 95/5 and 50/50 frequency treatments only (1 bird died before this test). We put 30 cryptic prey items, 15 squares and 15 stars, on the background grid of 80 cells in a random order. The birds were allowed to kill 15 prey items (completing the test took on average ca. 28 min), and the total number of each prey type killed was recorded. All the prey items were palatable so that the birds could not strengthen their avoidance of the mimetic prey but could have relearned to accept them as edible prey. We did not detect such a trend when comparing the number of signaling prey items the birds took within the first and last 5 prey items they were allowed to kill during the memory test: a total of 13 birds increased their attacks on the signaling prey, and 13 birds decreased their attacks as the test progressed. The remaining 17 birds took equal numbers of signaling prey in the beginning and in the end of the test. Thus, we are not measuring the rate of memory extinction, but the attacks on models/perfect mimics and imperfect mimics in this simple test reflect the "level of learning" or the willingness of the birds to sample the signaling prey items after the preceding trials.

Data analysis

For the analyses, we calculated a relative predation risk for the models (learning trial) and mimics (mimic trial and memory test) by dividing the number of killed prey items of each type by the number expected by chance. For example, in the learning trial, when 100 models and 100 cryptic prey items were presented in the setup, the expected number of models killed was 15 because the birds were allowed to kill 30 prey items. Because the total number of prey presented in each treatment was constant, the relative risks give the same results as mortalities (proportion of each prey type killed) but on a different scale, where 1 denotes random predation. Random predation does not take into account the visibility difference of the aposematic and cryptic prey (i.e., likelihood of detection) but gives a reference point for comparing the calculated risks of the aposematic prey types. For some analyses, we similarly calculated relative predation risks for the models and mimics in the beginning and/or the end of a trial, that is, during the first or last 5 prey items killed. In those cases, the expected number killed by chance was calculated for each bird separately based on what was left in the setup for the bird to kill. The data were analyzed with SPSS 11.5 for Windows statistical program. All tests are 2 tailed, and we used ANOVA by ranks and nonparametric tests where the data did not meet the requirements of parametric tests.

RESULTS

Learning trial

In the learning trial, when only cryptic and model prey items were offered, the predation risk of the star models was lower
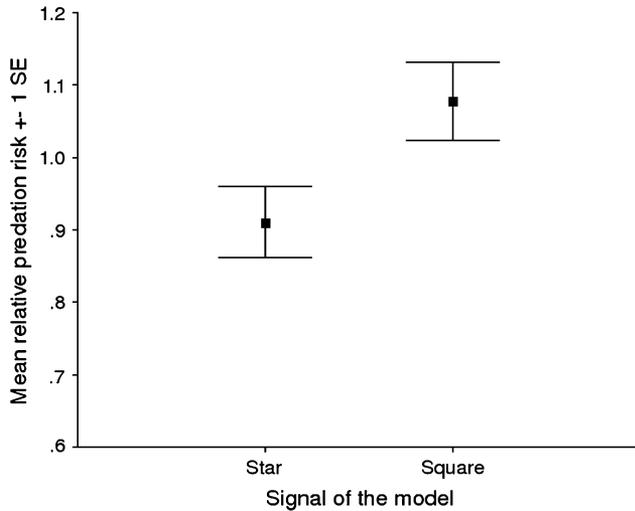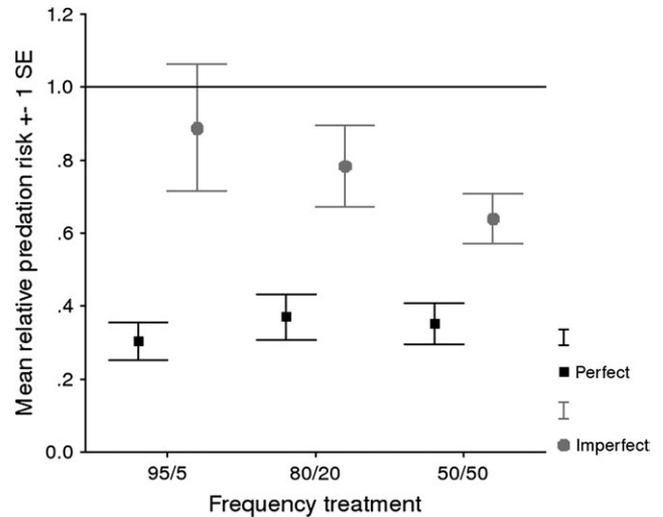
Figure 2
Mean relative predation risk (±SE) of star and square models in the learning trial.

