This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Rönkä, Katja; De Pasqual, Chiara; Mappes, Johanna; Gordon, Swanne; Rojas Zuluaga, Bibiana

Title: Colour alone matters: no predator generalization among morphs of an aposematic moth

Year: 2018

Version: Accepted version (Final draft)


Rights: CC BY-NC-ND 4.0

Rights url: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the original version:

Local warning colour polymorphism, frequently observed in aposematic organisms, is evolutionarily puzzling. This is because variation in aposematic signals is expected to be selected against due to predators’ difficulties associating several signals with a given unprofitable prey. One possible explanation for the existence of such variation is predator generalization, which occurs when predators learn to avoid one form and consequently avoid other sufficiently similar forms, relaxing selection for monomorphic signals. We tested this hypothesis by exposing the three different colour morphs of the aposematic wood tiger moth, *Arctia plantaginis*, existing in Finland to local wild-caught predators (blue tits, *Cyanistes caeruleus*). We designed artificial moths that varied only in their hindwing coloration (white, yellow and red) keeping other traits (e.g. wing pattern and size) constant. Thus, if the birds transferred their aversion of one morph to the other two we could infer that their visual appearances are sufficiently similar for predator generalization to
take place. We found that, surprisingly, birds showed no preference or aversion for any of the three morphs presented. During the avoidance learning trials, birds learned to avoid the red morph considerably faster than the white or yellow morphs, confirming previous findings on the efficacy of red as a warning signal that facilitates predator learning. Birds did not generalize their learned avoidance of one colour morph to the other two morphs, suggesting that they pay more attention to conspicuous wing coloration than other traits. Our results are in accordance with previous findings that coloration plays a key role during avoidance learning and generalization, which has important implications for the evolution of mimicry. We conclude that, in the case of wood tiger moths, predator generalization is unlikely to explain the unexpected coexistence of different morphs.

**Keywords.**

learning, polymorphism, predator generalization, predator–prey interactions, warning signals, wood tiger moth

Aposematic organisms display warning signals that predators learn to associate with their unprofitability (Poulton, 1890). The survival of such prey is thus highly dependent on a predator’s ability to learn, remember and generalize their learned avoidance to other individuals sharing the same warning signal (reviewed in Ruxton, Sherratt, & Speed, 2004). Signal sharing among aposematic prey benefits both the prey and their potential predators: (1) a given individual has a lower risk of predation when more individuals share the same warning signal, and (2) predators benefit from not having to sample as many unprofitable or toxic prey and can more easily remember one and not multiple signals (Ghirlanda & Enquist, 2003; Guilford & Dawkins, 1991; Müller, 1878; Rowland, Ihalainen, Lindström, Mappes, & Speed, 2007; ten Cate & Rowe, 2007). Therefore, local polymorphism in warning coloration is expected to be selected against (Chouteau,

Despite the predicted disadvantages, warning signal polymorphisms are present in several aposematic taxa, such as frogs (Amézquita, Castro, Arias, González, & Esquivel, 2013; Rojas & Endler, 2013), ladybirds (O’Donald & Majerus, 1984; Průchová et al. 2014) and butterflies (Jiggins & McMillan, 1997). In fact, they seem to be more common than expected considering that warning signals are predicted to be under positive frequency-dependent selection (Müller, 1878; Ruxton, Sherratt, & Speed, 2004). One possible explanation for the co-occurrence of several warning signal forms within the same population is predator generalization. This refers to a predator’s ability to transfer its learned avoidance of one signal to other signal(s) that share common characteristics (Gamberale-Stille & Tullberg, 1999; Lindström, Alatalo, Mappes, Riipi, & Vertainen et al. 1999b; Guilford & Dawkins, 1991; Mappes & Alatalo 1997). Generalization can be symmetric, meaning that once one colour is learned it is equally possible to transfer the learned aversion to other similar colours, or asymmetric, implying that transferring a learned avoidance from one colour to other(s) depends on the signal salience (Aronsson & Gamberale-Stille, 2008; Exnerová et al., 2006; Gamberale & Tullberg, 1996; Gamberale-Stille & Tullberg, 1999; Ham, Ihalainen, Lindström & Mappes, 2006; Ruxton, Franks, Balogh, & Leimar, 2008; Waldron et al., 2017).

Predator learning involves different cognitive processes that establish the association between warning coloration and unprofitability, and aid the memorization of this association once established. This learning process may vary between predators even at intraspecific levels (e.g. Adamová-Ježová et al., 2016; Endler & Mappes, 2004; Exnerová et al., 2010; 2015; Karlíková et al., 2016; Lindström, Alatalo, & Mappes, 1999a; Sherratt & Macdougall, 1995; Skelhorn, Halpin, & Rowe, 2016). Predators may also vary in their ability to cope with defended prey, due for example to dietary conservatism (Marples & Kelly, 1999; Mettke-Hofmann et al., 2002; Webster &
Lefebvre, 2000; Turini, Veselý & Fuchs, 2016). Therefore, investigating how predators learn to associate the appearance of prey with the noxious effects of their unprofitability is crucial to understanding how signal variation can be maintained within a population. During the learning process predators acquire information about the nutrient and toxin content of aposematic prey. Thus, individual predators are expected to make different decisions on how to use the information gathered from an encounter with aposematic prey (Exnerová et al., 2003; Exnerová et al., 2007; Halpin, Skelhorn, & Rowe, 2014; Lynn, 2005; Skelhorn et al., 2016; Trimmer et al., 2011), and modify their ingestion of toxic prey according to their toxic burden (Skelhorn & Rowe, 2007).

Generalized avoidance should be broad and persist for a relatively long time to offer protection to different warningly coloured prey morphs. On the other hand, naïve predators can also avoid warningly coloured prey due to innate wariness, neophobia or dietary conservatism (Exnerová et al., 2007; Lindström, Alatalo & Mappes, 1999; Marples & Kelly, 1999; Marples & Mappes 2011), which could be further reinforced by the short-term effects of negative experience with other aposematic prey. It has been suggested that multiple modalities of warning signals can help predators discriminate between palatable and unpalatable prey (Siddall & Marples, 2008, Kazemi, Gamberale-Stille & Leimar, 2015). However, generalized avoidance of aposematic prey can also be based on cues of different sensory modalities, such as odour, sound, colour or pattern or combinations of these. Depending on the cognitive processes of predators, they could also associate their negative experience with certain stimuli to any other stimuli encountered simultaneously (Mackintosh, 1975; Pavlov, 1927). These results emphasize the importance of studying how multiple cues and separate signal components influence a predator’s decision to attack prey (Kikuchi, Mappes, Sherratt & Valkonen, 2016; Rowe & Halpin, 2013).

