



UNIVERSITY OF LEEDS

This is a repository copy of *Field estimates of survival do not reflect ratings of mimetic similarity in wasp-mimicking hover flies*.

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/80034/>

Version: Accepted Version

---

**Article:**

Easley, JL and Hassall, C (2014) Field estimates of survival do not reflect ratings of mimetic similarity in wasp-mimicking hover flies. *Evolutionary Ecology*, 28 (2). 387 - 396. ISSN 0269-7653

<https://doi.org/10.1007/s10682-013-9678-3>

---

**Reuse**

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Field estimates of survival do not reflect ratings of mimetic similarity in wasp-mimicking hoverflies<sup>1</sup>

Jennifer Easley, Christopher Hassall

## ABSTRACT

The evolution of mimicry, and particularly the persistence of undefended Batesian mimetic forms that are imperfect copies of their defended models, remains a central question in evolutionary biology. Previous work has demonstrated that variation in mimetic fidelity in artificial prey can alter survival. However, no studies have validated the assumption that detailed laboratory-based measurements of mimetic fidelity are actually reflected in survival in natural field experiments. Here, we demonstrate that, in line with previous studies, the mimetic similarity of 77 hover fly (Diptera: Syrphidae) species to the common wasp *Vespula alascensis* is strongly related to the number of abdominal stripes exhibited by the flies. We then produce three artificial pastry baits: (i) a “model” which is chemically defended and has two stripes, (ii) a one-stripe mimic, and (iii) an unstriped mimic. Based on the ratings study, we predicted that the one-stripe mimic would exhibit survival intermediate between the unstriped mimic and the model. Baits were deployed in experiments each involving 81 baits (27 of each kind), at 3 sites, with experiments replicated 10 times at each site for a total deployment of 2430 baits. Proportional hazards models show that both one-striped and model baits survived equally well and significantly better than the unstriped baits, suggesting categorical prey identification rather than the use of stripe number as a continuous trait, as was suggested by the laboratory study. These findings suggests that, while humans and avian predators can distinguish mimics and models in the laboratory using a range of traits, behaviour in the field may not reflect this ability. This absence of a link between continuous measures of mimetic fidelity and prey selection may contribute to the maintenance of imperfect mimicry, but more studies using near-natural experimental paradigms are needed to investigate the phenomenon further.

**Keywords:** Batesian mimicry, imperfect mimicry, pastry baits, hoverflies, Diptera, Syrphidae, Hymenoptera

## INTRODUCTION

Evolutionary biologists have described countless examples of exquisite adaptations, whether they be extreme sexual ornaments in male widowbirds (Andersson 1982), physiological adaptations to harsh environments in extremophiles (Rothschild and Mancinelli 2001), or unique life histories to exploit unusual niches such as the piophilid flies that live entirely within discarded moose antlers (Bonduriansky 1995). Some of the most vivid demonstrations of natural selection occur during the co-evolutionary arms race between predators and prey. Prey exhibit a remarkable array of anti-predator traits, which can be nullified through concomitant evolution (at a cost) by the predator (Tien and Ellner 2012). The weight of evidence suggests that prey are more likely to evolve traits in response to a predator than *vice versa*, and that those defensive traits show a greater degree of refinement than the offensive traits exhibited by predators (Abrams 2000). Anti-predator traits can be classified into four broad categories that act at different points along the predation event: (i) traits that reduce detection (e.g. the camouflage of moths against trees, Webster et al. 2009), (ii) traits that reduce recognition (e.g. snake-mimicking caterpillars, Hossie and Sherratt 2012), (iii) traits that reduce capture (e.g. protean escape trajectories, Domenici et al. 2011), and (iv) traits that reduce consumption (e.g. the death-feigning posture of some crickets, Honma et al. 2006).

---

<sup>1</sup> Version of record can be viewed at the publisher and should be cited as [Easley, J. L. and Hassall, C. \(2014\) Field estimates of survival do not reflect ratings of mimetic similarity in wasp-mimicking hover flies, \*Evolutionary Ecology\*. 28, 387-396.](#)

Perhaps due to predominance of vision in humans, the disruption of predator recognition through morphological adaptations is among the most celebrated biological phenomena, and this mimicry of non-prey animals can take a number of forms. A palatable prey item can “masquerade” as an object that is undefended but inedible to the predator (Skelhorn et al. 2010), a phenomenon that could also be considered a form of camouflage. Acoustic mimicry occurs in moths, where bats learn to associate certain clicking patterns with chemically defended species and ignore palatable species that make similar sounds (Conner and Corcoran 2012), but not in hover flies (Diptera: Syrphidae) that resemble stinging Hymenoptera (Rashed et al. 2009). More perceptible to humans, due to our highly developed vision, is the morphological similarity between certain species that vary in whether they possess a defence. This phenomenon, known as Batesian mimicry (Bates 1862) where an undefended species mimics some aspect of a defended species, has resulted in some remarkable convergences of morphology between distantly-related species. However, far less emphasis has been placed on the *limits* of natural selection, and those organisms that appear “imperfect” or maladapted (Nesse 2005). Despite this, there is a great deal to learn from (apparently) suboptimal traits – indeed some have stated that it is impossible to establish a thorough understanding of the nature of phenotypic variation without considering adaptational limits and maladaptations (Crespi 2000). This concept of limits to selection has provided an understanding of how genetic variability persists despite strong selection during the historical domestication of certain plants (Wang et al. 1999).