Figure 3
Mean relative predation risks (±SE) of perfect (black squares) and imperfect mimics (gray circles) (±SE) in the mimic trial. The frequency treatments are 95 perfect and 5 imperfect mimics presented, 80 perfect and 20 imperfect mimics presented, and 50 perfect and 50 imperfect mimics presented. Reference line shows random predation (see Data analysis).

than that of the square models ($F_{1,62} = 5.213$, $P = 0.026$) showing that the star was more efficient a warning signal (Figure 2). There was neither effect of frequency treatment ($F_{2,58} = 0.067$, $P = 0.935$ as the strength of avoidance between the frequency treatments was controlled for, see Methods) nor 2-way interaction between the signal and frequency treatments ($F_{2,58} = 0.342$, $P = 0.712$). The birds learned to avoid the models: the mean predation risk for the star models was 1.28 (standard error [SE] = 0.09) at the beginning of the trial (during first 5 prey items killed) and decreased to 0.64 (SE = 0.06) at the end of the trial (during last 5 prey items killed) (paired samples t-test $t_{31} = 5.698$, $P < 0.001$). The mean risk of square models decreased from 1.54 (SE = 0.07) to 0.70 (SE = 0.07) ($t_{31} = 8.235$, $P < 0.001$).

Mimic trial

In the mimic trial, the predation risk of the perfect mimics (that were identical to the models presented in the learning trial) did not differ between the signal treatments (ANOVA $F_{1,58} = 2.538$, $P = 0.117$). The frequency treatment did not affect the risk of perfect mimics ($F_{2,58} = 0.391$, $P = 0.678$). Thus, familiar looking prey did not suffer from increased frequency of imperfect mimics and/or their own reduced frequency. There was also no 2-way interaction between the treatments ($F_{2,58} = 0.158$, $P = 0.855$).

The predation risk of the imperfect mimics was higher than that of the perfect mimics in all frequency treatments (paired samples t-test, all $P < 0.007$; Figure 3), which indicates that familiar aposematic prey had a benefit over unfamiliar prey; educated birds selected against the imperfect mimics. The predation risk of the imperfect mimics did not depend on their frequency (ANOVA by ranks $H_{2,58} = 1.267$, $P = 0.531$) contrary to what would have been expected if selection on the imperfect mimics was antiapostatic. Signal treatment had no effect on the risk of imperfect mimics ($H_{1,58} = 0.103$, $P = 0.748$), and there was no 2-way interaction between the frequency and signal treatments ($H_{2,58} = 3.020$, $P = 0.221$). However, when compared with random predation, only the imperfect mimics with a frequency of 50 were clearly avoided, that is, their risk was lower than 1.0 (1-sample t-test $t_{19} = -5.284$, $P = 0.001$). Imperfect mimics with a frequency of 20 were slightly avoided ($t_{19} = -1.942$, $P = 0.067$), but those with a frequency of 5 were not (Wilcoxon signed-ranks test:

$Z = -1.452$, $P = 0.147$), which indicates that the unfamiliar morph was at some disadvantage when rare (Figure 3).

The imperfect mimics seemed to gain protection from some generalization to the models. We compared the initial risk (relative predation risk within the first 5 prey items killed) of the imperfect mimics with the initial risk the models had in the beginning of the learning trial (i.e., when they were unfamiliar to the birds and likely attacked according to their visibility). We compared star mimics with star models and square mimics with square models because the signals differed in mortality in the learning trial; the comparison was thus made between birds in different signal treatments (frequency treatments were combined). The initial risk of the imperfect mimics was lower than that of the models (Mann–Whitney U test $Z = -2.221$, $P = 0.027$ for stars and $Z = -3.858$, $P < 0.001$ for squares; Figure 4).

