

Signal categorization by foraging animals depends on ecological diversity

David William Kikuchi^{1*}, Anna Dornhaus¹, Vandana Gopeechund², Thomas N Sherratt²

¹Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, United States; ²Department of Biology, Carleton University, Ottawa, Canada

Abstract Warning signals displayed by defended prey are mimicked by both mutualistic (Müllerian) and parasitic (Batesian) species. Yet mimicry is often imperfect: why does selection not improve mimicry? Predators create selection on warning signals, so predator psychology is crucial to understanding mimicry. We conducted experiments where humans acted as predators in a virtual ecosystem to ask how prey diversity affects the way that predators categorize prey phenotypes as profitable or unprofitable. The phenotypic diversity of prey communities strongly affected predator categorization. Higher diversity increased the likelihood that predators would use a 'key' trait to form broad categories, even if it meant committing errors. Broad categorization favors the evolution of mimicry. Both species richness and evenness contributed significantly to this effect. This lets us view the behavioral and evolutionary processes leading to mimicry in light of classical community ecology. Broad categorization by receivers is also likely to affect other forms of signaling.

DOI: <https://doi.org/10.7554/eLife.43965.001>

Introduction

Signals between species can evolve whenever selection favors both the evolution of a signal display by a 'sender' species, and a response by a 'receiver' species (*Bradbury and Vehrencamp, 2011; Maynard Smith and Harper, 2003*). However, signal evolution is mediated not only by economics, but also by the psychology of receivers (*Endler and Basolo, 1998; Guilford and Dawkins, 1993; Rowe, 2013; Ryan et al., 1990*). In ecological communities, animals are faced with a diverse panoply of stimuli. How they categorize stimuli as worth responding to or ignoring will influence when and how signals evolve. Here, we explore how signal evolution is affected by the set of stimuli present in communities of different levels of complexity.

Warning signals are one of the best studied examples of interspecific communication – they advertise prey defenses to potential predators, reducing negative interactions for both predator and prey (*Wallace, 1867*). Mimicry occurs when warning signals coevolve among multiple prey species. Mimics can vary in their resemblance to models, with low fidelity ('imperfect') mimics representing something of a paradox (*Cuthill and TD, 1993; Dittrich et al., 1993; Kikuchi and Pfennig, 2013; Sherratt and Peet-Paré, 2017*). Understanding variation in the extent of mimicry is a problem that spans evolution, ecology, and cognitive psychology (*Guilford and Dawkins, 1993; Mallet, 2001; Rowe, 2013; Ruxton et al., 2018*), since selection on mimetic resemblance is mediated by the way that predators categorize prey as profitable or unprofitable (*Beatty et al., 2004; Gamberale-Stille et al., 2012; Getty, 1985; Ihalainen et al., 2012; Kazemi et al., 2014; Kikuchi and Sherratt, 2015; Oaten et al., 1975; Sherratt, 2002; Sherratt and Peet-Paré, 2017; Speed and Ruxton, 2010*).

The diversity of a community affects predator decisions about prey. For example, in an experiment with artificial prey, diversity affected how predators made decisions in response to

*For correspondence:
dwkikuchi@gmail.com

Competing interests: The authors declare that no competing interests exist.

Funding: See page 10

Received: 28 November 2018

Accepted: 24 April 2019

Published: 25 April 2019

Reviewing editor: Bernhard Schmid, University of Zurich, Switzerland

© Copyright Kikuchi et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

warning signals that varied within a single, continuous dimension (*Ihalainen et al., 2012*). However, warning signals are often multicomponent, that is to say, complex – they consist of many different traits in concert (*Bradbury and Vehrencamp, 2011; Hebets and Papaj, 2005; Maynard Smith and Harper, 2003*). Indeed, genetic studies of mimicry complexes have revealed discrete variation among multiple mimetic traits (*Clarke and Sheppard, 1963; Jiggins, 2017; Kunte, 2009; Dasmahapatra et al., 2012*). Consequently, mimetic precision depends not only on how predators generalize within traits, but also on which traits they evaluate, and how they combine them to form higher-level categories. We use the terms ‘categorization’ and ‘generalization’ in the sense that categorization behavior results from using generalizations to make decisions (*Seger and Peterson, 2013*).

The use of ‘key’ traits is one simple way to classify prey (*Balogh and Leimar, 2005; Beatty et al., 2004; Gamberale-Stille et al., 2012*) – for example, using the rule ‘avoid yellow prey’ would mean that a predator would have to focus on the key trait of color (*Figure 1A*). However, the advantage of using key traits or any other form of categorization depends on the community in which these decisions are made (*Beatty et al., 2004; Ihalainen et al., 2012; Lindström et al., 2004*). In this study we examined the effects of communities on categorization, showing that different components of diversity have critical effects on which decision rules are used, and thus selection on mimetic signals.

Results and discussion

The simplest, most widely studied component of diversity is species richness, the number of species found in a community (*Magurran, 1988*). It might be difficult for predators to identify and remember the properties of individual prey types in rich communities with a large variety of prey. Predators could be limited by memory capacity (*Beatty et al., 2004; MacDougall and Dawkins, 1998*), or by the substantial risks of sampling unfamiliar species of prey, some of which might be highly unprofitable to attack (*Cohen et al., 2007; Houston et al., 2012; Sherratt and Peet-Paré, 2017*). If predators do not remember the characteristics of each discrete prey type but instead use rules (such as avoid yellow), then they could reduce the difficulty of deciding what to eat in a rich community. Therefore, it has been hypothesized that, as richness increases, predators will be more likely to use a key trait to make decisions (*Beatty et al., 2004; Wilson et al., 2013*). Indeed, *Beatty et al. (2004)* found that predators *could* use a key trait to make decisions in diverse communities; however, in their experiment, if predators did not use the key trait, no discrimination was possible at all. To make strong inference that increased richness causes predators to use a key trait for decisions, it helps to include the choice to use either the key trait, or a specific, reliable trait that has more values (by values, we mean unique states or versions). This way, it is possible to determine if predators would actually switch their behavior if they did not have to.

We designed virtual prey communities where predators could either use a completely reliable trait that had many (2-8) different values to perfectly classify prey as ‘good’ or ‘bad’, or simplify decision-making by using an unreliable key trait (binary, with only two values) at the price of committing more errors (*Figure 1A*). In our virtual communities, prey always had two traits (color and shape – which one was reliable and which one was unreliable was randomized). Both traits were discrete, meaning that they could take on different values that did not grade continuously into one another. If the predator learned to identify good prey based on values of the reliable trait R , it could forage without errors. That is to say $P(\text{good}|R_i^+) = 1$, where + indicates that the value R_i is positively correlated with profitability (for example, circle is always good in *Figure 1A*). The unreliable trait U was binary, having values of U^+ and U^- (for example, prey each have one of two colors in *Figure 1A*). This binary trait only predicted whether an individual was ‘good’ with a probability of 0.78 (i.e. $P(\text{good}|U^+) = 0.78$). When only two values of each trait existed in the community (e.g. circle vs. star and blue vs. yellow; Experiment 1 in *Figure 1A*), the same number of individuals could be classified using either shape or color. However, using the unreliable trait (e.g. color) would carry the cost of committing more errors. Cognitive psychology experiments suggest that in this situation, the reliable trait will be used to the exclusion of the unreliable one due to a phenomenon called the relative validity effect (*Hall et al., 1977; Wagner et al., 1968*). In the relative validity effect, when an animal can learn to associate two cues with an outcome, it will learn to use the one that is more reliable (valid).