While the neotropical butterflies in which Bates originally described this form of mimicry (Bates 1862) remain the classic exemplar, there are a number of other model-mimic complexes that follow the same pattern but to varying extents. For example, the hover flies that supposedly mimic stinging Hymenoptera appear to consist of relatively few “good mimics” and a much larger number of species that seem to bear little resemblance to their models (Gilbert 2005). A large number of hypotheses have been presented to explain the persistence of low fidelity mimics when there exists such a clear benefit to close mimetic similarity. It has been proposed that apparent imperfect mimicry is an artefact of the human visual system (Cuthill and Bennett 1993), that imperfect mimics are partially mimicking a range of models rather than just one (Edmunds 2000), that perfect mimicry removes the benefit to the mimic (Johnstone 2002), or that less profitable prey experience lower selection pressures (Sherratt 2002). In a recent review of this array of hypotheses, Penney et al. (2012) demonstrated that, for the hover flies, at least, there was substantial evidence for an association between body size (as a proxy for prey profitability) and mimetic similarity, but no evidence for any other hypothesis. It appears, then, that imperfect mimicry can persist when selection is relaxed.

However, while a powerful macroevolutionary analysis that brought together diverse empirical approaches, the key limitation of the comparative study by Penney et al. (2012) is that it relies on paradigms that are removed from the real ecological world. While birds and humans may agree upon levels of mimetic fidelity in a laboratory setting, there have been few experimental tests of natural levels of predation by wild birds on imperfectly mimetic prey (cf Morrell and Turner 1970), and there have been no experimental tests that use (i) an ecologically valid trait, and (ii) link field survival to laboratory ratings of mimetic similarity. In this study, we provide the first test of the hypothesis that variation in mimetic fidelity as rated by humans in the laboratory (which correlates with avian ratings, Penney et al. 2012) is correlated with field survival of prey. We test this hypothesis by exploring variation in a single trait: the number of yellow abdominal stripes on hover flies and their hymenopteran models. This trait has been shown to be a salient feature of hover fly mimicry when judged by avian subjects (Bain et al. 2007), and shows considerable variation within the hover flies. First, we evaluate the relationship between stripe number and mimetic similarity to a common wasp model using human participants. We then conduct an extensive field experiment where we experimentally manipulate stripe number in artificial, pastry prey to investigate whether ratings of mimetic fidelity are reflected in natural rates of predation by wild birds.

## METHODS

### *Human experiment*

Photographs of three different pinned individuals from each of 77 hoverfly species were taken from specimens stored at the Canadian National Collection of Insects and Arachnids, Ottawa, Canada. Also photographed were three individuals of five separate hymenopteran species that are thought to include the Batesian models of most hoverflies (the honeybee *Apis mellifera*, the common wasp *Vespula alascensis*, the buff-tailed bumblebee *Bombus terrestris*, the eastern bumblebee *Bombus impatiens*, and the bald-faced hornet *Dolichovespula maculata*). All of the photographs were taken using a Canon EO5 50D with a Canon macro lens (100 mm). Standardised illumination was provided by an 80 LED microscope ring light (KD-200).

Three separate blocks of human participants were used to evaluate mimetic similarity ( $n=24$ ,  $n=10$ , and  $n=10$ ). Human participants were shown a slideshow in which each mimic photograph was shown with five photographs corresponding to the five models (see Figure 1 for an example). Each block used a different photograph of the mimics and models, presented in a random order. Participants were asked to rate each hoverfly for its similarity to each of the models (1=very poor, 10=very good). For the present study, we are only concerned with the similarity of each species to the striped vespid wasp, *V. alascensis*, to which we calculated the mean similarity for each of the 77 hoverfly mimics. Each mimic was then scored for the number of discrete yellow stripes on the abdomen (see Table S1 for stripe numbers). Ethical approval was granted by the University of Leeds Biological Sciences Faculty Research Ethics Committee (reference BIOSCI 12-027) prior to the start of the study.

### *Field experiment*

Field work was carried out in three locations. Two of these were sections of lawn situated on the University of Leeds campus: Saint George's Field (53.809°N 1.557°W) and the Miall Building (53.805°N 1.555°W). These sites were used between 6 November 2012 and 6 December 2012. Magpies (*Pica pica*), carrion crows (*Corvus corone*) and pigeons (*Columba livia*), were all known to be present at these sites. The final location was situated off campus, approximately 70 miles from Leeds in Marske (54.592°N 1.011°W) where data was collected between the dates of 7 December 2012 and 7 January 2012. This site was used by a different species: the blackbird *Turdus merula*.