Despite the milder taste of the imperfect mimics (see Methods), their overall predation risk during the whole mimic trial was not higher than the overall risk of the models in the learning trial: the risk of imperfect star mimics did not differ from the risk of star models (Mann–Whitney U test $Z = -1.434$, $P = 0.152$), and the squares survived better as imperfect mimics than as models ($Z = -2.547$, $P = 0.011$).

Memory test

In the following, the signaling prey are still termed perfect and imperfect mimics according to the treatments in the mimic trial even though all prey were palatable in this test. Similarly, frequency treatment also refers to the frequencies in the preceding mimic trial because the frequency of both co-mimics was 15 in this test.

The frequency treatment had no subsequent effect on the predation risks of the perfect (ANOVA by ranks $H_{1,39} = 0.866$, $P = 0.352$) or the imperfect mimics (ANOVA $F_{1,39} = 0.530$, $P = 0.471$; Figure 5). This indicates that the birds that had been presented with only 5 imperfect mimics avoided that signal in the memory test as much as birds that had been presented with 50 imperfect mimics. The risks of the mimics were not affected by the signal treatment ($H_{1,39} = 1.075$,

a)

b)
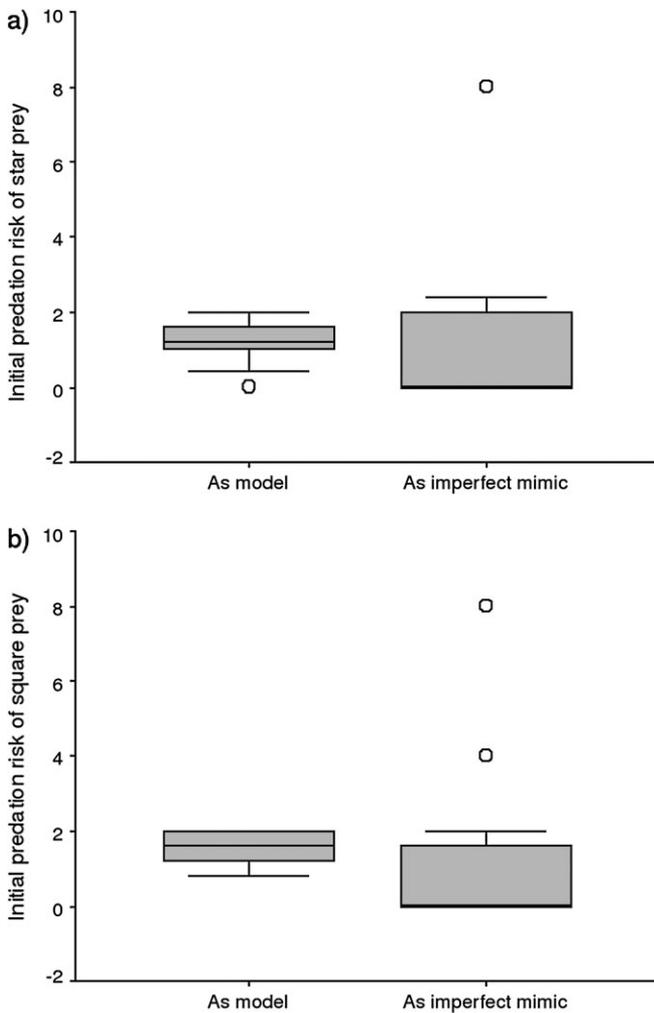
Figure 4
Initial predation risks of (a) the star prey as models and as imperfect mimics and (b) the square prey as models and as imperfect mimics. Boxplots show minimum, maximum, and the upper and lower quartiles around the median. The outliers denoted by open circles were included in the analyses.

$P = 0.300$ for perfect and $F_{1,39} = 0.916$, $P = 0.344$ for imperfect mimics), and there were no 2-way interactions between the signal and frequency treatments ($H_{1,39} = 0.640$, $P = 0.424$ for the risk of perfect and $F_{1,39} = 0.005$, $P = 0.942$ for the risk of imperfect mimics).

Overall, perfect mimics survived better than imperfect mimics (paired samples t-test $t_{42} = -2.862$, $P = 0.007$ for all treatments combined), but both signaling prey types were killed less than randomly (Wilcoxon signed-ranks test against random predation of 1.0, $P < 0.001$, for both mimic types; Figure 5).