Here, we tested the hypothesis that the hindwing colour polymorphism of an aposematic moth is enabled by predator generalization, and investigated whether or not that generalization is symmetric. We exposed paper models of the different hindwing colour morphs of the wood tiger
moth, *Arctia plantaginis* (formerly known as *Parasemia plantaginis*) to natural predators (blue tits, *Cyanistes caeruleus*), and examined whether, once they learned to avoid one of the colour morphs, they would generalize this aversion to the two unlearned colour morphs, which would allow multiple morphs to coexist. A lack of generalization among colour morphs would mean that birds pay more attention to colours than to other cues of the moth wings.

**METHODS**

The wood tiger moth (Rönkä, Mappes, Kaila, & Wahlberg, 2016) is an aposematic diurnal moth with a Holarctic distribution (Hegna, Galarza, & Mappes, 2015). They have two different chemical defences, one of which is secreted from the prothoracic glands. Although the chemical composition is not fully known, these fluids contain two types of methoxypyrazines, which are produced de novo (Burdfield-Steel, Pakkanen, Rojas, Galarza & Mappes 2016) and make them a deterrent to birds. Experiments with bird predators suggest that the fluids of yellow males have a more repulsive odour (Rojas et al. 2017), while those of white males taste worse (Rojas, Burdfield-Steel & Mappes 2015). Individuals vary in the degree of melanization and black patterning of the wings, as well as in levels of chemical defence, but the most striking feature of the wood tiger moth is its local hindwing colour polymorphism (Hegna et al., 2015). In Europe, its forewings present a black and white pattern in both males and females, whereas the hindwing colour combined with black pattern differs between the sexes (e.g. Galarza, Nokelainen, Ashrafi, Hegna, & Mappes, 2014; Hegna & Mappes, 2014). The distinct white and yellow male morphs are genetically determined by one autosomal locus and at least three alleles, dominant white, recessive white and intermediate yellow (Galarza, Nokelainen & Mappes 2016), while female hindwing coloration varies continuously from yellow to red (Lindstedt et al. 2017; Fig. 1). In Finland, for example, yellow and white males may occur within one population (Nokelainen, Valkonen, Lindstedt, & Mappes, 2014) whereas female hindwing coloration is mostly red (Hegna et al., 2015).
To study the reaction of bird predators (see below for details on procedure) to the different hindwing colour morphs, we used artificial moth models. The usage of artificial prey allows for the controlled manipulation of one or more warning signal components at a time, while accounting for how predators (i.e. birds) would see them (Endler & Mielke 2005). In this way, other components can be kept constant and independent of prey qualities, such as the variation in the level of chemical defence or behaviour (Karlíková et al., 2016; Lindström et al., 1999a; Veselý & Fuchs 2009). Here, our artificial moth models eliminated individual variation in moth size, shape, degree of melanization, wing pattern, wing posture, behaviour, smell or taste. Model wings were constructed with the software GIMP (2.8.16; http://www.gimp.org/) from pictures of a real male wood tiger moth specimen collected in Finland. Pictures of one forewing and one hindwing of a typical white moth were duplicated to obtain a symmetric pattern for the whole model. The melanization pattern of the moths used was a representative sample of a wing pattern in Finland (Fig. 1). To control for the amount and shape of melanized (mainly black) pattern of the wings, yellow and red models were created from the same wing picture, changing the hue of the white parts of the hindwing towards yellow or red. Finished models were printed double sided (HP Color LaserJet CP2025) on waterproof (Rite In The Rain, Tacoma, WA, U.S.A.) paper. To ensure that the model colours resembled the real wood tiger moth morphs, colour reflectance was measured with an Ocean Optics Maya2000 Pro spectrometer and average reflectance curves from three spots in the model hindwing coloration were compared to average reflectance curves of white, yellow and red moth hindwings (Fig. 1). Models were then cut out from the paper and completed with a body made of rolled pastry, composed of two parts of lard, six parts of coarse wheat flour and one part of water to make them edible. The total body weight was $0.04 \pm 0.005$ g. Bodies were dyed on top with black food colouring, to make models resemble the real moths as accurately as possible. Finally, bodies were glued on the paper models with nontoxic glue (UHU stick).
**Bird predators**

Blue tits were chosen as predators for several reasons: (1) they are visual foragers and their visual capabilities are well known (Hart, Partridge, Cuthill, & Bennett, 2000; Hart & Vorobyev, 2005), ensuring that they are able to distinguish all of the wood tiger moth’s colour morphs; (2) they have been used in several experiments on coloration (e.g. Dimitrova & Merilaita, 2010; Exnerová et al., 2007; Kikuchi et al., 2016) and wood tiger moths (Nokelainen, Hegna, Reudler, Lindstedt, & Mappes, 2012), and also with similar moth models (Rojas, Burdfield-Steel & Mappes 2015); (3) tits are likely to be important natural predators of wood tiger moths in Finland (Nokelainen, Valkonen, Lindstedt, & Mappes, 2014); and (4) blue tits are common in central Finland, and easy to capture and keep in captivity for a short period of time.

The birds used for the experiment were caught from Konnevesi Research Station and City of Jyväskylä (central Finland), maintained individually in plywood cages with a perch, water bowl and food ad libitum, and kept on a 12:12 h light:dark cycle. Each bird was weighed before and after the experiment, ringed, and its sex and age were determined before being released to the same place of capture. Birds were used with permission from the Central Finland Centre for Economic Development, Transport and Environment and licensed from the National Animal Experiment Board (ESAVI/9114/04.10.07/2014) and the Central Finland Regional Environment Centre (VARELY/294/2015). All experimental birds were used according to the ASAB/ABS Guidelines for the treatment of animals in behavioural research and teaching.

**Experimental procedures**

The experiment consisted of three phases: a preference test, a learning test and a generalization test (see details below). Each bird was tested individually and only once for each part of the experiment. The experiment was conducted between November 2015 and March 2016 at Konnevesi Research Station.
Station, in central Finland and lasted, on average, 3 days for each individual, depending on how long the bird took to complete the different tests.