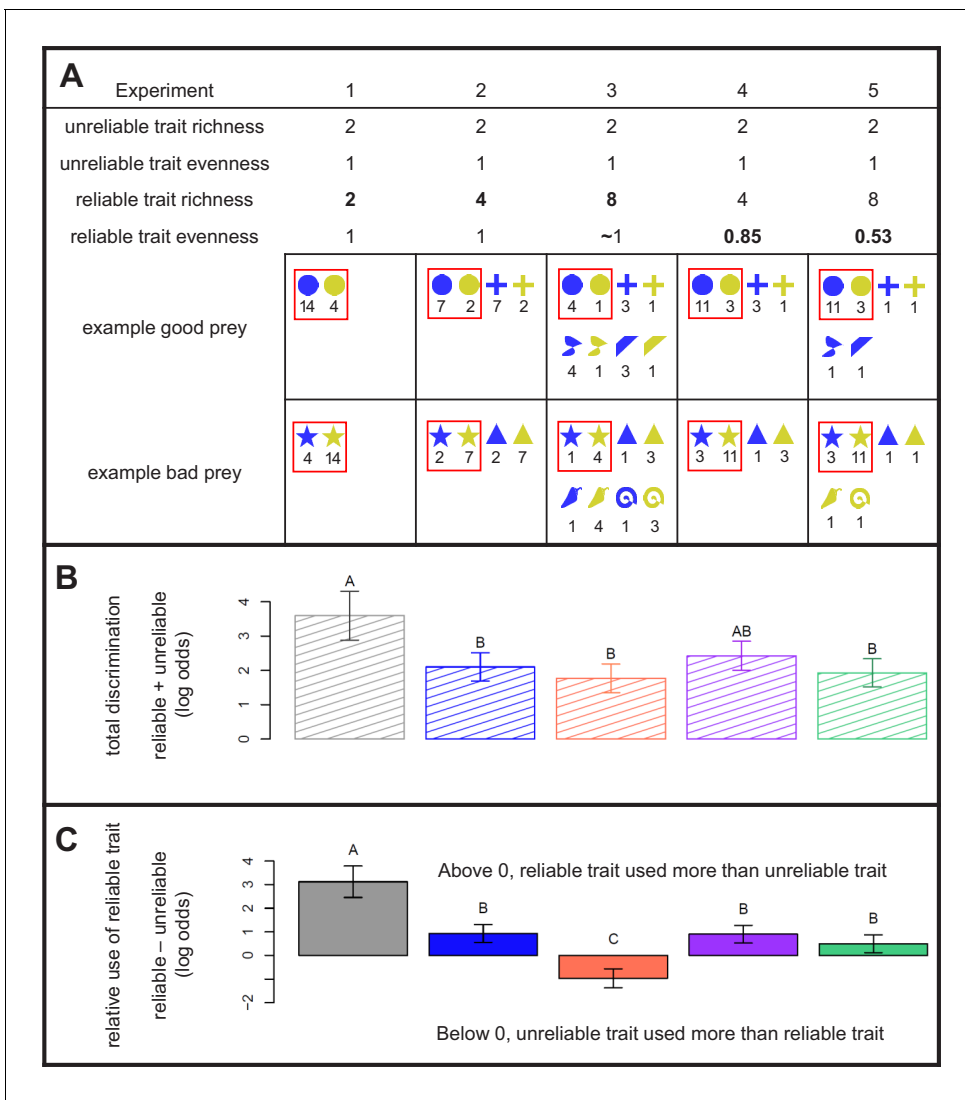


Figure 1. Design and results of Experiments 1 - 5. (A) Properties of the experimental prey communities used in this study, with examples. All communities had a 1:1 ratio of 'good' prey to 'bad' prey. A reliable trait allowed perfect discrimination. The richness and evenness of its values varied between experiments. An unreliable trait that did not vary between experiments yielded less accurate discrimination. The exact distribution of prey in each community is given below its richness and evenness statistics, with numbers to indicate the abundance of each prey. As drawn here, shape is the reliable trait (e.g. circle = good, star = bad), whereas color is the unreliable trait (blue = good 78% of the time, yellow = bad 78% of the time). Red boxes indicate the focal prey that were compared across experiments in panels B and C (their actual colors and shapes differed among treatments). (B) Total discrimination subjects exhibited towards focal prey, that is the summed influence of both reliable and unreliable traits. (C) Subjects' relative use of the reliable trait compared with the unreliable trait for discrimination, that is the difference between the effect of reliable and unreliable traits. The y-axis indicates the difference in the ability of the reliable trait to predict behavior compared to the unreliable trait. In (B) and (C), estimates are grouped using the Bonferroni correction for multiple pairwise comparisons, and 95% confidence intervals are shown. See Methods for details on interpreting log-odds.

DOI: <https://doi.org/10.7554/eLife.43965.002>

The following source data and figure supplements are available for figure 1:

Source data 1. Data used to generate **Figure 1** and its supplements.

DOI: <https://doi.org/10.7554/eLife.43965.005>

Figure supplement 1. Tabulated attack rates for prey of different types.

DOI: <https://doi.org/10.7554/eLife.43965.003>

Figure supplement 2. Tabulated attack rates for prey of different types.

DOI: <https://doi.org/10.7554/eLife.43965.004>

We recruited undergraduate student volunteers to serve as predators on our virtual prey communities. Each subject learned to forage on a grid of 36 prey during a training trial where they were allowed to attack up to 18 of the prey, and received feedback on whether each was ‘good’ or ‘bad’ to eat in the form of a smiley face with a chirp or an X with a gong sound. Their ‘life bar’ would also rise or fall accordingly (subjects lost twice as much life for attacking ‘bad’ prey than they gained for eating ‘good’ prey). After subjects finished the training trial, they took a test trial where they could choose as many prey as they liked, but received no feedback (**Figure 2**). The test trial served two purposes: 1) it allowed us to measure subjects’ categorization behavior without changing it by providing feedback, and 2) because the test trial was always the same, it allowed us to compare subjects’ categorization behavior after foraging in different training communities. Subjects participated in five experiments presented in random order (**Figure 2**; **Figure 2—figure supplement 1**). In Experiment 1, our control to see if the relative validity effect held with our design, we found that the reliable trait was used almost exclusively (**Supplementary file 1**).

We tested three mutually exclusive hypotheses for how predators will classify their prey as its phenotypic richness increases. Predators had to choose how much to rely on the key trait U at the price of committing some errors, or the completely reliable trait R at the price of learning about and memorizing multiple values. The first hypothesis was that they should select the former when the price of information (e.g. memory, costs of exploration) limits the profitability of using the reliable trait, so that as richness increases, they should use the unreliable trait to a greater degree (e.g. use color more and begin to ignore shape as the richness of shapes increases; **Figure 3A**). The second hypothesis was that if the relative validity effect were an invariant aspect of predator psychology, then predators should persist in using the reliable trait across different levels of richness (e.g., learn all of the shapes across Experiments 1–3, always ignoring color; **Figure 3B**). Indeed, associative learning experiments on the relative validity effect do not show a difference in which trait is used as the number of its values increases – that is, subjects always use the most reliable trait (**Baetu et al., 2005**; **Murphy et al., 2001**). However, the number of trait values in these experiments has been low. A third hypothesis is that at high levels of richness, predators may not be able to parse all of the information that they are confronted with and will guess randomly with respect to the reliable and unreliable traits (**Figure 3C**).