DISCUSSION

In the present study, birds with prior experience on a warning signal selected against aposematic prey with an unfamiliar signal; imperfect mimics had a higher predation risk than the familiar-looking perfect mimics (Figures 3 and 5). Similar results have been found by releasing aposematic butterflies with a locally common or with a locally novel wing pattern in the wild where predators likely had experience of the locally common warning signal: the butterflies with the novel wing pattern had higher mortality than ones with the com-
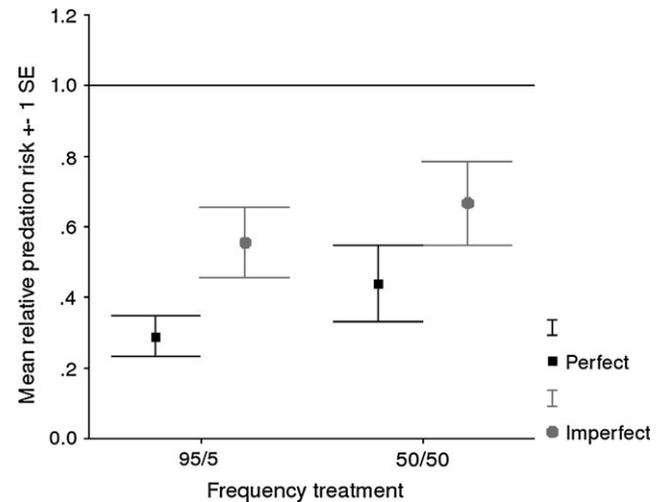
Figure 5
Relative predation risks (±SE) of perfect (black squares) and imperfect mimics (gray circles) in the memory test. The 2 original frequency treatments included in the test are 95 perfect and 5 imperfect mimics presented in the mimic trial and 50 perfect and 50 imperfect mimics presented in the mimic trial. Reference line shows random predation (see Data analysis).

mon pattern (Benson 1972; Mallet and Barton 1989; Kapan 2001). These results support the principle of Müllerian mimicry (Müller 1879) that a shared warning signal protects aposematic prey better than different warning signals, although they do not emphasize predator learning as the mechanism selecting for mimicry.

There are 2 factors that could have caused the difference in predation risk between the perfect and imperfect mimics in this experiment; the imperfect mimics looked different to the familiar signal, and they were less unpalatable than the models and perfect mimics. Consequently, the birds could have been more motivated to sample the milder imperfect mimics as they were recognizable (Lindström et al. 2006; Ihalainen et al. 2007; see Pearce 1997, p. 56–59). The imperfect mimics also seemed to have some disadvantage of rarity (predation risk was less than random only when the imperfect mimics were common; Figure 3), which is in line with the observation that novel aposematic prey face antiapostatic selection (Lindström et al. 2001b).

Antiapostatic selection was possibly attenuated because the birds did generalize between the models and the imperfect mimics to some extent. In the mimic trial, the initial predation risk of the imperfect mimics was lower than the initial risk of models in the learning trial when the models were unfamiliar to the birds and were likely attacked according to their high visibility (Figure 4). This indicates that the birds may have had altogether less interest in sampling the unfamiliar, imperfect mimics after their experience with the models. The models and imperfect mimics also resembled each other considerably more than they resembled cryptic prey, and generalization between them may be a result of simple categorization to edible and inedible prey (see Wallace 1867; Sherratt and Beatty 2003). Additionally, experience on the cryptic prey alone could have decreased the birds' willingness to attack the mimics in the mimic trial (Lindström et al. 2001a), and consequently, strong antiapostatic selection was not observed when comparing the predation risks of imperfect mimics directly between the frequency treatments.