Trials took place in experimental custom-built plywood cages (50x50 cm and 70 cm high) illuminated with a daylight lamp (Exo Terra Repti Glo 10.0 UVB, http://exo-terra.com/). Each aviary had a perch and a water bowl (access ad libitum). Birds were observed through a small mesh-covered window situated on the front of the cage, and filmed with a Canon Powershot S120 camera. The experiment took place in a dark room to minimize observers disturbing the birds.

Food and experimental models were offered on a green platform through a moveable tray behind a visual barrier, allowing us to estimate the exact time when the bird first saw the model and thus started the trial (see details in Nokelainen et al. 2012). A standard green background was used, because wood tiger moths rest on green leaves in nature (Hegna et al. 2013, Nokelainen et al., 2012). All colours used in the moth models are easily distinguished from the background by birds: Hegna et al. (2013) reported just noticeable difference (JND) values in colour contrast ranging from 8.6 to 11.6 for white and yellow artificial moth models and real wood tiger moths against the green background used also in this experiment, and Lindstedt et al. (2011) calculated JND values above 27.27 for orange and red females on natural green leaves of *Alnus incana*.

During pretraining, birds were allowed to habituate to the experimental cages and learned to eat three sunflower seeds from the green platform. To motivate the birds to attack the moth models during the experiment (see below), they were food deprived for 2 h before the preference test, 1 h before the learning test and 1 h before the generalization test. After food deprivation, bird motivation was tested with a sunflower seed; if eaten, the bird was considered ready to begin the test.

### Phase 1: Preference test
A preference test was included in the experimental protocol for two reasons. As we used wild-caught birds, we first tested whether they had any pre-existing biases towards white, yellow or red moth morphs. Second, by offering palatable morphs several times to birds we ensured that any potential unlearned or learned biases disappeared, allowing us to test the effect of the coloration on learning and generalization (Ghirlanda & Enquist, 2003).

All three morphs (white, yellow and red) were offered simultaneously on the green platform for 5 min, starting from when the bird first saw them. If the bird did not attack (i.e. grab or peck) any of the edible model pastry bodies during the 5 min, the models were taken away and presented again after a break. Once the first attack was made, the models were kept in the cage until the bird finished eating all the pastry bodies. To ensure that all birds had an equally rewarding experience with all the colours, we let the birds finish eating the pastry bodies of all models in three consecutive trials during the preference test. Between the trials, the presentation (order) of the models on the platform was always changed (Fig. 2).

As birds were hesitant to attack the moth models for the first time (hesitation times varying from 17 s to 2 h consisting of 5 min presentations), we did not use time to attack in analyses. Instead, we recorded the order in which the models were attacked and eaten during the three consecutive trials. We compared the order of attacks between the first and the last preference test to be sure that all the birds got rid of any potential bias in preferences before the learning phase.

Preference test presentations were continued for a maximum of 2 days. Eight of 53 birds did not attack or finish eating the artificial moth models offered during the preference test and were, therefore, excluded from further tests.

Phase 2: Learning test

In the second phase of the experiment we tested whether blue tits learn to avoid white, yellow and
red models differently, and established learned avoidance towards one of the colour morphs before
the following generalization test. Birds that completed the preference test were divided into three
groups for avoidance learning: 15 birds were offered white models, 15 yellow models and 16 red
models as unpalatable. Groups of birds were selected as similar as possible (i.e. similar sex, age and
size distribution) and birds from all groups were tested simultaneously. All models were made
unpalatable by replacing the water in pastry bodies with 15% chloroquine diphosphate solution
(Sigma Life Science, St Louis, MO, U.S.A.). As the pastry bodies were coloured with black dye on
top, we also added 15% chloroquine diphosphate solution on top of the bodies and let it dry before
the following trials. Chloroquine solution was used because it is odourless (Hong 1976) and thus all
qualities other than palatability (i.e. taste) of the prey items remained the same throughout the
experiment.

During the learning test, unpalatable models were presented individually in consecutive
trials alternating with sunflower seeds (Fig. 2). Sunflower seeds were offered to monitor the birds’
motivation to forage and avoid unnecessary starvation. If the bird did not attack the sunflower seed,
it got a 10 min break without food and was then offered a sunflower seed again. If the bird attacked
the sunflower seed, the next unpalatable model was offered 2 min after the bird finished eating. As
long as the bird attacked the models, trials were continued alternating with sunflower seeds. If the
bird did not attack the unpalatable model, but ate the sunflower seed, it was considered to reject the
model. After a bird did not attack the moth model the second time in a row, a small live mealworm
(< 20 mm *Tenebrio molitor* larva) was offered instead of the sunflower seed to test the bird’s
motivation to attack insect prey and increase its motivation to forage. If the bird now attacked the
unpalatable model offered after the mealworm, trials were continued again alternating with
sunflower seeds, but if it rejected the unpalatable model, it got another mealworm (Fig. 2). We
considered the bird to have learned to avoid the unpalatable models when it refused to attack three
models in a row, but consumed the sunflower seeds and mealworms offered in between and after
Presentation time was set to 5 min from when the bird first saw the model for the first three trials to make sure that each bird had the opportunity to attack and taste the model. To keep the overall duration of the generalization experiment within the permitted 4-day limit, a maximum of 30 presentations divided into 2 days was set for the avoidance learning. Furthermore, we reduced the presentation time to 2 min from the fourth to the sixth trial, and to 1 min for the rest of the trials.

Based on our observations during a pilot experiment with six birds and fixed durations of trials, birds were unlikely to attack the model and did not consume it within 5 min if they did not attack within the first minute. Sunflower seeds were usually attacked quickly, and hesitation time declined to a few seconds as the trials proceeded, implying that 1 min was sufficient to test the bird’s willingness to attack the models. Two of the 46 birds did not stop attacking (white and red) models within 30 presentations and were therefore excluded from the following generalization test.

**Phase 3: Generalization test**

In the third phase, we tested whether the 44 birds that had learned one of the colour morphs as unpalatable would avoid attacking the other two colour morphs. When birds had completed the learning test, half of them had a break of at least 2 h with food and 1 h of food deprivation before the last phase of the experiment, and half were tested the following day. The generalization test started after the bird had consumed a sunflower seed offered to test its motivation to attack. Birds were tested for the generalization with the colours that they did not learn as unpalatable: yellow and red for those that learned white as unpalatable, white and red for those that learned yellow, and white and yellow for those that learned red (Fig. 2). The two colours tested were presented simultaneously on the green platform in alternating positions for three trials lasting 5 min each. This allowed us to test the repeatability of bird behaviour. The trials were interspersed with sunflower seed presentations to make sure that birds were not attacking the models due to lack of motivation.
Moreover, offering alternative food ensured that birds were not forced to eat the models simply because of hunger. Models used for the generalization test were palatable.