Based on pilot data, we modified Speed et al's (2000) recipe for pastry baits (360g flour, 120g lard and 20ml water to make 500g of pastry) to 310g flour, 160g lard, and 30ml of water to make the pastry easier to work with. 10ml of yellow food colouring (Silver Spoon) was added to each batch of 500g. Models were created by adding 3ml of 5% Quinine solution and 3ml of 5% Bitrex solution, after pilot data showed that concentrations used in other studies (1.5 ml of 1% Quinine and 1.5ml of 1% Bitrex Marples 1993; Speed et al. 2000; Skelhorn and Rowe 2006) were insufficient to deter birds. This may be because other studies have worked with small-bodied species such as *Parus major* and *Sturnus vulgaris*, while species at our sites were much larger. Pastry baits were made 40mm length and 10mm diameter, and shaped using a clay extruder to ensure a consistent cross-section. Three types of bait were produced: (i) a "model" that was chemically defended was painted with two stripes using black non-toxic paint (Marabu Decorlack), (ii) a one-striped mimic that was undefended and possessed a black single stripe, and (iii) an unstriped mimic that was undefended and had no stripes. Twenty-seven baits of each kind were randomly deployed on the vertices of an 8m x 8m grid, giving 81 baits per experimental run. The baits were placed out at sunrise then censused at 11:00, 14:00, sunset and finally at sunrise the next day. They were then collected at the end of this 24 hour cycle and fresh baits were deployed for the next experimental run. This procedure was repeated twice per week between the dates previously specified (see Table S2 for specific dates). A total of 10 experimental runs were completed at each of the three sites, leading to 30 experimental runs and the deployment of 2430 baits.

### *Statistical analysis*

Human raters were tested for agreement using two-way intraclass correlations in the irr (Gamer et al. 2012) package in R (R Development Core Team 2013). In order to test for a correlation between the number of stripes and the human ratings, a Spearman's rank test was used. The data were further analysed using ANOVA with a Tukey posthoc test to investigate differences in ratings between species with varying numbers of stripes. To investigate the effect of stripe number and chemical defence on bait survival, we used a Cox proportional hazards model implemented using the Survival package (Therneau 2013) in R. Due to significant variation in weather over the course of the experiments, we expect there to be substantial variation in baseline hazard rates between experimental runs. As a result, we stratified the analysis according to date. The resulting model (survival ~ Bait type + strata(Date)) was checked to ensure that it conformed to the assumptions of proportional hazards using the cox.zph function in the Survival package.

## RESULTS

Human raters showed highly significant agreement on mimetic similarity in the hover fly species (intraclass correlation coefficient = 0.418,  $p < 0.001$ ). Mean ratings for each species were highly significantly correlated with the number of abdominal stripes that each species possessed (Spearman's  $\rho = 0.783$ ,  $p < 0.001$ ). A Tukey posthoc test showed strong differences between groups of species possessing different numbers of stripes, with unstriped flies rated as being significantly less wasp-like than flies with two or more stripes, and flies with more than three stripes rated as being significantly more wasp-like than flies with fewer than three stripes (Figure 2).

Of the 810 baits deployed of each bait type, 396 (48.9%) unstriped baits, 362 (44.7%) one-striped baits, and 352 (43.5%) model baits were eaten across all experiments. The Cox proportional hazards model conformed to the assumptions of proportional hazards ( $\chi^2 = 3.011$ ,  $p = 0.222$ ). The model confirmed that when baits were deployed in the field we saw a significantly greater survival in the model ( $z = -2.106$ ,  $p = 0.035$ ) and the one-stripe mimic ( $z = -2.734$ ,  $p = 0.006$ ) relative to the unstriped mimic. This equates to a reduction in the hazard rate of 14.3% in the model, and 18.4% in the one-stripe mimic relative to the unstriped mimic (Figure 3), but there was no significant difference in hazard between the model and the one-striped mimic. When we looked for evidence of learning over time, there were no obvious patterns in relative mortality over the course of the 10 experimental runs at each site (Figure 4). Hazard rates varied considerably over the course of the experiments, and the hazards for each bait type were correlated across time periods.

## DISCUSSION

We provide an extensive test of the assumption that laboratory (human) ratings of mimetic fidelity correspond to variation in survival among imperfect mimics in the field. We find that increasing stripe number is associated with higher ratings of mimetic fidelity in experiments with humans as raters. In particular, hover flies with zero stripes are rated as the poorest mimics, not significantly poorer than one-striped hover flies, but significantly poorer than two-striped mimics. In these rating experiments, one-striped mimics are rated as having intermediate levels of mimetic fidelity – not significantly different from unstriped or two-striped hover flies. However, the field study does not confirm this finding. Instead, the one-striped mimic was attacked significantly fewer times than the unstriped mimic, and at a level that was statistically indistinguishable from the two-striped model. These findings suggest that real world prey choice and learning by avian predators may not reflect variations in mimetic fidelity as they are quantified in the laboratory.