The memory test where all prey was palatable showed the same general pattern that the imperfect mimics had higher predation risk than the perfect mimics (Figure 3). This too

could be either due to the relatively unfamiliar signal of the imperfect mimics or due to the birds being more willing to attack prey they expect to be only moderately unpalatable (i.e., moderately unpalatable prey could have a higher attack asymptote compared with highly unpalatable prey, see, e.g., Speed 1993; Speed 1999). These effects cannot be reliably separated from these data, but some aspects of the data question the importance of moderate unpalatability: the birds had very little actual experience of the mild taste of the imperfect mimics. We compared how many models/perfect mimics (these were combined as they were identical) and imperfect mimics the birds had tasted or eaten during the learning and mimic trials. In the beginning of the memory test, an average bird from the 95/5 frequency treatments had experience of (had killed) 19.38 (SE = 1.22) models/perfect mimics and 0.67 (SE = 0.13) imperfect mimics, whereas an average bird from the 50/50 treatment had experience on 17.25 (SE = 1.30) models/perfect mimics and 4.8 (SE = 0.51) imperfect mimics. Whether the birds could have learned from 0.67 or 4.8 imperfect mimics that their unpalatability is only moderate is questionable.

Furthermore, birds from the different frequency treatments seemed to have the same "level of learning" in the memory test (Figure 5): the different numbers of perfect and imperfect mimics that had been presented to them in the mimic trial had no effect on their readiness to attack these signaling prey in the future (see also Lindström et al. 2001b). Although frequencies affected the survival of the prey in the mimic trial, they seemed to have less effect from the birds' perspective: the birds' "actual experience" of models/perfect mimics in the different treatments during the learning and mimic trials was comparable and the absolute numbers of killed imperfect mimics generally very low. Considering how little negative experience the birds had with the imperfect mimics, their predation risk in the memory test is perhaps surprisingly low, but here too, the availability of palatable alternative (cryptic) prey is likely to be important (see also Kokko et al. 2003; Sherratt et al. 2004) in addition to signal generalization.

During the learning trial, when the birds were trained to avoid the models, the models with the star signal had a lower predation risk than those with the square signal (Figure 2; see also Ihalainen 2006; Lindström et al. 2006), that is, the 2 warning signals differed in efficacy. Despite this initial difference, the weaker signal did not encourage the birds to keep attacking the prey because in the mimic trial and in the memory test the difference in predation between the signal treatments had disappeared. Additionally, the risk of the imperfect mimics did not depend on their signal when they were introduced. This is interesting because generally the same things that promote learning should also promote memory (Shettleworth 1998, p. 246) so that the characteristics of prey that enhance learning are also beneficial in the long term (see, e.g., Gittleman and Harvey 1980; Roper and Redston 1987; Roper 1994 on conspicuousness) at least if foraging decisions are seen simply as a result of learning and forgetting.

The present results suggest that experience of the predators makes an important difference in the previous laboratory and field studies and also that they shed light on slightly different aspects of mimicry: experiments that compare learning rates of predators facing 1 signal or 2 signals simultaneously are perhaps more mechanistic than studies where very small numbers of prey with an unfamiliar signal are under interest. Experienced birds seem to select for Müllerian mimicry in the sense that a common signal is a better protection against them. On the other hand, it is likely that mimicry evolves through imperfect stages (see Turner 1977; Balogh and Leimar 2005). If imperfect mimics never survive, interspecific mimicry is unlikely to evolve. Our results suggest that new

imperfect mimics can initially have similar problems in spreading than aposematic morphs in general (see Mallet and Joron 1999). However, the magnitude of such problems is likely to be very context dependent: the spread of a morph within a species depends not only on its fitness relative to the wild type but also on whether predators generalize the new mimic morph to a familiar candidate model, which depends not only on the degree of visual similarity but also on other aspects.

For example, bird predators may generalize more readily from negative than positive experience (Ham et al. 2006), show peak shifts (Gamberale and Tullberg 1996), or generalize asymmetrically between 2 different signals (Goodale and Sneddon 1977). It has been suggested that predators should generalize more broadly between Müllerian comimics that are all unpalatable than between unpalatable models and their edible (Batesian) mimics (Fisher 1958; Huheey 1988). Also, the severity of punishment from eating a model affects how broadly the predator generalizes to other signals (Goodale and Sneddon 1977; Darst and Cummings 2006). Some species can be better models for Müllerian mimicry complexes than others due to, for example, abundance or earlier emergence (Mallet 1999), and evolving Müllerian mimics with resemblance to such species can be spread more successfully than other morphs (see Mallet and Joron 1999; Beatty et al. 2004).