To test these hypotheses, we performed two experiments (2 and 3) that had higher richnesses than Experiment 1. Subjects used the reliable and/or unreliable traits to make decisions in all experiments (**Figure 1B**), which allowed us to reject the hypothesis that they would not use either trait at high diversities. In Experiment 2, where there were four values of the reliable trait (two associated with profitability, two with unprofitability), subjects decreased their use of the reliable trait

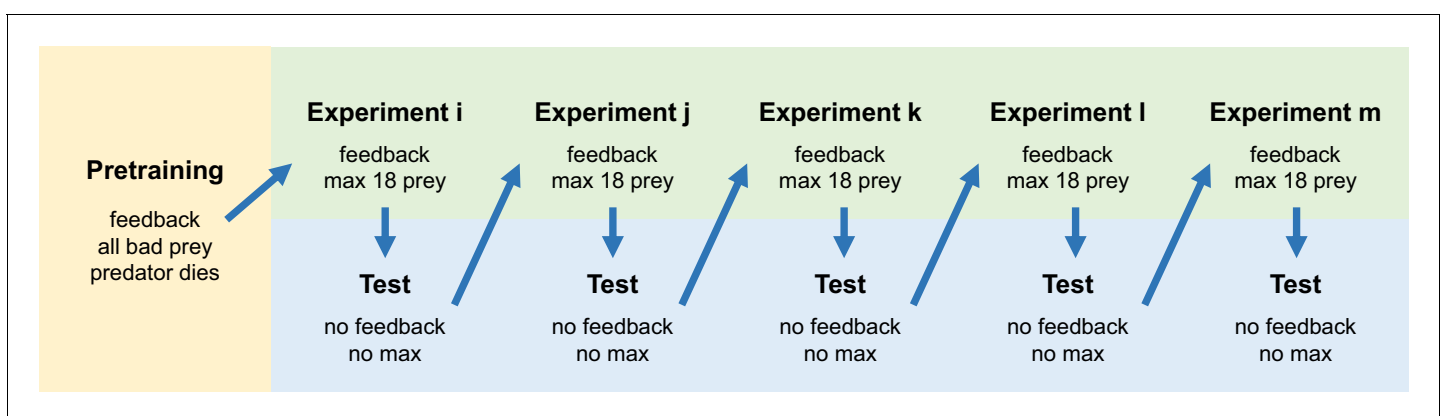


Figure 2. General experimental procedures. Subjects were randomly assigned to one of four different treatments within each of five experiments. Between treatments, colors and shapes were shuffled with respect to ‘good’ and ‘bad’ prey to prevent subjects from generalizing across experiments.

DOI: <https://doi.org/10.7554/eLife.43965.006>

The following figure supplement is available for figure 2:

Figure supplement 1. All experimental treatments used in this study.

DOI: <https://doi.org/10.7554/eLife.43965.007>

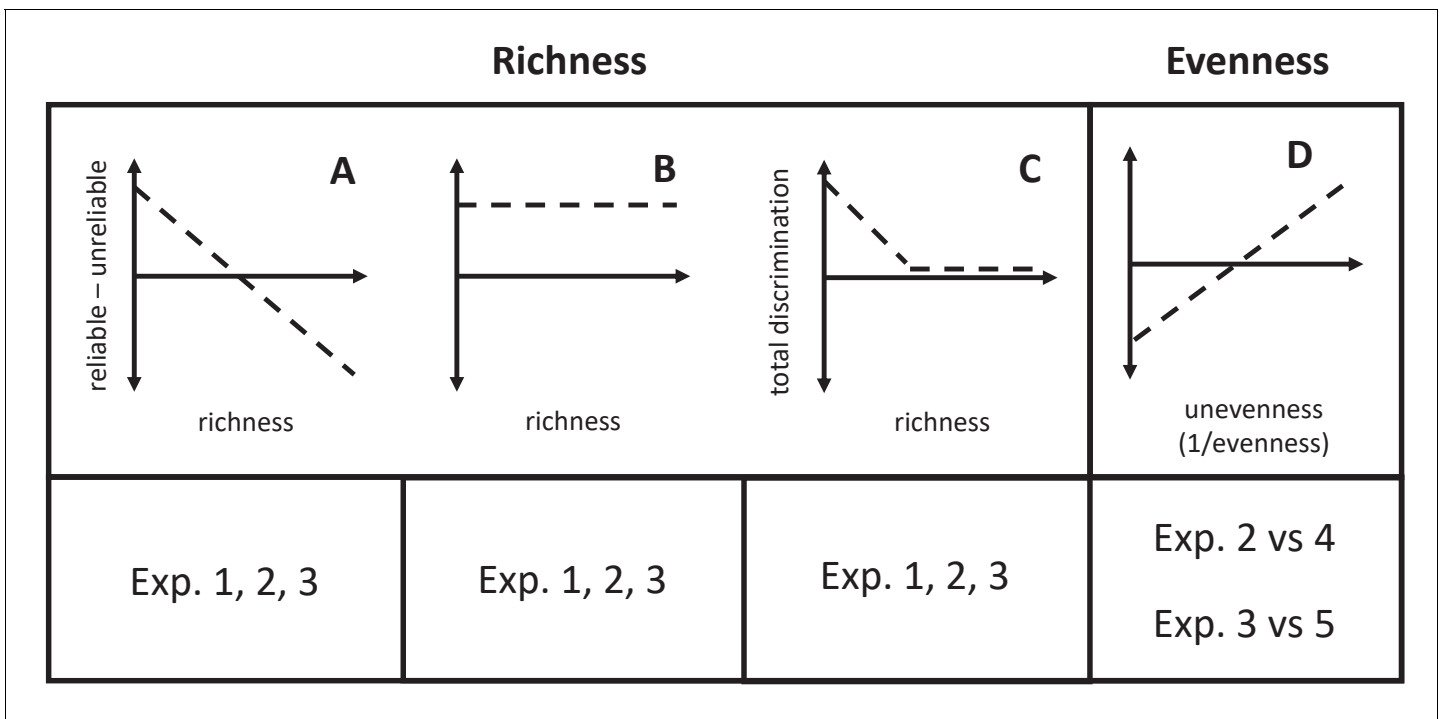


Figure 3. Graphical predictions of hypotheses described in the text. (A) The relative use of the reliable trait will decrease if increased prey richness causes predators to prefer the unreliable trait. (B) If the relative validity effect is robust to changes in richness, predators will always use the reliable trait. (C) If predators cannot process all of the information available in diverse communities, they will guess randomly. (D) If the reduced effective richness of prey in uneven communities reduces the costs of information, then use of the reliable trait will increase.