Statistical analysis

Phase 1: Preference test

The potential colour bias of blue tits was analysed separately for the first trial (Fig. 2) and all three trials pooled. Colour biases are most likely to be detected reliably by checking the order of attacks on the white, yellow and red models in the first trial ($N=53$), when the birds first encountered the models. The number of moths of each colour taken first, as well as left last, were compared by means of a chi-square test. Additionally, all three trials were pooled in another analysis to find out whether the potential biases disappeared as the birds learned to eat all the models. In the pooled data, each colour was scored based on the order of choice by the bird in each trial; the colour chosen first was scored 1, that chosen second was scored 2 and the colour chosen last was scored 3. Thereby, the minimum score expected for a preferred colour was 3 (i.e. always chosen first), and the maximum score expected for an avoided colour was 9 (i.e. always chosen last). To study the population level bias to all colours, the scores of each colour in each of the three trials were summed and compared to an even distribution by means of a chi-square test. The potential influence of the first colour chosen on the subsequent choice was checked with a binomial exact test.

Phase 2: Learning test

Potential differences in learning rate between the three colour morphs were analysed using a mixed-effect Cox regression model, using the ‘coxme’ package (version 2.2-5; Therneau, 2015) in RStudio (v. 0.99.902; RStudio, 2015). The response variable was the probability that the presented model
was attacked in each trial; time was represented as number of trials. Model colour was added as an explanatory factor and bird individual as a random effect.

## Phase 3: Generalization test

If birds generalized their learned aversion of a given colour to the two nonlearned colours, we would expect them to refrain from attacking models offered during the generalization test but eat the alternative prey offered between trials. Hence, the probability of attack on palatable models is expected to be significantly lower than random (< 0.5). If, in contrast, birds were unable to generalize their learned avoidance, we would expect the attack probability to be significantly higher than 0.5. High attack probability is expected (in the case of no generalization) since the birds had attacked and eaten similar palatable models in the preference test and did attack the models presented first in the learning test within the 5 min presentation. Thus, to test whether the birds generalized and the attack probability on the models was lower (or higher) than random, we built two generalized linear mixed models (GLMM 1 and 2) with a logit link and binomial distribution, including whether the prey was attacked (1) or not (0) as the dependent variable. Bird ID and bird ID nested within trial in GLMM 2 were added as random factors using package lme4 (Bates et al. 2015) in R. GLMM 1 was used to test for generalization in the first trial only, and GLMM 2 in all three trials.

To test for asymmetric generalization, we divided the birds into six treatment classes by the colours they learned (white, yellow and red) and were offered (yellow and red, white and red, white and yellow, respectively). This classification was then used as the explanatory variable (‘colour combination’) in two GLMM models separately for the first trial only (Table A1 in the Appendix) and all three trials (Table A2 in the Appendix) of generalization (again with a logit link and binomial distribution, including whether the prey was attacked (1) or not (0) as the dependent variable, and bird ID nested within trial and/or bird ID as random factors). A chi-square test was
used to check whether the birds attacked one colour morph first more frequently between the two
colour morphs offered, both in the first trial and in the first three trials pooled (Table A3 in the
Appendix). Birds tested the same or the following day after avoidance learning were pooled in all
analyses, as there were no differences in the number of attacks between birds tested the same or the
following day after avoidance learning in the first trial (unpaired two-sample Wilcoxon test:
\( W=276, N=44, P=0.21 \)) or in the three trials pooled (\( W=262, N=44, P=0.59 \)). We also checked
whether the rate of learning correlated with the number of attacks in the generalization test with a
Spearman correlation.

\textbf{RESULTS}

\textbf{Preference test}

At the population level, birds did not show any preferences (Table 1) or aversion (chi-square test:
\( \chi^2=2.577, P=0.28 \)) towards any of the colours (white, yellow or red) during the first trial. Birds
chose the second colour to attack with the same probability between the two colours left,
irrespective of the first colour chosen (binomial exact test: \( P>0.05 \) for all comparisons).

At the individual level, 35 birds (85.4\%, \( N=41 \)) chose at least one colour in the same order
for two different trials (for instance, the same bird chose the yellow morph as last choice in two
trials out of three). Two birds showed a strong preference for one of the colour morphs, choosing
the same colour (yellow and red, respectively) first for all three trials. Three birds showed
avoidance for one colour morph (one for white, two for red), leaving the same colour as last in all
the trials. All other birds changed their order of choice during the three trials, showing that they got
rid of potential biases towards the colours during training. When we tested the overall scores for
each colour morph during the three trials, birds did not show differences between the colour morphs
(chi-square test: \( \chi^2=0.789, P=0.67 \)).
Apart from two individuals, all birds (N=44) learned to avoid their moth model according to the
criterion of no attack over three subsequent trials. The number of trials needed to learn to avoid the
unpalatable model varied between 2 and 23 among the birds (mean=7). The Cox regression model
(Fig. 3) showed that birds learned to avoid the red colour morph significantly faster than the yellow
(Z=2.17, P=0.03), but showed no significant differences between the yellow and white morphs
(Z=0.87, P=0.38).

Overall, blue tits did not generalize their learned avoidance from one colour morph to the other two,
as the attack probabilities were significantly higher than 0.5 in the first trial (GLMM 1: Z=4.33,
P<0.001; Fig. 4) and the three trials pooled (GLMM 2: Z=6.42, P<0.001). Only three of 44
individuals did not attack any of the palatable models during the generalization test, showing
generalized avoidance.

We did not find clear evidence of asymmetric generalization. The estimated attack
probabilities did not differ significantly between the combinations of colour learned and colour
offered in generalization trials (Tables A1 and A2 in the Appendix), and no differences were found
in which colour the birds attacked first during the first trial of the generalization test (chi-square
test: P>0.05 in all cases; Table A3 in the Appendix). In the first trial, however, birds that learned
yellow attacked fewer white models compared to the other colour combinations, and the effect is
near the 0.05 significance level (Table A1 in the Appendix). Also, when the three trials were
pooled, we found that birds that learned to avoid the white morph attacked the red morph first
significantly more often than the yellow one (chi-square test: χ^2_1=5.9, P=0.02).
The rate of learning did not correlate significantly with the proportion of attacked models in the first generalization trial (Spearman correlation $r_s=0.12$, $N=44$, $P=0.45$) or the total number of attacks in the three generalization trials (Spearman correlation: $r_s=0.27$, $N=44$, $P=0.07$), thus allowing us to compare the effect of the colour learned on generalization despite different learning rates of red versus the other colours.