From this perspective, it seems plausible that although a new comimic morph that resembles, for example, a more common species would benefit from reduced mortality due to predator education, generalization by experienced predators will also have an effect on what kind of signal mutations will survive. Despite the ongoing discussion (e.g., Mallet 1999; Sherratt et al. 2004) and recent data (Langham 2006; Rowland et al. 2007; present results), the relative importance of naive and experienced predators as selective agents in Müllerian mimicry remains unsolved because there are no comprehensive data from the wild on how much they kill unfamiliar aposematic prey. It may be that in seasonal environments predation pressure on aposematic prey increases considerably when naive predators start to feed on their own (Ojala 2006), whereas in environments where breeding is not as seasonal and the ratio of naive and older predators is different, experienced individuals may play an essential role in selecting for signal similarity in Müllerian mimicry.

REFERENCES

Alatalo RV, Mappes J. 1996. Tracking the evolution of warning signals. Nature. 382:708–710.
Balogh ACV, Leimar O. 2005. Müllerian mimicry: an examination of Fisher's theory of gradual evolutionary change. Proc R Soc Lond B. 272:2269–2275.
Beatty CD, Beirinckx K, Sherratt TN. 2004. The evolution of müllerian mimicry in multispecies communities. Nature. 431:63–67.
Benson WW. 1972. Natural selection for Müllerian mimicry in Heliconius erato in Costa Rica. Science. 176:936–938.
Darst CR, Cummings ME. 2006. Predator learning favours mimicry of a less-toxic model in poison frogs. Nature. 440:208–211.

Dumbacher JP, Fleischer RC. 2001. Phylogenetic evidence for colour pattern convergence in toxic pitohuis: Müllerian mimicry in birds? Proc R Soc Lond B. 268:1971–1976.

Endler JA. 1988. Frequency-dependent predation, crypsis and aposematic coloration. Philos Trans R Soc Lond B Biol Sci. 319:505–523.

Endler JA, Mappes J. 2004. Predator mixes and the conspicuousness of aposematic signals. Am Nat. 163:532–547.

Fisher RA. 1958. The genetical theory of natural selection. 2nd rev. ed. New York: Dover Publication.

Gamberale G, Tullberg BS. 1996. Evidence for peak-shift in predator generalization among aposematic prey. Proc R Soc Lond B. 263:1329–1334.

Gittleman JL, Harvey PH. 1980. Why are distasteful prey not cryptic? Nature. 286:149–150.

Gittleman JL, Harvey PH, Greenwood PJ. 1980. The evolution of conspicuous coloration: some experiments in bad taste. Anim Behav. 28:897–899.

Goodale MA, Sneddon I. 1977. The effect of distastefulness of the model on the predation of artificial Batesian mimics. Anim Behav. 25:660–665.

Greenwood JJD, Cotton PA, Wilson DM. 1989. Frequency-dependent selection on aposematic prey: some experiments. Biol J Linn Soc. 36:213–226.

Ham AD, Ihalainen E, Lindström L, Mappes J. 2006. Does colour matter? The importance of colour in avoidance learning, memorability and generalisation. Behav Ecol Sociobiol. 60:482–491.

Huheey JE. 1988. Mathematical models of mimicry. Am Nat. 131:S22–S41.

Ihalainen E, Lindström L, Mappes J. 2007. Investigating Müllerian mimicry: predator learning and variation in prey defences. J Evol Biol. 20:780–791.

Ihalainen E. 2006. Experiments on defensive mimicry: linkages between predator behaviour and qualities of the prey [dissertation]. Jyväskylä (Finland): University of Jyväskylä. p. 37.

Jansson L, Enquist M. 2003. Receiver bias for colourful signals. Anim Behav. 66:965–971.

Kapan DD. 2001. Three-butterfly system provides a field test of müllerian mimicry. Nature. 409:338–340.

Kokko H, Mappes J, Lindström L. 2003. Alternative prey can change model-mimic dynamics between parasitism and mutualism. Ecol Lett. 6:1068–1076.

Langham GM. 2006. Rufous-tailed jacamars and aposematic butterflies: do older birds attack novel prey? Behav Ecol. 17:285–290.

