



## Hierarchical overshadowing of stimuli and its role in mimicry evolution



Thomas N. Sherratt<sup>\*</sup>, Emilee Whissell, Richard Webster, David W. Kikuchi

Department of Biology, Carleton University, Ottawa, ON, Canada

### ARTICLE INFO

#### Article history:

Received 8 April 2015

Initial acceptance 18 May 2015

Final acceptance 29 June 2015

Available online 22 August 2015

MS. number: A15-00289

#### Keywords:

discrimination  
generalization  
imperfect mimicry  
learning  
neophobia  
overshadowing  
salience

An elegant study by Kazemi et al. (2014, *Current Biology*, 24, 965–969) found that blue tits, *Cyanistes caeruleus*, that had been trained to distinguish between rewarding and nonrewarding prey subsequently avoided novel prey that had the same colour (but not the same shape or pattern) as the nonrewarding prey. Their work suggested that certain high-salience discriminatory traits can overshadow other informative traits, allowing imperfect mimics to establish. Here we began by replicating their ambitious experiment by evaluating the behaviour of 320 human subjects foraging on computer-generated prey. However, to fully understand how overshadowing might facilitate the establishment of incipient mimics, we significantly extended their protocol to a full factorial design involving mimics with colour, pattern and/or shape in common with their former models. As Kazemi et al. reported, participants placed more weight on colour similarity than on shape similarity when rejecting prey. Two-trait mimics with the same colour and pattern or the same colour and shape as their former models were avoided as frequently as perfect mimics, while colour-only mimics were avoided more frequently than shape-only mimics. Nevertheless, novel prey with no traits in common with nonrewarding models were avoided at high rates, possibly in part due to their dissimilarity to familiar rewarding prey. The implications of these findings for mimicry evolution are discussed.

© 2015 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Researchers have long debated how a palatable species can evolve to resemble an unpalatable species ('Batesian mimicry', Bates, 1862) when it involves changes in so many different appearance dimensions (Kikuchi & Pfennig, 2013). In the two-step model of mimicry evolution it is assumed that a mutation with a large effect on the phenotype first produces an approximate resemblance of the prey to the model, which may then be gradually improved by selection for resemblance-enhancing mutations of smaller phenotypic effect (Nicholson, 1927; Poulton, 1912; Turner, 1984). However, this argument implicitly assumes that predators would tend to overlook important discriminative traits in the incipient mimic and it is not entirely obvious why predators would do this (Gamberale-Stille, Balogh, Tullberg, & Leimar, 2012). Indeed in many well-known mimicry complexes the extent of mimicry is often far from perfect (see Kikuchi & Pfennig, 2013, for a recent review) and yet often predators do not appear to exploit these differences.

Predators ultimately determine the success of any defensive trait and so to fully understand how antipredator adaptations evolve and are maintained, we must evaluate how predators respond to rudimentary versions of these defences (Chittka & Osorio, 2007). In particular, if predators learn to discriminate between profitable and unprofitable prey by identifying the distinctive features of the two object types, then imperfect mimicry might be initially selected for and subsequently maintained if predators learn to associate prey defences with certain salient discriminative features and effectively overlook others, a phenomenon known as 'overshadowing' (Mackintosh, 1976; Shettleworth, 2010). Building on earlier studies on the degree of mimetic protection gained by new partial mimics (Ford, 1971; Schmidt, 1958, 1960; Terhune, 1977), there has recently been a surge of interest in understanding how objects are categorized and what traits are attended to when predators make their discriminative decisions (Aronsson & Gamberale-Stille, 2008, 2012; Bain, Rashed, Cowper, Gilbert, & Sherratt, 2007; Balogh, Gamberale-Stille, Tullberg, & Leimar, 2010). One of the most ambitious and systematic studies to date on this topic was recently conducted by Kazemi, Gamberale-Stille, Tullberg, and Leimar (2014), who trained wild-caught blue tits, *Cyanistes caeruleus*, to discriminate between rewarding prey (symbol-

<sup>\*</sup> Correspondence: T. N. Sherratt, Department of Biology, Carleton University, Ottawa, ON K1S5B6, Canada.