**DISCUSSION**

**No generalization based on hindwing colour**

Generalized avoidance by local predators from one warning signal to another has been proposed to contribute to the maintenance of local warning signal polymorphism in aposematic species (Amézquita et al., 2013; Exnerová et al., 2006; Gamberale & Tullberg, 1996; Gamberale-Stille & Tullberg, 1999; Ham et al., 2006; Hegna & Mappes, 2014; Rojas, Rautiala, & Mappes, 2014; Ruxton et al., 2008; Waldron et al., 2017). Here we studied in more detail how bird predators learn and generalize the warning colours of a polymorphic (red, yellow, white) wood tiger moth population using artificial moth models. Attack rates during the generalization test were in general very high. Indeed, the birds did not generalize their learned avoidance among the wood tiger moth morphs, but instead treated them as different prey types based on the differing hindwing colour alone, as the morph models used did not differ in size, shape, pattern, taste or smell.

**The importance of colour**

Our findings are in line with previous experiments showing that colour is of foremost importance in avian predator learning, contributing especially to the discrimination between palatable and unpalatable prey (Aronsson & Gamberale-Stille, 2008, Kazemi, Gamberale-Stille, Tullberg, & Leimar, 2014). A large body of research has demonstrated birds’ ability to learn to avoid
conspicuous, unpalatable prey (Aronsson & Gamberale-Stille, 2008; Rowe, Lindström, & Lyytinen, 2004; Svádová et al., 2009). This is because conspicuous warning coloration enhances prey recognition (Guilford, 1986; Sherratt & Beatty, 2003), speed of avoidance learning and memorability (e.g. Roper & Redston, 1987). Different predators may use different components of the warning signal as a primary cue depending on their sensory systems (Aronsson & Gamberale-Stille, 2012; Endler, 1992; Guilford & Dawkins, 1991) and disregard others. Studies done with birds have demonstrated that colour seems to be a more important feature in warning signals than size or pattern (Aronsson & Gamberale-Stille, 2008; Exnerová et al., 2006; Sillén-Tullberg, 1985; Terhune, 1977).

As predators can associate palatability or unpalatability with several different kinds of prey traits, it is convenient to compare the relative importance of those traits with how much they facilitate associative learning. The expectation is that more salient signals are learned faster (Kazemi et al. 2014). Our results indicate that red was the most salient warning colour: birds learned to avoid the red morph faster than the other morphs. This is in accordance with Lindstedt et al. (2011), who found that the red female morph of the wood tiger moth was better protected against bird predators, suffering fewer attacks than its orange or yellow counterparts. Indeed, red has been shown to be a very efficient warning signal compared to other warning colours such as orange, yellow or white, and other colours such as violet, blue, green and brown, at least for some bird predators (Cibulková, Veselý, & Fuchs, 2014; Exnerová et al., 2006; Gamberale-Stille & Tullberg, 1999; Lindstedt et al., 2011; Svádová et al., 2009).

In the present study prey items were made to resemble real wood tiger moth morphs as closely as possible, keeping all traits other than hindwing colour constant. This allowed us to compare the effects of warning coloration of hindwings only. Changing the warning colour hue altered not only the internal contrast on model hindwings, but also the contrast between the model and the green background. Although all colours in our experiment were clearly conspicuous to the
birds, red had the highest colour contrast against the green background whereas white had the lowest. This might explain why red seems to be the most salient signal. Aronsson and Gamberale-Stille (2009) found similar results using domestic chicks, *Gallus gallus domesticus*, which learned to avoid red prey faster if presented on a contrasting background compared to a background of similar hue. In another experiment, however, red prey colour was found to influence predator avoidance independent of background colour (Sillén-Tullberg, 1985). Thus, it seems that both prey coloration per se and its contrast against the background can contribute to predator avoidance, but it is still relatively unclear which properties of prey coloration, chromatic or achromatic, play the most important role. Previous work with wood tiger moths has shown that the achromatic contrast against a green background is highest for white morphs, which are the most luminous of the three (Lindstedt et al. 2011; Henze, Lind, Mappes, Rojas & Kelber 2017). Luminance has not been found to affect predator responses towards the wood tiger moth, while the chromatic contrast in hue seems to be very important (Nokelainen et al., 2012).

Generalization has been suggested to stabilize selection towards aposematic signals via a peak shift phenomenon (Leimar, Enquist, & Sillen-Tullberg, 1986; Lindström et al. 1999b). The minimum (and maximum) responses of predators (i.e. peaks of the generalization gradient) have been found to be displaced from the negative (and positive) stimulus (Gamberale & Tullberg, 1996; Hanson, 1959), such as yellow, towards a similar, but more salient novel stimulus, such as red. Overall, we did not find strong evidence of asymmetric generalization, but there were some trends between the colours tested. Birds that learned to avoid red models attacked almost all the white and yellow models in the generalization trials, whereas birds that learned the less salient colours yellow and white generalized more, hinting at a tendency to generalize from the less salient signals towards the more salient signal.

Svádová et al. (2009) found asymmetric generalization using great tits, *Parus major*, which did not generalize from red firebugs, *Pyrrhocoris apterus*, to white or yellow mutants, but did
generalize from yellow mutants to red firebugs. Interestingly, in our experiment, four blue tits that
learned to avoid the yellow morph (N=15) refrained from attacking white models, while only two
did so for the red ones. Birds that learned to avoid the white morph attacked both unlearned morphs
equally, but chose red models first more often than yellow ones. This indicates that birds tended to
generalize more between the white and yellow than between white and red. The yellow morph
seems to benefit least from the other colours, since only between 7 and 34% of yellow models were
left unattacked (Table 2, Fig. 4).