Lindström L, Alatalo RV, Lyytinen A, Mappes J. 2001a. Predator experience on cryptic prey affects the survival of conspicuous aposematic prey. Proc R Soc Lond B. 268:357–361.

Lindström L, Alatalo RV, Lyytinen A, Mappes J. 2001b. Strong antiapostatic selection against novel rare aposematic prey. Proc Natl Acad Sci. 98:9181–9184.

Lindström L, Lyytinen A, Mappes J, Ojala K. 2006. Relative importance of taste and visual appearance for predator education in Müllerian mimicry. Anim Behav. 72:323–333.

Mallet J, Joron M. 1999. Evolution of diversity in warning color and mimicry: polymorphisms, shifting balance, and speciation. Annu Rev Ecol Syst. 20:201–233.

Mallet JLB. 1999. Causes and consequences of a lack of coevolution in Müllerian mimicry. Evol Ecol. 13:777–806.

Mallet JLB, Barton NH. 1989. Strong natural selection in a warning-color hybrid zone. Evolution. 43:421–431.

Müller F. 1879. Ituna and Thyridia: a remarkable case of mimicry in butterflies. Trans Entomol Soc Lond. XX–XXIX.

Ojala K. 2006. Variation in defence and its fitness consequences in aposematic animals—interactions among diet, parasites and predators [PhD dissertation]. Jyväskylä (Finland): University of Jyväskylä.

Pearce JM. 1997. Animal learning and cognition. An introduction, 2nd ed. Hove (UK): Psychology Press.

Plowright RC, Owen RE. 1980. The evolutionary significance of bumble bee color pattern: a mimetic interpretation. Evolution. 34:622–637.

Poulton EB. 1890. The colours of animals, their meaning and use, especially considered in the case of insects. London: Kegan Paul, Trench, Trübner, & CO. Ltd.

Roper TJ. 1994. Conspicuousness of prey retards reversal of learned avoidance. Oikos. 69:115–118.

Roper TJ, Redston S. 1987. Conspicuousness of distasteful prey affect the strength and durability of one-trial avoidance learning. Anim Behav. 35:739–747.

Roper TJ, Wistow R. 1986. Aposematic colouration and avoidance learning in chicks. Q J Exp Psychol. 38B:141–149.

Rowe C, Lindström L, Lyytinen A. 2004. The importance of pattern similarity between Müllerian mimics in predator avoidance learning. Proc R Soc Lond B. 271:407–413.

Rowland HM, Ihalainen E, Lindström L, Mappes J, Speed MP. 2007. Co-mimics have a mutualistic relationship despite unequal defences. Nature. 448:62–67.

Ruxton GD, Sherratt TN, Speed MP. 2004. Avoiding attack—the evolutionary ecology of crypsis, warning signals and mimicry. 1st ed. Oxford: Oxford University Press.

Sherratt TN, Beatty CD. 2003. The evolution of warning signals as reliable indicators of prey defence. Am Nat. 162:377–389.

Sherratt TN, Speed MP, Ruxton GD. 2004. Natural selection on unpalatable species imposed by state-dependent foraging behaviour. J Theor Biol. 228:217–226.

Shettleworth SJ. 1998. Cognition, evolution and behaviour. New York: Oxford University Press.

Sillén-Tullberg B. 1985. Higher survival of an aposematic than of a cryptic form of a distasteful bug. Oecologia. 67:411–415.

Speed MP. 1993. Müllerian mimicry and the psychology of predation. Anim Behav. 45:571–580.

Speed MP. 1999. Batesian, quasi-Batesian or Müllerian mimicry? Theory and data in mimicry research. Evol Ecol. 13:755–776.

Symula R, Schulte R, Summers K. 2001. Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hyphothesis. Proc R Soc Lond B. 268:2415–2421.

Turner JRG. 1977. Butterfly mimicry: the genetical evolution of an adaptation. Evol Biol. 10:163–206.

Wallace AR. 1867. Journal of Proceedings of the Entomological Society of London. In: Transactions of the Entomological Society of London. 3rd Series (years 1864–1869); lxxxviii–lxxxi.