DOI: <https://doi.org/10.7554/eLife.43965.008>

significantly (**Figure 1C**). In Experiment 3, where there were eight values of the reliable trait, subjects again significantly decreased their use of the reliable trait - in fact, they used the unreliable trait more (**Figure 1C**). These results allow us to reject the hypothesis that the relative validity effect is constant across levels of richness. Instead, they support the hypothesis that species-rich communities carry a high price of information, either in memory constraints or the risks of acquiring information, which increases predators' tendency to use the simpler (binary) yet unreliable trait in decision-making. This is strong evidence supporting the hypothesis that in rich communities, mimicry could evolve easily on the basis of key features that predators use for identification (**Beatty et al., 2004**). It also supports important theoretical models of how mimicry evolves that depend upon key features (**Balogh et al., 2010; Balogh and Leimar, 2005; Gamberale-Stille et al., 2012**). Furthermore, this result is critical to the stability of warning signals that are parasitized by Batesian mimics because it implies that predators will not immediately switch to using more reliable traits simply because they are available. High species richness could still favor broad categorization.

The other component of community diversity is species evenness, or relative abundance (**Magurran, 1988; Tuomisto, 2012**). In every community, some species are common, while other species are rare. This 'overrepresentation' of some species and underrepresentation of others reduces the effective number of species in the community (i.e., lower evenness means effectively fewer species; **Jost, 2010**). Consequently, it is reasonable to hypothesize that unevenness will decrease predators' use of the key trait to form categories, reversing the effect of increasing richness. In fact, prior work has shown that both higher frequencies of profitable, non-mimetic prey and lower frequencies of Batesian mimics relaxes selection on mimicry (**Finkbeiner et al., 2018; Harper and Pfennig, 2007; Iserbyt et al., 2011; Lindström et al., 2004; Lindström et al., 1997; Pfennig et al., 2001**). Here, we ask more generally about the effects of evenness per se, where there is a distribution of relative abundance within both profitable and unprofitable prey.

In uneven communities, rarer prey types will be less important food resources. This led us to test the hypothesis that unevenness will decrease predators' tendency to categorize prey using an

unreliable key trait. It predicts that in uneven communities, predators will use the reliable trait more than in an evenly distributed community with the same phenotypic richness (**Figure 3D**).

To test this hypothesis, we conducted two experiments with uneven communities. Experiment 4 featured the same eight phenotypes as Experiment 2, and Experiment 5 involved the precisely the same shapes and colors as Experiment 3, but in Experiments 4 and 5, one of the focal 'good' values and one of the focal 'bad' values of prey were much more abundant than the others. The reliable trait was used to a significantly greater degree in Experiment 5 compared to Experiment 3, as evinced by their placement in different *post-hoc* groupings (**Figure 1C**). However, we observed no difference in the use of the reliable trait between Experiments 2 and 4 (**Figure 1C**).

The hypothesis was rejected; unevenness increased subjects' use of the reliable trait, but only at high levels of richness. We attribute this outcome to a simple cause. In both Experiments 4 and 5, where communities were uneven, predators mainly distinguished between the most abundant profitable value of the reliable trait, and all others. In other words, predators mainly attacked prey with the most abundant 'good' value, and categorized the other values as not worth attacking. Effectively, the most abundant, good, reliable value became a preferred alternative prey. In support of this interpretation, a model that included two different values for 'good' prey and two different values for 'bad' prey within the reliable trait R fit significantly better than one that lumped 'good' values together, and 'bad' values together (likelihood ratio test, $\chi^2_{24} = 70.1$, $P < 0.001$). Predators attacked the abundant good prey much more than any other kind of prey in Experiment 5 and exhibited this behavior to a lesser degree in Experiment 4 (**Figure 1—figure supplement 1**).

Results from the uneven communities contrast with the pattern from Experiments 1–3, which shows increasing reliance on the unreliable trait with increasing richness. We suggest that the unevenness of a prey community will be negatively correlated with predators' reliance upon key traits to form categories. This may make mimicry less likely to evolve in uneven communities. Furthermore, it connects evenness, a fundamental parameter of community ecology, to the concept of alternative prey from mimicry theory: when one species of profitable, non-mimetic prey is relatively abundant, selection on other, rarer prey to evolve mimicry will be relaxed (**Getty, 1985; Holling, 1965; Ihalainen et al., 2012; Kokko et al., 2003; Lindström et al., 2004**).

Very few studies from natural systems have collated the data that would be required to measure the relationship between community diversity and signaling systems. **Wilson et al. (2013)** argued that a negative relationship between mimetic precision and community diversity stems from increased generalization by predators in more diverse communities of velvet ants. Additionally, in experimentally manipulated communities of flowering plants, increased color diversity tended to increase visitation rates by pollinating insects (**Fornoff et al., 2017**). It would be interesting to know if this occurred because individual pollinators relied on coarser phenotypic categories in richer communities.

Our hypotheses might also be applicable to subsets of communities. Particularly, specialist predators might experience smaller prey communities than generalist predators, and specialist pollinators might visit fewer species of flowers than generalist pollinators. For example, different mimicry rings of *Heliconius* butterflies are segregated by microhabitat (**Elias et al., 2008**), and exposed to different suites of predators as a consequence; predation favors precise mimics within their preferred microhabitats (**Willmott et al., 2017**). Habitat specialization could reduce the size of the community about which a particular bird must learn, allowing them to select for precise mimicry (or none at all) because coarse categorization based on key traits would not occur.

Theoretical models suggest that other ecological conditions than those we explored here can also affect the number of traits that predators use to make decisions. Under some circumstances when the costs of attacking 'bad' prey are in a particular balance with the benefits of attacking 'good' prey, trusting only the most reliable trait may be most adaptive (**Rubi and Stephens, 2016**). Yet changing the cost:benefit ratio or underlying frequency of good prey can favor using multiple traits, or using no trait at all (**Sherratt and Hohen, 2018**).

Conclusions

Both prey richness and evenness affected predator categorization behavior. The result that richness will favor mimicry because predators use a key trait in categorization, even at the cost of decisional

accuracy, was not anticipated by literature on associative learning. Generally, our results show that ecological diversity predicts both the origins and maintenance of mimicry.

Our results are relevant to problems in signal evolution beyond defensive mimicry, for all communication systems evolve within the context of communities. Floral phenotypes have long been remarked to appear similar to one another, presumably to signal to shared pollinators (Ollerton *et al.*, 2009; Schaefer and Ruxton, 2010; Schaefer and Ruxton, 2009; Shrestha *et al.*, 2013). It follows from our study that in richer communities, flowers could benefit from sharing key traits that allow pollinators to categorize them together. Likewise, other signals between species such as pursuit deterrence (Caro, 1995), mobbing calls (Magrath *et al.*, 2015), herbivore-induced plant volatiles (Schuman and Baldwin, 2016), and ripening fruit (Willson and Whelan, 1990) may also evolve to share key traits in rich, even communities. Very few studies have collated the data that would be required to measure the relationship between community context and signaling systems. Yet clearly, communities have the potential to produce striking effects on signal evolution.

Materials and methods

Experiments

Human volunteers can readily be recruited to participate in short computer games that are completely harmless and yield large quantities of data. Experiments of similar design performed by other animals and humans have yielded qualitatively similar results (Alatalo and Mappes, 1996 vs. Beatty *et al.*, 2005; Cuthill *et al.*, 2005 vs. Fraser *et al.*, 2007; Kazemi *et al.*, 2014 vs. Sherratt *et al.*, 2015), although humans often learn faster. A major exception to this is XOR tasks (also called correlated features tasks), where no single dimension is helpful for categorization, but their combinations are. Humans rank XOR tasks as easier than several other kinds of tasks, but other primates rank them as more difficult (Smith *et al.*, 2004). Differences between humans and non-humans in this task are thought to be mediated by language (Smith *et al.*, 2011). Importantly, our experiments did not include XOR tasks, so human behavior is more likely to be representative of non-human species.