Limitations of testing generalization in the laboratory

Despite the majority of birds showing no generalization in our experiment, the possibility remains
that predators might generalize among morphs of the wood tiger moth under different
circumstances. Studying generalization in the wild is practically impossible due to the rareness of
predation events on aposematic prey as well as difficulties in observing the choices of individual
predators in natural conditions. A previous study aiming to explain the variability in the warning
signals of the harlequin poison frog, *Dendrobates histrionicus*, showed that predators avoided
attacking aposematic frog models but not cryptic ones in areas where aposematic frogs occur,
exhibiting some generalization among different frog colour morphs in the field. However, the same
study found no generalization by naïve chicks tested in the laboratory (Amézquita et al., 2013). This
might imply that naïve and experienced predators in the wild can use different generalization
strategies (see also Ihalainen, Lindström, Mappes & Puolakkainen, 2008). Birds might also be
prone to generalize more or less widely under different circumstances (Aronsson & Gamberale-
Stille, 2012), for example under physiological stress during winter months (Barnett, Bateson &
Rowe 2007; Chatelain et al. 2013; Veselý et al. 2017), limited food availability (Lindström, Alatalo,
Lyytinen & Mappes, 2004; Ihalainen, Rowland, Speed, Ruxton, & Mappes, 2012), limited time to
make decisions (Ings & Chittka, 2008), when the prey is dangerously toxic (Lindström, Alatalo, &
Mappes, 1997; Sherratt, 2002), when the prey community is complex versus simple (Ihalainen et al., 2007), or when the prey population has palatable Batesian mimics in addition to the unprofitable prey (Plowright & Owen, 1980).

Avoidance learning has been suggested to happen in two steps: first, the birds learn simple rules based on certain cues, and once the basic rules are formed, they then learn in more detail about prey quality (Chittka & Osorio, 2007). Recent studies indicate that birds are able to assess the nutritional benefits of unprofitable prey and use this information in subsequent encounters (Halpin et al., 2014). This ability could have affected not only bird learning rates, but also their decision to attack in the generalization phase of our experiment. As our models’ pastry bodies were of high nutritional value and the birds were hungry, it is possible that the birds were willing to take more risks and thus took more trials to learn to avoid the models than it would take them to learn to avoid defended prey in the wild. In addition, the 5 min presentations gave the birds plenty of time to decide whether to attack or not, and to make more sophisticated assessments of prey quality than might be possible in the wild. Birds were given alternative food between the presentations, but not enough for saturation, and would thus have benefited energetically from discriminating between the unprofitable and profitable models. Nevertheless, the cost–benefit relationship was exactly the same for all morphs in our experiment and, thus, we can safely compare the relative differences between morphs in their salience.

The avoidance learning was based on counterconditioning, where the colour signal was first associated with a positive reinforcement (i.e. palatability) and then with a negative reinforcement (i.e. unpalatability). Previous research has shown that in cases of single counterconditioning the associations learned second are forgotten at higher rates than those learned first (Speed, 2000 and references therein). Therefore, it is possible that the birds’ experience and learned association with palatability in the preference test exceeded the effect of generalized avoidance among the morphs for most of the birds, which could partly explain the low level of generalization observed. Offering
the models as palatable at the beginning of the experiment was necessary to get rid of any pre-
existing biases or neophobia prior to learning and testing generalization effects of the birds; this was
also necessary to motivate the birds to attack and taste the unpalatable models during the first
learning trials.

In the preference and generalization tests, simultaneous prey choice was used to decrease
the numbers of birds needed to accomplish the experiment. Simultaneous prey choice is also a very
powerful set-up to detect any potential predator biases but, obviously, this approach has
disadvantages too (Fig. 2). For example, it is possible that long hesitation delays during the first
presentation of the preference test were partly due to an aggregation effect, as aggregations of
conspicuous prey have been found to be aversive to predators (Gamberale-Stille, 2000; Riipi,
Alatalo, Lindstrom, & Mappes, 2001). On the other hand, Nokelainen et al. (2012) presented wood
tiger moths singly to birds, several of which also hesitated for a long time before attacking them.
Thus, it is difficult to say how much the simultaneous presentation influenced our results, but
during the flying season wood tiger moth morphs typically aggregate at the same sites. Males of
both morphs are often found near calling females, and thus all morphs can be visible and vulnerable
to predators simultaneously.

Lastly, if the wood tiger moths are able to survive bird attacks, the use of artificial models
does not necessarily give an accurate estimate of selection. A considerable proportion of attacks in
the generalization test were just a single peck, leaving the models uneaten, and thus not necessarily
‘killed’. The birds’ willingness to attack but reluctance to consume the models could stem from the
psychology of birds’ decision making (Marples & Kelly, 1999). Adamová-Ježová et al. (2016)
showed that for great tits and coal tits, *Periparus ater*, neophobia (i.e. the avoidance of novel prey
affecting the decision to attack), but not dietary conservatism (i.e. restriction of diet to certain prey
types affecting the decision to consume the prey), was deactivated during pretraining with a
palatable prey, but the initial hesitation of blue tits was not affected by earlier experience. Blue tits
have been found to show higher general aversion even towards palatable prey than, for example, great tits, probably because of higher dietary conservatism (Veselý et al. 2006, 2013, Prokopová et al. 2010, Turini et al. 2016). This indicates that predators’ decisions of whether to attack or not after avoidance learning might be species specific and, thus, not generalizable from blue tits to other predators. In our experiment, 45 of 53 blue tits overcame their initial avoidance during the preference test and attacked and consumed the models readily in the following avoidance trial. As those birds that did not overcome their hesitation during the preference test were not included in the following phases of the experiment, our results describe the generalization tendency of the less hesitant individuals, which are more likely to attack aposematic prey in the wild in the first place. Many of these less hesitant individuals, however, seemed to regain their dietary conservatism after they had learned avoidance, as they no longer consumed the palatable models attacked. In conclusion, whereas no generalization was found regarding the attack probabilities, we did find individual variation in avoidance learning and dietary conservatism, which could affect selection in the wild.