E-mail address: [Tom.Sherratt@Carleton.ca](mailto:Tom.Sherratt@Carleton.ca) (T. N. Sherratt).

bearing laminated cards over wells containing pieces of mealworm) and nonrewarding prey (symbol-bearing cards over empty wells). The symbols over rewarding and nonrewarding wells differed from one another in their colour, pattern and shape, and the blue tits learned to feed almost exclusively on the rewarding wells over four separate trials. Naturally, this high level of discrimination could have been achieved by using any given combination of colour, pattern and/or shape in the symbols. To evaluate what traits were being used in the discrimination process, the authors then presented the same birds with an identical collection of rewarding prey, along with nonrewarding prey that only had one trait (colour, pattern or shape) in common with the earlier models that the birds had learned to avoid (the former models were also used as a control). These nonrewarding prey were essentially Müllerian mimics, but following Kazemi et al. (2014), we hereafter simply refer to them as 'mimics'. The results of this test showed that rather than learning to avoid only prey with all three attributes, the birds generalized their avoidance to nonrewarding prey that had colour in common with the former models, to such an extent that they were not attacked any more frequently than the perfect controls. Collectively, this work suggests that high-salience discriminatory traits (in this case colour) can overshadow other informative traits, allowing incipient mimics a significant survival advantage even if they share only one trait with the model.

The work of Kazemi et al. (2014) was impressive (see review in Cuthill, 2014), even more so because the authors used three different combinations of model and mimic phenotypes (model variants A–C, see Fig. 1) to test whether their conclusions were robust to variation in the nature of the specific colours, shapes and patterns used in discrimination. However, as with all experiments aimed at delivering general insights, it is of interest to know whether similar results would also be generated by other species, including our own. The first aim of our study was therefore to repeat the ambitious experiments of Kazemi et al. (2014), this time using humans, which share with natural predators a finite capacity to process information (e.g. see Beatty, Beirinx, & Sherratt, 2004). However, there was an additional aim, namely to extend the authors' experiment to a full factorial design. Note that Kazemi and colleagues presented perfect 'mimics' and prey that had only one trait in common with the former models (colour, shape or pattern), yielding four different treatments in total per model variant (A–C). To broaden the scope and relevance of the approach to mimicry evolution, here we implemented the complete design, which allowed us to evaluate the success of prey types that had two traits in common (colour and pattern, colour and shape, or pattern and shape), as well as prey with nothing in common with the former models, that is, eight treatments in total ( $=2^3$ ).

A 'nothing in common' treatment is particularly informative for two reasons. First, it enables us to compare the success of non-mimics and incipient mimics with one, two or three traits in common, which is clearly an informative comparison when attempting to elucidate evolutionary pathways (Kazemi et al. were only able to compare the success of a perfect mimic with a one-trait mimic). More importantly, under the Kazemi et al. design, the nonrewarding prey presented in the generalization trials not only generally shared one feature with the former nonrewarding prey, but it also shared two features in common with the rewarding prey. Given that the same phenotypes of rewarding prey were presented throughout both stages of the experiment, it is quite possible that at least some of the birds learned to identify the unique features of rewarding prey, rather than unrewarding prey. In this case, any new variant phenotype might have been avoided not because it looked like the unrewarding prey, but because it looked less like the rewarding prey that the subjects had learned to exploit. In other words, it is possible that some

subjects might have succeeded in identifying the attributes of rewarding prey and simply carried this information through to the next stage of the experiment.

## METHODS

Our design closely followed that of Kazemi et al., this time using human volunteers as surrogate predators (see Fig. 1). No human subject participated more than once. In a pilot study following only the Kazemi et al. treatment conditions (eight replicates of three model variants (A–C) and four generalization tests (perfect, colour, pattern, shape); i.e. 96 human volunteers), we observed an unexpected response with model variant B, in which the stripes on the nonrewarding model subsequently appeared to be generalized (and hence avoided) particularly strongly in the generalization test. To explore this further, we included an additional model variant (D) in our full factorial design involving a striped yellow square as the nonrewarding prey (this extended complete design involved 10 replicates of four model variants and eight different generalization tests based on all possible permutations of model–mimic similarity; i.e. 320 human volunteers).

Images of the various forms of prey were extracted directly from Kazemi et al.'s pdf and saved as a jpegs for display in our computer program. No attempt was made to calibrate the targets' luminance or colour for human vision, as we were interested in the responses of humans to images similar to those used by the authors. The images of the nothing in common prey (those with no colour, pattern or shape in common with the nonrewarding model) were not part of Kazemi et al.'s experiment and so these images were our own design (see last column of Fig. 1). All of our human subjects were first presented with a short video explaining the nature of the experiment (neither the purpose, nor the specific prey types were introduced in this presentation). They were then presented with artificial prey and invited to forage.