We designed a computer game in Psychtoolbox-3 for Matlab R2017b where subjects were asked to hunt for artificial prey (Brainard, 1997; Kleiner *et al.*, 2007). The general format of the game was presented to subjects during the pre-training trial, which only occurred once at the very beginning of the game. Subjects saw a grid of 36 prey that constituted a prey community. They were told that they could sample as many as 18 of those prey (half), or terminate sampling prematurely. However, in the pretraining trial they did not have the option to stop sampling early. This forced them to participate long enough to understand the basics. When they attacked a 'bad' prey, it was immediately covered with an X while a gong noise played and the screen froze for two seconds. In addition, a life bar on the side of the screen would decrease. If subjects attacked a 'good' prey, it would be covered with a smiley face while a chirp sounded, no time penalty occurred, and the life bar immediately increased. The amount of life lost from attacking 'bad' prey was twice that gained from attacking 'good' prey, but the life bar did not change over time unless subjects attacked, that is they had no penalty for moving slowly or quickly. In the pretraining trial, all prey were 'bad', and the subject's life bar was set low to start so that they would die after attacking four prey. When they died, a lonesome whistle sounded and the screen froze for two seconds. We arranged the pretraining trial in this way so that all subjects would begin the experiments from the same starting point, and be more likely to pay attention to avoiding 'bad' prey in addition to finding 'good' prey.

Once subjects had completed pretraining, they were told that they would now proceed to the real game, which was the first experimental training trial. Subjects were randomly assigned to an experiment, and within that experiment, randomly assigned to a treatment. Treatments were designed so that within each experiment, shape would be the reliable trait in two treatments, and color the reliable trait in the other two. Which colors and shapes were associated with 'good' and 'bad' prey were permuted within the treatments in which they were the reliable trait. The full array of treatments is presented in **Figure 2—figure supplement 1**. Subjects began the game with their life bars at half full so that they could encounter eight 'bad' prey in a row before dying. This was intended to give them motivation to forage yet made it unlikely that they would actually 'die'. If they did die, the training trial ended, but the flow of the game did not otherwise change.

Once subjects completed the training trial, they were told that they had the opportunity to play a bonus round (the test trial) to increase their score. They were told that they could attack as many prey as they liked in the bonus round, but that they could stop whenever they wanted to. They were also told that they would receive absolutely no feedback until the bonus round was over. We designed the test trial without feedback so that subjects would not continue to learn (and hence change the categories they had formed) during the test trial. The prey in the test trial were always the same no matter what experimental treatment subjects experienced. These test prey always included prey with four values of the reliable trait (two good, two bad). After subjects finished the test trial, they were told their total score across both rounds. This was calculated as the sum of all 'good' prey attacked less all 'bad' prey attacked, but subjects were not informed of this formula – the only purpose of telling them these scores was to keep them motivated in both the training and test trials. By making it difficult to tell exactly how the score was calculated, though, we encouraged them to focus on their life bar during training trials, which continued to fluctuate with a cost:benefit ratio of 2:1 for bad:good prey.

Subjects experienced each of the five experiments in random order, taking the test trial immediately after completing each one (**Figure 2**). The only exceptions to this were a few subjects that completed fewer than five experiments to improve the balance of our design. A pseudo-random design might have made this unnecessary, but we did not want to unintentionally induce any bias in the order in which treatments were presented.

When subjects had completed all five experiments and the respective test trials, they were asked to take a color blindness test (Ishihara plates 6, 8, 13, and 23). This did not constitute a medically professional diagnosis of color blindness, so they were not informed of their results, but any subject failing the test was excluded from the final dataset. In total, we recruited 45 volunteers who passed this basic test from the Carleton University Student Union in Ottawa, Canada.

Data analysis

We designed our analysis of the test trial to answer the question of how well the reliable trait predicted subjects' behavior in each experiment, relative to the unreliable trait. This was critical to testing the predictions of the first two hypotheses about species richness (**Figure 3A & B**), and the hypothesis about evenness (**Figure 3D**).

There were three steps to this analysis: the first was to find, in each experiment, the estimated effects of the reliable and unreliable traits on subjects' decisions to attack or reject prey. We estimated these effects with a statistical model of subjects' decisions. Second, we found the *difference* between the effects of the reliable and unreliable traits in each experiment. Third, we performed pairwise comparisons of these differences between experiments. This required finding the differences of differences. Both of these difference calculations used parameter estimates that we obtained from the statistical model. We describe our methods below, and also refer readers to the RMark-down in **Supplementary file 1**.

It is necessary to understand the structure of the data. We analyzed subjects' attacks on focal prey, defined as the four most abundant prey present in the training trial (e.g. red boxes in **Figure 1A**). The focal prey were also always present in the test trial. The test trial included some trait values that were not present in the training – we eliminated them from analysis. To describe the effects of the reliable trait, which differed in its number of values between experiments, we recoded the values of the focal prey according to whether they were associated with profitability or unprofitability during training. That is to say, we combined R_i^+ and R_i^- into just two values, R^+ and R^- . Taking an example from **Figure 1A**, circle and cross in Experiments 2 – 5 were recoded as 'G', and star and triangle were recoded as 'B'. The end result of recoding was that the reliable and unreliable traits could be analyzed as factors with only two values, making them both binary (**Supplementary file 1**). This made their relative contributions easy to compare by simply looking at their effect sizes once they had been centered and incorporated into a suitable statistical model (**Schielzeth, 2010**).

We fit a model that was designed to find the effects of the reliable and unreliable traits and their standard errors, rather than to fit our data as well as possible. To do this, we fit the model without an intercept, and without main effects of the reliable trait and unreliable trait. This violates the principle of marginality, but our aim was not to test hypotheses with the model. Excluding the intercept and two main effects allowed us to directly find the effects of interest, instead of having to perform additional calculations (**Schielzeth, 2010**). We took into account potential confounding variables in

constructing our model, however, as they could affect estimates of the effects. Therefore, we included interactions with order, once it had been centered. Centering causes the estimates of lower-order terms to be made at the mean value of a variable, so that lower-order terms can be interpreted independently of interactions (*Schielzeth, 2010*). We also included subject identity as a main effect to control for variation among individuals (models that included it as a random effect did not converge). In R pseudocode, the model that we fit is shown below, with the specific quantities of interest in bold:

```
glm((attacked, not attacked) ~ 0 + experiment + subject ID + unreliable trait:
experiment + reliable trait:experiment + unreliable trait:experiment:order +
reliable trait:experiment:order
```

with a logit link function. All data are available in **Figure 1—source data 1**.