The importance of other cues

Somewhat surprisingly, the blue tits had no initial biases towards any of the hindwing colours. Earlier studies on the wood tiger moth have found differential predation pressure in the field (Lindstedt et al., 2011; Nokelainen et al., 2012; Nokelainen et al., 2014) and different hesitation times by local predators (Lindstedt et al., 2011; Nokelainen et al., 2012) towards the different colour morphs. As the differences in hesitation times were found using living moths (Nokelainen et al., 2012), it is possible that other cues, such as odour, influenced the results. In nature, the wood tiger moth relies on multiple signal components (i.e. odour, taste) in addition to the visual cues when exposed to potential predators (Rojas, Burdfield-Steel & Mappes 2015). Its chemical defence contains pyrazines (Rojas et al. 2017; Burdfield-Steel, Pakkanen, Rojas, Galarza & Mappes 2016).
a group of compounds with a characteristic aversive smell, which is effective against birds (Guilford, Nicol, Rothschild, & Moore, 1987; Rowe & Guilford, 1996). In fact, pyrazine is known to trigger hidden aversions to red and yellow colours (Rowe & Guilford, 1996) and conspicuous prey (Lindström, Rowe, & Guilford, 2001), and enhance both learning and memorability of yellow (Siddall & Marples, 2008) or red coloured prey (Barnea, Gvaryahu, & Rothschild, 2004) at least in domestic chicks (Siddall & Marples, 2008). Pyrazine odour has been associated with Müllerian mimicry rings of insects and suggested to function as a warning signal (Rothschild, 1961). On the other hand, pyrazine odour has also been shown to assist in discriminating prey and thus reduce avoidance generalization between differently coloured prey if the odour is present on only some of them (Siddall & Marples, 2008; Rowe & Guilford, 1996). The specific roles of different cues in predator–prey interactions are uncertain. It might be that odour is easy to associate with palatability, but only functions close up, whereas conspicuous colours aid in memorizing which prey to avoid even from a distance. Here we were interested in colour only. However, it has been shown that when colour is kept constant, predators can discriminate prey based on pattern (e.g. Prokopová et al. 2010; Veselý et al. 2013), and when both colour and pattern are equal, other visual and/or chemical features of the prey are used for prey recognition (Karlíková et al. 2016).

Conclusions

Overall, the colour polymorphism of the wood tiger moth in Finland seems unlikely to be maintained by generalized avoidance based on its warning coloration only. However, predators were hesitant to attack any of the aposmatic morphs in the first place, and if they were to encounter them in the wild sharing other warning cues such as pyrazine odour, general aversion seems likely to occur. More knowledge on how predators acquire and use information on prey qualities in different contexts is needed to conclude whether predator generalization contributes to the maintenance of multiple aposmatic morphs (see also Skelhorn et al. 2016). The possibility that
wild predators can generalize on the basis of the pyrazine odour or the combination of colour and odour requires further investigation. Alternative explanations for the occurrence of local warning signal polymorphism include negative frequency-dependent natural selection, sexual selection, frequency-dependent flight activity (Rojas, Gordon, & Mappes, 2015), signal efficacy trade-offs with other life history traits (Hegna, Nokelainen, Hegna, & Mappes, 2013; Nokelainen et al., 2012), predator species-specific mortality differences between morphs (Nokelainen et al., 2014) or combinations of these mechanisms (Gordon, Kokko, Rojas, Nokelainen, & Mappes, 2015). Colour polymorphism could also be explained by multiple-model mimicry (Edmunds, 2000), if the different morphs share warning colours with other defended prey species, and predators generalize their avoidance from one species to the other based on similar coloration. Thus, generalization of learned avoidance remains as a possible contributor to the maintenance of local polymorphism in wood tiger moth populations. In conclusion, we argue based on our results that although predator generalization could well contribute to the maintenance of different aposematic morphs under certain circumstances, it is unlikely to occur among distinct colour morphs of otherwise similar prey of visually oriented avian predators.

**Acknowledgments**

We are indebted to Helinä Nisu and Tuuli Salmi for help with the birds at Konnevesi Research Station, and to Andrés López-Sepulcre, Sebastiano De Bona and Janne Valkonen for their thoughtful advice on statistical analyses. We also thank the ‘Darwin group’ at the University of Jyväskylä for discussions and two anonymous referees and the editor for helpful comments on the manuscript. This study was funded by the Centre of Excellence in Biological Interactions (Academy of Finland, project no. 284666). C.D.P. was funded by the Erasmus Exchange Programme.
References


the aposematic moth *Arctia plantaginis*. Manuscript in preparation.


793 54–55.
795 between warning signal efficacy and mating success in the wood tiger moth. *Proceedings of
796 the Royal Society B: Biological Sciences*, 279(1727), 257-265. doi: 10.1098/rspb.2011.0880
798 structure shifts the efficacy of two warning signals in Arctiid moths. *Journal of Animal
801 frequency-dependent sexual selection *Biological Journal of the Linnean Society*, 23(2-3),
804 Plowright, R. C., & Owen, R. E. (1980). The evolutionary significance of bumble bee color
807 Kegan Paul, Trench, Trubner.
808 Prokopová, M., Veselý, P., Fuchs, R., & Zrzavý, J. (2010). The role of size and colour pattern in
810 protection of developmental stages of the red firebug (*Pyrrhocoris apterus*) against avian
811 predators. *Biological journal of the Linnean Society*, 100(4), 890-898. doi: 10.1111/j.1095-
812 8312.2010.01463.x
814 the ladybird Harmonia axyridis: the avian predators’ point of view. *Entomologia
815 Experimentalis et Applicata*, 151(2), 128-134. doi: 10.1111/eth.12454
817 cover detectability costs in aposematic aggregations. *Nature*, 413(6855), 512-514. doi:


Appendix
**Table 1.** Distribution of colours chosen first in the preference test trials (phase 1)

<table>
<thead>
<tr>
<th></th>
<th>White first</th>
<th>Yellow first</th>
<th>Red first</th>
<th>N</th>
<th>df</th>
<th>$\chi^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>34</td>
<td>32</td>
<td>34</td>
<td>53</td>
<td>2</td>
<td>0.04</td>
<td>0.98</td>
</tr>
<tr>
<td>Trial 2</td>
<td>44</td>
<td>22%</td>
<td>34</td>
<td>41</td>
<td>2</td>
<td>2.98</td>
<td>0.23</td>
</tr>
<tr>
<td>Trial 3</td>
<td>29</td>
<td>27%</td>
<td>44</td>
<td>41</td>
<td>2</td>
<td>2.10</td>
<td>0.35</td>
</tr>
</tbody>
</table>

The percentages of birds that chose white, yellow or red models first in the three trials of the preference test, and the corresponding chi-square comparison for preference for each trial, are shown.
Table 2. Proportions (± SE) of tested colour morphs attacked in the generalization test trials (phase 3) in relation to the colour learned