There are a number of potential explanations for the difference in the relationship between mimetic fidelity (even when experimentally manipulated) and field survival. Perhaps avian visual acuity is insufficient to be able to discriminate between mimics with similar patterning. Comparative analyses of a range of bird species have demonstrated that contrast sensitivity (CS) in birds is much lower than was previously thought – in particular, much lower than in mammals (Ghim and Hodos 2006). However, two facts appear to contradict this hypothesis: first, the mimetic signal is aposematic and

maximises contrast as far as possible through alternating bands of light and dark pigment (Endler 1992). This should reduce the impact of limited CS in avian visual systems. Second, we observe a range of hover fly species with different numbers of stripes which we assume are the product of selection by a visual predator.

A second alternative is that, despite our use of increased concentrations of “chemical defense”, our model prey were not sufficiently defended as to promote the association of stripe number with defense, but only the presence of stripes (i.e. categorical discrimination, Chittka and Osorio 2007). This is despite the observation that stripe number as a continuous variable was shown to be an important trait in discriminating wasps from hover flies (Bain et al. 2007). Modelling studies by Sherratt (2002) found that increasing the cost of attacking a model broadened the range of phenotypic variation that was protected in a population of mimics. This raises an interesting question: do predators switch either the traits that they are using to distinguish mimics and models, or the resolution at which they make that discrimination, based on the strength of the model’s defense? And what are the implications for the evolution of mimicry when mimics are attacked by a predator community within which predator species have a range of tolerances to a given chemical defense?

Our study differs from many others by using a combination of (i) a natural population of birds that are foraging freely within their habitat, and (ii) experimental manipulation of a discrete trait that is known to vary among mimetic species. Other studies have used wild caught birds in aviaries (e.g. great tits (*Parus major*), Rowland et al. 2007; starlings (*Sturnus vulgaris*), Halpin et al. 2013) or, when wild birds have been used in a natural setting, experimental manipulation of artificial prey has varied prey appearance in a way that does not reflect natural variation (Cuthill et al. 2005). However, the use of ecologically valid manipulations of artificial prey with birds foraging in a natural environment has provided great insights into the evolution of caterpillar eyespots (Hossie and Sherratt 2012; Hossie and Sherratt 2013), and other studies would benefit from adopting such an approach. An example of such a study that is much needed involves the manipulation of stripe number (mimetic fidelity) and prey size (profitability) using realistic ranges of each parameter to validate the relaxed selection hypothesis (where smaller prey are under weaker selection due to reduced profitability) that was supported by Penney et al.’s (2012) comparative analysis.

The findings we describe require a reconsideration of the nature of mimicry as we relate detailed measurements of mimetic fidelity to larger-scale evolutionary processes. We know that humans and birds *can* rate mimetic fidelity in quantitatively similar ways (Penney et al. 2012), contrary to the “eye of the beholder” hypothesis for imperfect mimicry (Cuthill and Bennett 1993). However, it is likely that other considerations aside from mimetic similarity influence avian prey selection under natural conditions, such as the availability (Lindström et al. 2004) or size of alternative prey (Halpin et al. 2013). Given the likely presence of publication bias in presenting mostly statistically significant findings in peer reviewed journals (Csada et al. 1996), it would be of interest to establish how many field tests of these principles have failed to detect a significant effect of mimicry under near-natural conditions. Such studies would provide valuable data on the relative strength of selection along the continuum from high ecological validity and low experimental control, to lower ecological validity but high experimental control.

## **CONCLUSIONS**

We tested a key assumption in mimicry theory: that observed variation in mimetic fidelity is reflected in natural levels of predation. We find that this is only partly true. While hover flies with a single stripe were rated as being intermediate between unstriped and two-striped mimics in a laboratory environment, field studies showed that survival of one- and two-striped prey was similar and higher than that for unstriped prey. We suggest that more studies of the effects of observed intra- or interspecific variation in traits involved in mimicry would yield a more realistic view of the strength of

selection for mimicry. The absence of a strong relationship between mimetic fidelity and selection in the field could contribute to the maintenance of imperfect mimicry.

## ACKNOWLEDGEMENTS

Adam Spencer, Joanne Ratcliffe, Hannah Dawson and Georgia Beattie provided assistance with experimental work. Tom Hossie and Phillip Spencer provided useful comments on the manuscript. Funding was provided by the University of Leeds.