The computer program presenting the artificial prey was developed in Microsoft® Visual Basic 6 and involved taking subjects through a series of discrimination learning trials and a series of generalization trials. In each discrimination learning trial (1–4), 16 artificial prey were presented in a  $4 \times 4$  grid, eight of them nonrewarding prey (all identical in appearance) and eight of them rewarding prey (eight different phenotypes), with nonrewarding and rewarding prey distributed at random across the grid (see Supplementary Fig. S1). To enhance motivation, subjects clicking their mouse on a rewarding prey item heard a pleasant cash register sound, their points tally was increased by 1 and the prey item was replaced by a green tick. Conversely, subjects clicking a nonrewarding prey item heard an unpleasant electric shock sound, their points were decreased by 1 and the prey item was replaced by a red cross. Following the general approach of Kazemi et al., subjects continued clicking on prey until all eight rewarding prey were attacked or 20 s had elapsed, whichever came first. At the end of a given trial, the remaining prey were rendered unavailable and subjects were allowed to move to the next screen. After four such discriminative learning trials, the program then seamlessly switched into the generalization test, presenting one each of the same eight rewarding prey types and two (or four) each of four (or two) different nonrewarding prey that differed from the former nonrewarding models in none, one, two or all three traits (colour, pattern and/or shape). Four screens (trials 5–8) containing the same mixtures of phenotypes were presented in this generalization test, with passage to the next screen governed by the same rules as above. At the end of our study, participants were evaluated for colour blindness using Ishihara plates (numbers 5 (plate 6), 15 (plate 8), 45 (plate 13), 42 (plate 23)). Three subjects were assessed to be colour blind on the basis of this test, and their data were

		Step 1: Discrimination learning				Step 2: Generalization tests				
Model variant		Perfect	Colour	Pattern	Shape	Colour and pattern	Colour and shape	Pattern and shape	None	
A	Rewarded stimuli									
	Unrewarded stimuli									
B	Rewarded stimuli									
	Unrewarded stimuli									
C	Rewarded stimuli									
	Unrewarded stimuli									
D	Rewarded stimuli									
	Unrewarded stimuli									

**Figure 1.** Schematic of the discrimination and generalization stages of the experiment, including a depiction of the symbols used. One of four model variants (A–D) was presented to human subjects, where they learned to avoid the model. Subjects were then presented with the same rewarding prey phenotypes but with nonrewarding phenotypes that shared either none, one (colour, pattern or shape), two (colour and pattern, colour and shape, pattern and shape) or all three ('perfect mimics') appearance characteristics with the former models. Each human subject was presented with only one model variant and one generalization test. The shaded area represents the original design of Kazemi et al., with the experiment now extended to cover an additional model variant (D) and phenotypes with none and two appearance dimensions in common.

discarded. In these cases additional subjects were recruited until a balanced design was achieved.

### Statistical Analysis

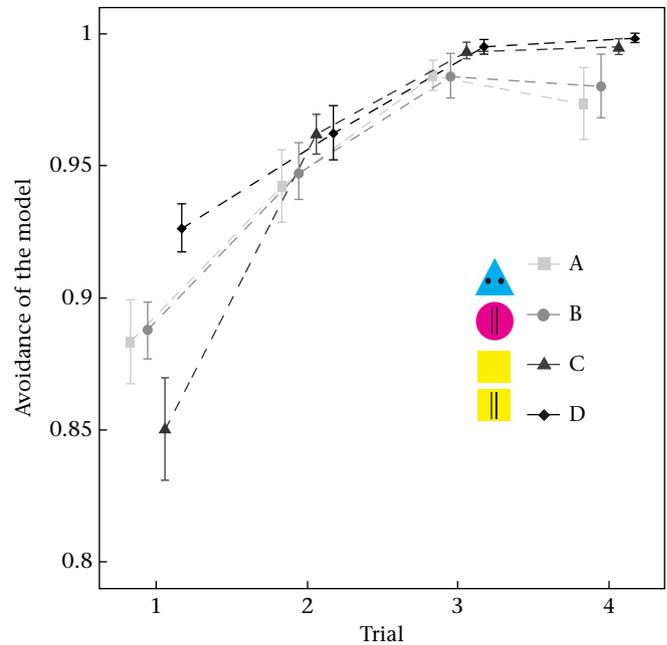
All analyses were performed in R 3.1.0 (R Core Team, 2013). Following Kazemi et al., we analysed the proportion of the first eight prey attacked per trial that were rewarding in the discrimination stage of the experiment using the mixed model 'glmer' function of the package lme4, assuming binomial error variance. In fitting the above model, the model variant (A–D depending on analysis) and trial number (1–4) were treated as fixed effects, while the individual human subject was treated as a random factor. We used Akaike's Information Criteria (AIC) to identify the most parsimonious fitted model, evaluating whether fitting an interaction between fixed effects was justified, and whether the random effect should be modelled as an intercept-only effect, or require a trial-dependent adjustment. Given the binomial error we had assumed, we evaluated the significance of predictors through an analysis of deviance (log likelihood ratio tests, LRT). Post hoc tests were conducted using the 'glht' function of the multcomp package.