The model provided the estimates of the effects of the reliable and unreliable traits in each experiment, with standard errors. We used these estimates to find their differences. The difference was appropriate because log-odds were the units for effect size estimates from the model (since it was fit with a logit link). Log-odds are an ideal metric of trait importance to decision-making because they are easily converted into the odds of attack for different kinds of prey. If β_{R1} is the log-odds estimate for $P(\text{attack}|R^+) - P(\text{attack}|R^-)$ and β_{U1} is the log-odds estimate for $P(\text{attack}|U^+) - P(\text{attack}|U^-)$ in Experiment 1, then $\exp(\beta_{R1} - \beta_{U1})$ gives the relative difference odds of attack due to the reliable trait compared to the unreliable trait in Experiment 1. For example, in Experiment 1 the effect size of the 'good' value of the reliable trait is 3.36, and the corresponding estimate for the unreliable trait is 0.24. This means R^+ prey have $e^{3.36} = 28.8$ times the odds of attack compared with prey that are R^- , but U^+ prey only suffer an increase in odds of attack of $e^{0.24} = 1.27$ compared to U^- prey. Finding their difference as $3.36 - 0.24 = 3.12$ means that in Experiment 1, R^+ prey have $e^{3.12} = 22.6$ times the attack risk of U^+ prey.

A function to find the difference in the effect of the reliable and unreliable traits in each experiment is simple subtraction (e.g., $R_1 - U_1$). However, finding the standard errors of the difference is more complicated. The delta method is one way of approximating the standard error of a function of estimated effects (*Bolker, 2008*). To implement the delta method, we used the function `deltaMethod` from the 'car' package in R 3.4.4 (see **Supplementary file 1; Fox et al., 2018**). This completed the first difference calculation.

We again used the delta method to perform pairwise comparisons between the relative importance of traits between experiments, using a Bonferroni correction for multiple comparisons. This allowed us to answer our question of how the use of the reliable versus unreliable traits changed across experiments. The statistical significance of our results depended on whether or not numerically calculated confidence intervals for the difference between estimates included zero or not, which is displayed in the groupings in **Figure 1B and C**.

We repeated the operations described above to find the difference of the *sums* of the effects of both traits, which is a way of describing the total discrimination of subjects in an experiment. We did this to test the prediction that high species richness results in random guessing (**Figure 3C**).

Finally, we tested the prediction that in uneven communities, predators would focus on the most abundant good prey. We modeled subjects' attack decisions just as we did above, but instead of using combined 'good' and 'bad' values of the reliable trait, we used the original focal prey values, for example circle, cross, star, triangle. Thus, there were up to two 'good' and two 'bad' values of the reliable trait. We used the likelihood ratio test to compare this model to the one we fit above. A significant difference in model fit would mean that subjects treated different values of good and/or bad prey differently, which is the qualitative pattern we observed in Experiments 4 and 5 (**Figure 1—figure supplement 2**). Note that although both models violated marginality, this does not matter for the comparison of fit that we performed between them.

Acknowledgements

We thank members of the Dornhaus and Sherratt labs for helpful comments and support. We thank members of Bob Wilson's lab for advice on coding in Matlab. We thank our reviewers, including Bernhard Schmid, for very helpful comments. Human subjects research was carried out with the permission of the Carleton University Research Ethics Board-B under permit number 13385 14-0276. Funding: DWK was funded by NIH-K12GM000708. TNS was funded by an NSERC Discovery Grant.

Additional information

Funding

Funder	Grant reference number	Author
National Institutes of Health	K12GM000708	David William Kikuchi
Natural Sciences and Engineering Research Council of Canada		Thomas N Sherratt

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

David William Kikuchi, Conceptualization, Software, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Anna Dornhaus, Writing—review and editing; Vandana Gopeechund, Data curation; Thomas N Sherratt, Conceptualization, Writing—review and editing

Author ORCIDs

David William Kikuchi  <http://orcid.org/0000-0002-7379-2788>

Ethics

Human subjects: Consent process is described in the Methods. Human subjects research was carried out with the permission of the Carleton University Research Ethics Board-B under permit number 13385 14-0276.

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.43965.012>

Author response <https://doi.org/10.7554/eLife.43965.013>

Additional files

Supplementary files

- Supplementary file 1. RMarkdown with full analysis, including code to reproduce all results. The file includes a legend to the columns of the **Figure 1—source data 1** with detailed explanation of the variable codings. It is preferable to use this file as a guide to the data.

DOI: <https://doi.org/10.7554/eLife.43965.009>

- Transparent reporting form

DOI: <https://doi.org/10.7554/eLife.43965.010>

Data availability

All data for this study are present in the supporting files, and source code to produce the figures from those files is included in the Supplementary RMarkdown file.

References

- Alatalo RV, Mappes J. 1996. Tracking the evolution of warning signals. *Nature* **382**:708–710. DOI: <https://doi.org/10.1038/382708a0>
- Baetu I, Baker AG, Darredeau C, Murphy RA. 2005. A comparative approach to cue competition with one and two strong predictors. *Animal Learning & Behavior* **33**:160–171. DOI: <https://doi.org/10.3758/BF03196060>
- Balogh AC, Gamberale-Stille G, Tullberg BS, Leimar O. 2010. Feature theory and the two-step hypothesis of müllerian mimicry evolution. *Evolution* **64**:810–822. DOI: <https://doi.org/10.1111/j.1558-5646.2009.00852.x>
- Balogh ACV, Leimar O. 2005. Müllerian mimicry: an examination of Fisher's theory of gradual evolutionary change. *Proceedings of the Royal Society B: Biological Sciences* **272**:2269–2275. DOI: <https://doi.org/10.1098/rspb.2005.3227>