## REFERENCES

- Abrams PA (2000) The evolution of predator-prey interactions: theory and evidence. *Annual Review of Ecology and Systematics* 31:79-105
- Andersson M (1982) Female choice selects for extreme tail length in a widowbird. *Nature* 299:818-820
- Bain RS, Rashed A, Cowper VJ, Gilbert FS, Sherratt TN (2007) The key mimetic features of hoverflies through avian eyes. *Proceedings of the Royal Society: Series B (Biological Sciences)* 274:1949-1954
- Bates HW (1862) Contributions to an insect fauna of the Amazon valley. Lepidoptera: Heliconidae. *Transactions of the Linnean Society of London*:495 - 566
- Bonduriansky R (1995) A new Nearctic species of *Protopiophila* Duda (Diptera: Piophilidae), with notes on its behaviour and comparison with *P. latipes* (Meigen). *Canadian Entomologist* 127:859-863
- Chittka L, Osorio D (2007) Cognitive dimensions of predator responses to imperfect mimicry. *PLoS Biology* 5:e339
- Conner WE, Corcoran AJ (2012) Sound strategies: the 65-million-year-old battle between bats and insects. *Annual Review of Entomology* 57:21-39
- Crespi BJ (2000) The evolution of maladaptation. *Heredity* 84:623-629
- Csada RD, James PC, Espie RHM (1996) The "file drawer problem" of non-significant results: does it apply to biological research? *Oikos* 76:591-593
- Cuthill IC, Bennett ATD (1993) Mimicry and the eye of the beholder. *Proceedings of the Royal Society: Series B (Biological Sciences)* 253:203-204
- Cuthill IC, Stevens M, Sheppard J, Maddocks T, Parraga CA, Troscianko TS (2005) Disruptive coloration and background pattern matching. *Nature* 434:72-74
- Domenici P, Blagburn JM, Bacon JP (2011) Animal escapology II: escape trajectory case studies. *The Journal of Experimental Biology* 214:2474-2494
- Edmunds M (2000) Why are there good and poor mimics? *Biological Journal of the Linnean Society* 70:459-466
- Endler JA (1992) Signals, signal conditions, and the direction of evolution. *The American Naturalist* 139:S125-S153
- Gamer M, Lemon J, Fellows I, Singh P (2012) irr: Various Coefficients of Interrater Reliability and Agreement, R package version 0.84. <http://CRAN.R-project.org/package=irr>
- Ghim M, Hodos W (2006) Spatial contrast sensitivity of birds. *Journal of Comparative Physiology A* 192:523-534
- Gilbert F (2005) The evolution of imperfect mimicry. In: Fellowes MDE, Holloway GJ, Rolff J (eds) *Insect Evolutionary Ecology*. CABI, Wallingford, UK, pp 231-288
- Halpin CG, Skelhorn J, Rowe C (2013) Predators' decisions to eat defended prey depend on the size of undefended prey. *Animal Behaviour* 85:1315-1321
- Honma A, Oku S, Nishida T (2006) Adaptive significance of death feigning posture as a specialized inducible defence against gape-limited predators. *Proceedings of the Royal Society: Series B (Biological Sciences)* 273:1631-1636
- Hossie TJ, Sherratt TN (2012) Eyespots interact with body colour to protect caterpillar-like prey from avian predators. *Animal Behaviour* 84:167-173

- Hossie TJ, Sherratt TN (2013) Defensive posture and eyespots deter avian predators from attacking caterpillar models. *Animal Behaviour* 86:383-389
- Johnstone RA (2002) The evolution of inaccurate mimics. *Nature* 418:524-526
- Lindström L, Alatalo RV, Lyttinen A, Mappes J (2004) The effects of alternative prey on the dynamics of Batesian and Müllerian mimicry. *Evolution* 58:1294-1302
- Marples NM (1993) Do wild birds use size to distinguish palatable and unpalatable prey types? *Animal Behaviour* 46:347-354
- Morrell GM, Turner JRG (1970) Experiments on mimicry: I. The response of wild birds to artificial prey. *Behaviour* 36:116-130
- Nesse RM (2005) Maladaptation and natural selection. *The Quarterly Review of Biology* 80:62-70
- Penney HD, Hassall C, Skevington JH, Abbott KR, Sherratt TN (2012) A comparative analysis of the evolution of imperfect mimicry. *Nature* 483:461-464
- R Development Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Rashed A, Khan MI, Dawson JW, Yack JE, Sherratt TN (2009) Do hoverflies (Diptera: Syrphidae) sound like the Hymenoptera they morphologically resemble? *Behavioral Ecology* 20:396-402
- Rothschild LJ, Mancinelli RL (2001) Life in extreme environments. *Nature* 409:1092-1101
- Rowland HM, Ihalainen E, Lindström L, Mappes J, Speed MP (2007) Co-mimics have a mutualistic relationship despite unequal defences. *Nature* 448:64-67
- Sherratt TN (2002) The evolution of imperfect mimicry. *Behavioral Ecology* 13:821-826
- Skelhorn J, Rowe C (2006) Prey palatability influences predator learning and memory. *Animal Behaviour* 71:1111-1118
- Skelhorn J, Rowland HM, Speed MP, Ruxton GD (2010) Masquerade: camouflage without crypsis. *Science* 327:51
- Speed MP, Alderson NJ, Hardman C, Ruxton GD (2000) Testing Mullerian mimicry: an experiment with wild birds. *Proceedings of the Royal Society B-Biological Sciences* 267:725-731
- Therneau T (2013) A Package for Survival Analysis in S. R package version 2.37-4, <URL: <http://CRAN.R-project.org/package=survival>>
- Tien RJ, Ellner SP (2012) Variable cost of prey defense and coevolution in predator-prey systems. *Ecological Monographs* 82:491-504
- Wang R-L, Stec A, Hey J, Lukens L, Doebley J (1999) The limits of selection during maize domestication. *Nature* 398:236-239
- Webster RJ, Callahan A, Godin J-GJ, Sherratt TN (2009) Behaviourally mediated crypsis in two nocturnal moths with contrasting appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364:503-510