<table>
<thead>
<tr>
<th>Colour learned</th>
<th>Colour tested</th>
<th>First trial</th>
<th>Second trial</th>
<th>Third trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>Yellow</td>
<td>0.85 ± 0.09</td>
<td>0.92 ± 0.07</td>
<td>0.64 ± 0.13</td>
</tr>
<tr>
<td>White</td>
<td>Red</td>
<td>0.85 ± 0.09</td>
<td>0.78 ± 0.11</td>
<td>0.78 ± 0.11</td>
</tr>
<tr>
<td>Yellow</td>
<td>White</td>
<td>0.73 ± 0.11</td>
<td>0.73 ± 0.11</td>
<td>0.73 ± 0.11</td>
</tr>
<tr>
<td>Yellow</td>
<td>Red</td>
<td>0.86 ± 0.09</td>
<td>0.73 ± 0.11</td>
<td>0.60 ± 0.13</td>
</tr>
<tr>
<td>Red</td>
<td>White</td>
<td>0.93 ± 0.06</td>
<td>0.86 ± 0.09</td>
<td>0.73 ± 0.11</td>
</tr>
<tr>
<td>Red</td>
<td>Yellow</td>
<td>0.93 ± 0.06</td>
<td>0.86 ± 0.09</td>
<td>0.86 ± 0.09</td>
</tr>
</tbody>
</table>
### Table A1. Test for asymmetric generalization in the first generalization trial

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>LRT</th>
<th>Pr(Chi)</th>
<th>Model AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept) + colour combination</td>
<td>5</td>
<td>10.70</td>
<td>0.058</td>
<td>37.4</td>
</tr>
<tr>
<td>(Intercept)</td>
<td></td>
<td></td>
<td></td>
<td>38.1</td>
</tr>
</tbody>
</table>

**Random effects**

<table>
<thead>
<tr>
<th></th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bird ID</td>
<td>4808</td>
<td>69.34</td>
</tr>
</tbody>
</table>

**Fixed effects**

<table>
<thead>
<tr>
<th>(Intercept: colour combination: yw)</th>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12.48</td>
<td>3.32</td>
<td>3.75</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

| Colour combination: yr                   | 14.85    | 5.22| 2.85 | 0.0044|
| Colour combination: rw                   | 2.27     | 8.04| 0.28 | 0.78  |
| Colour combination: ry                   | 2.27     | 8.19| 0.28 | 0.78  |
| Colour combination: wr                   | 1.38     | 6.16| 0.22 | 0.82  |
| Colour combination: wy                   | 1.38     | 6.15| 0.22 | 0.82  |

LRT: likelihood ratio test; y: yellow; w: white; r: red. Model selection was based on model fit, i.e. the model chosen was the one with the lowest Akaike information criterion (AIC).
value. The significance level of $\chi^2$ (Chi) indicates a change from the model with colour combination as an explanatory variable to the model below, with intercept only. Estimates of the best-fitting model (in bold) are shown below. Of the colour combinations, the combination of yellow learned and white offered had least attacks, and was thus set to the intercept.
### Table A2. Test for asymmetric generalization in all three generalization trials

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>LRT</th>
<th>Pr(Chi)</th>
<th>model AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept) + colour combination</td>
<td>5</td>
<td>4.32</td>
<td>0.50</td>
<td>161.9</td>
</tr>
<tr>
<td>(Intercept)</td>
<td></td>
<td></td>
<td></td>
<td>156.2</td>
</tr>
</tbody>
</table>

Random effects

<table>
<thead>
<tr>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial: Bird ID</td>
<td>167.82</td>
</tr>
<tr>
<td>Bird ID</td>
<td>58.01</td>
</tr>
</tbody>
</table>

Fixed effects

<table>
<thead>
<tr>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>10.05</td>
<td>1.57</td>
<td>6.42</td>
</tr>
</tbody>
</table>

LRT: likelihood ratio test. Model selection was based on model fit, i.e. the model chosen was the one with the lowest Akaike information criterion (AIC) value. The significance level of \( \chi^2 \) (Chi) indicates a change from the model with colour combination as an explanatory variable to the model below, with intercept only. Estimates of the best-fitting model (in bold) are shown below.
Table A3. Comparisons of colours attacked first in the generalization trials

<table>
<thead>
<tr>
<th>Colour learned</th>
<th>Colour tested</th>
<th>Attacked first</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yellow 6/14</td>
<td>White</td>
<td>1 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White 6/14</td>
<td>Red</td>
<td>1 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White 6/15</td>
<td>Red</td>
<td>1 0.13 0.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellow 6/15</td>
<td>White</td>
<td>1 0.75 0.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red 7/15</td>
<td>White</td>
<td>1 0.41 0.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red 8/15</td>
<td>White</td>
<td>1 0.75 0.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red 6/15</td>
<td>Yellow 6/15</td>
<td>1 0.13 0.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellow 12/42</td>
<td>White</td>
<td>1 5.89 0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White 12/42</td>
<td>Red 24/42</td>
<td>1 5.89 0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White 12/42</td>
<td>Red 24/42</td>
<td>1 5.89 0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White 20/45</td>
<td>Red 15/45</td>
<td>1 0.75 0.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red 15/45</td>
<td>White 22/45</td>
<td>1 0.41 0.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red 15/45</td>
<td>White 22/45</td>
<td>1 0.41 0.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red 15/45</td>
<td>Yellow 18/45</td>
<td>1 0.41 0.52</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chi-square comparisons of how many times each tested colour morph was attacked first in the generalization trials, by the colour morph learned. Models that were not attacked or attacked second are included in the total number of models offered. Bold indicates significant difference.
**Figure 1.** (a) Typical wood tiger moth colour morphs from central Finland (leg. Kari Kulmala), (b) artificial white, yellow and red moth wings used in the experiment and (c) reflectance curves of the white, yellow and red hindwings of real moths (darker colours) compared to reflectance curves from white, yellow and red model hindwings (lighter colours). Spectral measurements were taken from three wild-caught individuals of each colour from the spots marked with blue circles on the white moth. The same spots were used to measure the model hindwing colours. Model wings were set in a more natural posture, less spread than the spread collection samples, but unfolded to show the hindwing colour.

**Figure 2.** Schematic illustration of the experimental design. Each green circle represents a platform presented to a bird in one trial. Moth models as presented to a bird that learned to avoid the red morph are illustrated on the platforms and alternative food offered between the trials is shown above the platforms. For details of the experimental protocol see Methods (Experimental procedures; phases 1-3).

**Figure 3.** Proportion of models attacked during the learning trials for each colour model. The lines represent the cumulative attacks on unpalatable moth models of white (black line), yellow (yellow line) or red (red line) hindwing colour.

**Figure 4.** Proportions (± SE) of models that were attacked in the first trial of the generalization test. Symbol styles refer to white (black circle), yellow (yellow square) and red (red triangle) models offered in the test.