For comparability with Kazemi et al., we initially analysed the generalization stage of the experiment based on the proportion of the first four attacks on rewarding prey in the first generalization trial alone (trial 5). With only one response per human subject, we fitted a standard generalized linear model to these response data using the 'glm' function, again assuming binomial error variance. All relevant fixed factors and their interactions were initially included in the fitted model (model variants and the presence/absence of shared colour, shared pattern, shared shape), which was subsequently simplified using AIC to identify the most parsimonious model. As above, the significance of predictors was evaluated through log likelihood ratio tests. To evaluate the longer-term response of subjects, we also analysed the first four attacks in the last generalization trial alone (trial 8).

## RESULTS

After fitting the full factorial (including interaction) model to the discriminative learning component of our experiments, and comparing reduced candidate models using AIC, a main effects model with trial number and model variant, along with a subject-based adjustment for trial was considered the most parsimonious. Trial as a main effect explained significant variation in the proportion of the first eight prey attacked that were rewarding (LRT:  $\chi^2_3 = 322.56$ ,  $P < 0.001$ ), as did the model variant (LRT:  $\chi^2_3 = 2034.9$ ,  $P < 0.001$ ), but there was no compelling evidence for a trial number\*model variant interaction (LRT:  $\chi^2_9 = 6.105$ ,  $P = 0.729$ ). The high significance of the trial effect clearly indicates that learning was occurring (Fig. 2). Experimental group D clearly drove the model variant main effect, in that it was the only treatment level to differ significantly in Tukey HSD pairwise contrasts with other levels (A versus D:  $z = 2.638$ ,  $P = 0.041$ ; C versus D:  $z = 3.244$ ,  $P = 0.0064$ ). Here the unprofitable model was a black striped yellow square, and it was avoided at a high rate even from the first trial (see Fig. 2).

Stepwise fitting of different models to elucidate variation in the attack rate of rewarding prey (i.e. the avoidance of nonrewarding prey) in the first screen of the generalization test (trial 5) culminated in an overarching model with multiple interactions involving the presence/absence of shared colour of the novel prey with the former nonrewarding model, shared pattern, shared shape and model variant. The four-way interaction was nonsignificant in the complete fitted model and hence dropped; two three-way interactions were significant and so lower-level, two-way



**Figure 2.** Learning to discriminate rewarding from nonrewarding (model) phenotypes. The graph shows the mean proportion of all prey attacked per trial (generally 8) that were rewarding for each of the model variants (A–D). Error bars represent  $\pm$ SE. In each model variant group, the rewarding and nonrewarding prey differed from each other in three separate stimulus components (colour, pattern and shape).

interactions were retained for marginality. The roles of colour and pattern could be isolated as main effects, but the presence/absence of all features explained significant variation in avoidance rates as part of at least one interaction (Table 1). There was also an effect of model variant, but only as part of interactions. Collectively, it is clear that the nature of the feature(s) shared with the former nonrewarding prey influenced the degree of protection afforded to these novel prey, with colour and pattern being particularly influential. Reanalysis of the same data, this time based on the proportion of the first eight (as opposed to first four) prey items attacked that were rewarding in the first generalization trial, gave qualitatively similar results (although here the model variant also