## Figures

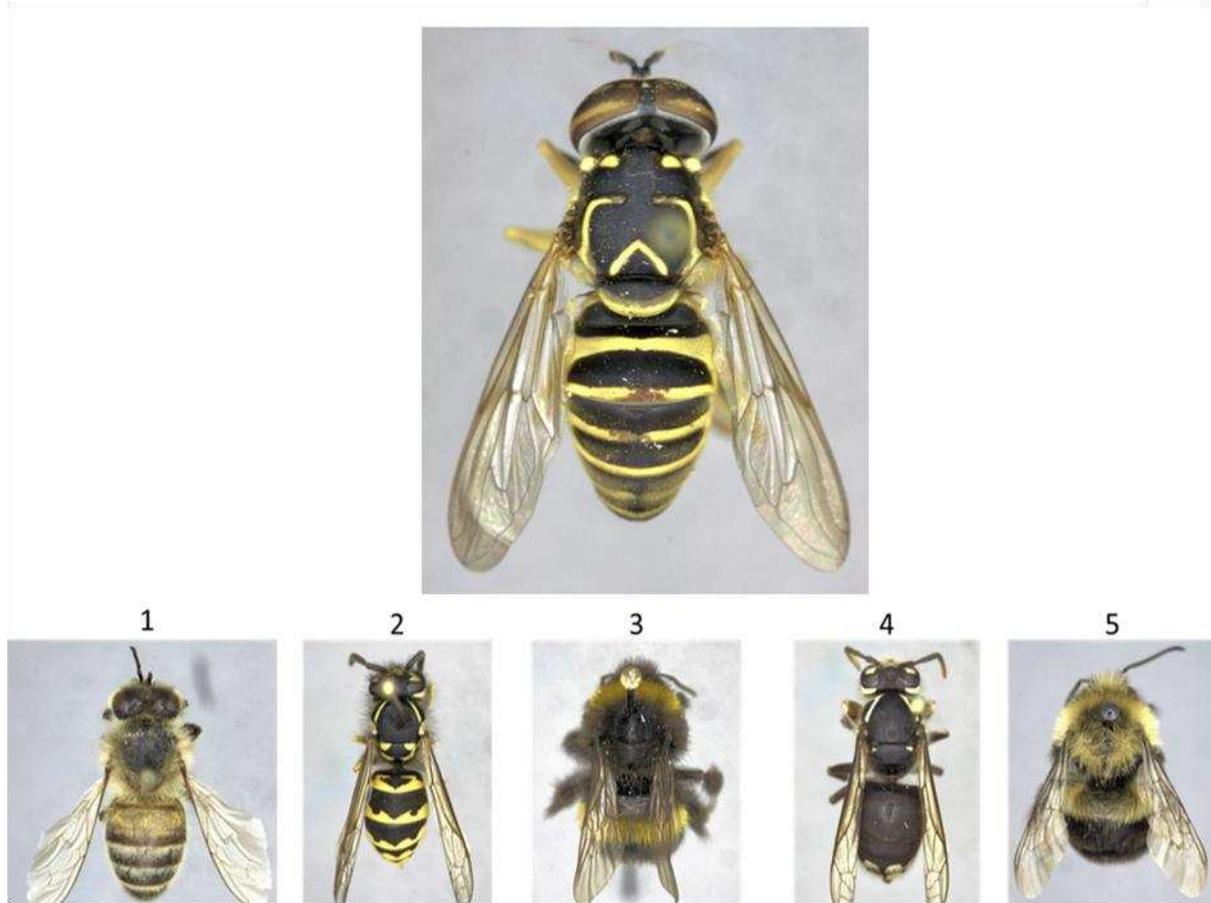


Figure 1 – Example of a screen from the presentation shown to human participants during the ranking of mimetic fidelity of hoverflies (large central image – in this example the hoverfly is *Spilomyia longicornis*, which is the only mimic with five stripes) and five putative Hymenoptera models. *Vespula alascensis*, the model of interest for the present study is labelled “2” on the slide.