- Beatty CD, Beirincx K, Sherratt TN. 2004. The evolution of müllerian mimicry in multispecies communities. *Nature* **431**:63–66. DOI: <https://doi.org/10.1038/nature02818>, PMID: 15343332
- Beatty CD, Bain RS, Sherratt TN. 2005. The evolution of aggregation in profitable and unprofitable prey. *Animal Behaviour* **70**:199–208. DOI: <https://doi.org/10.1016/j.anbehav.2004.09.023>
- Bolker BM. 2008. *Ecological Models and Data in R*. Princeton, NJ: Princeton University Press.
- Bradbury JW, Vehrencamp SL. 2011. *Principles of Animal Communication*. 2nd ed. Sunderland: Sinauer Associates, Inc.
- Brainard DH. 1997. The psychophysics toolbox. *Spatial Vision* **10**:433–436. DOI: <https://doi.org/10.1163/156856897X00357>, PMID: 9176952
- Caro TM. 1995. Pursuit-deterrence revisited. *Trends in Ecology & Evolution* **10**:500–503. DOI: [https://doi.org/10.1016/S0169-5347\(00\)89207-1](https://doi.org/10.1016/S0169-5347(00)89207-1), PMID: 21237125
- Clarke CA, Sheppard PM. 1963. Interactions between major genes and polygenes in the determination of the mimetic patterns of *Papilio dardanus*. *Evolution* **17**:404–413. DOI: <https://doi.org/10.1111/j.1558-5646.1963.tb03297.x>
- Cohen JD, McClure SM, Yu AJ. 2007. Should I stay or should I go? how the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**:933–942. DOI: <https://doi.org/10.1098/rstb.2007.2098>
- Cuthill IC, Stevens M, Sheppard J, Maddocks T, Párraga CA, Troscianko TS. 2005. Disruptive coloration and background pattern matching. *Nature* **434**:72–74. DOI: <https://doi.org/10.1038/nature03312>, PMID: 15744301
- Cuthill IC, TD B. 1993. Mimicry and the eye of the beholder. *Proc R Soc B* **253**:203–204.
- Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley AC, Nadeau NJ, Zimin A, Hughes DST, Ferguson LC, Martin SH, Salazar C, Lewis JJ, Adler S, Ahn S-J, Baker Da, Baxter SW, Chamberlain NL, Chauhan R, Counterman Ba, Dalmay T, et al. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**:94–98. DOI: <https://doi.org/10.1038/nature11041>, PMID: 22722851
- Dittrich W, Gilbert F, Green P, McGregor P, Grewcock D. 1993. Imperfect mimicry: a pigeon's perspective. *Proc R Soc B* **251**:195–200.
- Elias M, Gompert Z, Jiggins C, Willmott K. 2008. Mutualistic interactions drive ecological niche convergence in a diverse butterfly community. *PLOS Biology* **6**:e300–e309. DOI: <https://doi.org/10.1371/journal.pbio.0060300>
- Endler JA, Basolo AL. 1998. Sensory ecology, receiver biases and sexual selection. *Trends in Ecology & Evolution* **13**:415–420. DOI: [https://doi.org/10.1016/S0169-5347\(98\)01471-2](https://doi.org/10.1016/S0169-5347(98)01471-2), PMID: 21238370
- Finkbeiner SD, Salazar PA, Nogales S, Rush CE, Briscoe AD, Hill RI, Kronforst MR, Willmott KR, Mullen SP. 2018. Frequency dependence shapes the adaptive landscape of imperfect batesian mimicry. *Proceedings of the Royal Society B: Biological Sciences* **285**:20172786. DOI: <https://doi.org/10.1098/rspb.2017.2786>
- Fornoff F, Klein A-M, Hartig F, Benadi G, Venjakob C, Schaefer HM, Ebeling A. 2017. Functional flower traits and their diversity drive pollinator visitation. *Oikos* **126**:1020–1030. DOI: <https://doi.org/10.1111/oik.03869>
- Fox J, Weisberg S, Price B, Adler D, Bates D, Baud-bovy G, Bolker B, Ellison S, Firth D, Friendly M, Graves S, Heiberger R, Laboissiere R, Maechler M, Monette G, Murdoch D, Ogle D, Ripley B, Venables W, Walker S, et al. 2018. Companion to Applied Regression. Package 'car'. 3.4.4. <https://cran.r-project.org/web/packages/car/car.pdf>
- Fraser S, Callahan A, Klassen D, Sherratt TN. 2007. Empirical tests of the role of disruptive coloration in reducing detectability. *Proceedings of the Royal Society B: Biological Sciences* **274**:1325–1331. DOI: <https://doi.org/10.1098/rspb.2007.0153>
- Gamberale-Stille G, Balogh ACV, Tullberg BS, Leimar O. 2012. Feature saltation and the evolution of mimicry. *Evolution* **66**:807–817. DOI: <https://doi.org/10.1111/j.1558-5646.2011.01482.x>
- Getty T. 1985. Discriminability and the sigmoid functional response: how optimal foragers could stabilize Model-Mimic complexes. *The American Naturalist* **125**:239–256. DOI: <https://doi.org/10.1086/284339>
- Guilford T, Dawkins MS. 1993. Receiver psychology and the design of animal signals. *Trends in Neurosciences* **16**:430–436. DOI: [https://doi.org/10.1016/0166-2236\(93\)90068-W](https://doi.org/10.1016/0166-2236(93)90068-W)
- Hall G, Mackintosh NJ, Goodall G, Martello MD. 1977. Loss of control by a less valid or by a less salient stimulus compounded with a better predictor of reinforcement. *Learning and Motivation* **8**:145–158. DOI: [https://doi.org/10.1016/0023-9690\(77\)90001-7](https://doi.org/10.1016/0023-9690(77)90001-7)
- Harper GR, Pfennig DW. 2007. Mimicry on the edge: why do mimics vary in resemblance to their model in different parts of their geographical range? *Proceedings of the Royal Society B: Biological Sciences* **274**:1955–1961. DOI: <https://doi.org/10.1098/rspb.2007.0558>
- Hebets EA, Papaj DR. 2005. Complex signal function: developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology* **57**:197–214. DOI: <https://doi.org/10.1007/s00265-004-0865-7>
- Holling CS. 1965. The functional response of predators to prey density and its role in mimicry and population regulation. *Memoirs of the Entomological Society of Canada* **97**:5–60. DOI: <https://doi.org/10.4039/entm9745fv>
- Houston AI, Trimmer PC, Fawcett TW, Higginson AD, Marshall JA, McNamara JM. 2012. Is optimism optimal? functional causes of apparent behavioural biases. *Behavioural Processes* **89**:172–178. DOI: <https://doi.org/10.1016/j.beproc.2011.10.015>, PMID: 22085791
- Ihalainen E, Rowland HM, Speed MP, Ruxton GD, Mappes J. 2012. Prey community structure affects how predators select for müllerian mimicry. *Proceedings of the Royal Society B: Biological Sciences* **279**:2099–2105. DOI: <https://doi.org/10.1098/rspb.2011.2360>