**Table 1**

An analysis of deviance of the reduced model examining avoidance rates of nonrewarding prey by human subjects in the first generalization test (trial 5)<sup>a</sup>

	df	Deviance	Resid. df	Resid. dev.	$P(> \chi )$
Null			319	451.59	
Variant	3	0.102	316	451.48	0.992
Pattern	1	9.044	315	442.44	0.003**
Shape	1	0.912	314	441.53	0.34
Colour	1	66.395	313	375.13	3.691e-16***
Variant×pattern	3	2.197	310	372.94	0.532
Variant×shape	3	8.813	307	364.12	0.032*
Pattern×shape	1	1.580	306	362.54	0.209
Variant×colour	3	0.338	303	362.21	0.953
Pattern×colour	1	21.500	302	340.71	3.539e-06***
Shape×colour	1	31.070	301	309.64	2.489e-08***
Variant×pattern×shape	3	21.071	298	288.56	0.0001***
Pattern×shape×colour	1	10.870	297	277.70	0.001***

<sup>a</sup> Four-way interactions were removed (along with two three-way interactions), with factors (model variant (A–D), shared pattern (0/1), shared shape (0/1), shared colour (0/1)) added sequentially to evaluate whether their inclusion explained a significant proportion of the residual variability in the proportion of the first four prey attacked that were rewarding. All entered factors were significant ( $*P < 0.05$ ;  $**P < 0.01$ ;  $***P < 0.001$ ) either as main effects or as part of an interaction, or both. A reordered hierarchical model made little difference to qualitative conclusions.

became significant as a main effect, as did the pattern\*shape interaction). As might be expected, the avoidance rates of the nonrewarding mimics was much higher by the time human subjects reached their final generalization test (see [Supplementary Fig. S2](#)), but even here the colour, shape and pattern similarity influenced the overall avoidance rates of the first four prey attacked in this final (8th) trial (see [Supplementary Table S1](#)).

Given the complexity of the interactions, we can complement the above analysis by fitting a simpler model recognizing the eight categorical appearance combinations (perfect, colour and pattern (CP), colour and shape (CS), pattern and shape (PS), colour, shape, pattern, none) separately rather than in factorial form. Here, the number and type of shared features again explained significant variability in the proportion of the nonrewarding mimics avoided (LRT:  $\chi^2_7 = 140.28$ ,  $P < 0.001$ ). The model variant was not significant as a main effect (LRT:  $\chi^2_3 = 0.114$ ,  $P = 0.9901$ ) but was highly significant as part of an interaction with the number and type of shared features (LRT:  $\chi^2_{21} = 42.92$ ,  $P = 0.003$ ) (see [Fig. 3](#)). The presence of this interaction complicates the comparison of attack rates across the different treatment levels. However, identifying homogeneous subsets from Tukey HSD tests after fitting a model with model variant and the type of shared features as main effects alone provides some guidance ([Fig. 3](#)). Perfect, colour and shape (CS), colour and pattern (CP) mimics all achieved high levels of protection that did not differ significantly from one another (all  $P > 0.05$ ), indicating that these imperfect mimics achieve similar levels of protection to perfect mimics. The rate of avoidance of mimics with pattern and shape (but not colour) in common with the former nonrewarding model depended on the model variant tested. For example, while the plain square pattern and shape (PS) mimics of plain yellow squares were frequently avoided (model variant C), the dotted triangle pattern and shape mimics of dotted cyan triangles (model variant A) were not.

Like Kazemi et al., we found that colour-only mimics were avoided much more frequently than shape-only mimics in all four model variants, indicating that colour was generalized more readily than shape. However, even highly avoided colour-only mimics did

not receive as much protection as perfect mimics in the generalization phase. In addition, the avoidance of pattern-only mimics in particular depended on the specific model variant presented: striped mimics were frequently avoided when the original model was a striped magenta circle (model variant B), but plain mimics were far less frequently avoided when the original model was a plain yellow square (model variant C).

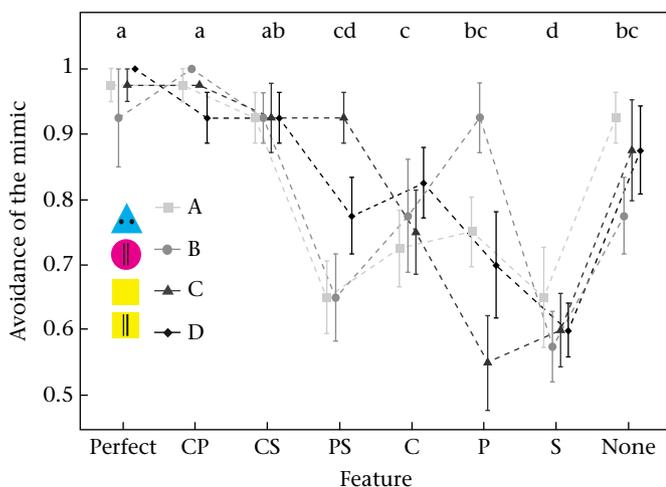
Perhaps the most informative aspects of the full factorial model were the rates of avoidance of unprofitable prey, which had no features in common with either the rewarding prey or the former nonrewarding prey. These 'nonmimics' were avoided at surprisingly high rates (60% of individuals did not attack any of them in the first four prey items of the generalization test). Note that while these nonrewarding nonmimics were entirely distinct from the former nonrewarding prey, they were also discriminable from the rewarding prey that the subjects had learned to attack.