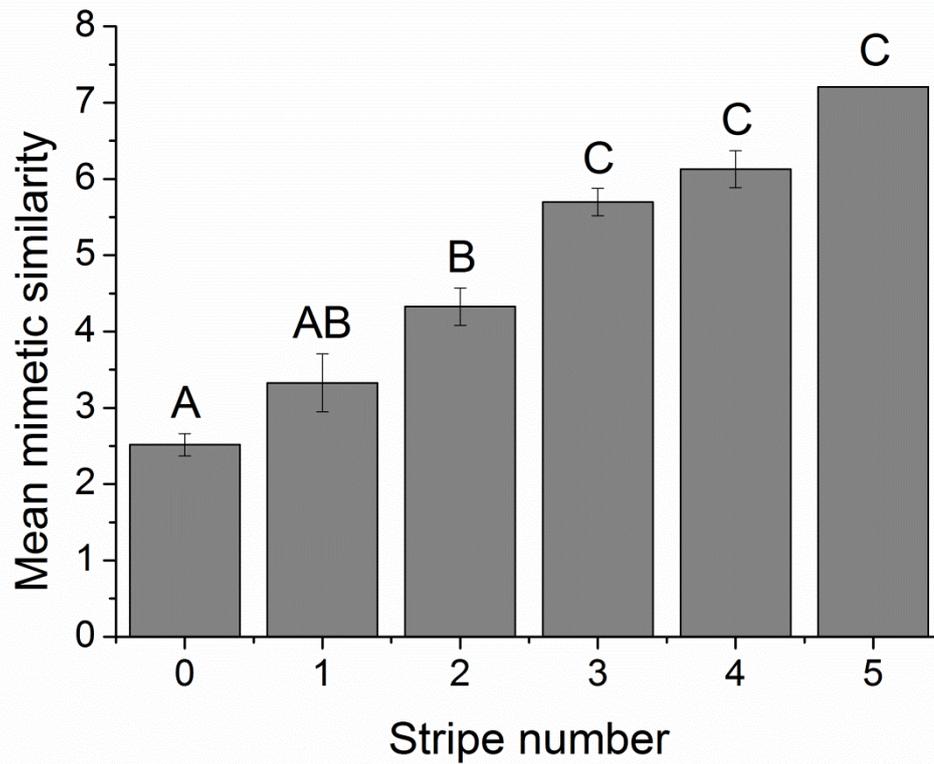


Figure 2 – Relationship between the number of yellow abdominal stripes and the ranking of mimetic similarity to *Vespula alascensis* for 77 species of hoverfly (Diptera: Syrphidae). Bars topped by the same letters are not significantly different according to a Tukey posthoc test. Error bars are 1 SE. Sample sizes from left to right are n=10, n=7, n=17, n=27, n=15, n=1. The single species indicated with five stripes is *Spilomyia longicornis* (pictured in Figure 1).