- Iserbyt A, Bots J, Van Dongen S, Ting JJ, Van Gossum H, Sherratt TN. 2011. Frequency-dependent variation in Mimetic fidelity in an intraspecific mimicry system. *Proceedings of the Royal Society B: Biological Sciences* **278**: 3116–3122. DOI: <https://doi.org/10.1098/rspb.2011.0126>
- Jiggins CD. 2017. *The Ecology and Evolution of Heliconius Butterflies*. Oxford: Oxford University Press.
- Jost L. 2010. The relation between evenness and diversity. *Diversity* **2**:207–232. DOI: <https://doi.org/10.3390/d2020207>
- Kazemi B, Gamberale-Stille G, Tullberg BS, Leimar O. 2014. Stimulus salience as an explanation for imperfect mimicry. *Current Biology* **24**:965–969. DOI: <https://doi.org/10.1016/j.cub.2014.02.061>, PMID: 24726157
- Kikuchi DW, Pfennig DW. 2013. Imperfect mimicry and the limits of natural selection. *The Quarterly Review of Biology* **88**:297–315. DOI: <https://doi.org/10.1086/673758>, PMID: 24552099
- Kikuchi DW, Sherratt TN. 2015. Costs of learning and the evolution of mimetic signals. *The American Naturalist* **186**:321–332. DOI: <https://doi.org/10.1086/682371>, PMID: 26655351
- Kleiner M, Brainard DH, Pelli DG. 2007. What's new in Psychtoolbox-3? *Perception* **36**:1–16.
- Kokko H, Mappes J, Lindström L. 2003. Alternative prey can change model-mimic dynamics between parasitism and mutualism. *Ecology Letters* **6**:1068–1076. DOI: <https://doi.org/10.1046/j.1461-0248.2003.00532.x>
- Kunte K. 2009. The diversity and evolution of batesian mimicry in papilio swallowtail butterflies. *Evolution* **63**: 2707–2716. DOI: <https://doi.org/10.1111/j.1558-5646.2009.00752.x>
- Lindström L, Alatalo RV, Mappes J. 1997. Imperfect batesian mimicry—the effects of the frequency and the distastefulness of the model. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **264**: 149–153. DOI: <https://doi.org/10.1098/rspb.1997.0022>
- Lindström L, Alatalo RV, Lyytinen A, Mappes J. 2004. The effect of alternative prey on the dynamics of imperfect batesian and müllerian mimics. *Evolution* **58**:1294–1302. DOI: <https://doi.org/10.1111/j.0014-3820.2004.tb01708.x>
- MacDougall A, Dawkins MS. 1998. Predator discrimination error and the benefits of müllerian mimicry. *Animal Behaviour* **55**:1281–1288. DOI: <https://doi.org/10.1006/anbe.1997.0702>, PMID: 9632511
- Magrath RD, Haff TM, Fallow PM, Radford AN. 2015. Eavesdropping on heterospecific alarm calls: from mechanisms to consequences. *Biological Reviews* **90**:560–586. DOI: <https://doi.org/10.1111/brv.12122>, PMID: 24917385
- Magurran. 1988. *Ecological Diversity and Its Measurement*. Princeton: Princeton University Press.
- Mallet J. 2001. Mimicry: an interface between psychology and evolution. *PNAS* **98**:8928–8930. DOI: <https://doi.org/10.1073/pnas.171326298>, PMID: 11481461
- Maynard Smith J, Harper D. 2003. *Animal Signals*. Oxford: Oxford University Press.
- Murphy RA, Baker AG, Fouquet N. 2001. Relative validity effects with either one or two more valid cues in pavlovian and instrumental conditioning. *Journal of Experimental Psychology: Animal Behavior Processes* **27**: 59–67. DOI: <https://doi.org/10.1037/0097-7403.27.1.59>
- Oaten A, Pearce CE, Smyth ME. 1975. Batesian mimicry and signal detection theory. *Bulletin of Mathematical Biology* **37**:367–387. DOI: <https://doi.org/10.1007/BF02459520>, PMID: 1182321
- Ollerton J, Alarcón R, Waser NM, Price MV, Watts S, Cranmer L, Hingston A, Peter CI, Rotenberry J. 2009. A global test of the pollination syndrome hypothesis. *Annals of Botany* **103**:1471–1480. DOI: <https://doi.org/10.1093/aob/mcp031>, PMID: 19218577
- Pfennig DW, Harcombe WR, Pfennig KS. 2001. Frequency-dependent batesian mimicry. *Nature* **410**:323. DOI: <https://doi.org/10.1038/35066628>, PMID: 11268195
- Rowe C. 2013. Receiver psychology: a receiver's perspective. *Animal Behaviour* **85**:517–523. DOI: <https://doi.org/10.1016/j.anbehav.2013.01.004>
- Rubi TL, Stephens DW. 2016. Should receivers follow multiple signal components? an economic perspective. *Behavioral Ecology* **27**:36–44. DOI: <https://doi.org/10.1093/beheco/arv121>
- Ruxton GD, Allen WL, Sherratt TN, Speed MP. 2018. *Avoiding Attack*. 2nd ed. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oso/9780199688678.001.0001>
- Ryan MJ, Fox JH, Wilczynski W, Rand AS. 1990. Sexual selection for sensory exploitation in the frog physalaemus pustulosus. *Nature* **343**:66–67. DOI: <https://doi.org/10.1038/343066a0>, PMID: 2296291
- Schaefer HM, Ruxton GD. 2009. Deception in plants: mimicry or perceptual exploitation? *Trends in Ecology & Evolution* **24**:676–685. DOI: <https://doi.org/10.1016/j.tree.2009.06.006>, PMID: 19683828
- Schaefer HM, Ruxton GD. 2010. Communication theory and the form of receiver-mediated selection. *Trends in Ecology & Evolution* **25**:383–384. DOI: <https://doi.org/10.1016/j.tree.2010.04.003>
- Schielzeth H. 2010. Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**:103–113. DOI: <https://doi.org/10.1111/j.2041-210X.2010.00012.x>
- Schuman MC, Baldwin IT. 2016. The layers of plant responses to insect herbivores. *Annual Review of Entomology* **61**:373–394. DOI: <https://doi.org/10.1146/annurev-ento-010715-023851>, PMID: 26651543
- Seeger CA, Peterson EJ. 2013. Categorization = decision making + generalization. *Neuroscience & Biobehavioral Reviews* **37**:1187–1200. DOI: <https://doi.org/10.1016/j.neubiorev.2013.03.015>, PMID: 23548891
- Sherratt TN. 2002. The evolution of imperfect mimicry. *Behavioral Ecology* **13**:821–826. DOI: <https://doi.org/10.1093/beheco/13.6.821>
- Sherratt TN, Whissell E, Webster R, Kikuchi DW. 2015. Hierarchical overshadowing of stimuli and its role in mimicry evolution. *Animal Behaviour* **108**:73–79. DOI: <https://doi.org/10.1016/j.anbehav.2015.07.011>
- Sherratt TN, Holen Øistein H. 2018. When should receivers follow multiple signal components? A closer look at the “flag” model. *Behavioral Ecology* **29**:e6–e8. DOI: <https://doi.org/10.1093/beheco/ary043>

- Sherratt TN**, Peet-Paré CA. 2017. The perfection of mimicry: an information approach. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**:20160340. DOI: <https://doi.org/10.1098/rstb.2016.0340>
- Shrestha M**, Dyer AG, Boyd-Gerny S, Wong BB, Burd M. 2013. Shades of red: bird-pollinated flowers target the specific colour discrimination abilities of avian vision. *New Phytologist* **198**:301–310. DOI: <https://doi.org/10.1111/nph.12135>, PMID: 23368754
- Smith JD**, Minda JP, Washburn DA. 2004. Category learning in rhesus monkeys: a study of the shepard, Hovland, and Jenkins (1961) Tasks. *Journal of Experimental Psychology: General* **133**:398–414. DOI: <https://doi.org/10.1037/0096-3445.133.3.398>
- Smith JD**, Coutinho MVC, Couchman JJ. 2011. The learning of exclusive-or categories by monkeys (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes* **37**:20–29. DOI: <https://doi.org/10.1037/a0019497>, PMID: 20718556
- Speed MP**, Ruxton GD. 2010. Imperfect batesian mimicry and the conspicuousness costs of mimetic resemblance. *The American Naturalist* **176**:E1–E14. DOI: <https://doi.org/10.1086/652990>, PMID: 20497052
- Tuomisto H**. 2012. An updated consumer's guide to evenness and related indices. *Oikos* **121**:1203–1218. DOI: <https://doi.org/10.1111/j.1600-0706.2011.19897.x>
- Wagner AR**, Logan FA, Haberlandt K, Price T. 1968. Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology* **76**:171–180. DOI: <https://doi.org/10.1037/h0025414>, PMID: 5636557
- Wallace AR**. 1867. (Untitled). *Proceedings of the Entomological Society of London*:lxxx–lxxxi.
- Willmott KR**, Robinson Willmott JC, Elias M, Jiggins CD. 2017. Maintaining mimicry diversity: optimal warning colour patterns differ among microhabitats in amazonian clearwing butterflies. *Proceedings of the Royal Society B: Biological Sciences* **284**:20170744. DOI: <https://doi.org/10.1098/rspb.2017.0744>
- Willson MF**, Whelan CJ. 1990. The evolution of fruit color in Fleshy-Fruited plants. *The American Naturalist* **136**:790–809. DOI: <https://doi.org/10.1086/285132>
- Wilson JS**, Jahner JP, Williams KA, Forister ML. 2013. Ecological and evolutionary processes drive the origin and maintenance of imperfect mimicry. *PLOS ONE* **8**:e61610–e61617. DOI: <https://doi.org/10.1371/journal.pone.0061610>, PMID: 23593490