## DISCUSSION

Just as Kazemi et al. found, our human predators rapidly learned to avoid nonrewarding prey and eventually concentrated almost exclusively on the rewarding prey phenotypes. Remarkably, several subjects attacked only rewarding prey on the first discrimination trial. It is possible that the high frequency of the nonrewarding phenotype (eight of the same type) compared to the rewarding phenotype (eight separate designs) allowed predators to recognize even more rapidly that there was something special about this particular phenotype and in rare cases avoid them outright. Yellow squares with black stripes were avoided at particularly high frequency even from the first trial, possibly because of their similarity to familiar warning signals, in particular the yellow and black stripes of many wasps (and road signs).

Like Kazemi et al., we found that some imperfect mimics gained a statistically indistinguishable degree of protection compared to perfect mimics, although in our case the imperfect mimics were based on sharing colour and pattern or colour and shape with the former models. The protection afforded to pattern and shape mimics was however more variable and dependent on the model variant tested. Many mimetic species of hoverfly have the yellow and black colour and pattern of their wasp models but lack the narrow waists, giving them a different shape. Our results therefore help explain why such imperfect mimicry can persist despite these differences in shape: in focusing on traits such as colour and pattern, predators tend to overlook them. Indeed [Bain et al. \(2007\)](#) found just such rules when attempting to elucidate the decisions of pigeons tasked with discriminating between mimetic hoverflies and their wasp models. Likewise, coral snake mimics may share the same colour and shape as their models, but predators tend to overlook the pattern ordering when deciding whether to attack or not ([Gamberale-Stille et al., 2012](#); [Kikuchi & Pfennig, 2010](#)).

In Kazemi et al., colour mimics alone gained equivalent protection to the perfect mimics, but as noted above, in our experiments, sharing two separate traits was necessary before similar levels of protection were achieved. Human predators are clearly highly motivated and strategic foragers, and it is possible that even closer mimicry was necessary to confer high degrees of protection with such refined decision making. Nevertheless, there are limits to the perfection required. It is well known that humans can be regularly duped by even imperfect mimicry (e.g. see [Golding, Ennos, Sullivan, & Edmunds, 2005](#)), while [Valkonen and Mappes \(2014\)](#) proposed that viper mimicry is disadvantageous to smooth snakes when encountering humans, who may kill them after mistaking them for venomous species. Moreover, the primacy of colour in overshadowing reported by Kazemi et al. was also observed in our experiments. Thus, subjects effectively overlooked



**Figure 3.** The initial extent (calculated from the first four prey attacked on the first generalization trial) of avoidance of nonrewarding 'mimics' that had all traits (perfect mimics), two traits (colour and pattern, CP; colour and shape, CS; pattern and shape, PS), one trait (colour, pattern or shape) or no traits (none) in common with the nonrewarding models that the human subjects had previously learned to avoid. The separate lines represent the different model variant groups that were separately compared (A–D). Lowercase letters across the top indicate statistically homogeneous subsets identified after fitting a main effects-only model (no interaction). Error bars represent  $\pm$ SE.

differences in shape or pattern in all four model experimental designs so long as the colour of the mimic (and an alternative characteristic) was the same as that of the former model. The importance of colour as a discriminative trait is further supported by the fact that (as Kazemi et al. likewise reported) colour-only mimics were avoided more frequently than shape-only mimics. Intriguingly, the high avoidance of pattern-only (striped) mimics in model variant B was not observed by Kazemi et al., but it was also seen in our preliminary experiment based on 92 human subjects.