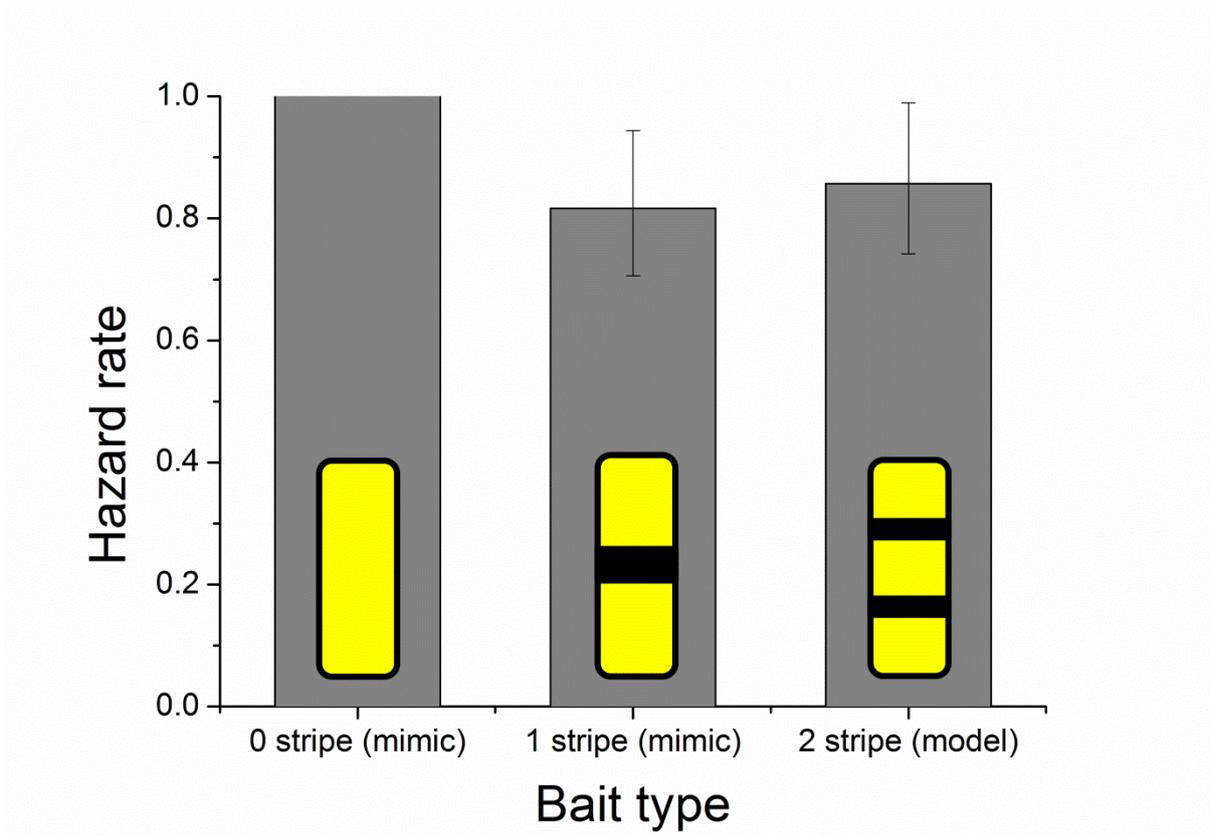


Figure 3 – Hazard rates of one- and two-striped baits relative to the hazard rate of baits without stripes. Error bars are 95% confidence intervals.

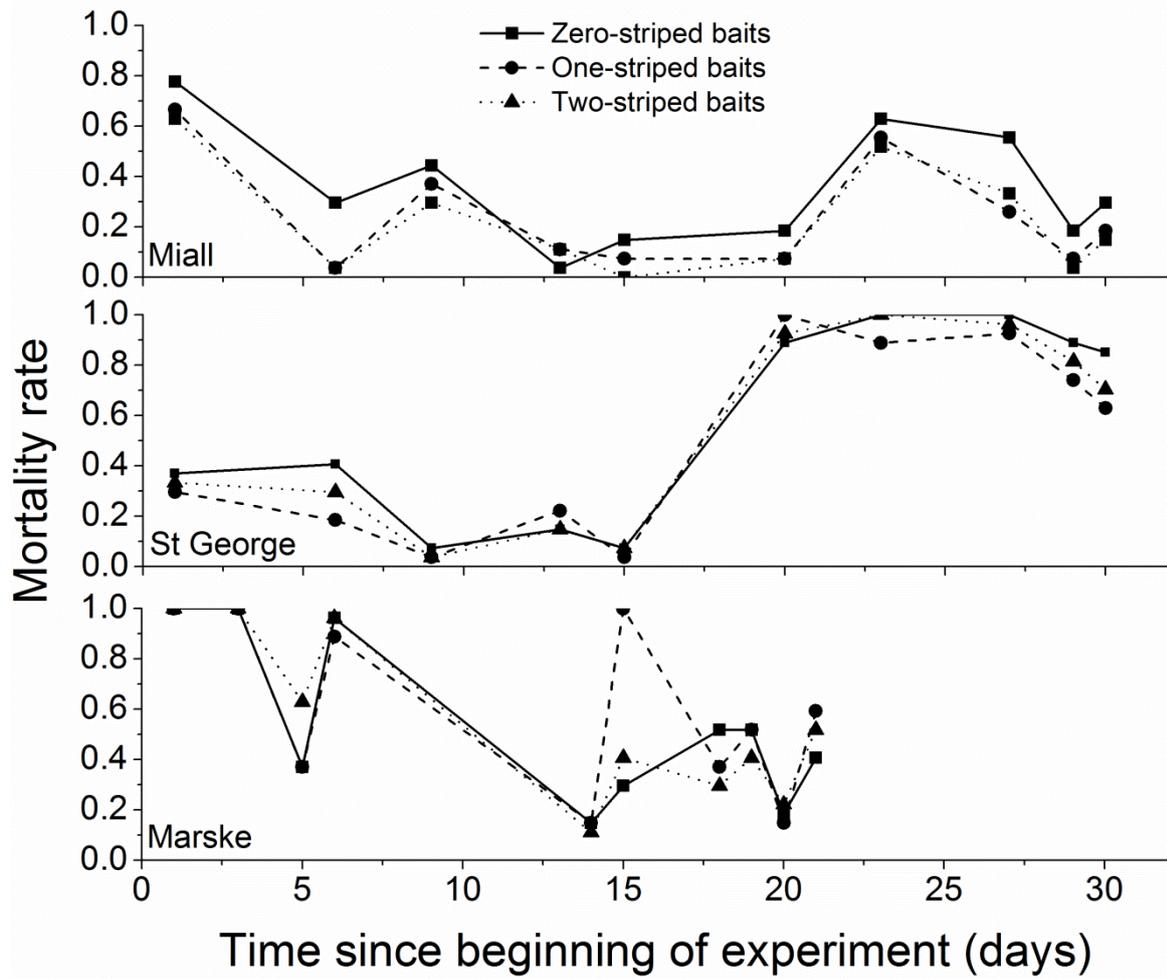


Figure 4 – Patterns of mortality in each bait type at each site after the start of the experiment. Experiments started on 9 November 2012 for Miall and St George, and on 11 December 2012 for Marske.