Clearly, the finding that colour (when coupled with other traits) has the capacity to overshadow other discriminative cues in both birds and humans suggests that Kazemi et al.'s broad findings are taxonomically robust. Currently we can only speculate why colour tends to play this role. Colour is a highly salient trait to both humans and birds (e.g. see Kikuchi, Malick, Webster, Whissell, & Sherratt, 2015; Osorio, Jones, & Vorobyev, 1999) and since overshadowing occurs when attentional resources are allocated to responding to one trait at the expense of others (Mackintosh, 1976), one might expect that priority would be given to reacting to traits that will allow the receiver to identify the object as quickly as possible. Colour constancy, or the ability to identify the same hue under different lighting conditions, is widespread among animals (including humans), despite differences in the evolution of their visual systems (Wolfe et al., 2009). By contrast, pattern perception can vary with distance (Barnett & Cuthill, 2014; Campbell & Green, 1965), and perception of shape is a complex process that can be affected by many factors (Wolfe et al., 2009). It is therefore unsurprising that shared colour should be a more important cue than shared shape or pattern to both human and bird subjects.

Novel nonmimics (which shared no attributes either with rewarding or the former nonrewarding prey) were avoided at high rates, possibly because subjects had learned the characteristics of rewarding prey and were continuing to use these rules to forage on precisely the same phenotypes in the second stage of experiment. That said, it is noteworthy that the nonmimics with noncircular dot-like patterns were avoided most frequently when the nonrewarding model had two circular dots (model variant A), suggesting that the superficial similarity of the 'nonmimics' to the former nonrewarding models may also have played some role. Likewise, we cannot rule out a role for neophobia. For example, Aronsson and Gamberale-Stille (2012) found that domestic chicks tended to avoid prey with novel traits after the first trial, effectively treating such prey as unprofitable.

Our factorial design may give a clearer indication of how mimicry might evolve in this system. We can begin by elucidating the conditions under which the one-trait mimics would be selected for. Unfortunately the answer is not entirely straightforward since it depends on context. Note that one-trait nonrewarding mimics elicited no higher avoidance rate than novel prey. By contrast, the rewarding prey were clearly attacked at high rates (reflecting the 'avoidance rate' of nonrewarding mimics of our human predators). Therefore, if we treat the one-trait (colour, pattern or shape) mimics as single-step mutants of the rewarding prey (albeit now also nonrewarding), then it is likely that these one-trait mimics would spread in the population. By contrast, if we treat the one-trait mimics as two-step mutants of the former nonrewarding prey, then, by lacking two traits in common, they are at a selective disadvantage compared to the original model (or perfect mimics).

Would imperfect two-trait mimics involving colour and one other trait continue to be avoided in the long term? Humans eventually learned to avoid all the nonrewarding prey types at high rates. Intriguingly, however, there remained significant residual variation in attack rates in the final trial due to shared colour, pattern and shape with the former model, so the effects of early experiences were not rapidly forgotten over the short duration of

the experiment. To evaluate whether two-trait mimics would continue to be avoided if they were rewarding, one would clearly need to introduce rewarding (Batesian) mimics in the generalization phase. In addition, while having separate discrimination and generalization phases helps to reveal the underlying mechanisms, perhaps a more natural version of the above experiment would be to introduce occasional 'mutants' without partitioning the experiment into these two phases.

Although the longer-term survival prospects of imperfect Batesian mimics remains to be established, these experiments confirm that traits differ in their importance for discrimination learning, in our case colour being more important than shape. Moreover, we confirm that predators (both birds and humans) can effectively be duped by imperfect mimics if the mimics share the key features of prey that the predators had learned to avoid in the discrimination process. This study therefore further establishes that the cognitive processes of predators are important to consider in the evolution of mimicry.

## Acknowledgments

We thank the numerous volunteers for their participation and happy smiles. The research was approved by the Carleton University Research Ethics Committee and conducted according to the guidelines set out in the Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans. Our work was facilitated by grants from the Natural Sciences and Engineering Research Council of Canada (T. N. S.) and the National Science Foundation (NSF-1306037 to D. W. K.).

## Supplementary Material

Supplementary Material associated with this article is available, in the online version, at <http://dx.doi.org/10.1016/j.anbehav.2015.07.011>.

## References

- Aronsson, M., & Gamberale-Stille, G. (2008). Domestic chicks primarily attend to colour, not pattern, when learning an aposematic coloration. *Animal Behaviour*, *75*, 417–423.
- Aronsson, M., & Gamberale-Stille, G. (2012). Colour and pattern similarity in mimicry: evidence for a hierarchical discriminative learning of different components. *Animal Behaviour*, *84*, 881–887.
- Bain, R. S., Rashed, A., Cowper, V. J., Gilbert, F. S., & Sherratt, T. N. (2007). The key mimetic features of hoverflies through avian eyes. *Proceedings of the Royal Society B: Biological Sciences*, *274*, 1949–1954.
- Balogh, A. C. V., Gamberale-Stille, G., Tullberg, B. S., & Leimar, O. (2010). Feature theory and the two-step hypothesis of Mullerian mimicry evolution. *Evolution*, *64*, 810–822.
- Barnett, J. B., & Cuthill, I. C. (2014). Distance-dependent defensive coloration. *Current Biology*, *24*, R1157–R1158.
- Bates, H. W. (1862). Contributions to an insect fauna of the Amazon valley. Lepidoptera: Heliconidae. *Transactions of the Linnean Society of London*, *23*, 495–566.
- Beatty, C. D., Beirincx, K., & Sherratt, T. N. (2004). The evolution of Mullerian mimicry in multispecies communities. *Nature*, *431*, 63–67.
- Campbell, F. W., & Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *Journal of Physiology*, *181*, 576–593.
- Chittka, L., & Osorio, D. (2007). Cognitive dimensions of predator responses to imperfect mimicry. *PLoS Biology*, *5*, 2754–2758.
- Cuthill, I. C. (2014). Evolution: the mystery of imperfect mimicry. *Current Biology*, *24*, R364–R365.
- Ford, H. (1971). The degree of mimetic protection gained by new partial mimics. *Heredity*, *27*, 227–236.
- Gamberale-Stille, G., Balogh, A. C. V., Tullberg, B. S., & Leimar, O. (2012). Feature saltation and the evolution of mimicry. *Evolution*, *66*, 807–817.
- Golding, Y., Ennos, R., Sullivan, M., & Edmunds, M. (2005). Hoverfly mimicry deceives humans. *Journal of Zoology*, *266*, 395–399.
- Kazemi, B., Gamberale-Stille, G., Tullberg, B. S., & Leimar, O. (2014). Stimulus salience as an explanation for imperfect mimicry. *Current Biology*, *24*, 965–969.
- Kikuchi, D. W., Malick, G., Webster, R. J., Whissell, E., & Sherratt, T. N. (2015). An empirical test of 2-dimensional signal detection theory applied to Batesian

- mimicry. *Behavioral Ecology*. <http://dx.doi.org/10.1093/beheco/arv072>. Advanced online publication.
- Kikuchi, D. W., & Pfennig, D. W. (2010). Predator cognition permits imperfect coral snake mimicry. *American Naturalist*, *176*, 830–834.
- Kikuchi, D. W., & Pfennig, D. W. (2013). Imperfect mimicry and the limits of natural selection. *Quarterly Review of Biology*, *88*, 297–315.
- Mackintosh, N. J. (1976). Overshadowing and stimulus intensity. *Animal Learning & Behavior*, *4*, 186–192.
- Nicholson, A. J. (1927). Presidential Address. A new theory of mimicry in insects. *Australian Zoologist*, *5*, 10–24.
- Osorio, D., Jones, C. D., & Vorobyev, M. (1999). Accurate memory for colour but not pattern contrast in chicks. *Current Biology*, *9*, 199–202.
- Poulton, E. B. (1912). Darwin and Bergson on the interpretation of evolution. *Bedrock*, *1*, 48–65.
- R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Schmidt, R. S. (1958). Behavioural evidence on the evolution of Batesian mimicry. *Animal Behaviour*, *6*, 129–138.
- Schmidt, R. S. (1960). Predator behaviour and the perfection of incipient mimetic resemblances. *Behaviour*, *16*, 149–158.
- Shettleworth, S. J. (2010). *Cognition, evolution and behaviour* (2nd ed.). Oxford, U.K.: Oxford University Press.
- Terhune, E. C. (1977). Components of a visual stimulus used by scrubjays to discriminate a Batesian model. *American Naturalist*, *111*, 435–451.
- Turner, J. R. G. (1984). Mimicry: the palatability spectrum and its consequences. In R. I. Vane-Wright, & P. Ackery (Eds.), *The biology of butterflies* (pp. 141–161). New York, NY: Academic Press.
- Valkonen, J., & Mappes, J. (2014). Resembling a viper: implications of mimicry for conservation of the endangered smooth snake. *Conservation Biology*, *28*, 1568–1574.
- Wolfe, J. M., Kluender, K. R., Levi, D. M., Bartoshuk, L. M., Herz, R. S., Klatzky, R. L., et al. (2009). *Sensation and perception*. Sunderland, MA: Sinauer.